Pedagogical Agents in Educational VR: An in the Wild Study

3 authors:

Gustav Bøg Petersen
University of Copenhagen
8 PUBLICATIONS **32** CITATIONS

SEE PROFILE

Aske Mottelson
University of Copenhagen
17 PUBLICATIONS **121** CITATIONS

SEE PROFILE

Guido Makransky
University of Copenhagen
64 PUBLICATIONS **1,264** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Instructional Design for Virtual Reality View project

Pedagogical Agents in Educational VR: An in the Wild Study

Gustav Bøg Petersen
gbp@psy.ku.dk
Department of Psychology
University of Copenhagen
Copenhagen, Denmark

Aske Mottelson
amot@psy.ku.dk
Department of Psychology
University of Copenhagen
Copenhagen, Denmark

Guido Makransky
gm@psy.ku.dk
Department of Psychology
University of Copenhagen
Copenhagen, Denmark

ABSTRACT

Pedagogical agents are theorized to increase humans' effort to understand computerized instructions. Despite the pedagogical promises of VR, the usefulness of pedagogical agents in VR remains uncertain. Based on this gap, and inspired by global efforts to advance remote learning during the COVID-19 pandemic, we conducted an educational VR study in-the-wild ($N = 161$). With a $2 \times 2 + 1$ between subjects design, we manipulated the appearance and behavior of a virtual museum guide in an exhibition about viruses. Factual and conceptual learning outcomes as well as subjective learning experience measures were collected. In general, participants reported high enjoyment and had significant knowledge acquisition. We found that the agent's appearance and behavior impacted factual knowledge gain. We also report an interaction effect between behavioral and visual realism for conceptual knowledge gain. Our findings nuance classical multimedia learning theories and provide directions for employing agents in immersive learning environments.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; Empirical studies in HCI; • **Applied computing** → *Interactive learning environments*.

KEYWORDS

Immersive Virtual Reality; Educational Technology; Learning; Cognitive Load, Pedagogical Agents

1 INTRODUCTION

Imagine visiting your favorite museum in the confines of your own home through the technology of virtual reality (VR); not having to stand in line to explore exhilarating exhibitions, and at the same time having your own private museum guide at hand. The goal of this study is to examine the role of a virtual museum guide (henceforth, a pedagogical agent) and its design in producing learning during a tour in an educational VR museum.

VR has recently received increased attention as an educational tool [14]. Theoretically, the contents of VR experiences are only bounded by the limits of our imagination [46]. In practice, however, educational VR often takes the shape of simulations or depictions of reality, and VR is typically combined with face-to-face instruction when used in the classroom (e.g., [42]). Such use of VR is in line with the blended learning systems approach [10].

Blended learning systems are in certain situations not ideal, because of cost or practicalities of mixing digital and classical learning. There are situations where real-life circumstances force us to rely solely on remote technology for instruction. This, for instance, is evident during the coronavirus disease 2019 (COVID-19) pandemic, where lock down has forced many teachers and students to transition into a remote learning style [6]. New approaches to remote teaching, including use of pedagogical agents and VR for remote learning, are in such situations desired.

Unlike avatars, that are controlled by humans [8], pedagogical agents are anthropomorphous computer-controlled virtual characters that are used in online learning environments to serve various instructional goals [49]. The purpose of using pedagogical agents is to mimic the social processes that usually take place in real-life teaching. According to multimedia learning theories, people are inclined to treat computerized agents as social partners if they exhibit social cues [33]. This, in turn, should motivate the learner to make sense of the presented material [33]. However, there is disagreement with regards to the usefulness of pedagogical agents. Opponents claim that the visual appearance of a pedagogical agent adds unnecessary distraction to the learning experience [45]. This view is consistent with the notion that the human working memory is limited and that embellished materials therefore impede learning [17]. For VR, where it is possible to render pedagogical agents in realistic 3D with higher behavioral realism compared to traditional media, their effect on learning is even less supported.

Inspired by the efforts to reorganize teaching into online formats during the COVID-19 pandemic, we created a virtual museum on the topic of viral diseases. The museum was accessed by participants 'in the wild' using their own Oculus Quest head-mounted display (HMD), an approach which has been proven feasible in previous VR research [23, 37, 47]. We were interested in how the addition of a social entity to the museum impacted participants' learning outputs and subjective ratings of the learning experience. Specifically, we investigated the role of the pedagogical agent's appearance and behavior. The hypothesis behind this was that a more realistic

pedagogical agent would induce a stronger sense of interacting with a social partner, and in turn enhance learning.

Our results show a large effect on learning (Cohen's d = 1.4). They also show that participants enjoyed the learning (M = 3.9/5). We report that the presence of a pedagogical agent diminished factual knowledge acquisition; yet, for learning conceptual knowledge, an interaction effect between visual and behavioral realism shows that the presence of pedagogical agents in educational VR might be worthwhile.

2 BACKGROUND AND RELATED WORK

2.1 VR in Education

Recent work in educational psychology and HCI emphasize the potentials of learning in immersive environments (e.g., [2, 19, 29, 31, 42]); these works together show how instructional and visual design of virtual environments are imperative for learning outcomes.

Jensen and Konradsen [14] recently conducted a review of research on education and training with the current generation of HMDs. They argue that VR does not automatically cause learning but rather provides a possibility for accessing simulations that might induce learning. Furthermore, with regard to cognitive skills acquisition, immersive experiences risk overwhelming learners unless they are designed specifically for the purpose and not overly interactive [14]. According to Jensen and Konradsen [14], the biggest barriers to adopting VR in education and training are a lack of content as well as the technical skills necessary, which challenge many instructors. Radianti et al. [43] similarly reviewed research on VR in higher education, showing how VR is used in many different fields, but for teaching engineering and computer science in particular. Furthermore, VR is used for teaching various types of knowledge, including procedural, practical, and declarative knowledge [43].

Studies that compare the relative effectiveness of media in promoting educational outcomes can be used to illuminate the advantages of VR in education compared to traditional media. Findings show that VR is an effective medium for promoting quality experiences, yet with mixed results for objective learning outcomes compared to traditional media [25, 29, 35].

Comparisons of the learning effects of immersive and non-immersive learning environments, respectively, show that immersion is associated with higher self-efficacy, enjoyment, and interest [28, 35]. Parong and Mayer [40] compared the effects of administering a biology lesson in VR or as a slideshow on a PC. They found that the VR group reported significantly higher ratings of motivation, interest, engagement, and affect than the group who used a PC. The VR group, however, scored worse on a post-test of factual knowledge [40]. Similarly, Makransky et al. [29] found that learners gained more knowledge from a lesson when the material was presented via a PC than via VR. Parong and Mayer [39] showed that lower retention scores when learning via VR, as opposed to desktop, is related to extraneous cognitive load. These findings indicate that VR may tax the cognitive resources of learners heavily. Careful attention to the instructional design of educational VR lessons is therefore necessary.

2.2 Social agency theory

Social agency theory is a frequently used theoretical framework from multimedia learning that explains the use of pedagogical agents [32]. According to social agency theory, social cues in multimedia lessons can prime a feeling of social presence in learners, which leads to deeper cognitive processing and more learning [33]. Social presence refers to a psychological state in which virtual social actors are experienced as actual social actors in either sensory or non-sensory ways [21]. Social agency theory states that people are attentive to social cues when interacting with computerized agents, and that these may induce a feeling of interacting with another social being. This will activate social rules such as the *cooperation principle*, meaning that the learner will try to make sense of the instructional message; consequently, with a social agent, the learner will make a deeper effort to understand and process the computerized instructional message [33].

According to Mayer [33], several kinds of social cues can induce a feeling of social presence in the learner. Two such cues are specifically relevant to educational VR: image cues (i.e., displaying a pedagogical agent who narrates the material) and embodiment cues (i.e., making the pedagogical agent display human-like behavior such as gesturing, movement, eye contact, etc.). Taken together, embodied pedagogical agents should, theoretically, give rise to social presence and, therefore, deeper processing and better learning [33]. Pedagogical agents have also been criticized for adding unnecessary complexity to learning environments. According to this view, the visual presence of a pedagogical agent is merely a seductive detail, and what really matters is the narration it provides [45].

