# Immersive virtual environments and embodied agents for e-learning applications

**Isabel S. Fitton** [Corresp., 1] , **Daniel J. Finnegan** [2] , **Michael J. Proulx** [3]

[1] Department of Computer Science, University of Bath, Bath, United Kingdom

[2] School of Computer Science & Informatics, Cardiff University, Cardiff, United Kingdom

[3] Department of Psychology, University of Bath, Bath, United Kingdom

Corresponding Author: Isabel S. Fitton
Email address: isf21@bath.ac.uk

Massive Open Online Courses are a dominant force in remote-learning yet suffer from persisting problems stemming from lack of commitment and low completion rates. In this initial study we investigate how the use of immersive virtual environments for Power-Point based informational learning may benefit learners and mimic traditional lectures successfully. We examine the role of embodied agent tutors which are frequently implemented within virtual learning environments. We find similar performance on a bespoke knowledge test and metrics for motivation, satisfaction, and engagement by learners in both real and virtual environments, regardless of embodied agent tutor presence. Our results raise questions regarding the viability of using virtual environments for remote-learning paradigms, and we emphasise the need for further investigation to inform the design of effective remote-learning applications.

# Immersive Virtual Environments and Embodied Agents for e-Learning Applications

**Isabel Fitton**[1]**, Daniel J. Finnegan**[2]**, and Michael Proulx**[3]

[1]**Department of Computer Science, University of Bath**
[2]**School of Computer Science and Informatics, Cardiff University**
[3]**Department of Psychology, University of Bath**

Corresponding author:
Isabel Fitton[1]

Email address: isf21@bath.ac.uk

## ABSTRACT

Massive Open Online Courses are a dominant force in remote-learning yet suffer from persisting problems stemming from lack of commitment and low completion rates. In this initial study we investigate how the use of immersive virtual environments for Power-Point based informational learning may benefit learners and mimic traditional lectures successfully. We examine the role of embodied agent tutors which are frequently implemented within virtual learning environments. We find similar performance on a bespoke knowledge test and metrics for motivation, satisfaction, and engagement by learners in both real and virtual environments, regardless of embodied agent tutor presence. Our results raise questions regarding the viability of using virtual environments for remote-learning paradigms, and we emphasise the need for further investigation to inform the design of effective remote-learning applications.

## INTRODUCTION

Technological advancements have played a vital role in accommodating vast numbers of students through the growth of distance learning applications and e-learning platforms (Kaplan and Haenlein, 2016; Kauffman, 2015). The predominant form of distance learning applications are Massive Open Online Courses (MOOCs). MOOCs offer access to teaching and material on a large scale via internet-based virtual learning environments for a limitless number of participants, making education more accessible (Freitas and Paredes, 2018). Modern MOOCs involve a video captured recording of a human lecturer who delivers the learning content, facilitating the completion of homework or exams, and discussion via forums (Feng et al., 2015). However, despite the potential of MOOCs to deliver teaching materials and content at a global scale, existing platforms suffer from issues with drop-out and learner motivation (Yang et al., 2013). In parallel to e-learning platforms gaining popularity (Sneddon et al., 2018), VR technology has increasingly been adopted in the classroom as a teaching aid for 'hands-on' skills-based teaching partly due to reductions in cost. For example, in medicine, digital models are much cheaper compared to physical anatomical models for training students (Rajeswaran et al., 2018). Using digital models in a virtual reality scenario is a cost-effective way to educate students on a large scale and as a result there is growing excitement regarding the potential of VR to revolutionize education and e-learning (Greenwald et al., 2017).

While VR is regarded as beneficial to students as a practical teaching aid, its application to formal, lecture style teaching–which e-learning platforms tend to deliver–is less common (Korallo, 2010). The use of Immersive Virtual Environments (IVEs) for corporate and higher-education purposes have only recently begun to emerge. Due to such applications being in their infancy there is very little empirical evaluation of their efficacy or research available to inform their design. A key component of many IVEs is the presence of an Embodied Agent (EA) which, in the context of learning, may serve as a virtual guide or tutor. The use of EAs as virtual tutors within educational IVEs is critical for effective pedagogy (Soliman and Guetl, 2010). Previous research suggests that the representation of artificial agents affects learners'

46 motivation (Maldonado and Nass, 2007). For example, an EA may be customized by the learner to suit
47 their preference - such customization has been shown to improve performance for some cognitive tasks
48 (Lin et al., 2017). In another study that tested male versus female pedagogical agents, the female seemed
49 to be preferred overall (Novick et al., 2019). A recent systematic review of pedagogical agents noted that
50 positive results have been found in numerous studies, yet different combinations of features and different
51 outcome variables have not been systematically studied to clarify which features work best or when
52 (Martha and Santoso, 2019). However, it is unclear how the presence of an EA and learner motivation
53 interact. A clear and robust analysis of these factors and their impact on the learning experience is critical
54 for future application of IVEs as engaging platforms for distance learning.

55 Furthermore, the recent novel SARS-COV-2 (COVID-19) pandemic has had huge repercussions for
56 higher education across the world. As millions of people were restricted to not leaving their homes for
57 extended periods of time, many institutions also shifted to remote delivery of learning material for the
58 academic year 2020/21. Although this presents challenges around blended learning and flipped classroom
59 design, our focus remains on technology and how it may act as a medium for learning material delivery as
60 opposed to what content such material should contain.

61 Our main contribution in this work is an empirical investigation into factors which impact the overall
62 student experience when learning in IVEs. Specifically we report how the presence of an embodied
63 teacher and students' sense of presence in the environment impact learning retention, satisfaction and
64 engagement, and student motivation to engage with learning material presented in an IVE. Our results
65 demonstrate how learning in IVEs is comparable to real classroom learning, yet can scale far beyond the
66 limits of traditional classrooms with constraints such as staff-student ratio and classroom size. Finally, we
67 emphasize implications for future work in designing and assessing IVEs for remote learning purposes.

