Learning and Instruction 84 (2023) 101716

Contents lists available at ScienceDirect

Learning and Instruction

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

Learning by teaching in immersive virtual reality – Absorption tendency increases learning outcomes

Sandra Chiquet [a,*], Corinna S. Martarelli [b], David Weibel [a], Fred W. Mast [a]

[a] Department of Psychology, University of Bern, 3012, Bern, Switzerland
[b] Faculty of Psychology, UniDistance Suisse, 3900, Brig, Switzerland

ARTICLE INFO

Keywords:
immersive Virtual reality
Learning techniques
Media in education
Interactive learning environments
Distance education

ABSTRACT

We investigated the learning outcome of teaching an agent via immersive virtual reality (IVR) in two experiments. In Experiment 1, we compared IVR to a less immersive desktop setting and a control condition (writing a summary). Learning outcomes of participants who had explained the topic to an agent via IVR were better. However, this was only the case for participants who scored high on absorption tendency. In Experiment 2, we investigated whether including social cues in the task instructions enhances learning in participants explaining a topic to an agent. Instruction manipulation affected learning as a function of absorption tendency: Low-absorption participants benefitted most from being instructed to imagine they were helping a student peer pass an upcoming test, while high-absorption participants benefitted more when they were to explain the text to a virtual agent. The findings highlight the crucial role of personality traits in learning by teaching in IVR.

1. Learning by teaching in immersive virtual reality – absorption tendency increases learning outcomes

Meaningful learning occurs when learners engage in cognitive processing, such as selecting the most relevant information, organizing it meaningfully, and integrating it with prior knowledge, which is also referred to as "generative processing" (Mayer, 2014, 2020). Teaching others is a technique suggested to promote generative processing and therefore to enhance learning (Coleman, Brown, Rivkin, & Coleman, 1997; Fiorella & Mayer, 2013, 2014; Hoogerheide, Loyens, & van Gog, 2014). In the present study, we examine the efficacy of learning by teaching others in immersive virtual reality (IVR) in two experiments.

Recently, immersive virtual reality (IVR) has been attracting attention due to its potential impact on education and the science of learning (e.g., Lui, McEwen, & Mullally, 2020; Makransky & Lilleholt, 2018; Mulders, Buchner, & Kerres, 2020; Radianti, Majchrzak, Fromm, & Wohlgenannt, 2020). This is not surprising because, with IVR, one can simulate realistic and interactive environments for use with learning and teaching strategies. By means of a head-mounted display (HMD), IVR creates a fully immersive sensory experience in which one feels a sense of being present in a mediated environment (Slater, 2009). However, it remains an open question whether IVR per se encourages learners to engage in meaningful learning. Previous studies on learning in IVR have reported mixed results (Makransky, Terkildsen, & Mayer, 2019; Markowitz, Laha, Perone, Pea, & Bailenson, 2018; Meyer, Omdahl, & Makransky, 2019; Parong & Mayer, 2018; Webster, 2016) which can partially be explained by differences in specific elements of IVR (e.g., cognitive load, Makransky et al., 2019) and/or user characteristics (e.g., immersive tendency, Weibel, Wissmath, & Mast, 2010). In the present study, we aim to investigate whether there are specific elements of IVR and/or user characteristics that promote meaningful learning when students explain learned content to a computer-based agent.

1.1. Learning by teaching

There is considerable evidence for the positive impact of learning when students teach students face-to-face (for reviews, see Duran, 2017; Kobayashi, 2019; Roscoe & Chi, 2007). Such interactive learning by teaching settings activate generative processing (i.e., selecting the most relevant information, organizing it meaningfully and integrating it with prior knowledge) and thus foster learning by mutual questioning and answering (Roscoe & Chi, 2008). However, there is also evidence for learning gains without the additional advantage of interactive settings: The mere social presence of someone who is listening has been shown to

---

* Corresponding author.
E-mail addresses: sandra.chiquet@unibe.ch (S. Chiquet), corinna.martarelli@fernuni.ch (C.S. Martarelli), david.weibel@unibe.ch (D. Weibel), fred.mast@unibe.ch (F.W. Mast).

https://doi.org/10.1016/j.learninstruc.2022.101716
Received 22 February 2022; Received in revised form 14 September 2022; Accepted 7 November 2022
Available online 24 November 2022

improve oral explanation quality and stimulate cognitive processes (Coleman et al., 1997; Ploetzner, Dillenbourg, Preier, & Traum, 1999; Rittle-Johnson, Saylor, & Swygert, 2008). Going one step further, the physical presence of another person is not a necessary condition because non-interactive teaching (e.g., teaching a fictitious or non-present recipient) was shown to be as effective (Fiorella & Kuhlmann, 2019; Fiorella & Mayer, 2013, 2014; Hoogerheide et al., 2014, 2016; Hoogerheide, Renkl, Fiorella, Paas, & van Gog, 2019; Lachner, Backfisch, Hoogerheide, van Gog, & Renkl, 2019, 2021). In two studies, Fiorella and Mayer (2013, 2014) found positive effects on learning when teaching involved creating an instruction for a fictitious recipient. Participants who taught the material by creating a brief, video-recorded lecture before being tested outperformed those who studied for a test (without teaching) on an immediate and a delayed comprehension test. Likewise, Hoogerheide, Deijkers, Loyens, Heijltjes, and van Gog (2016) found that explaining on video, but not in writing, enhanced learning compared to restudy. These findings suggest that the extent to which a communication partner is perceived as present and real in a mediated conversation (i.e., social presence) can be crucial for activating generative processing and thus enhancing learning by non-interactive teaching (Hoogerheide et al., 2016; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, Lachner, & van Gog, 2019; Jacob, Lachner, & Scheiter, 2020). However, contradicting results exist, see e.g., Jacob, Lachner, and Scheiter (2021) who found no effect of induced social presence during explaining in written form.

1.2. *Learning by teaching in immersive virtual reality*

Recent technological advancements (e.g., stand-alone devices) made IVR technology more available and accessible to the field of education. The adaptation of learning by teaching to IVR is promising since it provides a flexible learning opportunity suitable for various learning contents and can easily be implemented in an educational context. Contrary to real life, IVR enables the simulation of social experiences in the absence of another person. According to previous findings, social interactions with agents (i.e., computer-controlled virtual characters) in IVR can positively affect learning and training (Okita, Bailenson, & Schwartz, 2007; Schmid Mast, Kleinlogel, Tur, & Bachmann, 2018). Thus, IVR may add a social element to learning by teaching despite the absence of peers and thus be a useful tool in the context of distance education. In the context of IVR, effective learning by teaching may depend on whether students accept the virtual scenario (i.e., providing explanations for a person) as plausible. This is important because, when teaching an agent in IVR, the recipient is obviously not real. Nevertheless, IVR is characterized by immersion and spatial presence, which may support the learner's belief in teaching an actual person even though the recipient is controlled by a computer.

1.3. *Spatial presence in immersive virtual reality*

Virtual reality systems are typically characterized by their degree of immersion (Cipresso, Giglioli, Raya, & Riva, 2018; Makransky & Lilleholt, 2018). Low-immersion virtual reality (VR) systems (also referred to as *non-immersive systems* in the literature) use external displays to provide images or scenes and the user interacts with the simulated environment via a mouse or keyboard. In contrast, high-immersion VR (which we refer to as *immersive virtual reality*, IVR) completely surrounds the user via a head-mounted display (HMD) and the interactivity is based on head motion tracking. The term *immersion* is defined as an objective measure of how well a technology generates vivid virtual environments through sensory motor contingencies and eliminates sensory information from the outside world (Cummings & Bailenson, 2015; Slater & Wilbur, 1997).

In general, higher levels of objective immersion result in a stronger feeling of spatial presence in the virtual environment: This association has been shown in research comparing high- (IVR via HMD) and low-immersion VR systems (Gorini, Capideville, De Leo, Mantovani, & Riva, 2011; Makransky et al., 2019; Moreno & Mayer, 2002) and in assessments of different types of HMD (Buttussi & Chittaro, 2018; Rupp et al., 2019). The term *spatial presence*, often described as the feeling of being there in the environment (e.g., Slater & Sanchez-Vives, 2016; Wirth et al., 2007; Witmer & Singer, 1998), refers to the psychological state in which the experiences are related to a mediated environment (Minsky, 1980) rather than the surrounding physical environment (Hofer, Hartmann, Eden, Ratan, & Hahn, 2020; Steuer, 1992; Weibel, Wissmath, & Mast, 2011). According to Wirth et al. (2007), spatial presence is experienced when a user accepts the mediated environment as her/his primary egocentric frame of reference. Lombard and Ditton (1997) point out the perceptual illusion of non-mediation, in which the medium disappears from the subject's conscious attention. In their view, spatial presence strongly depends on sensory input from the underlying media system: Sensorimotor contingencies supported by the VR system give rise to the sensation of being surrounded by in the virtual reality.

Due to increased spatial presence in IVR, users are more likely to be emotionally engaged in such a simulated content. Several studies have reported positive correlations between spatial presence and emotional experiences in VR (Baños et al., 2004, 2008, 2012; Diemer, Alpers, Peperkorn, Shiban, & Muehlberger, 2015; Gorini et al., 2011; Price & Anderson, 2007; Weber, Mast, & Weibel, 2021), although some researchers did not find such an effect (e.g., Krijn et al., 2004). Riva et al. (2007) found stronger spatial presence in an emotional environment compared to a neutral environment and that spatial presence predicted the user's emotional state during IVR exposure, indicating a bidirectional positive relationship between spatial presence and emotion. Thus, IVR may support generative processing in a learning-by-teaching scenario because users are emotionally engaged with the agent.

1.4. *Individual differences and suspension of disbelief*

VR experiences are modulated by factors other than immersion and spatial presence (Coxon, Kelly, & Page, 2016; Gorini et al., 2011; Hofer, Wirth, Kuehne, Schramm, & Sacau, 2012; Sacau, Laarni, & Hartmann, 2008; Weibel et al., 2010, 2012): Individual differences and the user's belief in the simulated content can be associated with the positive effects on learning outcome of a social interaction in IVR.

Witmer and Singer (1998) developed the immersive tendency questionnaire (ITQ) to measure individual differences in involvement with and immersion in VR. Based on theoretical considerations, the authors postulated a scale consisting of three sub-dimensions: involvement, focus, and tendency to play video games. By testing the dimensionality of the ITQ, Weibel et al. (2010) suggested that immersive tendency is composed of two independent dimensions: Tendency towards emotional involvement and absorption.

Tendency towards emotional involvement is defined as an internal state of being affectively engaged in a virtual stimulus and experiencing emotions such as empathy, surprise, fear, and other emotions in virtual contexts (Wirth, Hofer, & Schramm, 2012). *Absorption* is defined as the user's ability to focus or redirect her/his attentional resources to engage in sensory experiences completely and ignore external distractors easily. Tellegen and Atkinson (1974) define *absorption* as the ability and readiness to assign meaning to the object of the attentional focus. High-absorption individuals are expected to have a heightened sense of reality: They perceive mental images, feelings, and visually simulated stimuli as present and real (Baños et al., 1999; Sas & O'Hare, 2003; Tellegen & Atkinson, 1974). Weibel et al. (2010) named the factors of the ITQ "emotional involvement" and "absorption" because the items pertain to a strong emotional reaction and focused attention, respectively, during media exposure.

Most previous studies have reported a positive association between ITQ scores and spatial presence (Ling, Nefs, Brinkman, Qu, & Heynderickx, 2013; Wallach, Safir, & Samana, 2010; Weibel & Wissmath, 2011), although Murray, Fox, and Pettifer (2007) found no correlation between spatial presence and immersive tendency. According to the model by Wirth and colleagues, a high absorption tendency allows the user to suspend her/his disbelief in the virtual environment and to experience and act as if the virtual world was real (Wirth et al., 2007). This model is supported by the results of a previous study by Baños et al. (1999).

The term *suspension of disbelief* is defined as the intention to accept a mediated content as real and to suppress external stimuli and internal cognition that might contradict the illusion of the experience. Suspension of disbelief helps one to resist doubts about the simulation's authenticity (Hofer et al., 2012; Muckler, 2017). Therefore, suspension of disbelief is expected to be associated with spatial presence as well as with emotional responses to virtual stimuli (Gorini et al., 2011; Wirth et al., 2007, 2012). According to de Gelder, Kätsyri, and de Borst (2018), suspension of disbelief in IVR refers to a special state in which the belief in the virtual reality and the knowledge of its unreality coexist. It is a cognitive act of going along with the story by accepting the experience as real within the virtual reality and may play a key role in IVR applications based on interactions with computer-based agents. Previous findings suggest that the mere belief in a social interaction in VR can lead to behavioral and neural changes (Caruana, de Lissa, & McArthur, 2016, 2017; Gorini et al., 2011). For example, Okita et al. (2007) found increased learning in participants who were told they would interact with a virtual character controlled by a human compared to those instructed that the virtual character was controlled by a computer. In the same vein, accepting an interaction in IVR as real may support learning by teaching in IVR. That is, participants may engage more in generative processing when they suspend their disbeliefs.

1.5. *The present study*

According to the cognitive theory of multimedia learning (Mayer, 2009, 2014) meaningful learning depends on whether learners engage in generative processing. Based on this theoretical framework, we expect teaching to foster learning when generative processes are activated (cf. Coleman et al., 1997; Fiorella & Mayer, 2013, 2014; Hoogerheide et al., 2014). Spatial presence is associated with emotional reactions in IVR (Baños et al., 2004, 2008, 2012; Price & Anderson, 2007; Riva et al., 2007; Weber et al., 2021) and may result in participants engaging in generative processing when teaching the agent in IVR. We further expected that the beneficial effects of teaching an agent not only depend on increased spatial presence experience in IVR, but also on individual differences in the tendency of being affectively engaged and absorbed in media content. Previous research on learning by non-interactive teaching (i.e., generating explanations for fictitious or non-present others) suggests that the feeling of generating explanations to an actual person may be crucial (Hoogerheide et al., 2016). Individuals with a high absorption tendency may engage in generative processing more than their low-absorption counterparts since they attribute more reality to virtual experiences (Baños et al., 1999) and because they are motivated and able to become intensely involved with virtual environments (Wirth et al., 2007). Thus, it is conceivable that low-absorption individuals benefit most from teaching an agent in IVR when the task instructions contain social cues that serve to support their acceptance of the experience in IVR as real.

Based on these considerations, the present study examined the efficacy of learning by teaching in IVR in two experiments. In Experiment 1, we investigated whether increased spatial presence in IVR supports learning when teaching a computer-based agent. We compared participants' learning outcome after they taught an agent presented in IVR and after they taught the same agent presented on a computer screen. Most importantly, the learning by teaching conditions (IVR condition and desktop condition) differed in terms of mediation technology (IVR vs. desktop VR), however, they were identical in terms of the simulated content (learning environment and agent) and in terms of instruction. These two conditions were in oral form. We added a control condition, in which participants were asked to write a summary of the learned content, which is a study-relevant activity that does not involve teaching. In Experiment 2, all participants were asked to teach the learning content in IVR. Half of the participants were instructed to imagine an interaction with a student peer in which they were helping her pass an upcoming test (student instruction); the other half was instructed to teach a computer-based agent (agent instruction). In particular, we investigated whether (a) the instruction to imagine a concrete social interaction fostered learning and (b) the effect of instruction was moderated by participants' absorption tendency.

2. Experiment 1

In Experiment 1 we tested three hypotheses. First, we hypothesized that the learning outcome would be better in the IVR condition than in the desktop and control conditions (H1). Second, we expected spatial presence to mediate the effect of media type (desktop vs. IVR) on learning (H2). Third, we expected trait variables to moderate the relationship between conditions and learning outcome. More precisely, we expected the individuals scoring high on tendency towards emotional involvement and absorption tendency to benefit more from teaching the agent, while we expected tendency towards emotional involvement and absorption tendency to have no influence on learning outcome in the control condition (writing a summary) (H3). To test these hypotheses, we assigned participants to one of three conditions: 1) explaining new study material to an agent presented in IVR, 2) explaining new study material to an agent presented on the computer screen, or 3) creating a written summary of the study material (control condition). Fig. 1 shows our working models of these hypotheses.

2.1. *Method*

2.1.1. *Participants*

Seventy-three students from the University of Bern took part in the experiment. Participants were recruited through a participant recruitment portal. The data of seven participants were excluded due to technical problems ($n = 1$), language difficulties ($n = 1$), cybersickness ($n = 3$), and missing data for measurement timepoint t3 ($n = 2$). The final sample consisted of 66 participants (57 females, 9 males) ranging in age from 18 to 48 years ($M = 22.35$, $SD = 3.9$).

Participants were unaware to the purpose of the study and received course credit for participation. The study was approved by the ethics committee of the Faculty of Human Sciences of the University of Bern.

2.1.2. *Learning material*

We created two technical texts with 328 words and 297 words, respectively, to serve as the learning material. The texts described the functionality of a condenser microphone (text A) and a dynamic microphone (text B). Both texts started with the same general introduction to microphones (60 words), describing them as sensors that convert changes in air pressure into electrical signals. The texts differed with respect to the conversion principle, which varies depending on microphone type. Each text included a schematic representation of the described microphone. The assumption of comparability of the texts and study time (12 min) was based on the results of a pilot study.

2.1.3. *Measuring instruments*

Spatial presence. Spatial presence was assessed using Kim and Biocca's (1997) presence scale (which we refer to as "verbal presence"). The authors defined *presence* as the psychological state of (spatially) being in a mediated environment. The questionnaire was originally designed to measure the experience of presence in televised media experiences and has often been used in the context of video games and online environments and turned out to be reliable and valid (e.g., Nicovich, Boller, & Cornwell, 2005; Weibel & Wissmath, 2011). We adapted the items to the VR experience (sample item: "During the virtual reality

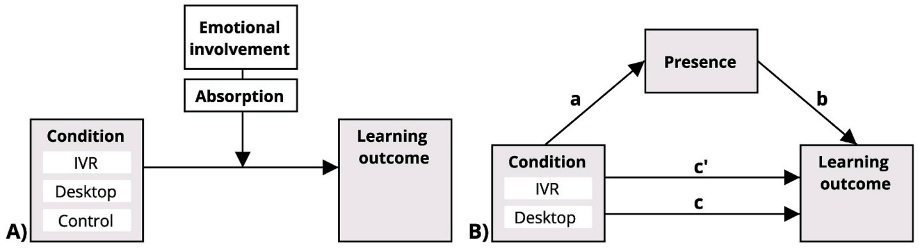

Fig. 1. Working models Experiment 1
Note. Our working model 1 predicted that the learning outcome – as measured by an immediate and a delayed test – in the three conditions (IVR vs. desktop vs. control) would differ. Tendency towards emotional involvement and absorption tendency were expected to moderate the relationship between condition and learning outcome. In our working model 2, Presence was expected to mediate the effect of media type (desktop vs. IVR) on learning outcome (path c represents the total effect, i.e., $c = ab + c'$).

experience, my mind was present in the real room and not in the virtual reality environment”). The eight items were rated on a 7-point Likert scale ranging from (1 = I do not agree at all, 7 = I agree very strongly). The scale has shown adequate internal consistency (Cronbach’s α = 0.75 Weibel & Wissmath, 2011). As regards our sample, the scale was found to be reliable (Cronbach’s α = 0.79).

In addition to the verbal questionnaire, we used Weibel et al.’s (2015) pictorial presence SAM scale (which we refer to as “presence SAM”). The scale uses self-assessment-manikins (SAM) to assess spatial presence as defined by Wirth et al. (2007) and was found reliable and valid to measure spatial presence. Participants rated four items by choosing 1 out of 5 pictograms for each item. In our sample, the Cronbach’s α estimate of reliability was 0.64, indicating a somewhat low level of internal consistency for the scale.

Immersive Tendency Questionnaire (ITQ). We used Weibel et al.’s (2010) adapted version of the ITQ (Witmer & Singer, 1998). The scale consists of two subdimensions: tendency towards emotional involvement (4 items; sample item: “Do you ever experience dreams that are so real that you are confused when you wake up?”) and absorption (5 items; sample item: “Are you sometimes so thrilled by a movie that you stop noticing things happening around you?”). The nine items were rated on a 5-point Likert scale (1 = I do not agree at all, 5 = I fully agree). The scale has shown adequate internal consistency (Cronbach’s α = 0.77; Parsons, Barnett, & Melugin, 2015). In this experiment, internal consistency for the ITQ was acceptable (Cronbach’s α = 0.74). The internal consistency for the sub-scale tendency towards emotional involvement (Cronbach’s α = 0.68) and absorption (Cronbach’s α = 0.59) were somewhat lower.

Control variables. We assessed experienced difficulty (“Compared to all the other activities I usually do, the current activity is…”), required skills (“I think my skills in this area are…”), and the balance between challenge and required skills (“For me personally, the current requirements are…”) with 9-point Likert scales ranging from 1 (easy) to 9 (hard) for experienced difficulty, from 1 (low) to 9 (high) for required skills, and from 1 (too low) to 9 (too high) for the balance between challenge and required skills. As in previous studies (e.g., Rodgers, Conner, & Murray, 2008), the single items were taken from Rheinberg et al.’s (2003) flow short scale.

Learning outcome. Learning outcome was measured using a test consisting of 16 questions per text (nine open questions and seven multiple-choice questions) about the content that was studied and explained (or summarized). Participants received one point for correct responses and zero points for incorrect responses. Two independent raters judged the accuracy of the responses and Cohen’s κ (weighted) indicated high agreement between the raters’ judgments (κ = 0.87).

2.1.4. *Equipment*

The 3D environment in the VR condition was presented by means of an *Oculus Rift Development Kit 2* (DK2) (http://www.oculus.com/en-us/dk2/). The DK2 is a head-mounted display (HMD) with a positional camera tracking system with six degrees of freedom head tracking (rotation and translation) and a maximum frame rate of 75 Hz at a display resolution of 960*1080 pixel per eye. We used PC hardware with a graphics card (*NVIDIA® GeForce GTX 970)* sufficient for the DK2 maximal refresh rate. In the desktop condition, we used a 17-inch monitor with a resolution of 1920*1080 pixels.

2.1.5. *Teaching environment and software*

We used the open-source program *3D creation suite Blender*, version 2.75a (https://www.blender.org/about/) to design the 3D environment. The environment was a simple room with an agent standing at a table (see Fig. 2). We animated the agent using the commercial virtual reality tool Worldviz Vizard (version 5.0): In both the IVR and desktop condition the agent blinked and slightly moved. In the IVR condition, the agent was able to maintain eye contact with the participant using the participant’s head tracking data (rotation and translation).

2.1.6. *Procedure*

The experiment was divided into a baseline and an intervention phase. During the baseline phase, the purpose of which was to control for individual differences in text-based learning, participants in all conditions were instructed to read and study a text for a later test with open and multiple-choice questions. Two different texts with associated tests on the same topic (how microphones work) were used in the baseline and intervention phases to prevent the baseline measurement from affecting participants’ focus in the intervention phase measurement (see Fiorella & Mayer, 2013; Hoogerheide, Renkl, et al., 2019). We counterbalanced the order in which the two texts were presented (text A in baseline, text B in intervention phase vs. text B in baseline, text A in intervention phase). The study phase lasted 12 min. Participants were allowed to take notes while studying the texts but were informed that they could not use their notes afterwards. After the study phase, the participants answered 16 questions concerning the content of the text studied (baseline).

During the intervention phase, participants were instructed to read and study a second text; they were not informed about the following task (teaching or writing a summary). They were allowed to take notes during the study phase. After the 12-min study phase, participants in the IVR and desktop conditions were to explain (maximum duration: 6 min) the content of the studied text to a female agent. They were asked to imagine that the agent would be tested afterwards and that her test performance would most likely depend on the quality of the explanation ("Please explain to the agent in [virtual reality/on the desktop screen] how a [condenser/dynamic] microphone works. Try not only to convey important facts to the agent, but also to show connections. Imagine, that the agent will be asked questions about the microphone afterwards. Her performance in this test depends on the explanations you have given her before. The agent will know no more and no less than what you have taught her”). In the VR condition, participants were wearing an HMD while explaining the content to the agent, whereas participants in the desktop condition viewed the agent on a computer screen. Participants in the control condition were to create a written summary of the text they had studied ("Please write a summary of the text. Try not only to note important facts, but also to show connections.”).

Participants in the VR and desktop conditions were in a standing position while explaining, whereas participants in the control condition sat on a chair. After participants provided the explanation or completed the written summary, they received a second set of 16 questions about

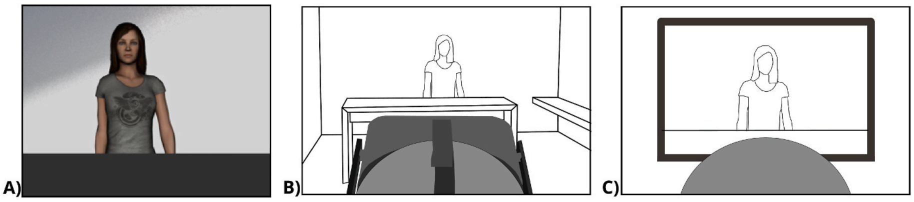

Fig. 2. Teaching environment
Note. Fig. 2. Participants view with an agent listening to the explanation in both the IVR and desktop condition (A) and schematic illustration of teaching an agent presented via headset (IVR condition, B) and on a computer screen (desktop condition, C).

the text they had explained or summarized before (immediate posttest). At the end of the intervention phase, all participants completed the ITQ and the spatial presence questionnaires and we assessed the control variables. To examine the stability of the learning outcome over time, participants returned to the laboratory 7 or 8 days later and they were again exposed to the second set of questions (delayed posttest).

2.1.7. *Data analysis*

Data preparation, analysis, and visualization was performed in R (version 3.5.1; R Core Team). We used the *bayestestR* (Makowski, Ben-Shachar, & Lüdecke, 2019) package for correlation analysis and the *brms* package for Bayesian (non)-linear mixed models (Bürkner, 2017) to analyze the data following a Bayesian approach. Parameter uncertainty was expressed using a 95% credible interval (CI). A 95% CI not containing zero is indicative of a 95% probability that the true (unknown) effect is not equal to zero given the observed data. We estimated posterior parameters using a Markov chain Monte Carlo (MCMC) method implemented with *Stan* (https://mc-stan.org/) with 4 chains of 2500 iterations. We visually inspected convergence of the chains using trace plots. Model notations including prior distributions and the results of model comparisons are available in the supplementary material.

We controlled for differences in perceived difficulty, required skills, and balance between challenge and required skills. The response variables (difficulty, skills, and balance) were estimated using a multivariate general linear model (GLM). The model (M_control) included condition as a fixed effect for the comparisons of IVR, control, and desktop.

Learning outcome was defined as the correctness of participants' responses to each of the 16 test questions (correct vs. incorrect). Due to the dichotomy of the variable, we used generalized linear mixed models (GLMM) with a binomial link function to analyze the probability of a participant giving a correct response to a test question. We were primarily interested in the difference between conditions (*H1*). The first model (M1 learning outcome) thus included time as a dummy-coded effect to compare t1 (i.e., the baseline, which was the reference category) with t2 (immediate posttest) and t3 (delayed posttest), condition as the dummy-coded effect to compare the IVR condition (the reference category) with the control condition and the desktop condition, and the interaction term between time and condition as fixed effects. Furthermore, we included random intercepts to account for repeated measurements.

We were further interested in whether the link between condition (IVR vs. desktop vs. control) and learning outcome was moderated by the participant's tendency to be emotionally involved in and absorbed by media content (*H3*). We added tendency towards emotional involvement, absorption tendency, and the interaction terms between time, condition, tendency towards emotional involvement, and absorption tendency as fixed effects to the first model (M1_learning outcome) and compared the model including all parameters (i.e., the full model) with the exclusion of either tendency towards emotional involvement or absorption tendency, based on leave-one-out cross validation (*LOO function in brms*). The best fitting model (M2_learning outcome) included time as a dummy-coded effect for t1 (defined as the reference category) vs. t2 vs. t3, condition as a dummy-coded effect for IVR (defined as the reference category) vs. control vs. desktop, absorption tendency, and the interaction terms between time, condition, and absorption tendency as fixed effects. We included random intercepts to account for repeated measurements.

We expected spatial presence in the IVR condition to result in a higher level of learning compared to the desktop condition. In two multivariate GLMM, we estimated spatial presence as measured with the verbal presence scale (M1_presence) and with the pictorial presence SAM (M2_presence) with a Gaussian link function. To compare desktop with IVR, we included media type (desktop vs. IVR) as a fixed effect in both models. We predicted the learning outcome with a binomial link function. Fixed effects included time as a dummy-coded effect (t1 as the reference category vs. t2 vs. t3), media type (desktop vs. IVR), verbal presence (M1_presence) or presence SAM (M2_presence), and the interaction terms between time, media type, and verbal presence (M1_presence) or presence SAM (M2_presence). We included random intercepts to account for repeated measurements.

2.2. *Results*

2.2.1. *Self-reported measures*

The multivariate GLM (M_control) revealed no differences between the IVR and control conditions in difficulty, $\beta_{difficulty\_control} = 0.37;\ SE = 0.51; CI = [-0.63; 1.35]$, required skills, $\beta_{skills\_control} = -0.09;\ SE = 0.53; CI = [-1.14; 0.97]$, or balance, $\beta_{balance\_control} = -0.13;\ SE = 0.34; CI = [-0.79; 0.53]$, as well as no differences between the IVR and desktop conditions in difficulty, $\beta_{difficulty\_desktop} = 0.22;\ SE = 0.53; CI = [-0.82; 1.25]$, required skills, $\beta_{skills\_desktop} = -0.01;\ SE = 0.54; CI = [-1.05; 1.05]$, or balance, $\beta_{balance\_desktop} = -0.13;\ SE = 0.34; CI = [-0.82; 0.52]$. After changing the reference level to control, we found comparable ratings between control and desktop in difficulty, $\beta_{difficulty\_desktop} = -0.14;\ SE = 0.52; CI = [-1.18; 0.87]$, skills, $\beta_{skills\_desktop} = 0.11;\ SE = 0.55; CI = [-0.96; 1.23]$, and balance, $\beta_{balance\_desktop} = -0.01;\ SE = 0.34; CI = [-0.68; 0.68]$.

Correlation analysis showed that verbal presence was positively related to presence SAM, $rho_{Presence:SAM} = 0.48; CI = [0.31; 0.66]$, and to tendency towards emotional involvement, $rho_{Presence:Emotional\ Involvement} = 0.32; CI = [0.10; 0.52]$. We found no correlation between verbal presence and absorption tendency. The 95% CI of the correlation between presence SAM and absorption tendency as well as between presence SAM and tendency towards emotional involvement included zero, indicating no relationships. Absorption tendency was positively related to tendency towards emotional involvement, $rho_{Absorption:Emotional\ involvement} = 0.43; CI = [0.28; 0.59]$. The means and standard deviations (SD) of all self-reported measurements (verbal presence, presence SAM, perceived difficulty, required skills, balance between challenge and skills, absorption tendency, and tendency towards emotional involvement) are reported in Table 1 and the correlation matrix of verbal presence, presence SAM, absorption tendency, and tendency towards emotional involvement is shown in Table 2.

2.2.2. *Learning outcome*

We expected participants who taught an agent presented in IVR to show better learning outcomes than those who taught an agent

Table 1
Means and standard deviations of self-reported measures (verbal presence, presence SAM, perceived difficulty, required skills, balance between challenge and skills, absorption, and emotional involvement).

	Mean (SD)		
	IVR	Desktop	Control
Self-reported measures			
Presence	3.83 (1.18)	2.53 (1.10)	
SAM	3.11 (0.90)	2.92 (0.82)	
Difficulty	5.14 (1.70)	5.36 (1.87)	5.50 (1.54)
Skills	3.45 (1.97)	3.45 (1.71)	3.36 (1.68)
Balance	6.05 (1.25)	5.91 (1.19)	5.91 (0.81)
Absorption	2.99 (0.60)	3.17 (0.70)	3.33 (0.76)
Emotional involvement	3.24 (0.69)	2.96 (0.61)	3.19 (0.61)

Note. Presence = verbal presence. SAM = presence SAM. Difficulty = perceived difficulty. Skills = required skills. Balance = balance between challenge and skills. Absorption = absorption tendency. Emotional involvement = **tendency towards emotional involvement**.

Table 2
Matrix of correlations [l-95% CI, u-95% CI].

	(1)	(2)	(3)	(4)
(1) Presence	1	**0.48** [0.31, 0.66]	0.07 [-0.14, 0.30]	**0.32** [0.10, 0.52]
(2) SAM		1	0.18 [-0.03, 0.40]	0.19 [-0.03, 0.41]
(3) Absorption			1	**0.43** [0.28, 0.59]
(4) Emotional involvement				1

Note. Presence = verbal presence. SAM = presence SAM. Absorption = absorption tendency. Emotional involvement = tendency towards emotional involvement. Estimates with credible intervals not including zero are indicated in bold.

presented on a desktop screen or those in the control condition (*H1*). The first model (M1_learning outcome) thus estimated the probability of a participant giving a correct response to a test question as a function of time (t1 vs. t2 vs. t3) and condition (IVR vs. desktop vs. control). Contrary to our expectation, the posterior estimates revealed no interactions between time and condition, suggesting that there was no overall difference in learning outcome when an agent was taught in IVR compared to the control or desktop conditions. Model parameters of M1_learning outcome are reported in Table 3.

The second model (M2_learning outcome) estimated learning outcome depending on time (t1 vs. t2 vs t3), condition (IVR vs. desktop vs. control), and a participant's tendency to be absorbed by media content.

Table 3
Regression coefficients (logit transformed posterior mean, standard error, and 95% credible intervals) of M1_learning outcome (probability of giving a correct response to a test question as a function of time and condition).

	Estimate	Est.Error	l-95% CI	u-95% CI
Group-Level Effects				
Participant (sd)	0.47	0.06	0.36	0.60
Population-Level Effects				
Intercept	0.40	0.15	0.11	0.70
T2	0.24	0.16	−0.06	0.55
T3	0.03	0.16	−0.28	0.33
Control	0.05	0.21	−0.37	0.46
Desktop	0.00	0.21	−0.41	0.41
T2:Control	−0.18	0.22	−0.62	0.26
T3:Control	−0.33	0.22	−0.76	0.09
T2:Desktop	−0.06	0.22	−0.49	0.37
T3:Desktop	−0.12	0.22	−0.56	0.31

Note. T2 = immediate posttest. T3 = delayed posttest. Desktop = desktop condition. Control = writing a summary.

The posterior distributions revealed an interaction between time, condition, and absorption tendency. Higher absorption scores were positively correlated with an increased probability of participants giving correct responses to the test questions from t1 to t2 in the IVR condition compared to the control condition, $\beta_{T2:control:absorption} = -0.85$; $SE = 0.36$; $CI = [-1.55; -0.14]$). Thus, absorption tendency moderated the link between condition and learning outcome from t1 to t2. Model parameters of M2_learning outcome are reported in Table 4 and visualized in Fig. 3.

We further analyzed whether spatial presence (verbal presence and presence SAM) mediates between media type (desktop vs. IVR) and learning outcome (*H2*). The first multivariate GLMM (M1_presence) revealed that verbal presence was predicted by media type (path a). As expected, compared to participants in the desktop condition, participants in the IVR condition experienced greater spatial presence as measured by the verbal presence scale $\beta_{presence\_IVR} = 1.30$; $SE = 0.20$; $CI = [0.92; 1.68]$. Contrary to our expectation, the 95% CI of the indirect effect included zero for the comparison between t1 and t2, $ab\_t2 = 0.26$; $CI = [-0.10; 0.69]$), as well as the comparison between t1 and t3, $ab\_t3 = 0.31$; $CI = [-0.06; 0.72]$. This is indicative of insufficient evidence of verbal presence mediating between media type and learning outcome in the immediate and the delayed test. The posterior distributions of the second model (M2_presence) revealed no effect of media type (path a) on presence SAM $\beta_{SAM\_IVR} = 0.19$; $SE = 0.15$; $CI = [-0.10; 0.49]$ and the 95% CI of the indirect effect included zero for the comparison between t1 and t2, $ab\_t2 = 0.03$; $CI = [-0.05; 0.19]$), as well as the comparison between t1 and t3, $ab\_t3 = 0.02$; $CI = [-0.06; 0.17]$. Model parameters of M1_presence and M2_presence are provided in the supplementary material.

2.3. Discussion of experiment 1

In Experiment 1, we compared participants' learning outcomes after they had taught an agent presented in IVR or on a computer screen or created a written summary. We found no overall differences between the three conditions. Thus, the results did not support the hypothesis of an improved learning outcome due to IVR per se (*H1*). However, participants who had taught an agent in IVR and who scored high on trait

Table 4
Regression coefficients (logit transformed posterior mean, standard error, and 95% credible intervals) of M2_learning outcome (probability of giving correct responses to the test questions as a function of time, condition, and absorption tendency).

	Estimate	Est.Error	l-95% CI	u-95% CI
Group-Level Effects				
Participant (sd)	0.44	0.06	0.32	0.56
Population-Level Effects				
Intercept	0.44	0.15	0.14	0.75
T2	**0.42**	0.18	**0.09**	**0.77**
T3	0.11	0.17	−0.22	0.44
Control	−0.01	0.21	−0.42	0.41
Desktop	−0.04	0.21	−0.45	0.38
Absorption	0.21	0.25	−0.27	0.71
T2:Control	−0.35	0.24	−0.82	0.10
T3:Control	−0.40	0.23	−0.85	0.05
T2:Desktop	−0.24	0.23	−0.70	0.22
T3:Desktop	−0.21	0.23	−0.66	0.24
T2:Absorption	**0.78**	0.28	**0.23**	**1.35**
T3:Absorption	0.42	0.27	−0.10	0.93
Control:Absorption	−0.10	0.32	−0.72	0.53
Desktop:Absorption	−0.36	0.32	−1.00	0.28
T2:Control:Absorption	**−0.85**	0.36	**−1.55**	**−0.14**
T3:Control:Absorption	−0.50	0.34	−1.16	0.16
T2:Desktop:Absorption	−0.57	0.37	−1.28	0.15
T3:Desktop:Absorption	−0.18	0.35	−0.86	0.51

Note. T2 = immediate posttest. T3 = delayed posttest. Control = control condition. Desktop = desktop condition. Absorption = absorption tendency. Estimates with credible intervals not including zero are indicated in bold.

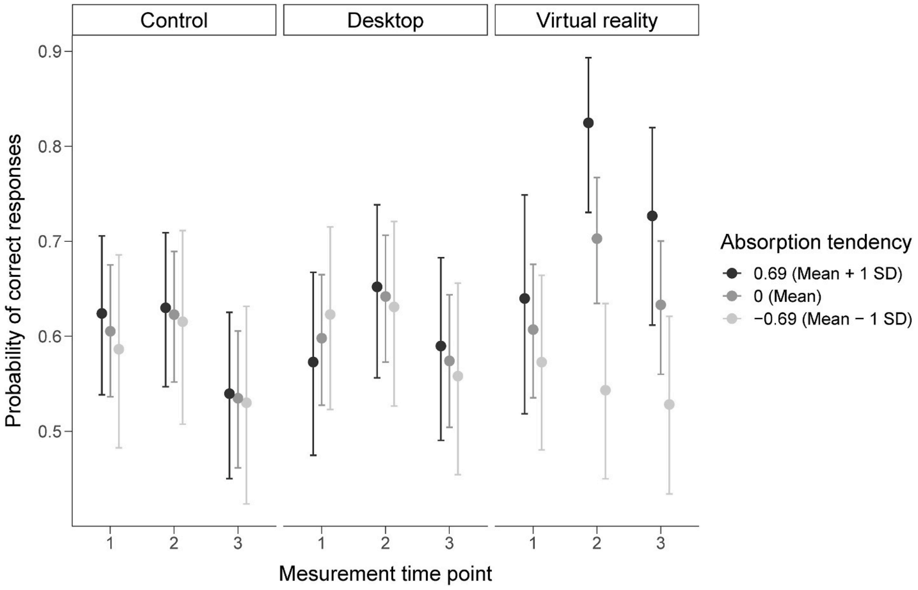

Fig. 3. Probability of correct responses as a function of condition, time, and absorption tendency
Note. The figure shows improved learning from t1 (= baseline) to t2 (= immediate posttest) in IVR compared to the control condition (writing a summary) as a function of absorption tendency (centered around the grand mean).

absorption did demonstrate better learning outcomes. This result suggests that whether one's learning outcome benefits from one's teaching a computer-based agent in IVR depends on one's absorption tendency (*H3*).

Teaching is directly linked to knowledge building by activating generative processing, such as selecting relevant information, organizing it meaningfully, and monitoring whether the explanation is accurate and understandable (Fiorella & Mayer, 2013, 2014). However, whether or not teaching is beneficial for learning depends on certain conditions (see Kobayashi, 2019). Our results suggest that generative processing is activated when a learner teaches a computer-based agent in IVR and that learner has a high tendency to be absorbed by media content. Individuals who had taught an agent presented in IVR showed a better learning outcome if they had a high absorption tendency.

Absorption tendency is considered to be associated with the acceptance of virtual contents as real through suspension of disbelief (see Hofer et al., 2012; Muckler, 2017). Regarding our results, it is likely that high-absorption participants were motivated and cognitively able to accept the computer-based agent as a real recipient and thereby suppress task-irrelevant stimuli and inner thoughts, while focusing more on task-relevant cognitive processing (e.g., monitoring whether their explanations were meaningful, accurate, and clear).

In line with previous research (Gorini et al., 2011; Hofer et al., 2020; Makransky et al., 2019; Moreno & Mayer, 2002), we found increased spatial presence (as measured with the verbal presence scale) in participants in the IVR condition compared to those in the desktop condition. However, the effect of media type (desktop vs. IVR) on learning outcome was not mediated by spatial presence (*H2*). Consistent with Murray et al. (2007), we found no relationship between absorption tendency and spatial presence, thus indicating that the feeling of being present in IVR is not determined by an individual's tendency to accept a mediated content as real. Contrary to our expectation, the results suggest that tendency towards emotional involvement has no influence on whether individuals benefit from learning by teaching. This might be due to the fact that teacher-learner interactions do not intrinsically involve strong emotions such as fear and tension.

Taken together, the results of Experiment 1 suggest that whether an individual's learning outcome is better when the individual teaches an agent in IVR depends on the individual's tendency to be absorbed by media content. High-absorption participants may have been more able to engage in generative processing because they were more able to suspend disbelief and therefore treated the computer-based agent as if it was real. Individuals with low-absorption tendency may be more dependent on external factors to accept a virtual content as real. For example, findings from previous research on social VR show that instructions can affect participants' behavior in a virtual experience (see Caruana et al., 2016, 2017; Gorini et al., 2011; Okita et al., 2007). Especially low-absorption individuals may benefit from a social relevant instruction when teaching a computer-based agent in IVR, as they do not suspend their disbeliefs on their own. Therefore, in a second experiment, we investigated whether the instruction to imagine an interaction with an actual person affects learning by teaching an agent in IVR, depending on participants' absorption tendency.

3. *Experiment 2*

Experiment 2 aims to extend the findings from Experiment 1 by testing whether the benefits of teaching an agent in IVR can be increased by instruction manipulation. We thus focused on learning by teaching in IVR. Half of participants were instructed to imagine that they were explaining the text to a female student peer to help her pass an upcoming test (student instruction), whereas the other half was instructed to explain the text to a virtual agent (agent instruction). In both conditions, participants generated explanations to a computer-based agent which was provided via HMD. In Experiment 2 we tested three hypotheses. First, we hypothesized that the learning outcome would be better in the student instruction condition than in the agent instruction condition (H1). We expected the participants in the student instruction condition to suspend their disbelief and thus activate generative processing, which in turn should improve their learning outcome. To test whether the experimental manipulation was successful, we added suspension of disbelief as an additional self-reported measure. Second, we hypothesized that the effect of condition (student instruction vs. agent instruction) on learning outcome would be mediated by suspension of disbelief (H2). Based on Experiment 1, we further hypothesized that participants' absorption tendency would moderate the influence of instruction on learning outcome (H3). We expected the effect of instruction condition on learning outcome to be stronger in low-absorption individuals. We reasoned that high-absorption individuals may be less dependent on external instructions since they easily suspend their disbelief due to their individual tendency to do so. Thus, as in Experiment 1, we assessed individual differences in absorption tendency. A working model visualizing the three hypotheses is shown in Fig. 4.

3.1. Method

3.1.1. Participants

Sixty-eight students from the University of Bern participated in the second experiment. The data of ten participants were excluded due to technical problems. The final sample consisted of 58 participants (45 female, 13 male) ranging in age from 19 to 29 years ($M = 21.9$, $SD = 2.13$). Participants were randomly assigned to one of two conditions: student instruction and agent instruction. Participants were fully informed about the experimental procedure and were given standardized written instructions. Informed consent was collected from all participants prior to the experiment. Participants received course credit for participation. The study was approved by the ethics committee of the Faculty of Human Science of the University of Bern.

3.1.2. Virtual environment, instruction manipulation, and equipment

In Experiment 2, we used the same virtual environment as in Experiment 1: a simple room with a female virtual character standing behind a table. The female virtual character was able to maintain eye contact with the participants while they were explaining the text using the participants' head tracking data (rotation and translation). All participants were immersed in the same virtual environment. Participants in the student instruction condition were to imagine an interaction with a student peer and were told that the student had failed the previous test and urgently needed the participant's help to prepare for the next test, which would be her last chance to pass the test ("You will now enter a virtual world via the HMD, where you will meet Anna. Please imagine that Anna is a student who failed the last statistics test. In view of the upcoming test, her last chance to pass the test, Anna urgently needs your help. Please explain to Anna, as best as you can, how the analysis of variance works. Try not only to convey important facts to Anna, but also to show connections so that she can go into the exam well prepared.") Participants in the agent condition were to explain the content of the text to a virtual agent ("Via HMD, you will now enter a virtual world where you will meet an agent. Please explain to the agent how the analysis or variance works. Try not only to convey important facts, but also to show connections."). The environment was presented by means of the Oculus Rift (https://www.vrbound.com/headsets/oculus/rift), an HMD with a maximum frame rate of 90 Hz at a display resolution of 2160*1200 pixel per eye.

3.1.3. Procedure

The procedure used in Experiment 2 was almost identical to that used in Experiment 1, except for the following exceptions: First, in Experiment 1, we controlled for individual differences in text-based learning by asking participants to read and study a text for a later test. In Experiment 2, we controlled for individual differences in the learning material by asking participants to take a pretest. Second, all participants were to read and study the same learning materials (viz., text explaining univariate analysis of variance, ANOVA). The study time lasted 15 min. All participants explained the content of the studied text while being immersed in IVR (maximal duration: 5 min instead of 6 min in Experiment 1) after they received one of two instructions according to the assigned condition (student instruction vs. agent instruction). Suspension of disbelief was assessed as an additional self-reported measure at the end of the immediate posttest (t2); absorption tendency was rated after the delayed posttest one week after the first session.

3.1.4. Learning material

We were interested in whether the effect of teaching an agent in IVR can be generalized to learning material of different domains. Thus, for Experiment 2, we created a text (1050 words) describing the statistical procedure of univariate analysis of variance (ANOVA). The text started with a general introduction, defining ANOVA in one sentence (29 words). The calculation and interpretation of an ANOVA was explained on the basis of a concrete example. The text included figures and formulas. The selected duration of study time (15 min) was based on the results of a pilot study.

3.1.5. Measuring instruments

Absorption tendency. As in Experiment 1, we used the sub-scale absorption from Weibel et al.'s (2010) adapted version of the immersive tendency questionnaire (ITQ) (Witmer & Singer, 1998). The Cronbach's α estimate of reliability was 0.61, indicating a moderate internal consistency for the scale.

Suspension of disbelief. We assessed suspension of disbelief using the sub-scale suspension of disbelief of the MEC Spatial Presence Questionnaire (MEC-SPQ) (Vorderer et al., 2004). The MEC-SPQ is based on a conceptual model of spatial presence formation, including different constructs. The eight items of suspension of disbelief were rated on a 5-point Likert scale (1 = I do not agree at all, 5 = I fully agree) (sample item: "I wondered whether the VR presentation could really exist like this"). The scale has shown high reliability (Cronbach's $\alpha = 0.83$; Vorderer et al., 2004). Internal consistency was acceptable in our sample (Cronbach's $\alpha = 0.79$).

Learning outcome. To assess the learning outcome, we constructed a test that consisted of 36 questions (1 open question and 35 multiple-choice questions) about ANOVA for each measurement timepoint (baseline, immediate posttest, delayed posttest). As in Experiment 1, participants received one point for each correct response and zero points for each incorrect response. The response to the open question had to include the terms "within groups" and "between groups" to be classified as correct.

3.1.6. Data analysis

As for Experiment 1, data preparation, analysis, and visualization was performed in R (version 3.5.1; R Core Team). Packages used for the analysis and details of the Bayesian analysis were consistent with those of Experiment 1. Model notations including prior distributions and the results of model comparisons are available in the supplementary material.

Learning outcome was defined as the correctness of participants' responses to each of the 36 test questions (correct vs. incorrect). Using a GLMM with a binomial link function, we estimated the probability of a

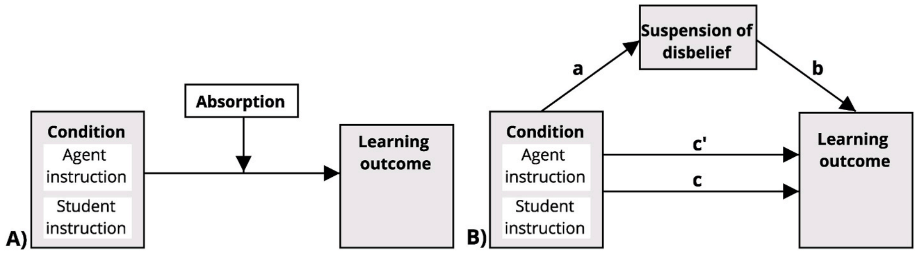

Fig. 4. Working models Experiment 2
Note. Working model A predicted that the learning outcome – as measured by an immediate and a delayed test – in the three conditions (IVR vs. desktop vs. control) would differ. Tendency towards emotional involvement and absorption tendency were expected to moderate the relationship between condition and learning outcome. In working model B, presence was expected to mediate the effect of media type (desktop vs. IVR) on learning outcome (path c represents the total effect, i.e., $\mathbf{c} = \mathbf{ab} + \mathbf{c}'$).

participant giving a correct response to a test question. We were primarily interested in the difference between conditions (*H1*). The first model (M1_learning outcome) thus included time as a dummy-coded effect to compare t1 (i.e., the baseline, which was the reference category) with t2 (immediate posttest) and t3 (delayed posttest), condition to compare the agent instruction condition with the student instruction condition, and the interaction term between time and condition as fixed effects. Furthermore, we included random intercepts to account for repeated measurements.

We further hypothesized that the effect of condition on learning outcome would be moderated by the participants' absorption tendency (*H3*). In the second model (M2_learning outcome), we thus added absorption tendency and the interaction terms between time, instruction, and absorption tendency as fixed effects to the first model (M1_learning outcome). Again, we included random intercepts to account for repeated measurement.

We expected to observe that participants in the student instruction condition would demonstrate better learning outcomes, due to a higher level of suspension of disbelief, than those in the agent instruction condition (*H2*). Using a multivariate GLMM, we estimated suspension of disbelief with a Gaussian link function. To compare the agent instruction condition with the student instruction condition, we included condition as a fixed effect in the model. We predicted the learning outcome with a binomial link function. We added time as a dummy-coded effect (t1 as the reference category vs. t2 vs. t3), condition (student instruction vs. agent instruction), suspension of disbelief, and their interaction terms as fixed effects. We included random intercepts to account for repeated measurements.

3.2. Results

3.2.1. Self-reported measures

Correlation analysis showed no relationship between absorption tendency and suspension of disbelief, $rho_{absorption:SoD} = 0.02; CI = [-0.09; 0.15]$. The means and standard deviations (SD) of the self-reported measures are reported in Table 5.

3.2.2. Learning outcome

We expected participants who had been instructed to imagine that they were explaining the text to a student peer to benefit more from teaching an agent in IVR than participants who had been instructed to explain the text to a virtual agent (*H1*). The first model (M1_learning outcome) thus estimated the probability that a participant gave a correct response to the test question as a function of time (t1 vs. t2 vs. t3) and condition (student instruction vs. agent instruction). The posterior estimates revealed an increased probability that a participant gave correct responses to the test questions at both t2, $\beta_{t2} = 0.61; SE = 0.09;\ CrI = [0.44; 0.78]$), and t3, $\beta_{t3} = 0.72; SE = 0.09;\ CrI = [0.54; 0.90]$, compared to t1. Contrary to our expectation, we found no interactions between time and condition. The instruction to imagine an interaction with a student peer did not result in a benefit in learning outcome from t1 to t2, $\beta_{t2:Agent} = -0.11; SE = 0.13;\ CrI = [-0.36; 0.13]$, or from t1 to t3, $\beta_{t3:Agent} = 0.13;\ SE = 0.13;\ CrI = [-0.13; 0.38]$, when compared to the agent instruction condition. Model parameters of M1_learning outcome are reported in Table 6.

Table 5
Means and standard deviations of self-reported measures (absorption tendency and suspension of disbelief).

	Mean (SD)	
	Student instruction	Agent instruction
Self-reported measures		
Suspension of disbelief	4.22 (0.10)	3.65 (0.16)
Absorption	3.34 (0.11)	3.25 (0.12)

Note. Suspension of disbelief = self-reported suspension of disbelief. Absorption = absorption tendency.

The second model (M2_learning outcome) estimated learning outcome depending on time (t1 vs. t2 vs t3), condition (student instruction vs. agent instruction), and participant's absorption tendency (*H3*). The posterior distributions revealed an interaction between time, condition, and absorption tendency. In the agent instruction condition, higher absorption scores were positively correlated with an increased probability of responding to the test questions correctly from t1 to t3, $\beta_{t3:agent:absorption} = 0.54;\ SE = 0.21; CI = [0.12; 0.95]$); learning gains at t3 depended on absorption tendency. In contrast, we found no such effect of absorption on learning gains from t1 to t2 $\beta_{t2:agent:absorption} = 0.15;\ SE = 0.21; CrI = [-0.26; 0.56]$. Model parameters of M2 learning outcome are reported in Table 7 and visualized in Fig. 5.

We further analyzed whether suspension of disbelief mediates between condition (student instruction vs. agent instruction) and learning outcome (*H2*). The multivariate GLMM (M_SoD) revealed that suspension of disbelief differed between conditions (path a). As expected, participants who were instructed to explain the text to a virtual agent reported a lower level of suspension of disbelief compared to those who were asked to imagine they were explaining the text to a student peer, $\beta_{SOD\_Agent} = -0.56;\ SE = 0.11; CrI = [-0.78; -0.35]$. Contrary to our expectation, we found no evidence that suspension of disbelief mediated between condition and learning outcome. The 95% CI of the indirect effect included zero for the comparison between t1 and t2, $ab\_t2 = 0.06; CI = [-0.12; 0.26]$), as well as the comparison between t1 and t3, $ab\_t3 = 0.14; CI = [-0.04; 0.34]$. Model parameters of M_SoD are provided in the supplementary material.

3.3. Discussion of experiment 2

In Experiment 2, in which all participants had been instructed to teach a computer-based agent in IVR, we compared the learning outcomes of those instructed to imagine they were explaining the text to a student peer to help her pass an upcoming test (student instruction) with those instructed to explain the text to a virtual agent (agent instruction). We found no overall difference between conditions with respect to learning outcome (*H1*). Interestingly, instruction manipulation was moderated by participants' absorption tendency: Low-absorption participants benefitted from being instructed to imagine they were explaining the text to a student peer, while high-absorption participants benefitted from being asked to explain the text to a virtual agent (*H3*). This finding supports our results from Experiment 1 (i.e., IVR fosters learning mostly in high-absorption individuals), but it further suggests that a simple change in instruction can support learning in low-absorption individuals.

The mere belief in a social interaction can change cognitive outcomes and behavior in virtual experiences (e.g., Okita et al., 2007). Regarding our result, it is likely that, in low-absorption individuals, the

Table 6
Regression coefficients (logit transformed posterior mean, standard error, and 95% credible intervals) of M1 learning outcome (probability of giving a correct response to a test question as a function of time and condition).

	Estimate	Est.Error	Q2.5	Q97.5
Group-Level Effects				
Participant (sd)	0.50	0.06	0.40	0.61
Population-Level Effects				
Intercept	−0.07	0.11	−0.28	0.15
T2	**0.61**	0.09	**0.44**	**0.78**
T3	**0.72**	0.09	**0.54**	**0.90**
Agent	−0.14	0.16	−0.46	0.16
T2:Agent	−0.11	0.13	−0.36	0.13
T3:Agent	0.13	0.13	−0.13	0.38

Note. Estimates with credible intervals not including zero are indicated in bold. T2 = immediate posttest. T3 = delayed posttest. Agent = instruction to explain the text to a virtual agent.

Table 7
Regression coefficients (logit transformed posterior mean, standard error, and 95% credible intervals) of M1 learning outcome (probability of giving a correct response to a test question as a function of time, condition, and absorption tendency).

	Estimate	Est.Error	Q2.5	Q97.5
Group-Level Effects				
articipant (sd)	0.50	0.06	0.40	0.61
Population-Level Effects				
Intercept	−0.07	0.11	−0.28	0.15
T2	**0.61**	0.09	**0.44**	**0.78**
T3	**0.72**	0.09	**0.54**	**0.90**
Agent	−0.14	0.16	−0.46	0.16
Absorption	0.10	0.18	−0.25	0.44
T2:Agent	−0.11	0.13	−0.37	0.14
T3:Agent	0.14	0.13	−0.11	0.40
T2:Absorption	−0.08	0.14	−0.37	0.21
T3:Absorption	0.02	0.26	−0.50	0.54
Agent:Absorption	−0.14	0.16	−0.46	0.16
T2:Agent:Absorption	0.15	0.21	−0.26	0.56
T3:Agent:Absorption	**0.54**	0.21	**0.12**	**0.95**

Note. T2 = immediate posttest. T3 = delayed posttest. Agent = instruction to explain the text to a virtual agent. Absorption = absorption tendency. Estimates with credible intervals not including zero are indicated in bold.

instruction to imagine that one is helping a student peer increased the feeling of interacting with an actual person and thus activated appropriate cognitive processing when teaching the agent. By contrast, high-absorption individuals may depend on external instructions less to believe in a social interaction because of their tendency to accept mediated contents as real (Baños et al., 1999).

Nevertheless, we found no correlation between absorption tendency and self-reported suspension of disbelief and suspension of disbelief was not related to learning outcome (*H2*). Thus, the interaction between instruction manipulation and absorption tendency might be explained better by other processes. For example, Okita et al. (2007) suggested that a simulated social interaction in IVR affects learning due to the belief in social relevance rather than the social belief per se. The authors found that increased arousal (as measured with skin conductance) correlated with better learning outcomes in a virtual question and answer session. Therefore, in low-absorption individuals, it is possible that the instruction to imagine they were helping a student peer pass a test benefitted their learning outcome by causing them to feel like they were undertaking a socially relevant cause within the virtual reality.

4 General discussion

High-immersion technologies enable the simulation of social interactions despite the absence of peers, creating innovative opportunities for remote learning. Across two experiments, we investigated whether IVR is effective in promoting learning when teaching a computer-based agent and whether factors such as immersion, presence, absorption tendency, tendency towards emotional involvement, and suspension of disbelief have an impact. Our most important finding is that the benefits of learning by teaching in IVR depend on the personality trait absorption tendency. More specifically, in participants who scored high on absorption tendency, learning outcomes were substantially better after they had taught an agent in IVR than they were after they had created a written summary. Moreover, in low-absorption participants, learning outcomes were better when the task instruction contained social cues.

Our results demonstrate that absorption tendency plays an essential role when it comes to learning by teaching a computer-based agent in IVR. Learning by teaching is inherently a social act and its efficacy is determined by the degree of interactivity (see Kobayashi, 2019). The benefits of learning by teaching are not the result of explaining per se, but depend on whether the learner is aware of the individual for whom the explanation is provided (Coleman et al., 1997; Fiorella & Mayer, 2013, 2014; Hoogerheide et al., 2016; Rittle-Johnson et al., 2008). Thus, whether learning by teaching a computer-based agent in IVR results in improved learning outcomes may depend on whether participants accept the agent as a real recipient. Our results suggest that the acceptance of a social interaction in IVR varies as a function of individuals' absorption tendency. High-absorption individuals may accept a virtual scenario with little effort because they are motivated and cognitively able to engage in sensory experiences such as virtual realities. Regarding our study, it is likely that high-absorption participants accepted the social interaction as real and thus behaved as if the agent was a real recipient that would really be affected by the quality of the explanation: They likely engaged in generative processing, such as monitoring whether their explanation was accurate and clear. In contrast, in low-absorption individuals, the acceptance of virtual scenarios may depend more on external factors. For example, the instruction to imagine one is helping a student peer pass her upcoming test could increase one's acceptance in a social interaction by creating a socially relevant context.

Taken together, we propose that possessing a high absorption

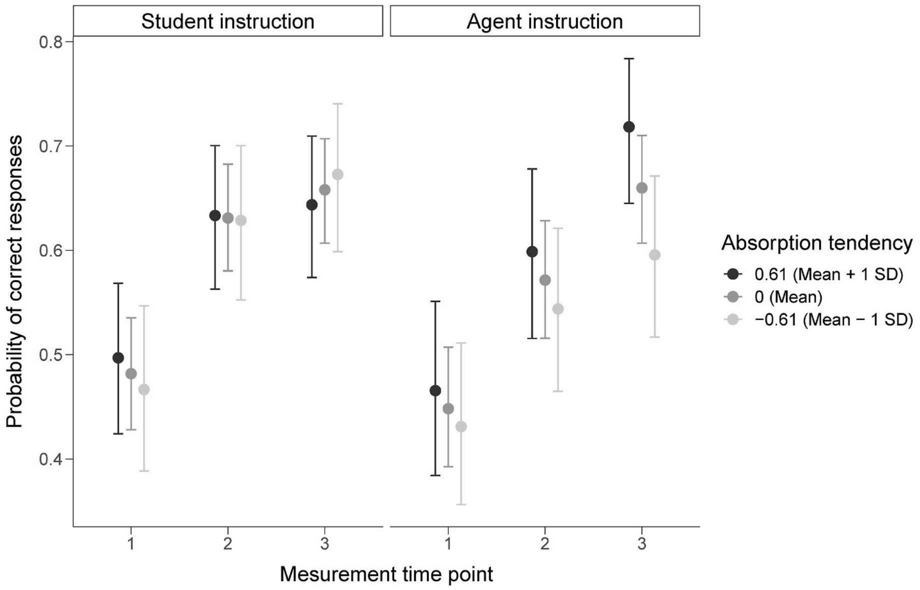

Fig. 5. Probability of correct responses as a function of condition, time, and absorption tendency
Note. The figure shows improved learning from t1 (= baseline) to t3 (= delayed posttest) in low-absorption participants who were instructed to imagine they were explaining the text to a student peer (student instruction condition), whereas high-absorption individuals benefited most when they were to explain the text to a virtual agent (agent instruction condition). Absorption tendency is centered around the grand mean.

tendency and instructions containing social cues for low-absorption individuals support the benefits of learning by teaching in IVR by triggering the acceptance of a social interaction within the virtual reality.

Nevertheless, although we found absorption tendency to be related to learning by teaching in IVR in both experiments, it is noteworthy that our results are based on rather small sample sizes. More research is needed to investigate whether and why IVR supports learning when teaching a computer-based agent. Moreover, it should be noted that we found no effect of suspension of disbelief on learning. Whether a person accepts a virtual scenario as real may depend on factors besides the active suppression of contradicting stimuli and inner thoughts. For example Slater (2009) proposed plausibility illusion, together with spatial presence (place illusion) to determine, realistic responses to virtual reality. By plausibility illusion the author refers to the sensation that the virtual scenario is actually taking place although the user knows that what he/she is experiencing is not real. Plausibility illusion, in turn, could be based on coherence (Skarbez, Neyret, Brooks, Slater, & C, 2017, see also Hartmann & Hofer, 2022), which refers to the extent to which a virtual scenario is consistent with the users' expectations. Thus, it is possible that in our experiment high absorption individuals perceived the scenario (i.e., teaching an agent) as plausible without much effort whereas low-absorption individuals may depend on external factors (i.e., task instruction including social cues) to increase coherence of such a scenario. However, more research is needed to investigate the role of plausibility and coherence in efficient learning by teaching in IVR. Future research should consider more implicit measurements, such as gaze behavior (e.g., joint attention) and communication behavior to investigate whether perceived plausibility determines realistic responses in IVR. Indeed, there is some evidence that using pronouns like *you* or *we*, rather than *she* or *it*, in interactions with teachable agents in 2D VR correlates with students' improved learning outcomes (Ogan et al., 2012). In addition, future studies will need to consider the impact of novelty on cognitive processes in VR technologies. For example Makransky et al. (2019) and Meyer et al. (2019) suggest cognitive overload and thus reduced learning when IVR technology is used for the first time. However, IVR could also lead to a reverse effect by reducing extraneous processing and surrounding participants with a sensory context that is related exclusively to the instructional goal of a task (e.g., teaching an agent). There is an ongoing debate on whether cognitive processes (e.g., learning) are affected by novelty or information richness of IVR or whether high-immersion technologies entail certain characteristics that are appropriate and support cognitive processing in students (Han, Zheng, & Ding, 2021; Liu, Bhagat, Gao, Chang, & Huang, 2017, pp. 105–130; Makransky et al., 2019; Meyer et al., 2019; Parong & Mayer, 2018). Adding conditions with varying degrees of pre-training in IVR would be important to investigate the relative influence of novelty on non-interactive learning by teaching in both IVR and desktop VR. Moreover, there is a general need for more investigations of long-term learning. Generative processing contributes to deep learning which might be especially visible in delayed comprehension tests (see Fiorella & Mayer, 2013). Whereas Experiment 2 supports this idea, Experiment 1 revealed a beneficial effect of IVR in high-absorption individuals on learning especially immediately after teaching rather than in the longer term (after one week).

In conclusion, the study provides evidence that IVR supports the benefits of learning by teaching in high-absorption individuals. This finding highlights the crucial role of personal traits in social IVR. Learning by teaching in IVR is not beneficial per se; teaching a computer-based agent may be a suitable learning technique for one person, but not for another. However, giving a learner the concrete instruction to imagine an interaction with an actual person can compensate for user characteristics in such a setting, thus supporting the idea that the belief in virtual social interactivity can activate generative processing despite the absence of peers. A common goal in future IVR learning applications should therefore be to create a belief in the virtual reality. This is especially important as distance learning has become more prevalent these days.

Funding

This research was funded by UniDistance Suisse.

Author statement

S.C. wrote the main manuscript text, prepared all tables and figures, and formally analyzed the data. All authors reviewed the manuscript.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.learninstruc.2022.101716.

References

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

Baños, R. M., Botella, C., García-Palacios, A., Villa, H., Perpiñá, C., & Gallardo, M. (1999). Psychological variables and reality judgment in virtual environments: The roles of absorption and dissociation. *CyberPsychology and Behavior, 2*(2), 143–148. https://doi.org/10.1089/cpb.1999.2.143

Baños, R. M., Botella, C., Rubió, I., Quero, S., García-Palacios, A., & Alcañiz, M. (2008). Presence and emotions in virtual environments: The influence of stereoscopy. *CyberPsychology and Behavior, 11*(1), 1–8. https://doi.org/10.1089/cpb.2007.9936

Baños, R. M., Etchemendy, E., Castilla, D., García-Palacios, A., Quero, S., & Botella, C. (2012). Positive mood induction procedures for virtual environments designed for elderly people. *Interacting with Computers, 24*(3), 131–138. https://doi.org/10.1016/j.intcom.2012.04.002

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

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

Caruana, N., de Lissa, P., & McArthur, G. (2016). Beliefs about human agency influence the neural processing of gaze during joint attention. *Social Neuroscience, 12*(2), 194–206. https://doi.org/10.1080/17470919.2016.1160953

Caruana, N., Mcarthur, G., Woolgar, A., & Brock, J. (2017). Detecting communicative intent in a computerised test of joint attention. *PeerJ, 5*(e2899). https://doi.org/10.7717/peerj.2899

Cipresso, P., Giglioli, I. A. C., Raya, M. A., & Riva, G. (2018). The past, present, and future of virtual and augmented reality research: A network and cluster analysis of the literature. *Frontiers in Psychology, 9*(NOV). https://doi.org/10.3389/fpsyg.2018.02086

Coleman, E. B., Brown, A. L., Rivkin, I. D., & Coleman, E. B. (1997). The effect of instructional explanations on learning from scientific texts. *The Journal of the Learning Sciences, 6*(4), 347–365.

Coxon, M., Kelly, N., & Page, S. (2016). Individual differences in virtual reality: Are spatial presence and spatial ability linked? *Virtual Reality, 20*, 203–212. https://doi.org/10.1007/s10055-016-0292-x

Cummings, J. J., & Bailenson, J. N. (2015). How immersive is enough ? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology, 1–38*. https://doi.org/10.1080/15213269.2015.1015740

Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y., & Muehlberger, A. (2015). The impact of perception and presence on emotional reactions: A review of research in virtual reality. In *Frontiers in psychology* (Vol. 6, p. 26). Frontiers Research Foundation. https://doi.org/10.3389/fpsyg.2015.00026.

Duran, D. (2017). Learning-by-teaching. Evidence and implications as a pedagogical mechanism. *Innovations in Education & Teaching International, 54*(5). https://doi.org/10.1080/14703297.2016.1156011

Fiorella, L., & Kuhlmann, S. (2019). Creating drawings enhances learning by teaching. *Journal of Educational Psychology, 112*(4), 811–822. https://doi.org/10.1037/edu0000392

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology, 38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Fiorella, L., & Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology, 39*(2), 75–85. https://doi.org/10.1016/j.cedpsych.2014.01.001

de Gelder, B., Kätsyri, J., & de Borst, A. W. (2018). Virtual reality and the new psychophysics. *British Journal of Psychology, 109*(3), 421–426. https://doi.org/10.1111/bjop.12308

Gorini, A., Capideville, C. S., De Leo, G., Mantovani, F., & Riva, G. (2011). The role of immersion and narrative in mediated presence: The virtual hospital experience. *Cyberpsychology, Behavior, and Social Networking, 14*(3), 99–105. https://doi.org/10.1089/cyber.2010.0100

Han, J., Zheng, Q., & Ding, Y. (2021). Lost in virtual reality? Cognitive load in high immersive VR environments. *Journal of Advances in Information Technology, 12*(4), 302–310. https://doi.org/10.12720/jait.12.4.302-310

Hartmann, T., & Hofer, M. (2022). I know it is not real (and that matters) media awareness vs. Presence in a parallel processing account of the VR experience. *Frontiers in Virtual Reality, 3*(April), 1–16. https://doi.org/10.3389/frvir.2022.694048

Hofer, M., Hartmann, T., Eden, A., Ratan, R., & Hahn, L. (2020). The role of plausibility in the experience of spatial presence in virtual environments. *Frontiers in Virtual Reality, 1*(April). https://doi.org/10.3389/frvir.2020.00002

Hofer, M., Wirth, W., Kuehne, R., Schramm, H., & Sacau, A. (2012). Structural equation modeling of spatial presence: The influence of cognitive processes and traits. *Media Psychology, 15*(4), 373–395. https://doi.org/10.1080/15213269.2012.723118

Hoogerheide, V., Deijkers, L., Loyens, S. M. M., Heijltjes, A., & van Gog, T. (2016). Gaining from explaining: Learning improves from explaining to fictitious others on video, not from writing to them. *Contemporary Educational Psychology, 44–45*, 95–106. https://doi.org/10.1016/j.cedpsych.2016.02.005

Hoogerheide, V., Loyens, S. M. M., & van Gog, T. (2014). Effects of creating video-based modeling examples on learning and transfer. *Learning and Instruction, 33*, 108–119. https://doi.org/10.1016/j.learninstruc.2014.04.005

Hoogerheide, V., Renkl, A., Fiorella, L., Paas, F., & van Gog, T. (2019). Enhancing example-based learning: Teaching on video increases arousal and improves problem-solving performance. *Journal of Educational Psychology, 111*(1), 45–56. https://doi.org/10.1037/edu0000272

Hoogerheide, V., Visee, J., Lachner, A., & van Gog, T. (2019). Generating an instructional video as homework activity is both effective and enjoyable. *Learning and Instruction, 64*. https://doi.org/10.1016/j.learninstruc.2019.101226

Jacob, L., Lachner, A., & Scheiter, K. (2020). Learning by explaining orally or in written form? Text difficulty matters. *Learning and Instruction, 68*, Article 101344. https://doi.org/10.1016/j.learninstruc.2020.101344

Jacob, L., Lachner, A., & Scheiter, K. (2021). Does increasing social presence enhance the effectiveness of writing explanations? *PLoS One, 16*(4 April 2021), 1–20. https://doi.org/10.1371/journal.pone.0250406

Kim, T., & Biocca, F. (1997). Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion. *Journal of Computer-Mediated Communication, 3*(2). https://doi.org/10.1111/j.1083-6101.1997.tb00073.x

Kobayashi, K. (2019). Interactivity: A potential determinant of learning by preparing to teach and teaching. *Frontiers in Psychology, 9*(JAN), 1–6. https://doi.org/10.3389/fpsyg.2018.02755

Krijn, M., Emmelkamp, P. M. G., Biemond, R., De Wilde De Ligny, C., Schuemie, M. J., & Van Der Mast, C. A. P. G. (2004). Treatment of acrophobia in virtual reality: The role of immersion and presence. *Behaviour Research and Therapy, 42*(2), 229–239. https://doi.org/10.1016/S0005-7967(03)00139-6

Lachner, A., Backfisch, I., Hoogerheide, V., van Gog, T., & Renkl, A. (2019). Timing matters! Explaining between study phases enhances students' learning. *Journal of Educational Psychology, 112*(4), 841–853. https://doi.org/10.1037/edu0000396

Lachner, A., Hoogerheide, V., van Gog, T., & Renkl, A. (2021). Learning-by-Teaching without audience presence or interaction: When and why does it work?. In *Educational psychology review*. Springer US. https://doi.org/10.1007/s10648-021-09643-4.

Ling, Y., Nefs, H. T., Brinkman, W. P., Qu, C., & Heynderickx, I. (2013). The relationship between individual characteristics and experienced presence. *Computers in Human Behavior, 29*(4), 1519–1530. https://doi.org/10.1016/j.chb.2012.12.010

Liu, D., Bhagat, K. K., Gao, Y., Chang, T.-W., & Huang, R. (2017). *The potentials and trends of virtual reality in education*. Singapore: Springer. https://doi.org/10.1007/978-981-10-5490-7_7

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2). https://doi.org/10.1111/j.1083-6101.1997.tb00072.x

Lui, M., McEwen, R., & Mullally, M. (2020). Immersive virtual reality for supporting complex scientific knowledge: Augmenting our understanding with physiological monitoring. *British Journal of Educational Technology, 51*(6), 2181–2199. https://doi.org/10.1111/bjet.13022

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the bayesian framework. *The Journal of Open Source Software, 4*(40), 1541. https://doi.org/10.21105/joss.01541

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

Makransky, G., 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

Markowitz, D. M., Laha, R., Perone, B. P., Pea, R. D., & Bailenson, J. N. (2018). Immersive Virtual Reality field trips facilitate learning about climate change. *Frontiers in Psychology, 9*(2364). https://doi.org/10.3389/fpsyg.2018.02364

Mayer, R. E. (2009). *Multimedia learning.* Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005.

Mayer, R. E. (2020). *Multimedia learning.* Cambridge University Press. https://doi.org/10.1017/9781316941355

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 in Education, 140*, Article 103603. https://doi.org/10.1016/j.compedu.2019.103603

Minsky, M. (1980). *Telepresence. OMNI, 2*(9), 45–51.

Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94*(3), 598–610. https://doi.org/10.1037/0022-0663.94.3.598

Muckler, V. C. (2017). Exploring suspension of disbelief during simulation-based learning. *Clinical Simulation in Nursing, 13*(1), 3–9. https://doi.org/10.1016/j.ecns.2016.09.004

Mulders, M., Buchner, J., & Kerres, M. (2020). A framework for the use of immersive virtual reality in learning environments. *International Journal of Emerging Technologies in Learning (IJET), 15*(24), 208. https://doi.org/10.3991/ijet.v15i24.16615

Murray, C. D., Fox, J., & Pettifer, S. (2007). Absorption, dissociation, locus of control and presence in virtual reality. *Computers in Human Behavior, 23*, 1347–1354. https://doi.org/10.1016/j.chb.2004.12.010

Nicovich, S. G., Boller, G. W., & Cornwell, B. T. (2005). Experienced presence within computer-mediated communications: Initial explorations on the effects of gender with respect to empathy and immersion. *Computer-Mediated C Ommunication, 10*(2). https://doi.org/10.1111/j.1083-6101.2005.tb00243.x

Ogan, A., Finkelstein, S., Mayfield, E., D'Adamo, C., Matsuda, N., & Cassell, J. (2012). Oh, dear Stacy!" Social interaction, elaboration, and learning with teachable agents. *Conference on Human Factors in Computing Systems - Proceedings*, 39–48. https://doi.org/10.1145/2207676.2207684

Okita, S. Y., Bailenson, J., & Schwartz, D. L. (2007). The mere belief of social interaction improves learning. *Proceedings of the Annual Meeting of the Cognitive Science Behaviour, 29*(29), 1355–1360.

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

Parsons, T. D., Barnett, M., & Melugin, P. R. (2015). Assessment of personality and absorption for mediated environments in a college sample. *Cyberpsychology, Behavior, and Social Networking, 18*(12), 752–756. https://doi.org/10.1089/cyber.2015.0336

Ploetzner, R., Dillenbourg, P., Preier, M., & Traum, D. (1999). Learning by explaining to oneself and to others. In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (pp. 103–121). Elsevier.

Price, M., & Anderson, P. (2007). The role of presence in virtual reality exposure therapy. *Journal of Anxiety Disorders, 21*(5), 742–751. https://doi.org/10.1016/j.janxdis.2006.11.002

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 in Education, 147*(November 2019), Article 103778. https://doi.org/10.1016/j.compedu.2019.103778

Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). *Die Erfassung des Flow-Erlebens* *.

Rittle-Johnson, B., Saylor, M., & Swygert, K. E. (2008). Learning from explaining: Does it matter if mom is listening? *Journal of Experimental Child Psychology, 100*, 215–224. https://doi.org/10.1016/j.jecp.2007.10.002

Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., et al. (2007). Affective interactions using virtual reality: The link between presence and emotions. *CyberPsychology and Behavior, 10*(1), 45–56. https://doi.org/10.1089/cpb.2006.9993

Rodgers, W. M., Conner, M., & Murray, T. C. (2008). Distinguishing among perceived control, perceived difficulty, and self-efficacy as determinants of intentions and behaviours. *British Journal of Social Psychology, 47*(4), 607–630. https://doi.org/10.1348/014466607X248903

Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research, 77*(4), 534–574. https://doi.org/10.3102/0034654307309920

Roscoe, R. D., & Chi, M. T. H. (2008). Tutor learning: The role of explaining and responding to questions. *Instructional Science, 36*(4), 321–350. https://doi.org/10.1007/s11251-007-9034-5

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

Sacau, A., Laarni, J., & Hartmann, T. (2008). Influence of individual factors on presence. *Computers in Human Behavior, 24*(5), 2255–2273. https://doi.org/10.1016/j.chb.2007.11.001

Sas, C., & O'Hare, G. M. P. (2003). Presence equation: An investigation into cognitive factors underlying presence. *Presence: Teleoperators and Virtual Environments, 12*(5), 523–537. https://doi.org/10.1162/105474603322761315

Schmid Mast, M., Kleinlogel, E. P., Tur, B., & Bachmann, M. (2018). The future of interpersonal skills development: Immersive virtual reality training with virtual humans. *Human Resource Development Quarterly, 29*(2), 125–141. https://doi.org/10.1002/hrdq.21307

Skarbez, R., Neyret, S., Brooks, F. P., Slater, M., & C, W. M. (2017). A psychophysical experiment regarding components of the plausibility illusion. *IEEE Transactions on Visualization and Computer Graphics, 23*(4), 1369–1378. https://doi.org/10.1109/TVCG.2017.2657158

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI, 3*(74). https://doi.org/10.3389/frobt.2016.00074

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 6*, 603–616. http://www.cs.ucl.ac.uk/staff/M.Slater.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication, 42*, 72–92.

Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences ("absorption"), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology, 83*(3), 268–277. https://doi.org/10.1037/h0036681

Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, L., et al. (2004). MEC spatial presence questionnaire (MECSPQ): Short documentation and instructions for application. In *Report to the European community, project presence: Mec (IST-2001-37661)*. http://www.ijk.hmt-hannover.de/presence.

Wallach, H. S., Safir, M. P., & Samana, R. (2010). Personality variables and presence. *Virtual Reality, 14*(1), 3–13. https://doi.org/10.1007/s10055-009-0124-3

Weber, S., Mast, F. W., & Weibel, D. (2021). Experiencing presence in a gaming activity improves mood after a negative mood induction. *International Journal of Gaming and Computer-Mediated Simulations, 12*(4). https://doi.org/10.4018/IJGCMS.2020100101

Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments, 24*(6), 1319–1333. https://doi.org/10.1080/10494820.2014.994533

Weibel, D., Schmutz, J., Pahud, O., & Wissmath, B. (2015). Measuring spatial presence: Introducing and validating the pictorial presence SAM. *Presence: Teleoperators and Virtual Environments, 24*(1), 44–61. https://doi.org/10.1162/PRES

Weibel, D., Stricker, D., & Wissmath, B. (2012). The use of a virtual learning centre in the context of a university lecture: Factors influencing satisfaction and performance. *Interactive Learning Environments, 20*(1), 77–87. https://doi.org/10.1080/10494821003755502

Weibel, D., & Wissmath, B. (2011). Immersion in computer games: The role of spatial presence and flow. *International Journal of Computer Games Technology, 2011*. https://doi.org/10.1155/2011/282345

Weibel, D., Wissmath, B., & Mast, F. W. (2010). Immersion in mediated environments: The role of personality traits. *Cyberpsychology, Behavior, and Social Networking, 13*(3), 251–256. https://doi.org/10.1089/cpb.2009.0171

Weibel, D., Wissmath, B., & Mast, F. W. (2011). The role of cognitive appraisal in media-induced presence and emotions. *Cognition & Emotion, 25*(7), 1291–1298. https://doi.org/10.1080/02699931.2010.543016

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., et al. (). A process model of the formation of spatial presence experiences. *Media Psychology, 9*(3), 493–525. https://doi.org/10.1080/15213260701283079

Wirth, W., Hofer, M., & Schramm, H. (2012). The role of emotional involvement and trait absorption in the formation of spatial presence. *Media Psychology, 15*, 19–43. https://doi.org/10.1080/15213269.2011.648536

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*(3), 225–240. https://doi.org/10.1162/105474698565686