In reviewing the effect of adding image and embodiment cues to multimedia lessons, Mayer [33] concluded that there is moderate evidence that embodiment cues improve learning (d = 0.36), but less support for the effect of adding a speaker's image to the screen (d = 0.20). This can be formulated as the image principle: people do not necessarily learn more deeply from a multimedia lesson when the speaker's image is on the screen compared to not on the screen [33]; and the embodiment principle: people learn more deeply when pedagogical agents display human-like gesturing, movement, eye contact, and facial expressions [33]. Importantly, however, the principles are based on findings in classic multimedia learning environments (such as desktop computers), often with non-humanoid pedagogical agents or humanoid pedagogical agents of low realism. Consequently, there is a need to revisit the the image and embodiment principles with 3D and immersive media such as VR [34]. This is relevant since VR is known to induce feelings of social presence in learners [30] and thereby potentially deep cognitive processing. Here follows a description of the current state of research regarding agents in VR and other media, as well as related work on social presence in virtual environments.

2.3 Agents in less immersive media

The literature contains a number of relevant HCI studies using non HMD-based technology (e.g., projection systems or PC) that can inform the design of pedagogical agents in VR; this section provides a short description of a select few. Kartiko et al. [15] used a projection system to test the impact of virtual actors' visual complexity (i.e., amount of visual information) on science learning, and
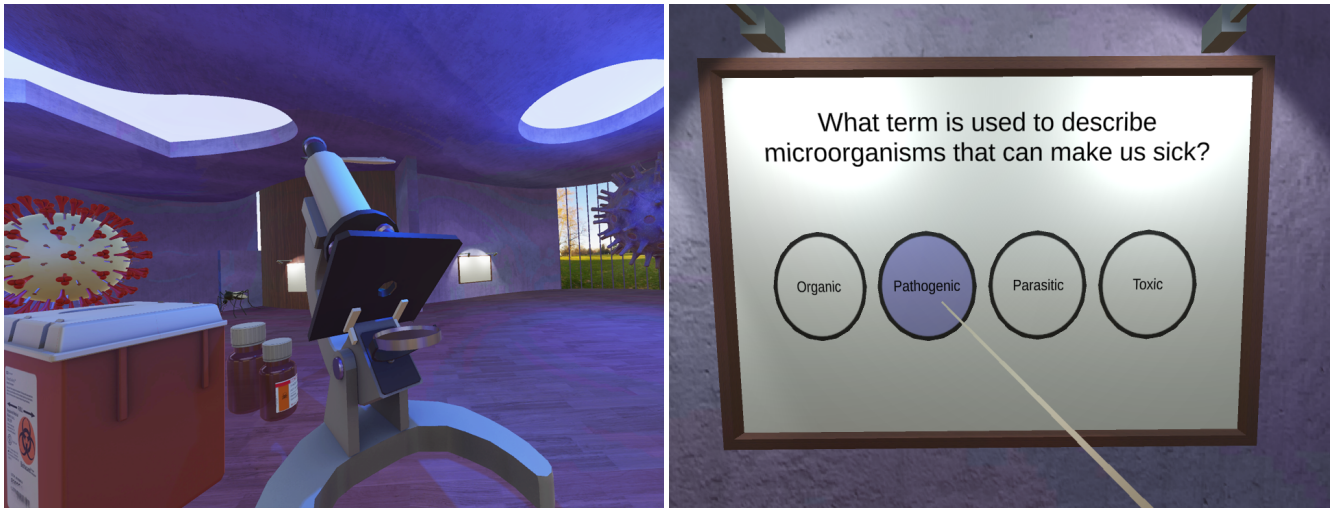
Figure 1: Participants in this study were immersed in a virtual museum with an exhibition on the topic of viruses (left). Participants could freely walk or teleport around the museum that would feature 3D objects relevant to the topic of the exhibition. A museum tour guide would lead the participant and use white boards with animated presentations to explain the topic. Before and after experiencing the museum, participants responded to a knowledge test on the subject of the exhibition (right).

found no effect of manipulation with regard to learning outcomes. Wang et al. [50] explored the impact of different virtual agents presented via augmented reality (AR) on a simple object finding task. Although there were no effect of manipulation on completion time, participants gazed more often at human-like agents compared to non-human. Kim et al. [16] investigated users' perceptions of AR agents acting as lab assistants, and reported that participants displayed most trust and social presence when an agent had a human body and was capable of speech, gestures, and locomotion. During a problem-solving task on PC, Groom et al. [11] found that participants liked a human virtual agent the most when it displayed inconsistent behavioral realism compared to agents either consistently low or high in behavioral realism (however, scores were generally low). The same pattern was reported with regard to participants' levels of comfort. Kizilcec et al. [18] compared the impact of presenting the instructor's face strategically (i.e., when learners should focus on spoken text independent of lecture slides) vs. constantly during video instruction. Results indicated that strategic presentation induced higher social presence relative to constant presentation but no difference in achieved course grade. Taking a more practice-oriented approach, Veletsianos et al. [48] provide a framework, 'EnALI', to enhance agent-learner interactions encompassing 15 guidelines, including suggestions regarding agent characteristics such as designing them to communicate in a polite and positive manner.

2.4 Pedagogical agents in VR

Only a few studies have been conducted concerning pedagogical agents in VR and their effect on learning outcomes and experiences (e.g., [30, 44]). Typically, these do not include a no-agent condition, making it difficult to be conclusive about the image and embodiment principles in VR; specifically if the presence of the agent has an effect on the learned subject. One study investigated the

effect of realism of agents on learning in VR exhibitions [44]. Interestingly, participants rated the absence of an agent as higher in 'humanness' compared to a realistic agent; the study, however, did not find a significant effect on learning, possibly because of low power. Makransky et al. [30] designed two pedagogical agents for VR and used them to teach middle school students about laboratory safety. A robot-like drone was intended to be more appealing to boys; a young female scientist was posited to be more inviting to girls. They demonstrated that boys learned better with the drone, and that girls learned better with the female scientist. This suggests that gender-specific design of pedagogical agents could be important in educational VR. The general lack of empirical studies on pedagogical agents in immersive learning environments makes it difficult to reason about the impact of virtual agents' appearance and behavior on learning outcomes.

2.5 Social presence in virtual environments

Social agency theory proposes that social presence during multimedia learning is a central mechanism that leads to deeper cognitive processing and consequently better learning outcomes. Oh et al. [38] recently conducted a systematic review of the predictors of social presence in virtual environments. Their findings emphasize the importance of visual representation of virtual communication partners. Specifically, they found that (i) people feel higher levels of social presence when a visual representation is available rather than not; (ii) behavioral realism is a powerful predictor of social presence; (iii) there are mixed results with regard to the effect of fidelity and human-likeness on social presence (some studies show an effect, others none); and (iv) that a 'consistency effect' possibly exists – level of behavioral realism should be consistent with level of visual fidelity to maintain high levels of social presence [38]. When applying these findings to pedagogical agents and social

agency theory, there are some similarities. First, there is empirical evidence for the effect of image cues as people generally feel higher social presence when the speaker is visible. Second, there is empirical evidence for the influence of embodiment cues, as people generally feel higher social presence when the speaker displays realistic behavior. Furthermore, Oh et al. [38] show the potential for fidelity (i.e., visual realism) and consistency between behavioral and visual realism in pedagogical agents to be important sources of social presence and thereby possibly learning.

3 EXPERIMENT

Based on social agency theory and research on social presence in virtual environments, four different pedagogical agents for the VR Museum were constructed. These agents varied by a combination of two levels of visual and behavioral realism, for a total of four agent conditions. Additionally, a control condition without a pedagogical agent (hence, only narration) was included. The participants signed up for the experiment online, and installed the experimental application on their Oculus Quest device. Self-report inside the VR application collected variables related to learning (i.e., knowledge gain and enjoyment), in addition to variables related to the pedagogical agent (i.e., humanness, attractiveness, and social presence). Lastly, free text feedback was collected after study completion in participants' web browser. See Table 1 for a full list of the included variables. All procedures performed during the study were approved by the institutional ethical committee.

3.1 Preregistration: Hypotheses and analyses

We preregistered the experimental study alongside hypotheses, study plan, and a statistical analyses plan (see https://osf.io/7wsya). The data collection, study design, and statistical analyses followed the preregistration, only with a minor deviation, as we, due to an unexpected high participation interest collected slightly more data than intended (162 participants instead of 150).

We preregistered the below five hypotheses. In summary, these speculate that pedagogical agents, and their realism, cause higher social presence which will lead to more learning.

H1 Participants in conditions with pedagogical agents of high visual realism, compared to low visual realism, will report higher social presence.

H2 Participants in conditions with pedagogical agents of high behavioral realism, compared to low behavioral realism, will report higher social presence.

H3 Participants who report higher social presence will have higher knowledge acquisition.

H4 Learning with a virtual pedagogical agent leads to higher knowledge acquisition compared to only learning with a voice.

H5 An interaction effect between visual and behavioral realism of the pedagogical agent exists, such that consistency (high/high or low/low) leads to more learning than inconsistency (high/low or low/high).

3.2 Participants

Following recommendations on conducting unsupervised VR studies [23, 37], participants were recruited to install our experimental application onto their own devices, and conduct the study at their discretion. A total of 162 participants, recruited on social media, participated in the experiment using their own VR headset over the course of 11 days. Most of the participants found our advertisement on Reddit (132), but some found it on Facebook (11) or Twitter (6). Participants were reimbursed with a gift certificate worth $15 USD (or the equivalent in their preferred currency). All of the demographics answers are nominal as participants answered the questions within VR by pointing (see Figure 1, right). Participants were mostly male (134 male, 24 female, 4 non-binary). Roughly half of the participants were between 18-29 (88), the rest were: 30-39 (41), 40-49 (22), 50-59 (10), and 60+ (3). Based on IP, participants were identified to be located in 23 different countries, among the most common: United States (78), United Kingdom (22), Canada (11), Dominican Republic (11), and Mexico (7). The participants' educational level ranged from 'High school or less' (73), 'Bachelor' (57), 'Master' (28), and 'PhD' (6). The resulting sample is, as the above shows, rather diverse. However, the majority were from a cohort with expert VR familiarity. This follows from limiting participation to only participants who own an Oculus Quest themselves, and who have the ability to install custom applications onto their device. This is also evident from self reports of VR experience, as the majority of participants had extensive VR experience, having been immersed more than 50 times (94); the remaining had mostly some experience (10-50 times, $N = 35$), or little experience (1-10 times, $N = 29$).

One participant was excluded for taking too long, as defined in the preregistration ($M + 3SD$). Eight participants were recorded with a negative knowledge acquisition; these were kept in the sample as an exclusion criteria based on learning outcomes was not established prior to data collection, and because of the relatively limited amount of participants who did not learn. The analyses presented are therefore conducted on 161 participants.

χ^2 tests were conducted to assess the equivalence of conditions on demographic variables. These were all non-significant: age ($p = .93$), gender ($p = .60$), education ($p = .27$), and English proficiency ($p = .92$). Hence, the assigned groups did not significantly differ based on demography.

3.3 Apparatus

The virtual environment was developed using Unity 2020. The application was targeted Oculus Quest only. The environment was an exhibition hall equipped with animated 3D models related to the topic of viruses (see Figure 1, left). Most of the 3D models were found on the Unity Asset store. The pedagogical agent was taken from the Microsoft Rocketbox repository [9]. An American female voice actor was employed for recording the manuscript. Ambient museum background sounds were present during the simulation.

3.4 Design

A 2×2 between subjects design was employed, with an additional control group for a total of five conditions. Condition was assigned randomly at run time on the device. The independent variables manipulated were (i) *behavioral realism*, with the levels *high* and *low* and (ii) *visual realism*, also with the levels *high* and *low*. For high behavioral realism, the pedagogical agent featured gesturing,
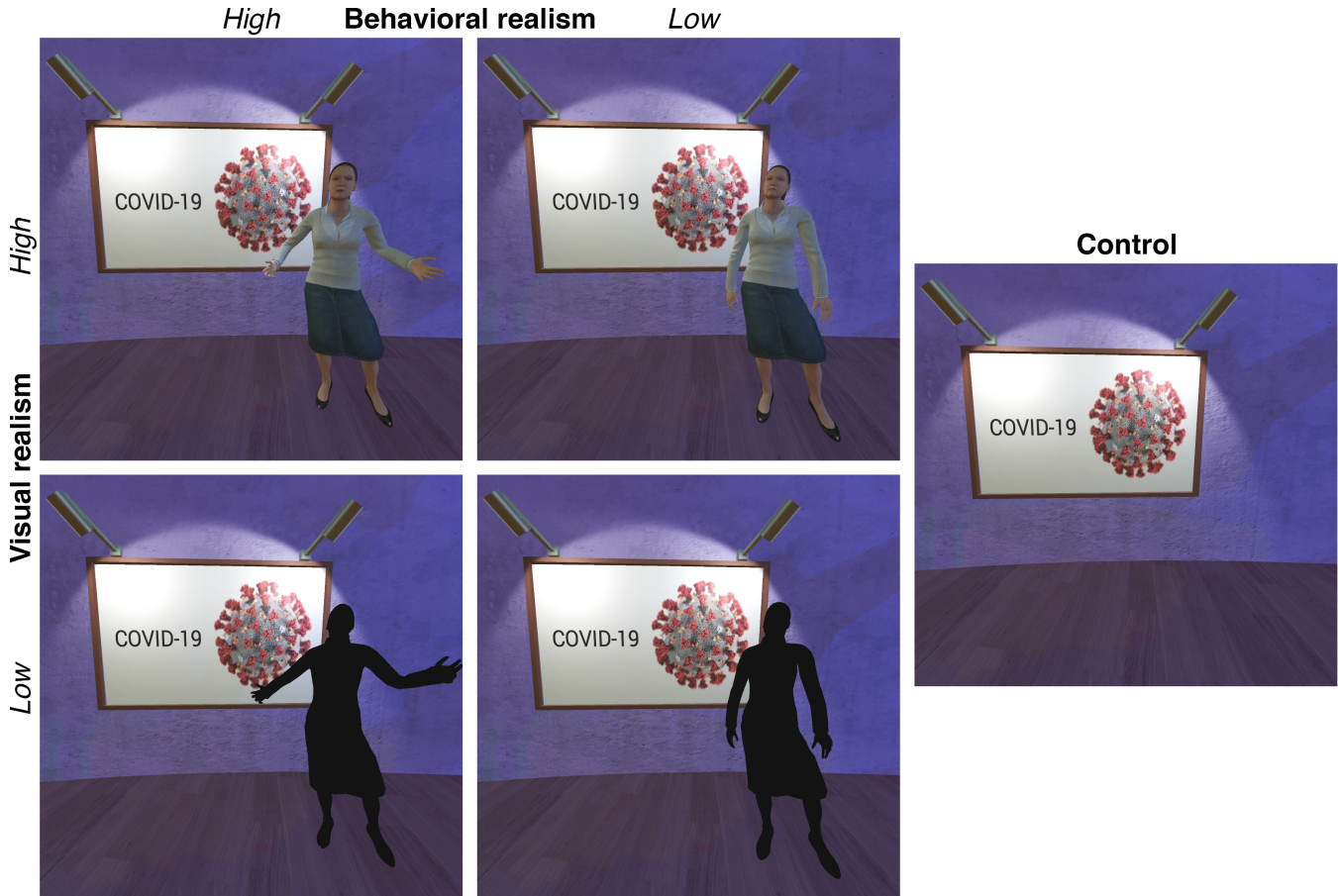
Figure 2: Experimental design of the study. High behavioral realism entailed eye contact, gesturing, lip sync, and natural movements (left). High visual realism entailed rendering the agent as a human (top) rather than in monochrome (bottom). The control group experienced the simulation without a pedagogical agent (far right).

eye contact, idle animations, speech and lip synchronization, and movement by walking (see Figure 2 left). Conversely, low behavioral realism entailed neither of these, and instead featured a static agent that would move by gliding over the floor (see Figure 2, right).

For high visual realism the pedagogical agent would look like a human female museum tour guide (see Figure 2, top); for low visual realism the same humanoid agent employed a black monochrome mesh (see Figure 2, bottom). The control condition did not feature a pedagogical agent, yet with the narration intact. This experimental design allowed an investigation of the importance of both behavior and appearance of pedagogical agents in virtual learning environments, also, it made comparisons between having an agent and no agent possible. This way, the design enabled verification of the image principle (concerning the appearance vs. absence of pedagogical agents), the embodiment principle (concerning the presence vs. absence of human-like behavior in pedagogical agents), as well as potential new principles derived from Oh et al. [38] concerning the effect of high vs. low visual realism in pedagogical agents and, finally, the effect of consistency vs. inconsistency between behavioral and visual realism in pedagogical agents. On a broader

level, the experimental design enabled an assessment of the two opposing views in the field: that pedagogical agents facilitate vs. impede learning.

3.5 Dependent measures

Eight variables were measured from a total of 40 questions; two of these (factual and conceptual knowledge [22]) were objective questions about the learning topic. Three variables relating to the interaction with the pedagogical agent were measured. Two measures, humanness and attractiveness, were from Ho and MacDorman's measures of the Uncanny Valley Effect [13] ('eeriness' was omitted to reduce the length of the within-VR questionnaire). Social presence [26] was included, which measures the subjective experience of being present with a 'real' person. Additionally, enjoyment and cognitive load were measured. Validated subjective scales were employed for all psychological variables. The knowledge questions were administered, in the same order, both before and after the study to study pre-to-post changes on learning. The subjective measures were only administered after the study. All questions were answered on a virtual screen within the VR application by pointing

Variable	Category	Questions	Type	Min/max	Reference
Factual knowledge	Learning outcome	10	Multiple choice	0-10	[22]
Conceptual knowledge	Learning outcome	10	Multiple choice	0-10	[22]
Perceived enjoyment	Experience	3	5-point Likert	1-5	[27]
Perceived humanness	Uncanny valley	5	Semantic differential	1-5	[13]
Attractiveness	Uncanny valley	5	Semantic differential	1-5	[13]
Social presence	Pedagogical agent	5	5-point Likert	1-5	[26]
Intrinsic cognitive load	Cognitive load	1	5-point Likert	1-5	[4]
Extraneous cognitive load	Cognitive load	1	5-point Likert	1-5	[4]

Table 1: Dependent measures used in the study.

and pulling the trigger on the controller (see Figure 1, right). For each knowledge question four possible answers were provided.

The collected variables were analyzed using analysis of variance (ANOVA), specifically the results presented are computed using the R function `car::Anova`.

3.6 Developing the learning material and outcome test

To underline the potentials of using home VR as a commodity educational tool, especially during the global health crisis, a virtual museum exhibition about viruses was chosen. In addition to a brief introduction to general virology, the exhibition progressed as a learning tour through three viral diseases: measles, Zika virus disease, and COVID-19.

The narration that accompanied the exhibition was developed with inspiration from a national biology teaching repository about epidemics and pandemics targeted 13-15 year-olds, as well as other relevant information sources such as the World Health Organization. The target group for such simulations is therefore potentially large. The environment features slides on virtual screens to supplement the narrations with relevant visuals such as a depiction of a baby suffering from microcephaly when learning about Zika virus disease.

To have a direct measure of participants' knowledge acquisition as a result of experiencing the simulation, a multiple choice test was developed, that contained questions about the information presented during the simulation. The test was developed with experts in educational psychology and psychometrics, and measured both factual and conceptual knowledge; that is, cognitive objectives of recalling and understanding, respectively [22]. The initial version of the test had a total of 10 questions. To estimate the difficulty of the test, 112 participants were recruited on Amazon Mechanical Turk (AMT) to conduct the test without any preparation (10 minutes; 1$ pay). The resulting median score was 7 out of 10, with roughly 10% of the participants answering all questions correctly. As a result of this test, the number and the difficulty of questions were increased. A next iteration of the test had 20 questions, and was also tested on AMT. This time, with 75 participants, the median score was 11/20. In this iteration, participants scored between 6 and 15 points, which attested that it was neither too easy or too difficult. The final learning outcome test therefore held 20 questions. It included ten questions on factual knowledge, such as *How many deaths has measles vaccination prevented?*, and ten questions on conceptual knowledge, such as *How do vaccines help the body develop immunity to diseases?*.

3.7 Procedure

For an overview of the study procedure, please consult Figure 3. Participants signed up to our study using an online survey. Upon giving informed consent to data collection and study participation, a brief guide to installing our experimental application followed. The open app store SideQuest[1] was utilized for the purpose of easing the installation burden. After completing the installation
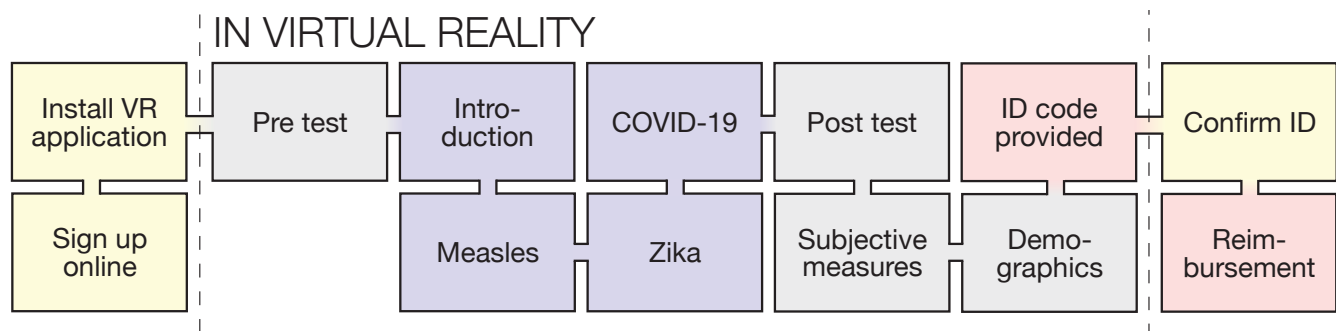
---

[1] http://sidequestvr.com

Figure 3: Visual overview of the study procedure. Yellow boxes denote PC-based user activity, gray are within-VR questionnaires, blue are core learning material, and the red shows meta activity. The entire procedure took about one hour to complete.

instructions, participants were instructed to launch the application, take the headset on, and follow within-VR guidelines. As such, guidelines were not provided before immersion. It was not possible to skip parts of the experience.

The virtual environment began with a brief introduction of the controls, the purpose, and the content. This was provided both in audio and on information displays blended into the museum environment. Participants could freely walk around or teleport themselves by pointing and clicking 'A' on the controller. The exhibition progressed as the participant followed the tour guide around to the different displays constituting the core learning material. A knowledge test was conducted, before beginning the guided tour. A general introduction to the topic of virology was followed by presentations of three viruses: Measles, Zika, and COVID-19. After completing the museum tour, participants conducted an identical knowledge test, in addition to a survey about subjective measures and demographics. Upon completion, a unique code emerged on a screen; this code had to be entered into the online form where participants signed up to ensure valid participation (and to qualify for reimbursement).

The mean duration of the immersion was 20.0 minutes ($SD = 5.4$).

4 RESULTS

Here, quantitative findings are reported. They relate to the knowledge acquisition and the effect of manipulations on subjective and objective measures. A visual inspection of the collected variables showed that data followed normal distributions (e.g., see Figure 4). The analyses are therefore based on parametric tests. Furthermore, the study collected subjective measures from scales that have previously been validated for parametric testing.

4.1 How much did they learn?

See Table 2 for an overview of differences between pre- and post scores. Out of a combined maximum of 20 points, the mean pre-score was 11.0 ($SD = 2.8$). The mean pre-score for factual knowledge was 3.5 ($SD = 1.5$); it was 7.6 ($SD = 2.1$) for conceptual knowledge. For the post test, the combined mean score was 15.1 ($SD = 2.8$). The mean post score for factual knowledge was 6.2 ($SD = 1.8$); for conceptual knowledge the mean was 8.9 ($SD = 1.5$). That shows that participants, on average, increased their tests scores with 4.0 points ($SD = 2.9$) after experiencing the virtual exhibition (see Figure 4); this difference was also significant, shown with a repeated measures ANOVA: $F(1, 160) = 305.9$, $p < .0001$, $d = 1.4$.

	Factual	Conceptual
Pre test, M (SD)	3.5 (1.5)	7.6 (2.1)
Post test, M (SD)	6.2 (1.8)	8.9 (1.5)
Difference	+2.7	+1.3
Welch's t-test, p	< .0001	< .0001
Cohen's *d*	1.6	0.7

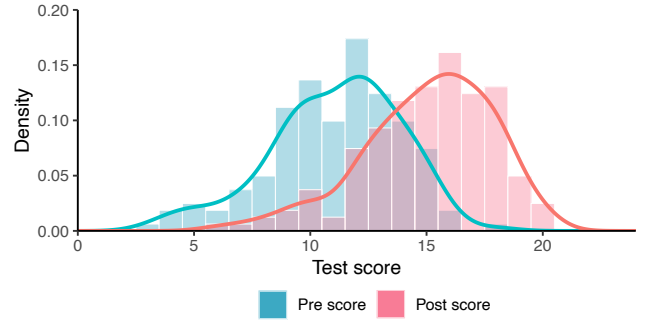**Table 2: Scores obtained in the pre- and post tests, respectively.**

Figure 4: Pre- to post test scores: participants performed better on the knowledge tests after experiencing the virtual exhibition.

4.2 The effect of agent on subjective measures

Figure 5 shows the mean reported scores on the subjective measures relating to interaction with the agent, divided by experimental manipulation.

For humanness (Figure 5, A), the absence of a pedagogical agent yields comparable humanness to an agent with high behavioral realism. Also, high behavioral realism resulted in higher humanness compared to low behavioral realism; $F(1, 156) = 8.2$, $p = .005$. In other words, behavior, but not appearance, of the virtual agent affected whether participants experienced the agent as human.

For attractiveness (Figure 5, B), the experimental manipulations of agent had less of an impact, and the effect of manipulation was not significant.

The high behavioral realism conditions showed significantly higher social presence (Figure 5, C); $F(1, 156) = 5.5$, $p = .02$. A comparable difference for appearance was not found.

In summary, our findings suggest that behavior of pedagogical agents (gesturing, eye contact, natural movements) impact subjective social accounts of the agent, while appearance to a lesser degree does. It should be noted that the absence of an agent results in comparable, or even higher, reports of humanness and attractiveness (but not social presence). This counter intuitive finding, that no agent leads to high reports of attractiveness and humanness, was also reported by Rzayev et al. [44].

4.3 The effect of agent on learning

Figure 6 shows the mean difference between pre- and post-test knowledge scores for each condition. Two learning outcomes were measured; factual and conceptual knowledge gain [22]. Factual knowledge relates to recalling (e.g., numbers, places, years) while conceptual knowledge relates to understanding (e.g., explaining, connecting, transferring).

Both the appearance and behavior of a pedagogical agent had an effect on factual knowledge gain (see Figure 6, A). One-way ANOVAs showed significant effects: $F(2, 158) = [5.5; 3.8]$, $p = [.005; .03]$. Post hoc Tukey's HSDs showed that high visual realism as well as high behavioral realism significantly differed with the control condition. This shows that, for learning facts, the presence of a pedagogical agent is not ideal; rather, the agent, and its visual and behavioral fidelity impede factual retention.
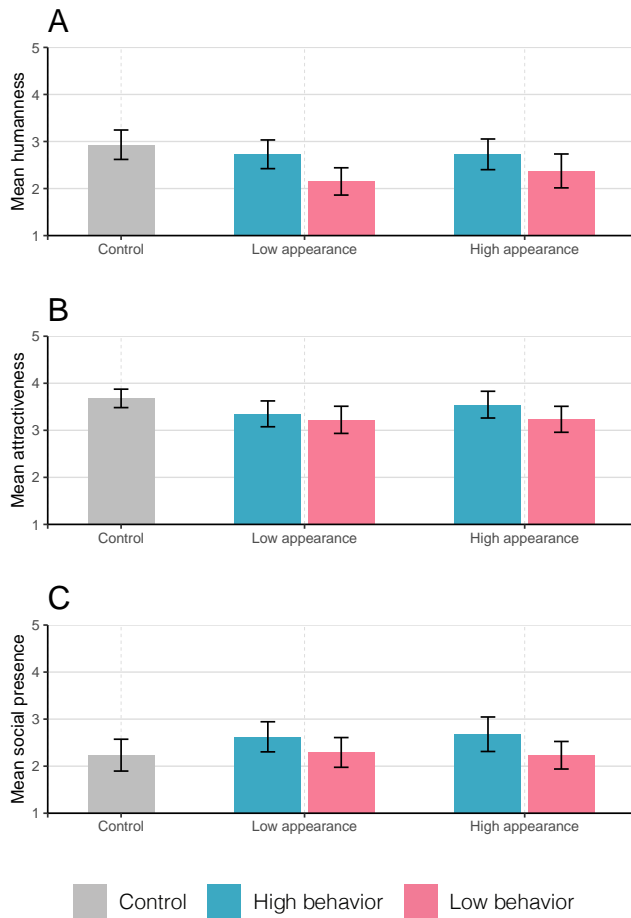
Figure 5: Means of subjective measures by experimental manipulation: humanness (A), attractiveness (B), and social presence (C). Error bars show 95% confidence intervals.

For conceptual knowledge, a two-way ANOVA showed a significant interaction effect between appearance and behavior (see Figure 6, B); $F(1, 155) = 9.2, p = .003$. Oh et al. [38] reported a 'consistency effect'; they state that, for learning, consistency of an agent is preferred (i.e., that fidelity of behavioral and visual realism ideally match). Our findings contradict this finding, as we observe a higher conceptual knowledge gain when behavior and appearance are incongruent.

4.4 Enjoyment

Participants generally reported high enjoyment rates for the learning experience, with a mean score of $M = 3.9$ ($SD = 0.8$) of a maximum 5. These enjoyment rates were consistent across experimental manipulations, see Figure 7. Together with the general positive knowledge acquisition, this tells us, that the virtual museum was received positively as a new form of remote learning during the global health crisis. This finding is consistent with previous research that associated VR with higher enjoyment compared to less immersive media [24].
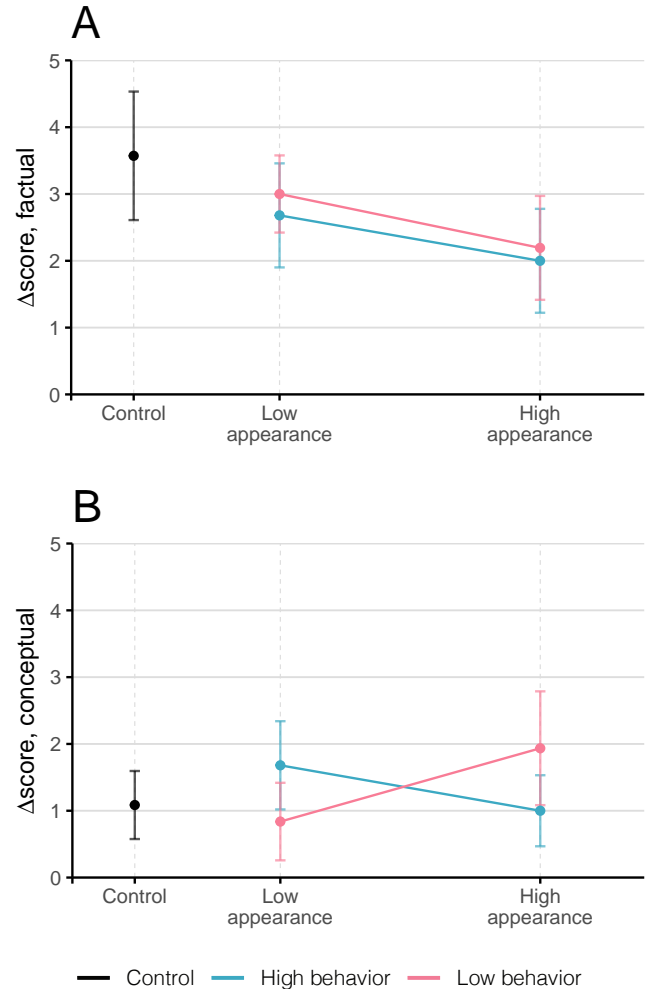

Figure 6: Mean differences between pre- and post test scores. For factual knowledge (A), absence of a pedagogical agent leads to higher pre- to post scores. For conceptual learning (B), we find a significant interaction effect between behavior and appearance, namely that incongruence (low/high or low/high) leads to better learning than congruence (high/high or low/low). Error bars show 95% CI.

4.5 Cognitive load

As previous findings in educational VR suggest that specific instructional designs in VR learning environments lead to increased cognitive load [2, 29], a subjective measure for both intrinsic (difficulty of subject) and extraneous (difficulty of instruction) cognitive load [4] were collected.

Participants reported comparable intrinsic and extraneous cognitive load for all conditions. As such, medians for all five conditions for both cognitive load measures were 2 out of 5. There were no significant differences between conditions. It should be noted that cognitive load was assessed via single items. Use of full scales could have provided further insights (e.g. [1]).
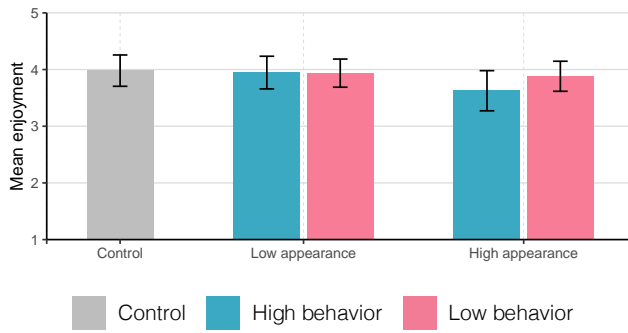
Figure 7: Means of enjoyment ratings: participants consistently experienced the virtual exhibition as enjoyable. Error bars show 95% confidence intervals.

4.6 Brief Summary of hypotheses and findings

We preregistered five hypotheses related to the realism of pedagogical agents, their effect on social presence, and in turn learning. We hypothesized that high visual realism would lead to high social presence (H1); that high behavioral realism would lead to high social presence (H2); that high social presence would lead to better learning (H3); that agents would be better than no agents for learning (H4); and that there would be an interaction effect between visual and behavioral realism, where consistency would be better than inconsistency for learning (H5). In relation to our findings, specifically, only H2 was confirmed, namely that behavior of an avatar significantly impacts social presence. On the contrary, we did not find support for H1; that is, that appearance of an agent impacts social presence. We did not find a significant correlation between social presence and factual learning (rather slightly negative, Pearson's $\rho = -0.11$, $p = .15$). Yet, for conceptual learning we do find it significantly correlated to social presence ($\rho = 0.25$, $p = .001$). Consequently, H3 has a more nuanced answer. Importantly, for H4, we find an opposite effect for factual learning. For conceptual learning we did find an interaction effect, but rather the reverse than hypothesized in H5. Consequently we find partial support for the opposite effect for H4 (for factual learning) and H5 (for conceptual learning).

5 DISCUSSION

We conducted a VR experiment 'in the wild' during the COVID-19 pandemic. In general, our findings show that unsupervised, remote learning in VR is feasible as all participants enjoyed the experience and improved on a knowledge test. Furthermore, we show that the design of pedagogical agents in educational VR impacts learning depending on the type of learning considered. Including a pedagogical agent leads to lower factual knowledge acquisition compared to only including a narration. Looking at conceptual information acquisition, a pedagogical agent may aid learning. These findings expand classical multimedia learning theory in the context of VR.

Our finding that pedagogical agents are useful for learning about concepts but not facts could be explained by the differing nature of the two types of information in combination with the human capacity for selectively attending to certain stimuli [7]. During the lesson, factual information, such as specific dates and numbers, was presented very quickly and therefore imposed large demands on attention at specific moments in time. In contrast, much of the conceptual information had broader explanations and therefore allowed short diversions in attention. Thus, the addition of a detailed agent would specifically have a negative effect on factual learning by stealing attention at critical moments. Although we did not use eye-tracking, other research corroborates people's tendency to gaze at human agents [50]. This echoes the issue raised by Veletsianos et al. [48] that pedagogical agents may be mesmerizing and misdirect attention from the task.

Our participants were expert VR users, hence a novelty effect most likely did not interfere with the results. The novelty effect refers to a heightened motivation to use something simply on account of its newness [20]. Novelty may confound the results of media studies as the increased effort and attention could result in achievement gains that would not occur if the learner was familiar with the medium [5]. In that light, and in combination with the relatively high participation count, our results are a reliable source of evidence on the efficacy of pedagogical agents in immersive learning environments.

5.1 Theoretical implications

The findings of this study has a number of implications for social agency theory and its derived learning principles in the context of educational VR.

5.1.1 The image principle. The image principle suggests that using visible pedagogical agents has a small effect on learning compared to only using narration. As some previous studies found negligible or even negative effects of presenting an image of the speaker, however, it led Mayer to conclude that people do not necessarily learn more from lessons when the speaker's image is on the screen compared to when it is not [33]. Importantly, the theory is based on relatively dated empirical evidence, some of which date back 20 years ago, where learning environments and pedagogical agents were less sophisticated. We examined the image principle with highly sophisticated pedagogical agents rendered using state of the art consumer VR technology. Our findings show that for factual learning, including a pedagogical agent of high visual or behavioral realism leads to less learning compared to using an 'invisible' speaker. Realistically looking agents presumably distract the learner, yet, we did not record a change in subjectively measured cognitive load. In contrast, learning about concepts was not hampered by the inclusion of pedagogical agents.

To sum up, our findings suggest a refinement of the image principle when applied to educational VR. When compared to only a narration, pedagogical agents do not lead to higher factual knowledge gain. On the contrary, realistic pedagogical agents may actually hamper learning of factual information. This was not the case when learning about conceptual information.

5.1.2 The embodiment principle. The embodiment principle focuses on the absence vs. presence of behavioral cues in pedagogical agents. Previous research points to the benefits of embodiment, which led Mayer to conclude that when pedagogical agents display

human-like gesturing, movement, facial expressions, etc. as opposed to appearing static, it leads to better learning. In the present study we did not find a positive effect of behavioral realism on factual knowledge gain. In fact, participants who learned from behaviorally realistic pedagogical agents (i.e., agents exhibiting gesturing, eye contact, speech and lip synchronization) increased their scores slightly less on the factual knowledge test than participants who learned from static pedagogical agents (although this warrants further investigation). Consequently, our findings do not corroborate the embodiment principle when learning about facts. When learning about concepts, however, our findings indicate that the embodiment principle exists if the agents are of low but not high visual realism. In the latter scenario, a reversed embodiment principle is actually found. This is an important addition to the embodiment principle in educational VR.

5.1.3 Visual realism and consistency. Oh et al. [38] reviewed determinants of social presence in virtual environments and found that agents' visual realism and consistency in appearance and behavior could be important sources of social presence. We built on these findings and tested if visual realism and consistency influenced learning of factual and conceptual information. Our results indicate that visual realism impacts factual learning negatively. One possible explanation for this could be the uncanny valley effect, which refers to the relation between the human-likeness of an entity and the perceiver's affinity for it [36]. The theory posits that affinity increases as a function of the human-likeness of artificial humans until it reaches a valley where affinity suddenly drops. This corresponds to a point where there is a relatively high degree of human-likeness in an entity combined with evidence that it is artificial, and this is accompanied by a creepy sensation [36]. The visually realistic agents might have caused a creepy sensation in the participants, lowering their motivation to understand the learning material. This would be consistent with the feedback reported by some of the participants learning with visually realistic agents: *"You could replace the guide with a robot. A clearly non-human robot guide wouldn't be so off putting (sic)"* or *"the guide looks creepy"*. In terms of consistency, our results, again, indicate an effect in the opposite direction of what we had hypothesized. There was a significant interaction effect between visual and behavioral realism for conceptual knowledge gain; knowledge gain was higher when these were incongruent.

5.1.4 Social presence as a learning mechanism. Social presence during multimedia learning is posited to be an important construct that leads to better learning outcomes [33]. Consistent with social agency theory, we found that behavioral realism had a significant effect on social presence. However, an increase in subjective social presence was not unequivocally associated with more learning. This shows that social presence does not necessarily lead to more learning, and that other constructs than those provided by social agency theory are important for learning with virtual agents.

5.2 Limitations and future research directions

Pedagogical agents can take on a variety of instructor roles during learning [12]. Similar to Baylor and Kim [3], the agents in the present study performed the role of an expert as their primary

function was to provide accurate and concise information. However, agents can also inhabit other roles, such as motivators whose primary function is to provide encouragement [3]. Future research should investigate learning with pedagogical agents inhabiting different roles.

The participants in this study primarily consisted of expert VR users. This was an advantage in terms of limiting novelty effects, but future research should investigate the efficacy of remote learning with pedagogical agents in a non-expert sample.

We investigated factual and conceptual learning as pre- to posttest changes on a multiple choice question test administered in VR before and after the lesson. However, we did not investigate transfer of learning (i.e., when learning in one context improves performance in another context [41]). Transfer is a key goal in education, and future research should therefore consider investigating transfer effects of learning with agents in VR.

Another limitation concerns the relevancy of the pedagogical agents and their behavior in terms of teaching the content of the VR simulation. For instance, it could be argued that an agent would be more relevant for teaching specific procedures compared to teaching facts, as its behavior could then be tied more easily to the learning material (e.g., if demonstrating how to perform a certain procedure). Theoretically, this would map onto classical psychological theories of model learning as put forward by scholars such as Albert Bandura. Future research could therefore examine procedural learning from pedagogical agents in VR. One promising direction could be to examine whether watching an agent perform a procedure produces stronger procedural learning compared to hearing about it.

While out-of-lab VR experimentation allows for a larger and more heterogeneous sample which gives higher statistical power and ecological findings at a lower cost [37], it comes with the cost of reduced internal validity. Specifically, we cannot control the size and constraints of the participants' surroundings or a strict adherence to the study protocol. Qualitative findings are also hard to collect using this paradigm. Some of the aspects discussed in this paper, concerning for instance creepiness of agents, might be beneficial to investigate further using traditional laboratory protocols.

Based on inquiry by some of the participants in regards to inviting lock downed family members and friends to participate in the study, we allowed multiple participation from the same IP (a total of 17 recurrent IP addresses were recorded). This resulted in a slightly more diverse participant pool than otherwise expected. We note that this makes it technically possible for a single participant to complete the study multiple times, we however, do not suspect this have influenced the reported data.

Furthermore we collected free form text answers as part of the debriefing by asking for any feedback. Although many wrote lengthy comments of mostly feature requests, the focus of the study was quantitative, and an extensive analysis of qualitative data was therefore neither significant nor the scope. If the aim is qualitative, a recommendation for future online user studies is therefore to be specific about any desired type of qualitative data and to formulate open-ended questions accordingly.

The general low number of non-males in the study was a limitation along with the fact that we did not manipulate the sex of

the agent. Participants' gender did not reveal considerable learning differences at post-test: M 15.4, 15.0, and 16.0 (F/M/X).

6 CONCLUSION

Remote learning in the confines of your own home through VR is an enjoyable and effective alternative to the real-life classroom, especially during global health crises. Caution should be taken, however, when it comes to incorporating a pedagogical agent into the virtual learning experience. Our findings expand upon classical multimedia learning theory in the context of VR, and provide new insights into the scholarly discussion about whether pedagogical agents are facilitators of learning or merely unnecessary distractions; our findings suggest that the answer depends on the type of knowledge in question. Behavior of a pedagogical agent did in fact increase social presence and humanness, yet the effects on learning were mixed. If the material to be learned is factual, a pedagogical agent may impede learning. Contrary, if the material to be learned is conceptual, a pedagogical agent may be worthwhile.

7 ACKNOWLEDGEMENTS

REFERENCES

[1] Martin S. Andersen and Guido Makransky. [n.d.]. The validation and further development of a multidimensional cognitive load scale for virtual environments. *Journal of Computer Assisted Learning* n/a, n/a ([n. d.]). https://doi.org/10.1111/jcal.12478 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/jcal.12478

[2] Sarune Baceviciute, Aske Mottelson, Thomas Terkildsen, and Guido Makransky. 2020. Investigating Representation of Text and Audio in Educational VR Using Learning Outcomes and EEG. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '20*). Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376872

[3] Amy L Baylor and Yanghee Kim. 2004. Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. In *International conference on intelligent tutoring systems*. Springer, 592–603.

[4] Gabriele Cierniak, Katharina Scheiter, and Peter Gerjets. 2009. Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior* 25 (03 2009), 315–324. https://doi.org/10.1016/j.chb.2008.12.020

[5] Richard E Clark and Terrance G Craig. 1992. Research and theory on multi-media learning effects. In *Interactive multimedia learning environments*. Springer, 19–30.

[6] Joseph Crawford, Kerryn Butler-Henderson, Jürgen Rudolph, Matthias Glowatz, et al. 2020. COVID-19: 20 Countries' Higher Education Intra-Period Digital Pedagogy Responses. *Journal of Applied Teaching and Learning (JALT)* 3, 1 (2020).

[7] Peter Dayan, Sham Kakade, and P Read Montague. 2000. Learning and selective attention. *Nature neuroscience* 3, 11 (2000), 1218–1223.

[8] Jesse Fox, Sun Joo (Grace) Ahn, Joris H. Janssen, Leo Yeykelis, Kathryn Y. Segovia, and Jeremy N. Bailenson. 2015. Avatars Versus Agents: A Meta-Analysis Quantifying the Effect of Agency on Social Influence. *Human–Computer Interaction* 30, 5 (2015), 401–432. https://doi.org/10.1080/07370024.2014.921494 arXiv:https://doi.org/10.1080/07370024.2014.921494

[9] Mar Gonzalez-Franco, Eyal Ofek, Ye Pan, Angus Antley, Anthony Steed, Bernhard Spanlang, Antonella Maselli, Domna Banakou, Nuria Pelechano, Sergio Orts Escolano, Veronica Orvalho, Laura Trutoiu, Markus Wojcik, Maria V. Sanchez-Vives, Jeremy Bailenson, Mel Slater, and Jaron Lanier. 2020. Importance of rigging for procedural avatars. Microsoft Rocketbox a public library. https://github.com/microsoft/Microsoft-Rocketbox. [Online; accessed April 1, 2020].

[10] Charles R Graham. 2006. *Blended learning systems: Definition, current trends, and future directions. In The handbook of blended learning: Global perspectives, local designs.* John Wiley & Sons.

[11] Victoria Groom, Clifford Nass, Tina Chen, Alexia Nielsen, James K. Scarborough, and Erica Robles. 2009. Evaluating the effects of behavioral realism in embodied agents. *International Journal of Human-Computer Studies* 67, 10 (2009), 842 – 849. https://doi.org/10.1016/j.ijhcs.2009.07.001

[12] Steffi Heidig and Geraldine Clarebout. 2011. Do pedagogical agents make a difference to student motivation and learning? A review of empirical research. *Educational Research Review* 6 (12 2011), 27–54. https://doi.org/10.1016/j.edurev.2010.07.004

[13] Chin-Chang Ho and Karl MacDorman. 2017. Measuring the Uncanny Valley Effect: Refinements to Indices for Perceived Humanness, Attractiveness, and Eeriness. *International Journal of Social Robotics* 9 (01 2017), 129–139. https://doi.org/10.1007/s12369-016-0380-9

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

[15] Iwan Kartiko, Manolya Kavakli, and Ken Cheng. 2010. Learning science in a virtual reality application: The impacts of animated-virtual actors visual complexity. *Computers I& Education* 55, 2 (2010), 881 – 891. https://doi.org/10.1016/j.compedu.2010.03.019

[16] K. Kim, L. Boelling, S. Haesler, J. Bailenson, G. Bruder, and G. F. Welch. 2018. Does a Digital Assistant Need a Body? The Influence of Visual Embodiment and Social Behavior on the Perception of Intelligent Virtual Agents in AR. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 105–114. https://doi.org/10.1109/ISMAR.2018.00039

[17] Paul A. Kirschner. 2002. Cognitive load theory: implications of cognitive load theory on the design of learning. , 10 pages. https://doi.org/10.1016/S0959-4752(01)00014-7

[18] René F Kizilcec, Jeremy N Bailenson, and Charles J Gomez. 2015. The instructor's face in video instruction: Evidence from two large-scale field studies. *Journal of Educational Psychology* 107, 3 (2015), 724.

[19] Sara Klingenberg, Maria L. M. Jørgensen, Gert Dandanell, Karen Skriver, Aske Mottelson, and Guido Makransky. 2020. Investigating the effect of teaching as a generative learning strategy when learning through desktop and immersive VR: A media and methods experiment. *British Journal of Educational Technology* 51, 6 (2020), 2115–2138. https://doi.org/10.1111/bjet.13029 arXiv:https://bera-journals.onlinelibrary.wiley.com/doi/pdf/10.1111/bjet.13029

[20] Michael Koch, Kai von Luck, Jan Schwarzer, and Susanne Draheim. 2018. The Novelty Effect in Large Display Deployments - Experiences and Lessons-Learned for Evaluating Prototypes.. In *ECSCW Exploratory Papers*. European Society for Socially Embedded Technologies (EUSSET). http://dblp.uni-trier.de/db/conf/ecscw/ecscw2018ep.html#0001LSD18

[21] Kwan Min Lee. 2006. Presence, Explicated. *Communication Theory* 14, 1 (01 2006), 27–50. https://doi.org/10.1111/j.1468-2885.2004.tb00302.x arXiv:https://academic.oup.com/ct/article-pdf/14/1/27/22294490/jcomthe0027.pdf

[22] Anderson LW, Krathwohl DR, Airasian PW, Cruikshank KA, Richard Mayer, Pintrich PR, J. Raths, and Wittrock MC. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.*

[23] Xiao Ma, Megan Cackett, Leslie Park, Eric Chien, and Mor Naaman. 2018. Web-Based VR Experiments Powered by the Crowd. In *Proceedings of the 2018 World Wide Web Conference* (Lyon, France) (*WWW '18*). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 33–43. https://doi.org/10.1145/3178876.3186034

[24] Guido Makransky, Niels Koch Andreasen, Sarune Baceviciute, and Richard Mayer. 2020. Immersive Virtual Reality Increases Liking but Not Learning with a Science Simulation and Generative Learning Strategies Promote Learning in Immersive Virtual Reality. *Journal of Educational Psychology* (02 2020). https://doi.org/10.1037/edu0000473

[25] Guido Makransky, Stefan Borre-Gude, and Richard E. Mayer. 2019. Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning* 35, 6 (2019), 691–707. https://doi.org/10.1111/jcal.12375 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/jcal.12375

[26] Guido Makransky, Lau Lilleholt Jensen, and Anders Aaby. 2017. Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior* 72 (2017), 276–285. https://doi.org/10.1016/j.chb.2017.02.066

[27] Guido Makransky and Lau Lilleholt. 2018. A Structural Equation Modeling Investigation of the Emotional Value of Immersive Virtual Reality in Education. *Educational Technology Research and Development* (02 2018). https://doi.org/10.1007/s11423-018-9581-2

[28] Guido Makransky, Gustav B Petersen, and Sara Klingenberg. 2020. Can an immersive virtual reality simulation increase students' interest and career aspirations in science? *British Journal of Educational Technology* (2020).

[29] Guido Makransky, Thomas S. Terkildsen, and Richard E. Mayer. 2019. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction* 60 (2019), 225 – 236. https://doi.org/10.1016/j.learninstruc.2017.12.007

[30] Guido Makransky, Philip Wismer, and Richard E. Mayer. 2019. A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *Journal of Computer Assisted Learning* 35, 3 (2019), 349–358. https://doi.org/10.1111/jcal.12335

[31] David M. Markowitz, Rob Laha, Brian P. Perone, Roy D. Pea, and Jeremy N. Bailenson. 2018. Immersive Virtual Reality Field Trips Facilitate Learning About Climate Change. *Frontiers in Psychology* 9 (2018), 2364. https://doi.org/10.3389/fpsyg.2018.02364

[32] Ati Suci Dian Martha and Harry B Santoso. 2019. The Design and Impact of the Pedagogical Agent: A Systematic Literature Review. *Journal of Educators Online* 16, 1 (2019), n1.

[33] Richard E Mayer. 2014. Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. *The Cambridge handbook of multimedia learning* 16 (2014), 345–370.

[34] Richard E Mayer. 2019. Thirty years of research on online learning. *Applied Cognitive Psychology* 33, 2 (2019), 152–159.

[35] Oliver A. Meyer, Magnus K. Omdahl, and Guido Makransky. 2019. Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers I& Education* 140 (10 2019). https://doi.org/10.1016/j.compedu.2019.103603

[36] Masahiro Mori, Karl F MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics Automation Magazine* 19, 2 (2012), 98–100. https://doi.org/10.1109/MRA.2012.2192811

[37] Aske Mottelson and Kasper Hornbæk. 2017. Virtual Reality Studies Outside the Laboratory *(VRST '17)*. Association for Computing Machinery, New York, NY, USA, Article 9, 10 pages. https://doi.org/10.1145/3139131.3139141

[38] Catherine S. Oh, Jeremy N. Bailenson, and Gregory F. Welch. 2018. A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI* 5 (2018), 114. https://doi.org/10.3389/frobt.2018.00114

[39] Jocelyn Parong and Richard E Mayer. [n.d.]. Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning* ([n. d.]).

[40] Jocelyn Parong and Richard E Mayer. 2018. Learning science in immersive virtual reality. *Journal of Educational Psychology* 110, 6 (2018), 785.

[41] David N Perkins, Gavriel Salomon, et al. 1992. Transfer of learning. *International encyclopedia of education* 2 (1992), 6452–6457.

[42] Gustav B. Petersen, Sara Klingenberg, Richard E. Mayer, and Guido Makransky. 2020. The virtual field trip: Investigating how to optimize immersive virtual learning in climate change education. *British Journal of Educational Technology* 51, 6 (1 11 2020), 2098–2114. https://doi.org/10.1111/bjet.12991

[43] Jaziar Radianti, Tim A. Majchrzak, Jennifer Fromm, and Isabell Wohlgenannt. 2020. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers I& Education* 147 (2020), 103778. https://doi.org/10.1016/j.compedu.2019.103778

[44] Rufat Rzayev, Gürkan Karaman, Katrin Wolf, Niels Henze, and Valentin Schwind. 2019. The Effect of Presence and Appearance of Guides in Virtual Reality Exhibitions. In *Proceedings of Mensch Und Computer 2019* (Hamburg, Germany) *(MuC'19)*. Association for Computing Machinery, New York, NY, USA, 11â€"20. https://doi.org/10.1145/3340764.3340802

[45] Noah L. Schroeder, Olusola O. Adesope, and Rachel Barouch Gilbert. 2013. How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research* 49, 1 (2013), 1–39. https://doi.org/10.2190/EC.49.1.a arXiv:https://doi.org/10.2190/EC.49.1.a

[46] Mel Slater. 2009. Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364 (12 2009), 3549–57. https://doi.org/10.1098/rstb.2009.0138

[47] A. Steed, S. Frlston, M. Lopez, J. Drummond, Y. Pan, and D. Swapp. 2016. An In the Wild Experiment on Presence and Embodiment using Consumer Virtual Reality Equipment. *IEEE Transactions on Visualization I& Computer Graphics* 22, 04 (apr 2016), 1406–1414. https://doi.org/10.1109/TVCG.2016.2518135

[48] George Veletsianos, Charles Miller, and Aaron Doering. 2009. Enali: A Research and Design Framework for Virtual Characters and Pedagogical Agents. *Journal of Educational Computing Research* 41, 2 (2009), 171–194. https://doi.org/10.2190/EC.41.2.c arXiv:https://doi.org/10.2190/EC.41.2.c

[49] George Veletsianos and Gregory S Russell. 2014. Pedagogical agents. In *Handbook of research on educational communications and technology*. Springer, 759–769.

[50] Isaac Wang, Jesse Smith, and Jaime Ruiz. 2019. Exploring Virtual Agents for Augmented Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300511