## Background

69 As distance learning continues to expand, catering for larger numbers of students across the globe, current
70 solutions provide inefficient delivery systems which are not immersive, engaging, or motivating to the
71 learner – often resulting in poor rates of completion (Wise et al., 2004; Yang et al., 2013; Chen, 2018).
72 For example, an investigation into the use of 'Accountable Talk: Conversation that Works', a MOOC
73 provided by the University of Pittsburgh, revealed that despite in excess of 60,000 students registering for
74 the programme less than half continued to access the course material through to completion (Rosé et al.,
75 2014). Attrition rates of learners in MOOCs is much higher than that in formal education (Clow, 2013; Joo
76 et al., 2018). Users also report that they do not perceive MOOCs as equivalent to traditional education, and
77 their engagement with them is less serious (Nemer and O'Neill, 2019). As a result, MOOCs are unable to
78 deliver educational experiences with the same rigour as formal educational institutions. While MOOCs
79 have several short-comings, the ability to study without being physically located in a certain space has
80 many advantages to both students unable to attend and universities who are coping with growing numbers
81 of students. Therefore, finding ways to improve the experience of distance learning and encouraging
82 greater levels of engagement with online courses is of great public interest and their efficacy in education
83 is of equal pedagogic interest.

84 Bringing learners into IVEs may overcome engagement issues experienced in MOOCs. Students
85 prefer to engage in traditional lectures over online courses because they lack self-discipline and they can
86 become too easily distracted during online learning (Crook and Schofield, 2017). Applying immersive
87 VR to education may engage students better than MOOCs, removing distractions outside of the learning
88 environment, mimicking the experience of traditional learning experiences (Lessick and Kraft, 2017;
89 Pirker et al., 2018). Existing examples of educational applications of VR have focused on non immersive
90 desktop-VR and have shown that simulating learning environments is highly effective. For example,
91 desktop-VR has been successfully used for social cognition training in children with Autism Spectrum
92 Disorders and for assessing procedural skills such as dissecting frogs in a laboratory study (Didehbani
93 et al., 2016; Merchant et al., 2014). IVE based learning environments have shown that learning using
94 VR results in better retention and improves learners' performance by up to a grade compared to simply
95 watching a lecture or reading (Sitzmann, 2011; Graesser et al., 2005).

96 While educational applications of desktop-VR have merit, research suggests IVEs lead to better
97 results as interaction with the environment is more intuitive, therefore users spend less time learning
98 how to use the computer interface and can focus their full attention on the task (Psotka, 1995). To date,
99 IVEs have been predominantly applied to procedural and skills-based education, successfully enhancing

learning outcomes. For example, the performance of a group of material science students on a series of questions about crystal structures improved when they were presented with virtual 3D diagrams of crystal structures via a head-mounted display (HMD), compared to using 2D diagrams (Caro et al., 2018). The ability to manipulate and rotate the crystals in the IVE helped students understand the relationships between atoms and perform better in assessment tasks than those who studied the textbook diagrams. Additionally, students reported that the IVE was easy to use and preferable over the 2D format. IVEs are also commonly used successfully to train complex psychomotor skills required by medical students. For example, AirwayVR provides a safe, immersive environment to practice endotracheal intubation procedures, leading to clear improvements in students' self-reported understanding of the procedure compared to their knowledge prior to using the application (Rajeswaran et al., 2018).

Recent research suggests applying IVEs to lecture-styled learning may provide distance learners with enriched learning experiences that are more immersive, enjoyable, and realistic (Chen, 2018). Preliminary research has shown the value in using IVEs to replicate classroom learning, finding that students perform better on a quiz about the topic after watching a virtual lecture compared to watching a video recording of a lecture, as is typically done in MOOCs (Tsaramirsis et al., 2016). Additionally, all learners reported that they preferred the IVE as it was more enjoyable, reinforcing the idea that IVEs are likely to successfully engage a larger number of distance learners than MOOC platforms. However, while educational applications of IVEs for distance learning in both higher education and corporate level training have begun to emerge, these applications are in their infancy. Recent work has explore the use of modern game development engines and HMD based environments for creating virtual lecture theatres and classrooms (Misbhauddin, 2018), but has not explored the how effective these environments are at improving learner performance, motivation, and satisfaction & engagement. As such, to the best of our knowledge there are currently no published findings regarding their effectiveness, resulting in very little robust evidence to inform how the design of an IVE impacts learning outcomes (Moro et al., 2017).

When designing IVEs, one does not consider just the aesthetic, but also the presence of other agents inside the environment. Within IVEs, embodied agents (EAs) are frequently used as pedagogical agents for virtual tutoring. For example, STEVE is a human-like EA used to teach engineers how to use complex machinery onboard ships (Johnson and Rickel, 1997). In addition to humanoid EAs there are non-human examples such as Herman the bug–a non-humanoid EA implemented in Design-A-Plant, a virtual environment used to teach children about plant biology and the environment (Lester et al., 1999). The appearance and behaviour of EA tutors influences learners' feelings of co-presence (Baylor, 2011; Baylor and Kim, 2009) – the perception that one is not alone but in the presence of others (Heeter, 1992; Short et al., 1976). Co-presence increases when an EA tutor has appearance and behavioural realism – a key point being that there is no mismatch between appearance and behavioural realism, as this results in very low levels of perceived co-presence (Bailenson et al., 2005).

Increasing a learner's perceived co-presence increases learner satisfaction and motivation to engage with material. For example, it has been shown that learner's spend approximately 25% more time learning and report that the learning experience is more enjoyable when an EA is present (Sträfling et al., 2010). A limitation of current distance learning platforms, such as MOOCs, is that learners must try to maintain enthusiasm and motivation to complete the course in the absence of an educator (Hasegawa et al., 2014). Implementing an appropriate EA which represents a lecturer within an IVE may maintain learner interest and motivation, positively impacting learning outcomes. The appearance of the virtual tutor impacts a learner's perception of the tutor's abilities. For example, human-like agents are perceived as more intelligent and helpful compared to non-human agents (King and Ohya, 1996; Lester and Stone, 1997), while familiar agents are rated more positively than unfamiliar agents (Bailenson et al., 2005). Previous research has demonstrated that virtual tutor realism influences learners' reported likability and motivation (Maldonado and Nass, 2007), in turn influencing performance. Thus, we expect that a realistic EA tutor which is familiar to the learner will be more likeable, improving the learning experience and motivation to learn (Maldonado and Nass, 2007; Scaife and Rogers, 2001).

While some evidence suggests that EAs play a substantial role in the learning experience, increasing learning efficiency and retention (Roussou et al., 2006), others have found minimal-to-no effect of EAs on learning outcomes. For example, in one study EAs were found to have no influence and prior knowledge was identified as the greatest contributing factor to learner performance (Sträfling et al., 2010). Therefore, to clearly establish the utility of EAs within IVEs researchers should aim to control this potentially extraneous variable to prevent participants' prior knowledge of the topic concealing any effects of the EA.

Overall, IVEs for educational purposes have the potential to mimic traditional learning experiences greater than MOOCs. By utilising IVEs to develop more engaging distance learning experiences, universities and corporate training bodies may cater for increasing student numbers. However, a major barrier to implementing IVEs compared to MOOCs is the higher relative cost of the equipment required. Therefore, it is essential that interdisciplinary research is conducted to establish whether IVEs, which can be run on low powered hardware such as smartphones, are able to provide a method of engaging more students, provide remote-learners with an experience which is more equivalent to formal education, and make a worthwhile contribution to higher education institutions looking to provide effective distance learning.

## User Study

Our study focuses on the educational applications of IVEs, specifically investigating the effectiveness of learning novel information in an IVE compared to a physical classroom, and the role of embodied agents as tutors within the IVEs. We devised the following hypotheses:

**H1:** Participants who learn inside an IVE will learn more effectively and outperform participants who learn in a physical classroom since prior research has shown that virtual learning environments result in better retention and improves performance (Sitzmann, 2011; **?**).

**H2:** Participants who learn in the presence of an EA tutor will outperform participants who learn without one because the presence of a virtual tutor influences motivation which may in turn influence performance (Sträfling et al., 2010).

**H3:** The presence of a humanoid EA tutor will be more likable and increase motivation in learners compared to an abstract EA tutor because more familiar and realistic tutors are more likeable and motivating (Bailenson et al., 2005; Maldonado and Nass, 2007).

# MATERIALS & METHODS

A between-participant design was used, whereby participants were randomly assigned to one of the four learning conditions. The independent variable was the learning environment (IVE with no tutor, IVE with a humanoid tutor, IVE with an abstract tutor, Non-virtual learning environment). The dependent variables were performance (test score), and reported motivation, satisfaction, and engagement (questionnaire). The experiment lasted approximately $45 \sim 60$ minutes.

## Design & Apparatus

Opportunity sampling was used to recruit participants from a local university. The target sample size for this initial study was 48 participants, split equally between the four learning conditions, based on the results of an a-priori power analysis, conducted using G*Power 3 (Faul et al., 2007), revealing a one-way ANOVA with 12 participants per group would provide 80% power to detect an effect size of 0.5 at a significance level of .05. Note that we instead used a more conservative Kruskal-Wallis test rather than the ANOVA to evaluate the results; prior work has shown that Kruskal-Wallis has greater statistical power than ANOVA under these conditions, so the analyses presented were indeed sufficiently powered (Hecke, 2012). In total, 48 participants were recruited (24 M, 24 F), aged 18 and over (M = 20.8 years, SD = 3.3 years). All participants were students from a variety of disciplines who reported normal or corrected to normal vision and hearing. Participants were incentivized through course credit (n = 8), £5 reward (n = 22), or simply volunteered to participate (n = 18). Statistical tests confirmed that there was no significant effect of the type of incentive received on participant performance (See Supplementary Material).

A machine running Windows 10 with a single Nvidia 970 GPU was adequate to drive the virtual environment since it is without cutting-edge graphics. To display the virtual environment, we used the HTC Vive head-mounted display (HMD). This HMD covers a 110°field of view, with two 1080 x 1200 pixel screens to render stereoscopic graphics to the viewer. Head position and orientation were tracked using the hardware base stations packaged with the HMD. However, the environment can also be demoed as a mobile application and we expect that in a larger cohort the set up could be easily scaled up using more consumer-friendly devices such as Smartphones and Google cardboards.
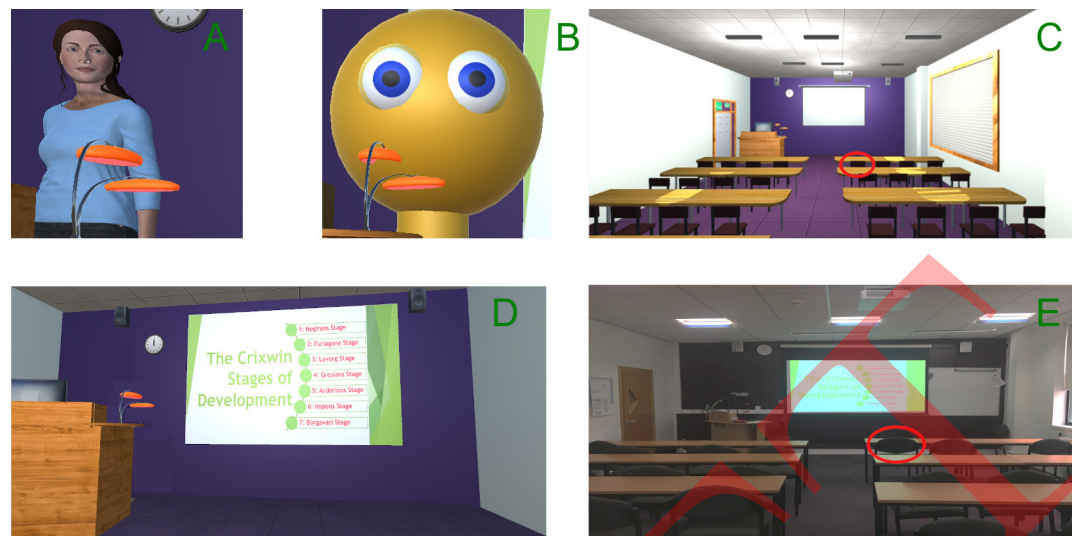
**Figure 1.** In row-order from the top-left: the humanoid embodied agent tutor (A), the non-human tutor (B), a view of the entire virtual classroom environment (C), a view from the perspective of participants in our experiment showing the novel learning material (D), and a view of the real world classroom (E). Participants sat in the same position in both real and virtual environments as shown by the red circles.

### Learning Environments & Material

A seminar room on a local university campus was used as the non-virtual learning environment (See Figure 1). A PowerPoint was projected onto the screen to display the learning material and the female experimenter represented the tutor, reading a script alongside each slide (See supplementary materials).

A virtual replica of the physical classroom, made to scale in order to minimise the number of extraneous variables (Figure 1 panel C) was created using Unity 2018.2.17. To replicate the appearance colours and textures were applied and generic classroom furniture were used to decorate the virtual environment. To display the PowerPoint slides in the IVEs, custom software applied images of the PowerPoint slides as textures to the virtual projector screen. The lecture slide changed to the next one in sequence when spacebar was pressed. Audio recordings of the female experimenter reading the script were automatically played with each slide to keep delivery of the lecture material consistent for all participants.

For the IVE with a humanoid tutor, a female avatar was created using AdobeFuse CC Beta and imported into the environment (Figure 1 panel A). The female avatar has an animator controller to loop an 'idle' and an 'eye blink' motion to appear more realistic. For the IVE with an abstract tutor, a block-shape representation was created from geometrically primitive shapes (Figure 1 panel B). The abstract tutor was animated using key frame animation which moved the body side-to-side and rotated the eyes to replicate the humanoid 'idle' and 'blink' motions. Novel information was created for this study about the developmental stages of a made-up alien species. This was used as the learning material in order to eliminate the possibility of prior knowledge becoming a confounding variable (See supplementary material).

### Questionnaire

All data were recorded using Qualtrics, a web browser interface that automatically recorded responses from participants. The first section of the questionnaire contained the knowledge test, composed of 29 questions designed to test participants' knowledge of the alien species. The majority of the questions were multiple choice in order to test retention, with some short answer questions to test comprehension (Schrader and Bastiaens, 2012) (See supplementary material). However, multiple choice tests have been critiqued as they 'feed' students the answers, making it possible to gain artificially high scores (Bush, 2001). Therefore to accurately reflect retention, the test was negatively marked (meaning correct answers were given a score of 1, incorrect answers scored -1, and any unanswered questions scored 0) to discourage guessing (Davies, 2002). The test was marked to produce a score to indicate participant performance.

Three blocks of questions followed: learner satisfaction and engagement (7 items); learner motivation (3 items); and virtual presence (5 items; see supplementary materials for the questionnaire).

***Learner Satisfaction and Engagement:***
The questions designed to measure satisfaction and engagement with the learning experience were 5-point Likert scales which asked participants how strongly they agreed or disagreed with the statements. For example, "The learning experience captured my interest". Each item was scored out of five (5 = Strongly Agree) and added together to produce a learner satisfaction and engagement score out of 35. The questions were created for the purpose of this experiment as it aimed to measure specifically how engaged 'students' were in this one experience. We had considered using an existing student satisfaction questionnaire but opted to develop our own so that questions could be focused on the experience in our study.

***Learner Motivation:***
Another block of questions was specifically tailored to investigate the effects of the EA tutor manipulation on learner motivation, for example, "The presence of the tutor increased my motivation to learn". Each item included a 5-point Likert scale which asked participants to what extent they agreed with each statement, with 'Strongly Agree' being scored as five. The item scores were added together to produce a learner motivation score out of 15.

***Virtual Presence:***
The virtual presence questions were taken from an existing questionnaire (Witmer and Singer, 1998). The most applicable items were selected, for example "To what degree did your experiences in the virtual environment seem consistent with your real-world experiences?", participants responded to each statement via a 5-point scale (1 = Not at all, 5 = Completely) to indicate how immersive the IVEs were, and this produced a virtual presence score out of 25.

Finally, to measure any potentially confounding effects participants in the IVE with a humanoid tutor were asked to report anything they found 'odd' about the human avatar, as a perceived mismatch between appearance and expected behaviour, for example 'speaking' with no changing facial expression or lip movement, could lead to disliking of the tutor and affect performance (Mori, 1970; Bailenson et al., 2005).

**Procedure**
This study was approved by the University of Bath Psychology Research Ethics Committee (17-292). Participants were provided with further information about the study before giving written informed consent. Only one participant took part in the experiment at a time. Each participant was randomly allocated to one of the four learning conditions.

Participants in the IVEs were seated at a computer in the laboratory, the experimenter would assist with fitting the headset and headphones to ensure the participant was comfortable. The experimenter would then launch the appropriate classroom application (i.e. with a humanoid/abstract/no tutor). Participants allocated to the non-virtual condition were seated in a seminar room on the university campus. Participants in the non-virtual condition took part in the experiment individually: the only other person present in the room was the experimenter. In both virtual and non-virtual conditions participants were shown the same PowerPoint presentation, and heard the same experimenter deliver the scripted information. The only difference being that in the virtual condition the voice-clips were pre-recorded and incorporated into the environment, whereas in the non-virtual condition the experimenter delivered the information in person. In all conditions, participants observed the full presentation with corresponding audio once, and were then allowed the remaining time to read through the slides themselves with no audio input. After 30-minutes the experimenter halted the learning part of the experiment, and those in the IVEs would be asked to remove the headset. All participants were then required to complete the online test and questionnaire.

Throughout the experiment, all participants remained naïve to the manipulation of the tutor and the environment. Afterwards all participants were fully debriefed, and the full aims of the study were revealed, participants then provided final consent for the data to be used.

# ANALYSIS & RESULTS

Frequentist null hypothesis significance testing and the associated *p*-value has many shortcomings, for example it relies on hypothetical data and can be easily manipulated - with larger sample sizes able to make small differences significant without any practical value (Jarosz and Wiley, 2014). Bayes Factor (BF) is a

**Table 1.** Kruskal-Wallis results summary covering the 3 core variables in our study: learner performance, satisfaction & engagement, and sense of presence.

| Measure | non-virtual | | virtual, no tutor | | virtual, humanoid tutor | | virtual, abstract tutor | | $H(3)$ | $p$ | $\eta^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | M | SD | | | |
| Performance | 30.75 | 11.67 | 28.92 | 10.22 | 28.33 | 9.20 | 26.33 | 8.17 | 2.806 | .422 | .06 |

| Measure | no tutor | | humanoid tutor | | abstract tutor | | | | $H(2)$ | $p$ | $\eta^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | | | | | |
| Satisfaction & Engagement | 27.42 | 5.35 | 25.75 | 5.91 | 25.67 | 5.77 | | | 2.954 | .399 | .063 |
| Presence | 14.42 | 2.27 | 14.67 | 2.46 | 14.42 | 2.39 | | | .066 | .968 | .001 |

**Table 2.** t-test results for the impact of tutor on motivation to learn.

| Measure | abstract tutor | | humanoid tutor | | $t(22)$ | $p$ | $d$ |
|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | | | |
| Motivation | 10 | 2.2 | 8.8 | 2.1 | -1.316 | .202 | -.537 |

ratio which indicates the likelihood of the observed data fitting under either of the two hypotheses (Jarosz and Wiley, 2014), meaning if the null hypothesis is more plausible than the alternative hypothesis the stronger (i.e. closer to 0) the BF becomes (Dienes, 2011). Therefore, Bayesian statistics were conducted using JASP version 0.9 to determine the relative strength of the support for the null versus alternative hypotheses. BF represents the likelihood that the evidence is explained by one hypothesis over another, for example a BF of 20 would indicate that one hypothesis is 20 times more likely to explain the data. BF can be given as $BF_{10}$ (evidence for the alternative hypothesis) or $BF_{01}$ (evidence for the null hypothesis) (Schut et al., 2018). We used $BF_{01}$ values as they are easier to interpret in relation to our findings. Based on this interpretation scheme, $BF_{01}$ values of 3 to 10 indicate moderate support for the null hypothesis, while values $< 3$ indicate weak support for the null hypothesis (Wagenmakers et al., 2018). Tables 1 & 2 presents a summary of these results.

Additional statistical analyses carried out using SPSS version 24.0 were evaluated against an alpha level of 0.05. An independent t-test was used to determine differences in learner motivation between the humanoid tutor IVE and the abstract tutor IVE. Assumption checking revealed that the data were normally distributed as assessed by the Shapiro-Wilk test (p > .05), and the variance in each group was approximately equal, as assessed by Levene's test for homogeneity of variances (p > .05). However, due to the small sample size three Kruskal-Wallis tests were conducted to compare learner performance, satisfaction and engagement, and virtual presence ratings across multiple learning conditions. Preliminary analyses confirmed that the data met the test assumptions as there were no extreme outliers, and there was homogeneity of variances.

**Learner Performance**

Mean scores for learner performance in each learning conditions are shown in Figure 2. On average, learner performance was similar across all learning environments, with only slight differences among conditions. With respect to **H1** and **H2**, we conducted a Kruskal-Wallis test and results report no statistically significant differences in performance scores between conditions, $H(3) = 2.806$, $p = .422$, $\eta^2 = .06$. , this result is moderately reinforced by the Bayes statistics which indicate that the data are six times more likely to be explained by the null hypothesis ($BF_{01} = 6.2$).

**Learner Satisfaction & Engagement**

Mean scores for learner satisfaction and engagement in each virtual learning condition are shown in Figure 3. Learner satisfaction and engagement levels, as measured by 7 items in the questionnaire (Cronbach $a = .84$), were similar across all virtual learning conditions, with only slightly higher levels measured in the IVE with no tutor. A Kruskal-Wallis test supported that learner satisfaction and engagement levels were not significantly different between the virtual learning conditions, $H(3) = 2.954$, $p = .399$, $\eta^2 = .063$ Furthermore, Bayes statistics indicate that the data are four times more likely to be explained by the null
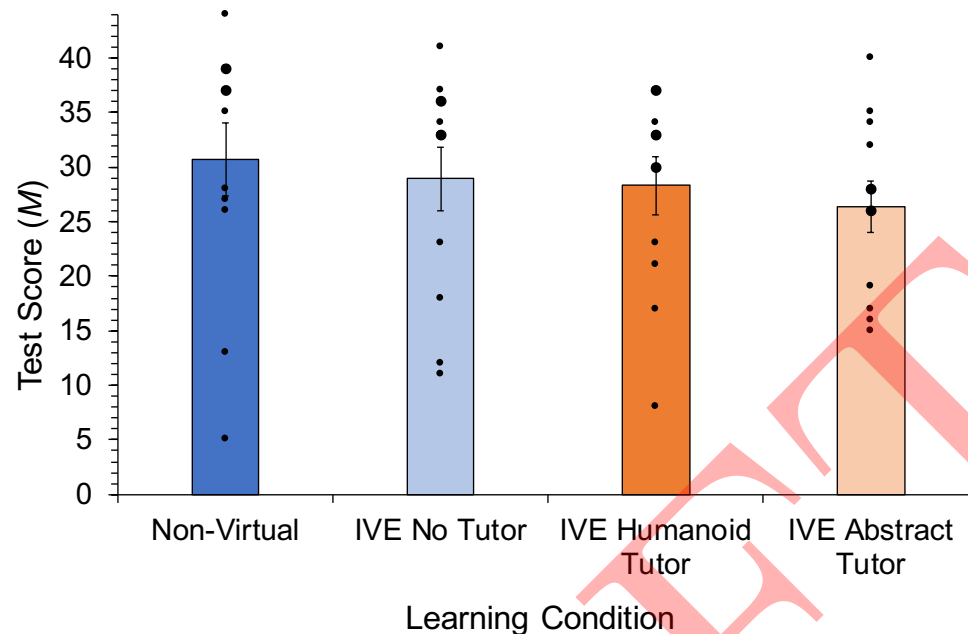
**Figure 2.** Mean test scores in the different learning environments. Error bars represent the standard error (SE). Dots show distribution of participant scores, with larger dots indicating multiple participants with the same score.

319   hypothesis ($BF_{01} = 4.1$).

320   **Learner Motivation**

321   Mean scores for learner motivation in the presence of humanoid and non-humanoid EAs are shown
322   in Figure 4. Learner motivation scores, measured using 3 items in the questionnaire (Cronbach $a$ =
323   .62), appeared higher in the IVE with the abstract tutor (M = 10.0, SD = 2.2) compared to learner
324   motivation scores in the IVE with the humanoid tutor (M= 8.8, SD = 2.1). With respect to **H3**, an
325   independent samples t-test revealed that these differences in learner motivation between conditions were
326   not significantly different, $t(22) = -1.316$, $p = .202$, $d = -.537$. However, Bayes statistics only indicate
327   very weak support for the null hypothesis in this case as it suggests that the null hypothesis is only one
328   times more likely to explain the data ($BF_{01} = 1.4$).

329   **Virtual Presence**

330   Virtual presence was measured using only a subset of Witmer and Singer's presence questionnaire so
331   measures of internal reliability were not conducted. Learner ratings of virtual presence were consistent
332   across the different IVEs (no tutor M= 14.4, SD= 2.3; human tutor M= 14.7, SD= 2.5; abstract tutor
333   M= 14.3, SD= 2.3) A Kruskal-Wallis test supported that virtual presence scores were not significantly
334   different between the virtual reality learning conditions, $H(2) = .066$, $p = .968$, $\eta^2 = .001$, furthermore
335   Bayes statistics provide moderate support for the null ($BF_{01}= 5$).

336   **Learner Perceptions of Avatar**

337   In response to the question "Did you notice anything odd about the human avatar?" 58% of the participants
338   in the IVE with the humanoid tutor reported that they did find the humanoid EA tutor strange. The reasons
339   for answering 'yes' to the question were that the avatar had strange or repetitive movement, no changing
340   facial expressions, and did not speak.

341   **DISCUSSION**

342   Previous research and educational applications of VR have focused on desktop-VR simulations for
343   skills-based tasks (Freina and Ott, 2015), neglecting the use of more immersive virtual environments
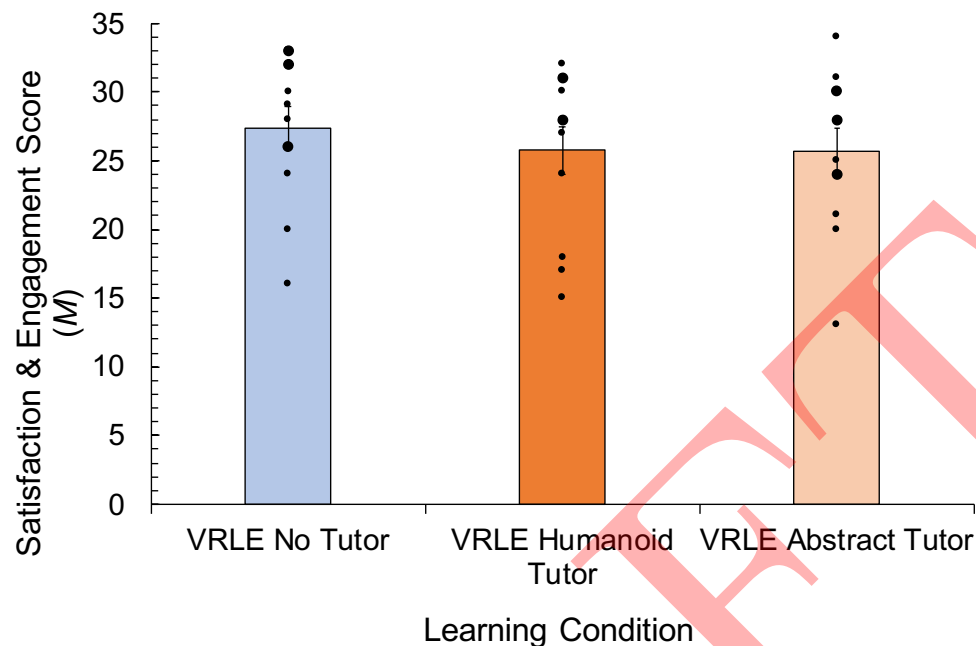
**Figure 3.** Mean satisfaction and engagement scores with error bars representing SE, for the predictor of learning condition within immersive virtual environments (no tutor, humanoid tutor, abstract tutor). The real environment was not modelled as a condition in this analysis and therefore means for three conditions are shown. Error bars and dots as in Figure 2
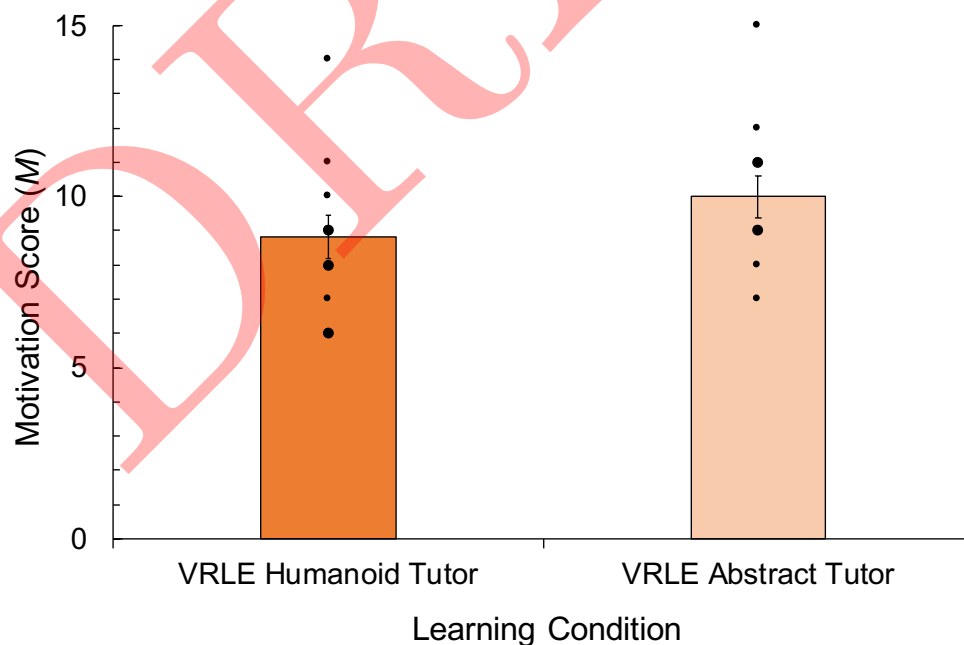
**Figure 4.** Results of a t-test conducted to analyse the impact of humanoid vs. abstract tutor representation on learner motivation scores, and therefore means for two conditions are shown. Error bars and dots as in Figure 2.

(IVEs) and their potential use for informational, lecture-styled learning experiences. Therefore, in this pilot study we investigated to what extent informational-learning within an IVE is effective compared to learning in a physical classroom. We created a virtual replica of a classroom and compared its use for informational learning to the traditional, real-world classroom.

Previous literature indicated that simulated learning environments are highly effective, enhance declarative knowledge and lead to better retention compared to conventional learning methods (Graesser et al., 2005; Merchant et al., 2014; Sitzmann, 2011). While desktop-VR has dominated the literature, it is argued that more immersive experiences may lead to even greater results (Psotka, 1995). Therefore, we hypothesized that participants who learned virtually would outperform participants who learned non-virtually (**H1**). However, results demonstrated that participants who learned virtually did not outperform participants who learned non-virtually. A Bayes Factor analysis provides moderate support for this finding as it suggests that the data are six times more likely to be explained by the null hypothesis. It is plausible that familiarity with learning material, which is known to impact learner engagement, performance, and motivation (Schönwetter et al., 2002), would have impacted our results: to combat this effect we used fabricated information to eliminate prior knowledge as a confounding variable. Thus our findings are robust, indicating there is no detriment to learning in an IVE compared to a conventional classroom setting (Madden et al., 2020).

The lack of a statistically significant difference in learner performance between the IVE and the non-virtual classroom is of particular importance as educational institutions are under increasing pressure to cater for large numbers of students and as such require effective distance learning applications (Kauffman, 2015). Current distance learning platforms suffer from poor student engagement and high levels of drop-out (Yang et al., 2013), however, IVEs have the potential to improve this. Previous research has demonstrated that IVEs provide a more engaging platform for distance learners than existing video-based applications (Tsaramirsis et al., 2016). In light of the novel SARS-COV-2 (COVID-19) pandemic and its impact on the higher education sector, namely creating situations where many institutions have closed their campus until further notice, IVEs may yield a better experience for distance learners. Furthermore, our research supports that distance learning applications would benefit from incorporating the use of IVEs, as distance-learner performance would be consistent with learners in traditional settings, yet IVEs can be used on a much larger scale, making them highly cost-efficient.

### The Role of Embodied Agents

Few studies have been conducted which inform how the design and use of EA tutors impacts the learner experience (Moro et al., 2017), so we investigated the role of EA tutors within IVEs. We manipulated the tutor in the IVE on two dimensions; its presence or absence, and its human-like representation. Previous research has suggested that co-presence influences the learning experience, with higher feelings of co-presence resulting in greater learning performance (Roussou et al., 2006; Wise et al., 2004). Therefore, we hypothesized that participants who learn in the presence of an EA tutor will outperform participants who learn in its absence.

We found no statistically significant difference in performance of learners who learned without a virtual tutor, with a humanoid tutor, or with an abstract tutor. Participants who learned in the IVE without an EA tutor were expected to perform worse on the post-learning test. Our findings fit into current discourse and debate on the utility of EAs and their impact on learning: while some research has concluded no impact on performance, consistent with ours (Sträfling et al., 2010), other research has found that EAs do impact learning performance (Baylor Amy L., 2009; Maldonado and Nass, 2007; Rosenberg-Kima et al., 2007). One explanation for our results is the age of participants in our study. Previous research in agreement with our results used young adults (Sträfling et al., 2010), while others have recruited young school children (Roussou et al., 2006). The presence of virtual avatars is known to positively impact learning in young children (Darves et al., 2002) and it is possible that the positive impact of EA tutors on learner performance may be confined to when IVEs are used by younger students (Baylor and Kim, 2004; Ashby Plant et al., 2009).

EA tutors impact learner satisfaction and engagement within IVEs by simulating the relationship between student and tutor (Alseid and Rigas, 2010). A learner's social judgement of interactions with an EA impact perceived co-presence and satisfaction, with human-like representations regarded as more social than non-human avatars (Nowak, 2004). Therefore, we hypothesized that a humanoid EA tutor would be preferred over an abstract EA. However, our results show no statistically significant differences

in learner satisfaction and engagement when comparing the IVE with no tutor, the humanoid tutor, or the abstract tutor. Bayes Factor analysis indicates that the data were four times more likely to be explained by the null hypothesis in this instance and therefore can be accepted with moderate confidence (Wagenmakers et al., 2018).

Although measures were taken to provide the humanoid tutor with a realistic appearance and behaviour, such as using deictic gestures and a natural human voice (Atkinson et al., 2005; Baylor, 2011; Baylor and Ryu, 2003; Janse, 2002), the avatar was not equipped with any animations which replicated changing facial expressions. This may have hindered the level of satisfaction and engagement the humanoid tutor was able to evoke in the learners, which is known to influence perceived realism (Atkinson, 2002). In our study, many participants exposed to the humanoid tutor commented on the absence of facial expression, with the majority of participants feeling as though it had 'strange' and 'repetitive movement'. In contrast, participants did not have pre-defined expectations of how the abstract tutor should behave and as such it was not susceptible to the uncanny valley effect, unlike the humanoid avatar (Bailenson et al., 2005; Mori, 1970). Therefore, rather than the humanoid tutor increasing motivation, participants may have found the abstract tutor more appealing. This mismatch between learner expectations of the tutor and reality may have been detrimental to the perceived co-presence (Bailenson et al., 2005). Therefore, it is possible that a lack of emotional expression in the humanoid avatar contributed to the absence of significantly greater learner satisfaction and engagement.

Previous research has indicated that EA tutors affect learning outcomes indirectly by influencing learner motivation (Baylor, 2011). Research suggests greater likeability, and ascribed intelligence when using a humanoid EA tutor (King and Ohya, 1996; Lester and Stone, 1997); therefore, we expected the humanoid tutor would increase levels of learner motivation. However, there was no statistically significant difference in learner motivation between the humanoid and abstract tutor groups immersed in the VLE. Bayes Factor analysis suggests that the data are almost equally likely to be explained by either the null or the alternative hypothesis. Thus, we do not rule out the possibility that tutor appearance can affect motivation.

A distinct strength of our study is the control for immersion as a factor which could influence the efficacy of IVEs, as the more immersive the environment the more comparable it is thought to be to real-world environments. To determine if learning outcomes are affected by immersion a virtual-presence questionnaire was used. The results indicated similarly high levels of immersion in all three IVEs, meaning that the IVEs are comparable to non-virtual learning (Peperkorn et al., 2015; Shin, 2018) and ensuring that environment quality was unlikely to produce any differences in learning outcomes.

**Future Work**

While this pilot study provides preliminary support for the use of IVEs by demonstrating that learning within an IVE is not significantly different to non-virtual learning, this is only demonstrated in the immediate short-term as the test and outcome measures were administered immediately after the learning experience. For effective distance learning applications, long-term outcomes need to be assessed, perhaps in the realm of a longitudinal study. Future work should consider incorporating additional follow-up assessment periods, in order to provide evidence for whether the performance outcomes observed in the IVE and the non-virtual classroom are maintained over a longer period of time. In this preliminary study recruitment was restricted to the student population at the university, however it is likely that large differences in learning styles and ability will vary within this population resulting in a large range of scores in all conditions. Future studies using a larger sample-size should consider the prior grades of all participants, and use random allocation to minimise the effects of individual differences.

Additionally, the present study highlights the need for further investigation into the impact of embodied agents in IVEs to understand the varying results surrounding their impact on learner performance, motivation, and satisfaction. Previous work has highlighted the impact of graphical realism on peoples' perceptions of avatars in virtual environments while engaging in various tasks in various scenarios (Tessier et al., 2019; Kang et al., 2008; Lugrin et al., 2015). Our goal was to assess the importance of a humanoid avatar, not necessarily the physical realization of said avatar. Future work may consider graphical fidelity and realism as a factor in learner motivation, presence, and satisfaction and engagement with the learning experience. A possible trend in the literature is based around the age of participants, with younger participants seemingly more likely to be influenced by the presence of an EA. Future research should seek to investigate this theory.

In the present study there was no verbal interaction allowed between participants and the tutor in all learning conditions in order to remove the likelihood of differing levels of social interaction between participants and the tutor becoming a confounding variable. Furthermore, the EA was not equipped with any facial animation to replicate changing expression, both of which likely had a negative impact on the perceived realism. Future work investigating learner satisfaction, engagement, and motivation should consider introducing EA's with changing expressions and allow verbal interaction, such as the ability to ask the tutor questions, as this may better simulate student-tutor relationships and have a greater impact upon perceived co-presence, producing more insightful results regarding the role of EA tutors within IVEs. It may also reduce the strangeness reported in Section  as the avatar's behaviour is improved.

Finally, we highlight a novel avenue for future research: whether the influence of an EA on learning outcomes are mediated by their relevance to the learning material itself. In our study, participants studied fabricated information about an alien species, hence the abstract tutor may be more salient in this context, promoting interest in the learning material to a greater extent than the humanoid tutor (Maldonado and Nass, 2007). Future work will assess the link between learning material and the EA tutor's appearance, as well as its contextual relevance and form within the IVE.

## CONCLUSIONS

To the best of our knowledge, this pilot study is the first to directly compare informational-learning in a traditional classroom to a virtual replica using immersive VR for groups of participants in a controlled, laboratory setting. Our findings suggest that learner performance is equivalent in both learning situations. It remains unclear how the design of the IVE might impact learning outcomes, in particular whether the presence and appearance of the virtual tutor plays a role in learning outcomes. We have discussed avenues for future work, building on our preliminary study and exploring other factors which may impact learning performance in IVEs as well as guidelines and recommendations for how to design future experiments which control for extraneous variables. There are important implications for developers of distance learning applications; by providing IVEs opposed to video-based applications, it is possible to reduce the issues with current distance learning platforms and achieve comparable performance levels with those who learn in a traditional classroom, making it possible to cater for increasing numbers of students. As academic and corporate education moves towards IVEs under increasing pressure to meet the demands of growing numbers of students (Kauffman, 2015), the scalability and significant financial incentives they provide while maintaining satisfactory learning outcomes make them an attractive alternative to current video-based distance learning platforms. In addition, the SARS-COV-2 (COVID-19) pandemic has also forced institutions to rethink their ability to provide effective blended learning and virtual learning environments for students. IVE technology can help to create effective learning environments that are safe for staff and students, and continue to provide high quality learning and teaching.

## REFERENCES

Alseid, M. and Rigas, D. (2010). Three different modes of avatars as virtual lecturers in e-learning interfaces: a comparative usability study. *The Open Virtual Reality Journal*, 2:8–17.

Ashby Plant, E., Baylor, A. L., Doerr, C. E., and Rosenberg-Kima, R. B. (2009). Changing middle-school students' attitudes and performance regarding engineering with computer-based social models. *Computers & Education*, 53(2):209–215.

Atkinson, R. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, 94(2):416–427.

Atkinson, R. K., Mayer, R. E., and Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology*, 30(1):117–139.

Bailenson, J. N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., and Blascovich, J. (2005). The Independent and Interactive Effects of Embodied-Agent Appearance and Behavior on Self-Report, Cognitive, and Behavioral Markers of Copresence in Immersive Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 14(4):379–393.

Baylor, A. L. (2011). The design of motivational agents and avatars. *Educational Technology Research and Development*, 59(2):291–300.

Baylor, A. L. and Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior*, 25(2):450–457.

Baylor, A. L. and Kim, Y. (2004). Pedagogical Agent Design: The Impact of Agent Realism, Gender, Ethnicity, and Instructional Role. In Lester, J. C., Vicari, R. M., and Paraguaçu, F., editors, *Intelligent Tutoring Systems*, Lecture Notes in Computer Science, pages 592–603. Springer Berlin Heidelberg.

Baylor, A. L. and Ryu, J. (2003). The Effects of Image and Animation in Enhancing Pedagogical Agent Persona. *Journal of Educational Computing Research*, 28(4):373–394.

Baylor Amy L. (2009). Promoting motivation with virtual agents and avatars: role of visual presence and appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3559–3565.

Bush, M. (2001). A Multiple Choice Test that Rewards Partial Knowledge. *Journal of Further and Higher Education*, 25(2):157–163.

Caro, V., Carter, B., Dagli, S., Schissler, M., and Millunchick, J. (2018). Can Virtual Reality Enhance Learning: A Case Study in Materials Science. In *2018 IEEE Frontiers in Education Conference (FIE)*, pages 1–4.

Chen, S. (2018). Research on the Application of Virtual Reality in Remote Education Based on the Example of MOOC. In *2018 15th International Conference on Service Systems and Service Management (ICSSSM)*, pages 1–4.

Clow, D. (2013). MOOCs and the Funnel of Participation. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, LAK '13, pages 185–189, New York, NY, USA. ACM.

Crook, C. and Schofield, L. (2017). The video lecture. *The Internet and Higher Education*, 34:56–64.

Darves, C., Oviatt, S., and Coulston, R. (2002). The impact of auditory embodiment on animated character design. In *Proceedings of the International Joint conference on autonomous agents and multi-agent systems, embodied agents workshop*.

Davies, P. (2002). "There's no Confidence in Multiple-Choice Testing, . . . . . .".

Didehbani, N., Allen, T., Kandalaft, M., Krawczyk, D., and Chapman, S. (2016). Virtual Reality Social Cognition Training for children with high functioning autism. *Computers in Human Behavior*, 62:703–711.

Dienes, Z. (2011). Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspect Psychol Sci*, 6(3):274–290.

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2):175–191.

Feng, Y., Chen, D., