Computers & Education 179 (2022) 104429

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

A study of how immersion and interactivity drive VR learning

Gustav Bøg Petersen *, Giorgos Petkakis, Guido Makransky

University of Copenhagen, Department of Psychology, Copenhagen, Denmark

A R T I C L E I N F O

Keywords:
Virtual reality
Media in education
Simulations
Adult learning
Online learning

A B S T R A C T

Even though learning refers to both a process and a product, the former tends to be overlooked in educational virtual reality (VR) research. This study examines the process of learning with VR technology using the Cognitive Affective Model of Immersive Learning (CAMIL) as its framework. The CAMIL theorizes that two technological features of VR, interactivity and immersion, influence a number of cognitive and affective variables that may facilitate or hinder learning. In addition, VR studies often involve media comparisons that make it difficult to disentangle the relative effects of technological features on learning. Therefore, this study also aims to provide insights concerning the unique and combined effects of interactivity and immersion on the cognitive and affective variables specified by CAMIL. We employed a 2×2 between-subjects design ($N = 153$) and manipulated the degree of interactivity and immersion during a virtual lesson on the topic of viral diseases. Analyses of variance (ANOVAs) were used to examine the effects of interactivity and immersion on our variables of interest, and structural equation modeling (SEM) was used to assess the process of learning as predicted by the CAMIL. The results indicated that the process of learning involves situational interest and embodied learning. Main effects of interactivity and/or immersion on cognitive load, situational interest, and physical presence are also reported in addition to interaction effects between immersion and interactivity on agency and embodied learning. The findings provide evidence for the CAMIL and suggest important additions to the model. These findings can be used to provide a better understanding of the process of learning in immersive VR and guide future immersive learning research.

1. Introduction

Multimedia presentations that include both words and pictures are an integral part of education in most societies. Proponents of using multimedia for instructional purposes claim that people learn more deeply from words and pictures than from words alone (Mayer, 2014b). Virtual reality (VR) is an example of a system capable of leveraging the instructional powers of multimedia. Having been described by early researchers in the field as a white knight in the arsenal of educational technologies (Hedberg & Alexander, 1994), VR has been subjected to increasing scholarly interest in recent years for its potential as a learning aid (Radianti, Majchrzak, Fromm, & Wohlgenannt, 2020). VR can be defined as "*a mosaic of technologies that* support *the creation of synthetic, highly interactive three dimensional (3D) spatial environments that represent real or non-real situations*" (Mikropoulos & Natsis, 2011, p. 769). This study focuses on immersive VR, which is often associated with the use of head-mounted displays and contrasted with non-immersive (desktop) VR (Makransky, Terkildsen, & Mayer, 2019).

Researchers have a long history of interest in the distinguishing characteristics of VR that provide unique opportunities for

* Corresponding author. University of Copenhagen, Øster Farimagsgade 2A, 1353, København K, Denmark.
E-mail address: gbp@psy.ku.dk (G.B. Petersen).

https://doi.org/10.1016/j.compedu.2021.104429
Received 30 April 2021; Received in revised form 16 December 2021; Accepted 30 December 2021
Available online 4 January 2022
0360-1315/

educational use. Almost three decades ago, Hedberg and Alexander (1994) discussed the efficacy of VR in education. They identified increased immersion, fidelity, and active learner participation as defining characteristics. In a similar fashion, Dalgarno and Lee (2010) argued that representational fidelity and learner interaction are unique characteristics of 3-D virtual learning environments; they, however, claimed that immersion is a result of representational fidelity and interactivity and, thus, not a unique property per se. In an experimental study, Lee, Wong, and Fung (2010) found evidence for an indirect effect of VR features, measured by representational fidelity and immediacy of control, on learning outcomes. Current research identifies high levels of immersion and interactivity as defining characteristics of VR (Johnson-Glenberg, 2018; Makransky & Petersen, 2021). Immersion can be understood as the objective level of sensory fidelity provided by a VR system (Bowman & McMahan, 2007). For instance, VR provides high levels of visual immersion through head-mounted displays. Interactivity refers to the amount of freedom the user is given to control the learning experience, often through handheld controllers and a virtual body (Makransky & Petersen, 2021). While these definitions imply that immersion and interactivity are objective characteristics, it should be noted that they have also been defined as subjective concepts (e.g., Mütterlein, 2018). These two features have been used to predict increased learning outcomes as a result of VR interventions (Makransky, Petersen, & Klingenberg, 2020). One issue with VR studies, however, is that they often involve media comparisons that make it difficult to disentangle the relative effects of technological features on learning outcomes. For instance, comparing learning from VR and video as in Meyer, Omdahl, and Makransky (2019) does not permit examining the unique contributions of immersion and interactivity. As a result, the first research objective of this study is to examine the isolated effects of immersion and interactivity on factors identified by recent research to be most central to learning in VR (Makransky & Petersen, 2021). These factors are specified in the Background section.

Another issue concerns the fact that there has been a tendency in the field of educational technology to focus on the novelty factor and technological capabilities rather than human learning processes (Chandler, 2009). Not considering the learner in technology-driven learning environments may result in the hindrance of learning goals (Chandler, 2009). In recent years, several papers concerning the process of learning with VR have emerged (e.g., Lee et al., 2010; Makransky & Lilleholt, 2018; Makransky & Petersen, 2019). The present study uses the newly published Cognitive Affective Model of Immersive Learning (CAMIL) as a point of departure (Makransky & Petersen, 2021). The CAMIL describes the process of learning with immersive VR and builds on prior VR learning process research as well as established psychological theories such as self-determination theory (Deci & Ryan, 2015), cognitive load theory (Sweller, 2011), the cognitive theory of multimedia learning (Mayer, 2014a), and embodied cognition (Wilson, 2002). Importantly, the CAMIL was proposed on the basis of theoretical arguments but has not been empirically tested as yet.

Although the CAMIL is not technology-specific, and applies to immersive learning technologies in general, it builds on a recent surge of media comparison studies involving VR. When it comes to learning, the CAMIL states that some of the advantages and disadvantages of VR originate in its two defining features: interactivity and immersion. Both of these features influence a range of affective and cognitive factors (depicted in Fig. 1 and described in detail below) that may facilitate or hinder learning (Makransky & Petersen, 2021). These relations are proposed as specific, testable paths. This leads to the second research objective of this study: To empirically verify the CAMIL.

To sum up, whereas the first research objective deals with the isolated effects of immersion and interactivity, the second research objective is concerned with verifying a theoretical model implicating immersion and interactivity in a broader structural model of

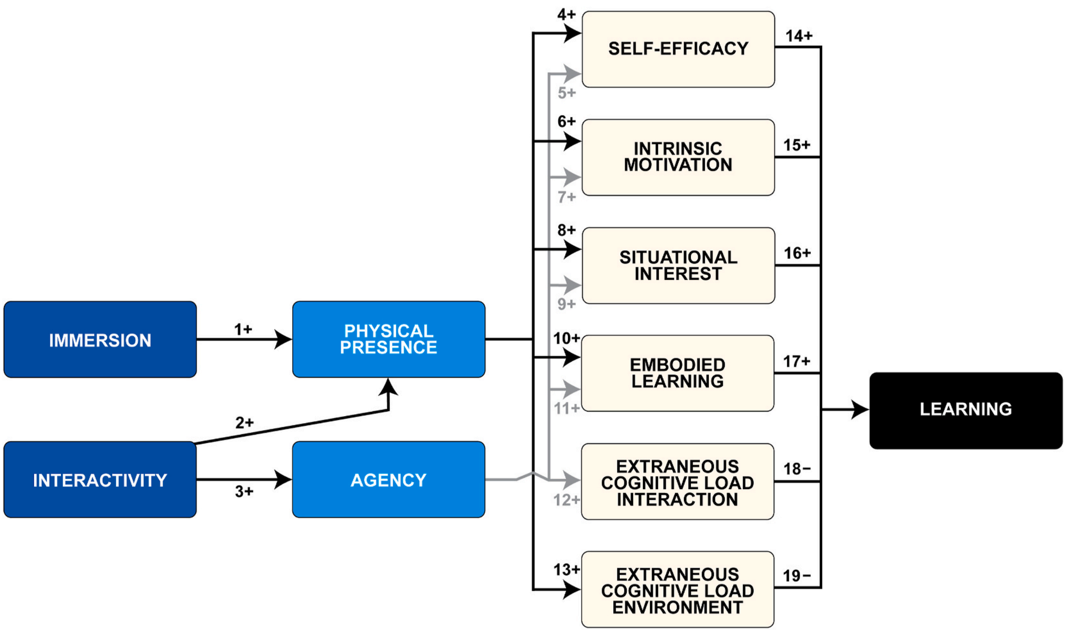

Fig. 1. The hypothesized model, based on the CAMIL.

learning.

1.1. Current study

This paper describes a media experiment featuring a virtual lesson on the topic of viral diseases in which the degree of interactivity and immersion was manipulated. The first objective was to investigate potential main effects of interactivity and immersion on factors identified in CAMIL to be most central to learning in VR. This was achieved by administering the virtual lesson via four different types of media of varying combinations of immersion and interactivity: video (low immersion/low interactivity), PC (low immersion/high interactivity), VR-video (high immersion/low interactivity), and VR (high immersion/high interactivity). The fact that we used head-mounted displays for manipulating immersion meant that it was the level of visual immersion that was influenced (Bowman & McMahan, 2007). The second objective was to empirically test the paths outlined in the CAMIL concerning the process of learning in immersive environments. This was done via structural equation modeling (SEM) on the measured constructs. The latter objective was motivated by a desire to empirically chart the process of learning in technology-driven learning environments as outlined by recent theoretical developments in the field, thereby providing up-to-date knowledge to inform instructional design and implementation decisions. In the following, we describe the background to the study along with our hypotheses and predictions.

2. Background

The CAMIL is a research-based theoretical model, which describes the process of learning in immersive environments with VR as an example. One of the reasons the model was developed, was to answer a need voiced by scholars in the field of educational VR regarding a lack of theories to guide research and development (e.g., Jensen & Konradsen, 2018; Radianti et al., 2020). In the following, an outline of the model and its assumptions is provided (please see Makransky and Petersen (2021) for an extensive description). This is accompanied by the specific predictions made in present study based on the model.

The CAMIL builds on several research papers reporting on the process of learning in VR; this section summarizes the most central ones. Based on a theoretical framework proposed by Salzman, Dede, Loftin, and Chen (1999), Lee et al. (2010) used a SEM approach to investigate the effects of desktop VR on learning outcomes. They found that the path from VR features to learning outcomes was mediated by non-cognitive outcomes such as motivation, presence, and usability. Makransky and Lilleholt (2018) built on those results, focusing on the role of immersion when learning in VR. Using the control value theory of achievement emotions (Pekrun, 2006), Makransky and Lilleholt uncovered an affective and a cognitive path that led to perceived learning in VR. These cognitive and affective paths were further investigated by Makransky and Petersen (2019). In their desktop VR study, they used SEM to build pathways from VR features to knowledge, motivation, and self-efficacy, as mediated by affective and cognitive variables. Their results confirmed that both affective and cognitive pathways lead to learning outcomes when those are measured through pre-to-post-test scores.

Theoretically, CAMIL can be placed within a body of literature that integrates cognitive and affective factors in models of learning from instructional media. Notable existing theoretical frameworks in this area include: Cognitive-Affective Theory of Learning with Media (CATLM; Moreno & Mayer, 2007) and Integrated Model of Multimedia Interactivity (INTERACT; Domagk, Schwartz, & Plass, 2010). We briefly summarize the former, as it is can be regarded as an underlying theory with regard to CAMIL. CATLM is a theory of how people learn from interactive multimodal environments such as VR. Being based on cognitive and motivational research, it places importance on factors such as cognitive load, motivation and affect, as well as metacognitive – self-regulative – factors during learning (Moreno & Mayer, 2007). Of practical use, CATLM proposes several instructional design principles that can improve learning from interactive multimodal environments (Moreno & Mayer, 2007). Since CATLM was published, a plethora of VR studies have surfaced, and more knowledge on the defining characteristics of VR has emerged. Based on a review of these studies, CAMIL provides an up-to-date framework consisting of the cognitive and affective factors identified as being most relevant to learning in VR in present times, such as interest and cognitive load (Makransky & Petersen, 2021). In this way, CAMIL can be said to extend the central ideas contained in CATLM. However, since many of the identified factors are made up of both affective and cognitive components, the general label, *affective and cognitive factors*, is used to describe them (Makransky & Petersen, 2021). One example is situational interest, which contains both focused attention (cognitive) and positive emotions (affective; Knogler, Harackiewicz, Gegenfurtner, & Lewalter, 2015).

2.1. The CAMIL's underlying conception of learning

It can be argued that the CAMIL adopts a constructivist view of learning by emphasizing the central role of immersion and interactivity during VR instruction. While different types of constructivist paradigms exist, what they have in common is an emphasis on active construction of knowledge as opposed to passive copying (Moshman, 1982). For instance, immersion and interactivity enable the use of endogenous constructivist methods (Moshman, 1982) where learners are immersed in realistic problem-solving scenarios that prime transformations of existing schemas through possibilities for interaction with the environment. As follows, the affective and cognitive variables that the CAMIL places as central to the process of learning with VR (e.g., self-efficacy and situational interest) should be understood in relation to the given (constructivist) learning content. To give an example, a full-blown interactive virtual laboratory would likely promote situational interest in the laboratory work, thereby influencing knowledge construction via promoting attention and engagement. It is worth noting that the aforementioned CATLM also opts for a knowledge construction view of learning, where the learner selects, organizes, and integrates new material with existing knowledge (Moreno & Mayer, 2007). On a related note, Johnson-Glenberg, Bartolomea, and Kalina (2021) understand learning in VR from the perspective of constructionism – a learning theory, which shares constructivism's view of learning but focuses especially on the construction of public entities (Papert & Harel, 1991).

As Fig. 1 shows, the variables included in the CAMIL can be partitioned into different sections. Starting from the left, we review each of these sections in the following. Fig. 1 also illustrates proposed paths between the variables; these are labeled with numbers (corresponding to specific predictions) and posited directions of effect (positive or negative). These paths function as predictions in the present study and are highlighted below.

2.2. *VR features*

The leftmost section contains VR features and includes visual immersion and interactivity (these have previously been defined). Please note that the original version of the CAMIL includes immersion, control factors, and representational fidelity as VR features. In this study, interactivity should be viewed as a control factor. Furthermore, we chose to keep representational fidelity constant to limit the amount of conditions necessary.

2.3. *Presence and agency*

The next part of the model contains presence and agency: the main psychological results of experiencing the prior technological features (Johnson-Glenberg, 2018). In this study, we focus on physical presence, which is defined as "*a psychological state in which virtual (para-authentic or artificial) physical objects are experienced as actual physical objects*" (Lee, 2004, p. 44). Much research on presence during multimedia instruction have compared PC and VR (e.g., Buttussi & Chittaro, 2018; Makransky & Lilleholt, 2018) or video and VR (e.g., Yeo et al., 2020) and found VR to induce the highest levels of presence.

Agency refers to the experience of controlling one's own actions (Haggard & Chambon, 2012). The variable is therefore closely connected to the concept of autonomy, which plays a key role in self-determination theory (Deci & Ryan, 2015). According to Kilteni, Groten, and Slater (2012), agency is effortlessly provided in VR when the actions of the participant are mapped to a virtual body in real-time (or near real-time). Piccione, Collett, and De Foe (2019) compared active and passive VR skills training simulations by having participants either actively enact a procedure or watch a 360° video of someone else doing it. The authors reported significantly higher agency scores in the active condition.

According to the CAMIL, presence is influenced by immersion and control factors (i.e., interactivity), and agency is influenced by control factors (i.e., interactivity) exclusively. Based on this, we expect that physical presence will be predicted by visual immersion (Prediction 1) and interactivity (Prediction 2), and that agency will be predicted by interactivity (Prediction 3).

2.4. *Affective and cognitive variables*

The subsequent part of the model contains a range of affective and cognitive variables: self-efficacy, intrinsic motivation, situational interest, embodied learning, extraneous cognitive load: interaction, and extraneous cognitive load: environment.

Self-efficacy refers to people's beliefs in their capabilities to produce particular attainments (Bandura, 2006). Empirical studies indicate that VR can have beneficial effects on learners' efficacy beliefs in comparison with video presentation (Meyer et al., 2019) or textbook presentation (Makransky, Borre-Gude, & Mayer, 2019).

Intrinsic motivation occurs when one engages in an activity because of its inherent satisfactions as opposed to doing it for external reasons (Ryan & Deci, 2000). Randomized, controlled research comparing VR to other presentation formats indicate that VR leads to higher levels of intrinsic motivation (e.g., Makransky & Lilleholt, 2018; Taranilla, Cózar-Gutiérrez, González-Calero, & Cirugeda, 2019).

Situational interest describes the focused attention and affective reaction that is triggered in the moment by particular environmental stimuli (Hidi & Renninger, 2006). We focus on the initial phase of situational interest, which is referred to as triggered situational interest, or the 'catch' component of situational interest, in the literature (Knogler et al., 2015). Randomized, controlled research comparing VR to video presentation indicates that VR leads to the largest increases in interest (Makransky, Petersen, & Klingenberg, 2020).

Embodied learning refers to the notion that we come to know ourselves and the world around us "as a 'lived body' subject that senses and does the sensing in a meaningful way" (Stolz, 2015, p. 483). VR enables embodied learning by virtue of its ability to render modifiable virtual environments where users' actions are mapped to a virtual body. A useful way of thinking about embodied learning is presented in Lindgren and Johnson-Glenberg (2013, p. 446), who posit that "*increasing an individual's repertoire of conceptually grounded physical movement will provide fertile areas from which new knowledge structures can be developed.*"

Extraneous cognitive load can be defined as cognitive load imposed by the way information is presented (as opposed to load stemming from the intrinsic nature of the information; Sweller, 2011). Based on recent research on extraneous cognitive load in virtual environments, the focus here is on extraneous cognitive load related to the interaction technique as well as extraneous cognitive load stemming from the complexity of the environment (Andersen & Makransky, 2021). Extraneous cognitive load interaction refers to the load imposed by the method with which one engages with the environment such as how you pick up artefacts or advance the lesson (Andersen & Makransky, 2021). Extraneous cognitive load environment refers to the load imposed by elements in the virtual environment such as 3D objects or virtual people (Andersen & Makransky, 2021). Research suggests that VR induces higher levels of extraneous cognitive load compared to other presentation formats (Makransky, Terkildsen, & Mayer, 2019; Parong & Mayer, 2021).

The CAMIL posits that presence and agency influence the abovementioned affective and cognitive factors. Since a distinction between extraneous cognitive load related to interaction and environment is made in this study, it is assumed that physical presence especially influences extraneous cognitive load environment, and that agency especially influences extraneous cognitive load interaction. It is therefore expected that self-efficacy will be predicted by physical presence (Prediction 4) and agency (Prediction 5); that intrinsic motivation will be predicted by physical presence (Prediction 6) and agency (Prediction 7); that situational interest will be predicted by physical presence (Prediction 8) and agency (Prediction 9); and that embodied learning will be predicted by physical presence (Prediction 10) and agency (Prediction 11). Finally, we expect that extraneous cognitive load interaction will be predicted by agency (Prediction 12), and that extraneous cognitive load environment will be predicted by physical presence (Prediction 13). The different paths for extraneous cognitive load interaction and environment are based on ideas contained in CAMIL (Makransky & Petersen, 2021) and research on the importance of separating extraneous cognitive load into sub-dimensions when dealing with virtual learning environments (Andersen & Makransky, 2021).

2.5. Learning

The rightmost part of the model contains learning; here defined as positive change in long-term memory (Sweller, 2011). In this study, learning is measured as pre-to post-test changes in number of correct responses on a declarative knowledge test. A recently published meta-analysis found that VR is more effective than non-immersive approaches in terms of learning (Wu, Yu, & Gu, 2020). Furthermore, the literature is quite consistent in demonstrating an interaction effect between media and methods on learning, with VR benefitting the most from the implementation of various instructional methods; this has been demonstrated with the pre-training principle (Meyer et al., 2019), the learning strategy of teaching (Klingenberg et al., 2020), and enactment (Makransky, Andreasen, Baceviciute, & Mayer, 2020).

According to the CAMIL, the abovementioned affective and cognitive factors influence knowledge acquisition positively (apart from extraneous cognitive load, which has a supposed negative influence). Based on this, it is expected that learning will be predicted by self-efficacy (Prediction 14), intrinsic motivation (Prediction 15), situational interest (Prediction 16), embodied learning (Prediction 17), extraneous cognitive load interaction (Prediction 18), and extraneous cognitive load environment (Prediction 19).

2.6. Hypotheses regarding immersion and interactivity

As mentioned previously, current research indicates that a unique combination of immersion and interactivity is what separates VR from other types of educational multimedia (Makransky & Petersen, 2021). Although many of the media comparison studies cited above have not isolated immersion and interactivity, this study's design makes it possible to do so. Based on the research listed above concerning the variables of interest, which indicates a superior effect of VR when compared to other presentation formats, it is predicted that immersion and interactivity will have a positive effect on all nine variables. Consequently, a total of nine hypotheses split in *a* (signifying immersion) and *b* (signifying interactivity) are proposed: Immersion and interactivity will have a positive effect on agency (Hypothesis 1a; 1b), physical presence (Hypothesis 2a; 2b), intrinsic motivation (Hypothesis 3a; 3b), self-efficacy (Hypothesis 4a; 4b), extraneous cognitive load interaction (Hypothesis 5a; 5b), extraneous cognitive load environment (Hypothesis 6a; 6b), situational interest (Hypothesis 7a; 7b), embodied learning (Hypothesis 8a; 8b), and learning (Hypothesis 9a; 9b). Furthermore, an open research question concerns whether there are any interaction effects between immersion and interactivity on the variables of interest. Hypotheses 1 through 9 are displayed in Table 1.

3. Material and methods

3.1. Sample

A total of 185 psychology students from a large European university participated in the experiment as part of a mandatory undergraduate course in educational psychology. Due to the COVID-19 situation, 32 students were not able to participate physically in the experiment, and instead participated from their own homes. However, as participants at home were not subjected to the same experimental control as those who participated physically, their data was omitted, and the analyses presented are therefore conducted on 153 participants. The final sample mostly consisted of females (122 females, 31 males, and zero non-binary). The vast majority were between 18 and 29 years old (139); the rest were 30–39 (5) or 40–49 (9). Most indicated that they spoke English well (66) and very

Table 1
Overview of hypotheses 1 through 9.

Hypothesis 1	There will be a positive effect of immersion (1a) and interactivity (1 b) on agency
Hypothesis 2	There will be a positive effect of immersion (2a) and interactivity (2 b) on physical presence
Hypothesis 3	There will be a positive effect of immersion (3a) and interactivity (3 b) on intrinsic motivation
Hypothesis 4	There will be a positive effect of immersion (4a) and interactivity (4 b) on self-efficacy
Hypothesis 5	There will be a positive effect of immersion (5a) and interactivity (5 b) on extraneous cognitive load interaction
Hypothesis 6	There will be a positive effect of immersion (6a) and interactivity (6 b) on extraneous cognitive load environment
Hypothesis 7	There will be a positive effect of immersion (7a) and interactivity (7 b) on situational interest
Hypothesis 8	There will be a positive effect of immersion (8a) and interactivity (8 b) on embodied learning
Hypothesis 9	There will be a positive effect of immersion (9a) and interactivity (9 b) on learning

well (78); there were a few native speakers (8) and only one participant did not speak English well. Participants in the two VR conditions were relatively inexperienced VR users: most had used VR zero times before (43), some had used it once (23), few had used it from two to ten times before (10), and two had used it more than ten times. With a power level of 0.80, a sensitivity analysis conducted using the *G*Power* software (Faul, Erdfelder, Lang, & Buchner, 2007) suggested that the study design amounted to minimum detectable effects of $f = 0.23$ ($\eta_p^2 = 0.05$) for main and interaction effects. This corresponds to a medium effect (Richardson, 2011). Sensitivity analyses are used to provide information about the effect size a study is able to detect with certain power given its sample size and the chosen alpha level (Faul et al., 2007). A power of .80 was chosen as it is the generally recommended standard in psychological research (Cohen, 1992).

3.2. Procedures

An overview of the study procedure is given in Fig. 2. A total of nine different seminar groups of approximately 20 students per group participated. The procedure was identical for all nine seminar groups. Prior to randomization, the participants read and signed a consent form, which gave a general description of the experiment and stated their rights as research participants. Participants were then randomly allocated to one of four different conditions: VR (39), VR-video (39), PC (40), or video (35). The different media conditions constituted the experimental manipulation of interactivity and immersion (see Table 2). Following randomization where each participant obtained a unique ID, they responded to a pre-test, which assessed prior knowledge and demographic characteristics. The students were then escorted to one of four different locations (one for each condition) on the university campus. This made students blind to the fact that there were different conditions. Once students were at the designated location, they were given a virtual lesson on the topic of viral diseases, making up the core learning material. Participants in each condition viewed the virtual lesson via their respective medium. Afterward, a post-test assessed their knowledge again and collected subjective measures. Participants were not allowed to take notes or interact with each other during the experiment.

3.3. Practical considerations

This laboratory study was conducted during the COVID-19 pandemic. Even though the infection rate was under control at the time of experimentation, a number of safety measures had to be taken. This section deals with our insights from conducting laboratory experiments amidst a pandemic, with the hope that other researchers can learn from them.

Due to the low infection rate at the time of experimentation, all students were allowed to attend classes in small, predefined groups of approximately 20 students. The experiment was therefore conducted with one group of students at a time to avoid mixing students from different groups. To ensure adequate distance between students, spacious rooms were employed for each condition. An option to participate from home was added as not every student was comfortable with physical attendance (these students were not included in the final analyses). After each participant completed the experimental procedure, all equipment they had used was disinfected using disinfecting wipes. In addition, VR headsets were decontaminated using a Cleanbox CX1. There were no COVID-19 cases reported as a consequence of this study.

We note that the use of learning material on viral diseases during the COVID-19 pandemic could have resulted in larger engagement during learning, as the topic was personally relevant to the participants.

3.4. Materials

3.4.1. Virtual learning material

A virtual museum exhibition about viruses was developed in Unity 2020 and used as learning material. The lesson progressed as a tour through an exhibition hall, starting with an introduction to general virology, which was followed by presentations of three viral diseases: measles, zika virus disease, and COVID-19. A narration recorded by an American female voice actor accompanied the exhibition. The manuscript for the narration was based on existing learning material about epidemics and pandemics targeted a 6th grade reading level, as well as other relevant information sources (e.g., the World Health Organization). The degree of difficulty of the learning content was targeted adults. Virtual screens were positioned around the exhibition hall to supplement the narrations with appropriate visuals during each presentation (e.g., a baby suffering from microcephaly was shown during learning about zika virus

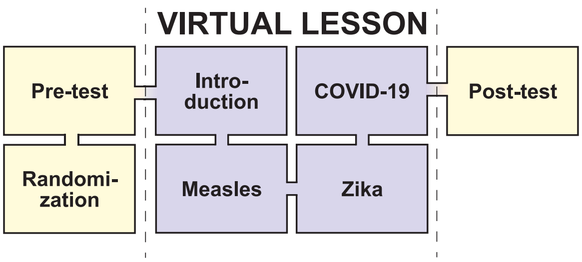

Fig. 2. Overview of study procedure.

Table 2
Level of interactivity and immersion for each condition.

	Interactivity	Immersion
Video	Low	Low
PC	High	Low
VR	High	High
VR-video	Low	High

disease). In addition to ambient museum background sounds, the environment was equipped with animated 3D models related to the topic of viruses, most of which were found on the Unity Asset store (see Fig. 3). Two different versions of the museum exhibition were created. In the first version, users were able to teleport around the museum and actively initiate the exhibition presentations (that is, a high level of interactivity was provided). In the second version, the user watched a pre-recorded video of an optimal tour of the exhibition (that is, a low level of interactivity was provided). This type of interactivity resembles Moreno and Mayer's (2007) concept of interactivity by controlling, whereby the learner controls the pace and/or order of a learning episode. The museum format was chosen due to its natural connection with providing declarative knowledge (the object of study), and it was designed with inspiration from real-life museum display technology. It was not possible to pick up artefacts in the interactive version: Learners could walk around and instigate presentations (as in many real-life museums). The topic of viral diseases was chosen due to its relevance at the time of study.

The video recording of an optimal run through the application was 10 minutes long. In other words, both the video group and the VR-video group were exposed to the learning material for approximately the same time. We do not have data for the VR and PC groups, but they spent approximately ten to 15 min in the virtual environment due to being able to explore the environment freely.

3.4.2. Devices

A total of four different presentation formats with different combinations of interactivity and immersion (see Table 2) were used to display the virtual learning material. The high immersion conditions, VR and VR-video, used head-mounted displays. For the VR condition (high immersion/high interactivity), Oculus Quest headsets with controllers were employed. For the VR-video condition (high immersion/low interactivity), Lenovo Mirage Solo headsets were used. For the PC condition (low immersion/high interactivity), Dell Latitude 7490 laptops with a computer mouse were used. Finally, in the video condition (low immersion/low interactivity), a canvas and projector were utilized along with speakers. Participants used headphones in the PC, VR and VR-video conditions.

3.4.3. Pre-test

The pre-test collected demographic information, and assessed participants' prior levels of intrinsic motivation, self-efficacy, and knowledge regarding viral diseases. The intrinsic motivation scale consisted of five items adapted from Makransky and Petersen (2019; e.g., It's fun to perform activities related to the topic of viral diseases). The self-efficacy scale consisted of four items adapted from Meyer et al. (2019) and Pintrich, Smith, Garcia, and McKeachie (1991; e.g., I'm confident I could do an excellent job on assignments and tests about viral diseases). The knowledge test was based on the narration used during the virtual lesson and consisted of nine multiple choice questions with four possible answers (e.g., What term is used to describe the birth defect caused by Zika virus where a baby's head is smaller than expected? a) Microcephaly, b) Anotia, c) Microtia, d) Omphalocele). All knowledge items assessed declarative memory (Ten Berge & Van Hezewijk, 1999) and were designed with experts in educational psychology as well as

Fig. 3. Screenshot of the virtual museum exhibition.

psychometrics.

3.4.4. Post-test

The post-test assessed participants' posttreatment levels of intrinsic motivation, self-efficacy, and knowledge (learning was defined as pre-to post-test changes in knowledge), using the same items as above. In addition, it measured agency, physical presence, extraneous cognitive load interaction, extraneous cognitive load environment, situational interest, and embodied learning. All items are shown in Appendix A. The agency scale consisted of three items adapted from Polito, Barnier, and Woody (2013; e.g., During the lesson, my experiences and actions were under my control). The physical presence scale consisted of five items from Makransky, Lilleholt, and Aaby (2017; e.g., The virtual environment seemed real to me). The extraneous cognitive load interaction scale consisted of four items from Andersen and Makransky (2021; e.g., The interaction technique used in the simulation was very unclear). The extraneous cognitive load environment scale consisted of four items from Andersen and Makransky (2021; e.g., The elements in the virtual environment made the learning very unclear). The situational interest scale consisted of six items from Knogler et al. (2015; e.g., Did the lesson capture your attention?). The embodied learning scale consisted of three items based on Skulmowski and Rey (2018; e.g., During the lesson, my movements and bodily activities were an integrated part of learning).

3.4.5. Statistical analyses

All statistical analyses were conducted in R version 4.0.2. Specifically, the 'lavaan' package was used for SEM. The full information maximum likelihood approach was used to deal with missing values (this was true for the self-efficacy scores of 21 participants due to a measurement error). A p-value less than 0.05 was considered statistically significant.

4. Results

4.1. Quality of measures

For all measures except learning, Cronbach's alpha was used as an indicator of scale quality. As recommended by DeVellis (2012), alpha levels between 0.80 and 0.90 were considered indicative of very good scale reliability. The intrinsic motivation scale had a Cronbach's α of 0.88 at pre-test and 0.90 at post-test. The self-efficacy scale had a Cronbach's α of 0.86 at pre-test and 0.84 at post-test. The agency scale had a Cronbach's α of 0.90. The physical presence scale had a Cronbach's α of 0.89. Regarding the extraneous cognitive load interaction scale, the results indicated that the deletion of one item would raise the internal consistency ('ECL_I_4' in Appendix A), so it was omitted, and the final scale had a Cronbach's α of 0.82. Looking at the extraneous cognitive load environment scale, one item was shown during confirmatory factor analysis to have a low standardized factor loading (0.58; 'ECL_E_4' in Appendix A), so it was omitted, and the final scale had a Cronbach's α of 0.80. Regarding the situational interest scale, the results indicated that the deletion of one item would raise the internal consistency ('SI_3' in Appendix A), so it was omitted, and the final scale had a Cronbach's α of 0.92. The embodied learning scale had a Cronbach's α of 0.89. It should be noted that modifications to the scales strictly served to improve their psychometric properties.

The knowledge test was edumetric, meaning that it was designed to measure the knowledge gain of individuals from pre-to post-test (i.e., learning). It was therefore evaluated using the edumetric principle of average item increase from pre-to post-test (Carver, 1974). The results indicated that all items had average pre-to post-test increases over 0.2 on a possible range of −1 to 1.

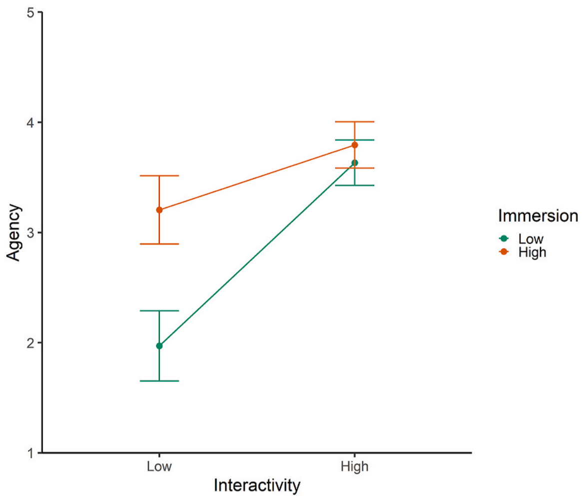

Fig. 4. Plot of the effects of interactivity and immersion on agency with 95% confidence intervals.

4.2. Group differences at pre-test

χ^2 test and Fisher's exact test were conducted to assess the equivalence of conditions on demographic variables. These were all non-significant: age ($\rho = 0.21$), gender ($\rho = 0.66$), and English proficiency ($\rho = 0.96$). Therefore, we conclude that the assigned groups did not significantly differ based on demography.

Visual inspection of histograms indicated normality of data for intrinsic motivation and self-efficacy at pre-test. One-way ANOVAs were therefore performed to verify equivalence of conditions on these measures. The pre-test level of intrinsic motivation was $M = 3.34$ ($SD = 0.66$) for the video group, $M = 3.37$ ($SD = 0.77$) for the PC group, $M = 3.36$ ($SD = 0.59$) for the VR-video group, and $M = 3.35$ ($SD = 0.59$) for the VR group. The pre-test level of self-efficacy was $M = 3.08$ ($SD = 0.90$) for the video group, $M = 3.01$ ($SD = 0.83$) for the PC group, $M = 2.86$ ($SD = 1.07$) for the VR-video group, and $M = 3.11$ ($SD = 0.93$) for the VR group. Both one-way ANOVAs were non-significant: intrinsic motivation ($p = .82$) and self-efficacy ($p = .52$). Hence, post-test measures of intrinsic motivation and self-efficacy were subjected to analysis.

4.3. Two-way ANOVAs of interactivity and immersion

Visual inspection of Q-Q plots of model residuals was used to assess normality before two-way ANOVAs (type II or III) were performed. Since the CAMIL pinpoints presence and agency as the main psychological catalysts of learning with VR, only graphs for agency and physical presence are shown in Fig. 4 and Fig. 5.

4.3.1 Hypotheses 1 through 9: Effect of interactivity and immersion on agency, physical presence, intrinsic motivation, self-efficacy, extraneous cognitive load interaction, extraneous cognitive load environment, situational interest, embodied learning, and learning.

Table 3 displays an overview of the results.

As shown in Fig. 4, there was a statistically significant interaction between the effects of immersion and interactivity on agency, $F(1, 149) = 17.00$, $p < .001$, $\eta_p^2 = 0.10$. A further examination of the interaction indicated that the effect of interactivity on agency is larger when immersion is low ($d = 2.08$) compared to high ($d = 0.71$). Furthermore, the main effects of interactivity and immersion on agency were both significant, $F(1, 149) = [79.95; 43.53]$, $p < .001$, $\eta_p^2 = [0.35; 0.23]$, respectively. Thus, hypothesis 1a, that immersion would have a positive effect on agency, and hypothesis 1b, that interactivity would have a positive effect on agency, were confirmed.

As shown in Fig. 5, the main effects of interactivity and immersion on physical presence were significant, $F(1, 149) = [18.88; 142.66]$, $p < .001$, $\eta_p^2 = [0.11; 0.49]$, respectively. Thus, hypothesis 2a, that immersion would have a positive effect on physical presence, and hypothesis 2b, that interactivity would have a positive effect on physical presence, were confirmed.

The analysis did not reveal any significant effects on intrinsic motivation although the main effect of immersion was close ($p = .056$). Thus, hypothesis 3, that immersion and interactivity would have a positive effect on intrinsic motivation, was not confirmed. Furthermore, the results did not reveal any significant effects on self-efficacy. Thus, hypothesis 4 was not confirmed. There were also no significant effects on extraneous cognitive load interaction although the main effect of interactivity was close ($p = .059$). Thus, hypothesis 5 was not confirmed. There was no significant effect of immersion on extraneous cognitive load environment. However, the main effect of interactivity on extraneous cognitive load environment was significant, $F(1, 149) = 5.14$, $p = .025$, $\eta_p^2 = 0.03$. Thus, hypothesis 6b was confirmed, however hypothesis 6a was not confirmed. More specifically, only the effect of interactivity was significant, and, surprisingly, it was in the opposite direction than predicted: Lower interactivity was associated with higher extraneous cognitive load environment. Furthermore, only the main effect of immersion on situational interest was significant, $F(1, 149) = 35.62$,

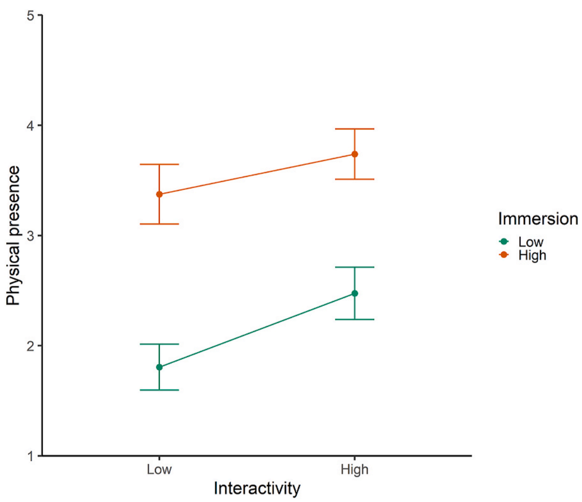

Fig. 5. Plot of the effects of interactivity and immersion on physical presence with 95% confidence intervals.

Table 3
Overview of results.

	Low im.		High im.		Main effect im.	Main effect int.	Inter-action
	Low int.	High int.	Low int.	High int.			
Agency	1.97 (0.93)	3.63 (0.64)	3.21 (0.96)	3.79 (0.65)	$p < .001$	$p < .001$	$p < .001$
Physical presence	1.81 (0.61)	2.48 (0.74)	3.37 (0.84)	3.74 (0.70)	$p < .001$	$p < .001$	$p = .199$
Intrinsic motivation	3.57 (0.71)	3.61 (0.84)	3.82 (0.65)	3.81 (0.62)	$p = .056$	$p = .923$	$p = .849$
Self-efficacy	3.74 (0.62)	3.67 (0.83)	3.41 (0.75)	3.69 (0.72)	$p = .254$	$p = .385$	$p = .171$
Extraneous cognitive load interaction	2.61 (0.85)	2.34 (0.87)	2.73 (0.90)	2.46 (0.85)	$p = .400$	$p = .059$	$p = .992$
Extraneous cognitive load environment	2.39 (0.73)	2.12 (0.77)	2.64 (0.92)	2.30 (0.92)	$p = .116$	$p = .025$	$p = .803$
Situational interest	3.17 (0.88)	3.42 (0.89)	4.06 (0.70)	4.16 (0.89)	$p < .001$	$p = .191$	$p = .592$
Embodied learning	1.30 (0.44)	2.38 (1.00)	2.68 (1.05)	2.73 (0.86)	$p < .001$	$p < .001$	$p < .001$
Learning	4.26 (1.72)	3.48 (1.55)	3.56 (1.89)	3.62 (1.82)	$p = .356$	$p = .210$	$p = .143$

Note. The table displays means with standard deviations in parentheses. im. = immersion, int. = interactivity. P values are shown for main effects and interaction effects, with significant effects in bold.

$p < .001$, $\eta_p^2 = 0.19$. Thus, hypothesis 7a, that immersion would have a positive effect on situational interest, was confirmed, but hypothesis 7b, that interactivity would have a positive effect on situational interest, was not confirmed. Moving to embodied learning, there was a statistically significant interaction between the effects of immersion and interactivity on embodied learning, $F(1, 149) = 12.80$, $p < .001$, $\eta_p^2 = 0.08$. A further examination of the interaction indicated that the effect of interactivity on embodied learning is larger when immersion is low ($d = 1.40$) compared to high ($d = 0.05$). Furthermore, the main effects of interactivity and immersion on embodied learning were both significant, $F(1, 149) = [27.63; 44.76]$, $p < .001$, $\eta_p^2 = [0.16; 0.23]$, respectively. Thus, hypothesis 8a, that immersion would have a positive effect on embodied learning, and 8b, that interactivity would have a positive effect on embodied learning, were confirmed. Finally, the analysis did not reveal any significant effects on learning. Thus, hypothesis 9 was not confirmed.

4.4. *Structural equation modeling (SEM)*

Fig. 1 depicts the hypothesized model, which builds on a body of theoretical and empirical literature and investigates the relations between the latent variables included in the study. Three common goodness-of-fit indicators were used to assess the fit between the model and data: Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and the Tucker Lewis Index (TLI). The model's fit was not acceptable at first (RMSEA = 0.085, CFI = 0.832, TLI = 0.816). Hence, modification indices were used to adjust the model based on data. This resulted in the addition of covariance between the following items from the same scales (not in the graphical illustration of the final model): S_2 and S_3, I_3 and I_4, SI_1 and SI_2, and P_1 and P_3 (for an overview of each item, please consult Appendix A). Additionally, paths were added between the following latent variables (shown in the final model): physical presence and extraneous cognitive load interaction, intrinsic motivation and situational interest, and extraneous cognitive load environment and extraneous cognitive load interaction. All adjustments based on modification indices were psychometrically and theoretically sound; these are addressed in more detail in the Discussion section. As a next step, non-significant paths were deleted by an iterative procedure. Each of these paths were evaluated and removed one at a time based on the greatest misfit until all the remaining paths were significant. This resulted in a simplified model containing only significant loadings, presented in Fig. 6. The final, simplified model obtained an acceptable fit (RMSEA = 0.061, CFI = 0.914, TLI = 0.906). When compared to the hypothesized model, the final model retained several of the initial paths, albeit theoretically meaningful modifications were made based on data to obtain a better fitting model (potential refinements to the original theoretical model are proposed in the Discussion). Below, the results are related to the initial predictions. The estimates reported are standardized with respect to both the observed and latent variables ('Std.

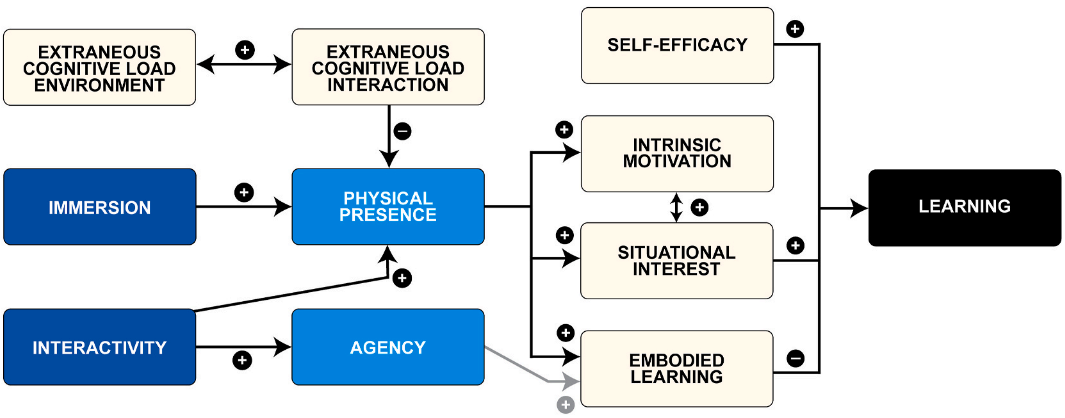

Fig. 6. The final model.

all' in *sem {lavaan}*).

4.4.1. Predictions 1 through 3: immersion and interactivity will predict physical presence, and interactivity will predict agency

Physical presence was predicted by immersion ($\beta = 0.770$, $p < .001$) and interactivity ($\beta = 0.203$, $p < .001$). Agency was predicted by interactivity ($\beta = 0.551$, p $< .001$). Thus, prediction 1, 2, and 3 were supported.

4.4.2. Predictions 4 through 13: physical presence and agency will predict self-efficacy, intrinsic motivation, situational interest, and embodied learning. Additionally, agency will predict extraneous cognitive load related to interaction, and physical presence will predict extraneous cognitive load related to the environment

Physical presence predicted intrinsic motivation ($\beta = 0.334$, $p < .001$), situational interest ($\beta = 0.692$, $p < .001$) and embodied learning ($\beta = 0.573$, $p < .001$). Agency only predicted embodied learning ($\beta = 0.214$, $p = .015$). The rest of the paths were not significant. Thus, only prediction 6, 8, 10, and 11 were supported.

4.4.3. Predictions 14 through 19: self-efficacy, intrinsic motivation, situational interest, embodied learning, and extraneous cognitive load interaction and environment will predict learning

Learning was predicted by self-efficacy ($\beta = 0.202$, $p = .019$), situational interest ($\beta = 0.207$, $p = .035$), and embodied learning ($\beta = -0.298$, $p = .002$). It should be noted that embodied learning impacted knowledge gain negatively contrary to what was expected (this finding is discussed in the Discussion section). The rest of the paths were not significant. Thus, only prediction 14 and 16 were supported. Furthermore, a relation in the opposite direction of prediction 17 concerning embodied learning was found.

4.4.4. Exploratory findings

There was covariance between extraneous cognitive load environment and interaction ($\beta = 0.905$, $p < .001$), and extraneous cognitive load interaction predicted physical presence ($\beta = -0.327$, $p < .001$). Furthermore, there was covariance between intrinsic motivation and situational interest ($\beta = 0.518$, $p < .001$).

5. Discussion

The aims of the present work were twofold: 1) to investigate the effects of VR's two defining features, interactivity and immersion, on people's learning experience and degree of learning and 2) to validate the CAMIL; a preexisting theoretical model concerning the process of learning with immersive learning technologies. Several substantial findings were made regarding these objectives. What follows is a description of the empirical contributions of present study.

5.1. Empirical contributions

Consistent with the predictions, both interactivity and immersion were shown to impact agency, physical presence, and embodied learning positively. This is consistent with Johnson-Glenberg et al.'s (2021) and Makransky and Petersen's (2021) positioning of presence and agency as "profound affordances" of VR. However, significant interaction effects between the two VR features, interactivity and immersion, on agency and embodied learning indicated that the effects of interactivity on agency and embodied learning is higher when immersion is low. That is, having a high level of interactivity is more important for achieving a sense of agency and embodied learning when the lesson has a low level of immersion. Although interactivity still influenced agency and embodied learning when lessons were highly immersive, the impact was not as strong as when the lesson was less immersive. Supposedly, being visually immersed is such a powerful experience in itself that interactivity provides less added value; under conditions of low visual immersion, however, interactivity is shown to its full advantage. In addition, it was found that only immersion influenced situational interest and only interactivity influenced extraneous cognitive load environment. Surprisingly, lower interactivity was associated with higher extraneous cognitive load stemming from the environment; perhaps watching a video of someone else experiencing a virtual museum exhibition led participants to consider the elements in the virtual environment irrelevant and cognitively taxing as personal exploration was not possible. This finding goes against Moreno and Mayer's (2007) notion that free exploration of complex multimodal environments could generate heavy cognitive load. No effects of interactivity or immersion on intrinsic motivation, self-efficacy, extraneous cognitive load interaction, or learning were found.

To recap, this study's experimental design made it possible to discern the effects of two defining features of VR learning, interactivity and immersion, on important learning experience variables. Both features positively influenced the levels of physical presence experienced by learners, while interaction effects indicated that interactivity is more important for the experience of agency and embodied learning under conditions of low immersion. Additionally, significant main effects suggest that high interactivity reduces extraneous cognitive load from the environment, and that high immersion leads to larger situational interest. The latter expands on previous research findings that VR leads to more interest compared to other instructional media (e.g., Makransky, Petersen, & Klingenberg, 2020).

SEM was applied to examine the relations between the variables investigated in present study. The CAMIL guided the structural model initially tested. Support for two indirect paths from VR features to learning was found: one involving situational interest and the other involving embodied learning. Hence, while interactivity and immersion did not directly affect learning, as mentioned above, there was evidence for indirect links. Note that self-efficacy also predicted learning but was not predicted by presence and agency as expected.

5.1.1. Situational interest path

As depicted in Fig. 6, learning was positively predicted by the degree of situational interest experienced by the participants. Situational interest, in turn, was predicted by participants' level of physical presence, which increased with higher objective levels of immersion and interactivity. Note that situational interest and intrinsic motivation covaried, and that physical presence was negatively influenced by extraneous cognitive load stemming from the interaction (which covaried with extraneous cognitive load stemming from the environment). Using a similar experimental design as the one employed in present study, Johnson-Glenberg et al. (2021) investigated paths from VR factors to learning. They manipulated immersion (they refer to this as platform) and interactivity (they refer to this as embodiment), and collected measures of knowledge, agency, presence, and engagement. What was found was that engagement, which resembles situational interest, predicted learners' level of knowledge at post-test (Johnson-Glenberg et al., 2021). Engagement, in turn, was predicted by presence and agency, which were predicted by immersion and interactivity respectively (Johnson-Glenberg et al., 2021). In other words, the situational interest path in the present study corroborates previous research.

5.1.2. Embodied learning path

As illustrated in Fig. 6, and contrary to what was expected, learning was negatively predicted by the levels of embodied learning expressed by the participants. Embodied learning, again, was predicted by both physical presence and agency, which increased as a function of higher immersion and interactivity. As in the situational interest path, note the negative influence of extraneous cognitive load interaction on physical presence.

In terms of the negative relation between self-reported embodied learning and participants' degree of objective learning, one plausible explanation is the fact that the content of the simulation was not ideal for embodied learning in the first place; the learning material was declarative, and memory of the presented information might therefore have been hampered by the experience of actively using one's body for learning at the cost of attention to the narration. A closer look at the results in Table 3 shows that the video group reported the lowest levels of embodied learning, as could be expected, while at the same time displaying slightly higher levels of learning, which explains the negative association. There is a body of literature on technology and embodied learning, which could further clarify this finding. In their paper on embodied learning through immersive technologies, Lindgren and Johnson-Glenberg (2013) contend that for effective learning to take place, actions must be structurally or analogically related to the learning content. In the present study, it was not possible to perform actions explicitly related to the learning content, which is why the experience of actively using one's body for learning might have done more harm than good.

Huang, Roscoe, Craig, and Johnson-Glenberg (2021) recently proposed a model of VR learning based on CATLM and evaluated it using SEM; their aim of integrating relevant VR learning research into CATLM bears resemblance to the ideas presented in the current manuscript. However, substantial discrepancies between Huang and colleagues' investigation and the one described here, are worth noting for future reference. Their results indicated that VR format (signifying three levels of sophistication in VR systems) had a direct effect on embodiment, which led to presence and motivation (Huang et al., 2021). Motivation, in turn, led to cognitive engagement, which predicted learning achievement (Huang et al., 2021). Furthermore, they found that cognitive load had a negative effect on cognitive engagement (Huang et al., 2021). In other words, the study by Huang and colleagues shows that there are other ways of modeling learning in VR.

5.2. Theoretical implications

This work provides the first empirical validation of the CAMIL: a recently published theoretical model concerning the process of learning with immersive technologies, which builds on a review of relevant theories and prior immersive educational research (Makransky & Petersen, 2021). Evidence for several of the assumptions made in the CAMIL is found, including the fact that physical presence and agency arise from immersive and interactive learning technologies and influence learning via the affective and cognitive factors of situational interest and embodied learning. Of note, elevated levels of embodied learning decreased knowledge acquisition, presumably because of the lack of congruency between bodily actions and learning content. While the hypothesis that self-efficacy is a predictor of learning is corroborated, there is no evidence to suggest that immersion or interactivity influence self-efficacy, either directly or indirectly. Conversely, it is found that intrinsic motivation is influenced indirectly by immersion and interactivity (via physical presence) but does not predict learning.

This study's findings also suggest refinements to the CAMIL. Specifically, cognitive load was found to predict the experience of presence, not the other way around as suggested by the theory. That is, elevated levels of extraneous cognitive load from the interaction decreased one's feelings of physical presence. It can be theorized that difficulties with the interaction technique led to breaks in presence by leading the participants' attention towards the technological setup instead of the virtual learning content (i.e., the real world became apparent; Brogni, Slater, & Steed, 2003). Furthermore, two-way ANOVAs indicated an interaction effect between immersion and interactivity on agency such that the effect of interactivity increases when immersion is low. Lastly, the data supported a link between interest and intrinsic motivation: This connection has previously been discussed in the educational psychological literature (e.g., Schraw, Flowerday, & Lehman, 2001). These are the major theoretical implications of this study and could serve as valuable additions to the CAMIL.

5.3. Limitations and suggestions for future research

The literature identifies a range of individual differences variables that could moderate the impact of technological learning interventions. These include factors such as cybersickness, spatial ability, or personality traits (Makransky & Petersen, 2021). To keep the questionnaire length as short as possible, participants' ratings on such traits were not assessed. Hence, the influence of individual differences on the process of learning with immersive technologies remains an important avenue for future research. Another important direction for future research could be to investigate the influence of implementing instructional techniques on the process of learning in VR (Mayer, 2014a). To give an example, the signaling principle (i.e., highlighting essential material) could enhance the process of learning by reducing extraneous cognitive load (Mayer, 2014a). Albus, Vogt, and Seufert (2021) recently investigated the signaling principle in VR and found that it increased recall performance. Another promising direction could be to include generative learning activities such as elaboration prompts, which have been found to improve recall of VR learning material (Vogt, Babel, Hock, Baumann, & Seufert, 2021).

It is also worth noting the limitations associated with the specific virtual learning environment chosen for this study. As mentioned previously, the main purpose of the lesson was to teach declarative knowledge: Participants were rated on how much they improved on a knowledge test related to viral diseases. While this allowed us to investigate the variables involved in the process of declarative learning, it also restricts the generalizability of the findings. For instance, it is possible that the paths leading to learning would have been different if the purpose of the lesson had been to teach a procedure such as how to operate a certain piece of equipment. Moreover, it is conceivable that presence and agency would predict one's self-efficacy for a procedure, in contrast to what was found in this study.

An often-encountered criticism of SEM, the statistical method used in this study for analyzing relations between variables, concerns its alleged aim of establishing causal relations from associations alone (Bollen & Pearl, 2013). In reality, tests of causality by means of SEM do not prove the validity of the causal assumptions; rather, they lend credibility to them (Bollen & Pearl, 2013). Furthermore, the structural model tested can be said to represent and rely upon the causal assumptions of the researcher, which, in turn, are derived from such sources as prior studies, scientific knowledge, and logical arguments (Bollen & Pearl, 2013). In this case, this study's assumptions are based on a well-founded, research-based theoretical model. Nevertheless, it remains for future research to replicate the results in order to strengthen the conclusions drawn.

It is worth mentioning that the study sample was not comprised of native English speakers despite the fact that the learning material was in English; this could potentially have impacted comprehension. In addition, applying the same knowledge test at pre- and post-test comes with potential disadvantages. A pre-test may, for instance, direct attention towards what the learner does not know during subsequent instruction, thereby enhancing performance on the post-test (Hartley, 1973). Lastly, there was an unequal gender composition of roughly 80 percent women in the sample, as is common among psychology students. There is no reason to assume that a larger amount of women in itself could have influenced the results. However, the fact that the sample was European psychology students could influence the external validity of the study as it can be argued that they constitute a special group of people.

6. Conclusion

The defining features of VR learning experiences – high degrees of immersion and interactivity – are more than just entertaining add-ons. The analyses indicated that these features both had unique and combined effects on a number of important variables relevant to VR learning. Specifically, both immersion and interactivity positively influenced the levels of physical presence experienced by learners, while interaction effects showed that interactivity is more important for the experience of agency and embodied learning under conditions of low immersion. Moreover, significant main effects suggested that high interactivity reduces extraneous cognitive load from the environment, and that high immersion leads to larger situational interest. As to the process of learning with VR, several of the predictions made on the basis of a recently published theoretical model, the CAMIL, were verified empirically (see Fig. 6). Specifically, there was evidence for two indirect paths from interactivity and immersion to learning: the first involving situational interest and the second involving embodied learning. Contrary to theoretical claims, self-reported embodied learning predicted participants' declarative memory of the learning topic negatively. It can be theorized that this negative relation stemmed from the lack of congruency between bodily actions and learning content. Hence, an important direction for future research in the field could be to examine the process of learning using simulations actively designed for embodied learning.

Credit author statement

Gustav Bøg Petersen: Formal analysis; Investigation; Writing – original draft; Writing – review & editing. Giorgos Petkakis: Software, Guido Makransky: Supervision; Writing – review & editing

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors have no conflicts of interest to declare.

Acknowledgements

We would like to thank all the students who participated in this research. Furthermore, we would like to thank Sara Klingenberg and Martin Stolpe Andersen for assisting with the experiment.

Appendix A

Construct	Item code	Items	Source
Agency	A_1	During the lesson, my experiences and actions were under my control.	Polito et al. (2013)
	A_2	During the lesson, I felt that my experiences and actions were not caused by me. (R)	
	A_3	During the lesson, my experiences and actions felt self-generated.	
Physical presence	P_1	The virtual environment seemed real to me.	Makransky et al. (2017)
	P_2	I had a sense of acting in the virtual environment, rather than operating something from outside.	
	P_3	My experience in the virtual environment seemed consistent with my experiences in the real world.	
	P_4	I had a sense of "being there" in the virtual environment.	
	P_5	I was completely captivated by the virtual world.	
Intrinsic motivation	I_1	I enjoy working with the topic of viral diseases.	Makransky and Petersen (2019)
	I_2	It's fun to perform activities related to the topic of viral diseases.	
	I_3	The topic of viral diseases is boring. (R)	
	I_4	The topic of viral diseases does not hold my attention at all. (R)	
	I_5	I would describe the topic of viral diseases as very interesting.	
Self-efficacy	S_1	I'm confident I can understand the basic concepts of viral diseases.	Meyer et al. (2019)
	S_2	I'm confident I can understand the most complex material regarding viral diseases.	Pintrich et al. (1991)
	S_3	I'm confident I could do an excellent job on assignments and tests about viral diseases.	
	S_4	I'm confident I could explain the basic causes of epidemics and pandemics to a friend.	
Extraneous cognitive load interaction	ECL_I_1	The interaction technique used in the simulation was very unclear.	Andersen and Makransky (2021)
	ECL_I_2	The interaction technique used in the simulation was, in terms of learning, very ineffective.	
	ECL_I_3	The interaction technique used in the simulation made it harder to learn.	
	ECL_I_4	The interaction technique used in the simulation was difficult to master.	
Extraneous cognitive load environment	ECL_E_1	The elements in the virtual environment made the learning very unclear.	Andersen and Makransky (2021)
	ECL_E_2	The virtual environment was, in terms of learning, very ineffective.	
	ECL_E_3	The virtual environment was full of irrelevant content.	
	ECL_E_4	It was difficult to find the relevant learning information in the virtual environment.	
Situational interest	SI_1	Did the lesson spark your curiosity?	Knogler et al. (2015)
	SI_2	Did the lesson capture your attention?	
	SI_3	Were you concentrated on the lesson?	
	SI_4	Was the lesson entertaining for you?	
	SI_5	Did you have fun during the lesson?	
	SI_6	Was the lesson exciting for you?	
Embodied learning	EL_1	During the lesson, my movements and bodily activities were an integrated part of learning.	Skulmowski and Rey (2018)
	EL_2	During the lesson, I performed bodily movements and interactions that were relevant for learning.	
	EL_3	Performing gestures/movements during the lesson helped me learn.	
Knowledge	MCQ_1	What term is used to describe the birth defect caused by zika virus where a baby's head is smaller than expected? a) Microcephaly b) Anotia c) Microtia d) Omphalocele	Manuscript
	MCQ_2	What term can be used to classify an infectious disease caused by a virus that has jumped from an animal to a human? a) Oncogenic b) Cytopathic c) Veterinary d) Zoonotic	
	MCQ_3	What percentage develop only mild or no symptoms after contracting COVID-19? a) 60 b) 70 c) 80 d) 90	
	MCQ_4	How many days are measles patients contagious after rashes appear? a) 0 b) 2 c) 4 d) 6	
	MCQ_5	When was the measles vaccination introduced? a) 1961 b) 1962 c) 1963 d) 1964	
	MCQ_6	Which genus is the virus that causes measles a member of? a) Mamastrovirus b) Machlomovirus	

(continued on next page)

(*continued*)

Construct	Item code	Items	Source
		c) Morbillivirus	
		d) Maculavirus	
	MCQ_7	How many deaths has measles vaccination prevented? a) 13 million	
		b) 23 million	
		c) 33 million	
		d) 43 million	
	MCQ_8	How many died in 2016 from measles? a) 60.000	
		b) 70.000	
		c) 80.000	
		d) 90.000	
	MCQ_9	What percentage require treatment in an intensive care unit after contracting COVID-19? a) 2	
		b) 3	
		c) 4	
		d) 5	

Note. (R) reverse coded; all items used in this study, except knowledge items, were measured on a five-point Likert scale.

References

Albus, P., Vogt, A., & Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. *Computers & Education, 166*. https://doi.org/10.1016/j.compedu.2021.104154

Andersen, M. S., & Makransky, G. (2021). The validation and further development of a multidimensional cognitive load scale for virtual environments. *Journal of Computer Assisted Learning, 37*(1), 183–196. https://doi.org/10.1111/jcal.12478

Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. Urdan, & F. Pajares (Eds.), *Self-Efficacy Beliefs of Adolescents* (pp. 307–337). Information Age Publishing (NO DOI).

Bollen, K. A., & Pearl, J. (2013). Eight myths about causality and structural equation models. In S. L. Morgan (Ed.), *Handbook of causal analysis for social research* (pp. 301–328). Springer Science. https://doi.org/10.1007/978-94-007-6094-3 15.

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer, 40*(7), 36–43. https://doi.org/10.1109/MC.2007.257

Brogni, A., Slater, M., & Steed, A. (2003). More breaks less presence. In *The 6th annual international workshop on presence* (pp. 1–4). International Society for Presence Research.

Buttussi, F., & Chittaro, L. (2018). 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. https://doi.org/10.1109/TVCG.2017.2653117

Carver, R. P. (1974). Two dimensions of tests: Psychometric and edumetric. *American Psychologist, 29*(7), 512–518. https://doi.org/10.1037/h0036782

Chandler, P. (2009). Dynamic visualisations and hypermedia: Beyond the "Wow" factor. *Computers in Human Behavior, 25*(2), 389–392. https://doi.org/10.1016/j.chb.2008.12.018

Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin, 112*(1), 155–159.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology, 41*(1), 10–32. https://doi.org/10.1111/j.1467-8535.2009.01038.x

Deci, E. L., & Ryan, R. M. (2015). Self-determination theory. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 486–491). Elsevier.

DeVellis, R. F. (2012). *Scale development: Theory and applications*. SAGE Publications.

Domagk, S., Schwartz, R. N., & Plass, J. L. (2010). Interactivity in multimedia learning: An integrated model. *Computers in Human Behavior, 26*(5), 1024–1033. https://doi.org/10.1016/j.chb.2010.03.003

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. https://doi.org/10.3758/BF03193146

Haggard, P., & Chambon, V. (2012). Sense of agency. *Current Biology, 22*(10). https://doi.org/10.1016/j.cub.2012.02.040

Hartley, J. (1973). The effect of pre-testing on post-test performance. *Instructional Science, 2*(2), 193–214. https://doi.org/10.1007/BF00139871

Hedberg, J., & Alexander, S. (1994). Virtual reality in education: Defining researchable issues. *Educational Media International, 31*(4), 214–220. https://doi.org/10.1080/0952398940310402

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest fevelopment. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Huang, W., Roscoe, R. D., Craig, S. D., & Johnson-Glenberg, M. C. (2021). Extending the cognitive-affective theory of learning with media in virtual reality learning: A structural equation modeling approach. *Journal of Educational Computing Research*, 1–36. https://doi.org/10.1177/07356331211053630

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies, 23*(4), 1515–1529. https://doi.org/10.1007/s10639-017-9676-0

Johnson-Glenberg, M. (2018). Immersive VR and education: Embodied design principles that include gesture and hand controls. *Frontiers in Robotics and AI, 5*(81), 1–19. https://doi.org/10.3389/frobt.2018.00081

Johnson-Glenberg, M. C., Bartolomea, H., & Kalina, E. (2021). Platform is not destiny: Embodied learning effects comparing 2D desktop to 3D virtual reality STEM experiences. *Journal of Computer Assisted Learning, 37*(5), 1263–1284. https://doi.org/10.1111/jcal.12567

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments, 21*(4), 373–387. https://doi.org/10.1162/PRES_a_00124

Klingenberg, S., Jørgensen, M., Dandanell, G., Skriver, K., Mottelson, A., & Makransky, G. (2020). Investigating the effect of teaching as a general learning strategy when learning through desktop and immersive VR: A media and methods experiment. *British Journal of Educational Technology, 51*(6), 2115–2138.

Knogler, M., Harackiewicz, J. M., Gegenfurtner, A., & Lewalter, D. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemporary Educational Psychology, 43*, 39–50. https://doi.org/10.1016/j.cedpsych.2015.08.004

Lee, K. M. (2004). Presence, explicated. *Communication Theory, 14*(1), 27–50. https://doi.org/10.1111/j.1468-2885.2004.tb00302.x

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

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher, 42*(8), 445–452. https://doi.org/10.3102/0013189X13511661

Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2020). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000473

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

Makransky, G., Lilleholt, L., & Aaby, A. (2017). Development and validation of the multimodal presence scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior, 72*, 276–285. https://doi.org/10.1016/j.chb.2017.02.066

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

Makransky, G., & Petersen, G. B. (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. https://doi.org/10.1007/s10648-020-09586-2

Makransky, G., Petersen, G. B., & Klingenberg, S. (2020). Can an immersive virtual reality simulation increase students' interest and career aspirations in science? *British Journal of Educational Technology, 51*(6), 2079–2097. https://doi.org/10.1111/bjet.12954

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. https://doi.org/10.1016/j.learninstruc.2017.12.007

Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). Cambridge University Press.

Mayer, R. E. (2014b). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (2nd ed., pp. 1–24). Cambridge University Press.

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*, 1–17. https://doi.org/10.1016/j.compedu.2019.103603

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education, 56*(3), 769–780. https://doi.org/10.1016/j.compedu.2010.10.020

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*, 309–326. https://doi.org/10.1007/s10648-007-9047-2

Moshman, D. (1982). Exogenous, endogenous, and dialectical constructivism. *Developmental Review, 2*(4), 371–384. https://doi.org/10.1016/0273-2297(82)90019-3

Mütterlein, J. (2018). The three pillars of virtual reality? Investigating the roles of immersion, presence, and interactivity. In *Hawaii International Conference on System Sciences*. https://doi.org/10.24251/HICSS.2018.174

Papert, S., & Harel, I. (1991). *Constructionism*. Ablex Publishing Corporation.

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

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Piccione, J., Collett, J., & De Foe, A. (2019). Virtual skills training: The role of presence and agency. *Heliyon, 5*(11). https://doi.org/10.1016/j.heliyon.2019.e02583

Pintrich, P. R., Smith, D., Garcia, T., & McKeachie, W. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. https://eric.ed.gov/?id=ED338122. Retrieved from.

Polito, V., Barnier, A. J., & Woody, E. Z. (2013). Developing the Sense of Agency Rating Scale (SOARS): An empirical measure of agency disruption in hypnosis. *Consciousness and Cognition, 22*(3), 684–696. https://doi.org/10.1016/j.concog.2013.04.003

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*. https://doi.org/10.1016/j.compedu.2019.103778

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review, 6*(2), 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Salzman, M. C., Dede, C., Loftin, R. B., & Chen, J. (1999). A model for understanding how virtual reality aids complex conceptual learning. *Presence: Teleoperators and Virtual Environments, 8*(3), 293–316. https://doi.org/10.1162/105474699566242

Schraw, G., Flowerday, T., & Lehman, S. (2001). Increasing situational interest in the classroom. *Educational Psychology Review, 13*(3), 211–224. https://doi.org/10.1023/A:1016619705184

Skulmowski, A., & Rey, G. D. (2018). Embodied learning: Introducing a taxonomy based on bodily engagement and task integration. *Cognitive Research: Principles and Implications, 3*(1), 1–10. https://doi.org/10.1186/s41235-018-0092-9

Stolz, S. A. (2015). Embodied learning. *Educational Philosophy and Theory, 47*(5), 474–487. https://doi.org/10.1080/00131857.2013.879694

Sweller, J. (2011). Cognitive load theory. In J. P. Mestre, & B. H. Ross (Eds.), *The psychology of learning and motivation* (pp. 37–76). Elsevier Academic Press.

Taranilla, R. V., Cózar-Gutiérrez, R., González-Calero, J. A., & Cirugeda, I. L. (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*. https://doi.org/10.1080/10494820.2019.1674886

Ten Berge, T., & Van Hezewijk, R. (1999). Procedural and declarative knowledge: An evolutionary perspective. *Theory & Psychology, 9*(5), 605–624. https://doi.org/10.1177/0959354399095002

Vogt, A., Babel, F., Hock, P., Baumann, M., & Seufert, T. (2021). Prompting in-depth learning in immersive virtual reality: Impact of an elaboration prompt on developing a mental model. *Computers & Education, 171*, 1–15. https://doi.org/10.1016/j.compedu.2021.104235

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–636. https://doi.org/10.3758/BF03196322

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

Yeo, N. L., White, M. P., Alcock, I., Garside, R., Dean, S. G., Smalley, A. J., et al. (2020). What is the best way of delivering virtual nature for improving mood? An experimental comparison of high definition TV, 360° video, and computer generated virtual reality. *Journal of Environmental Psychology, 72*. https://doi.org/10.1016/j.jenvp.2020.101500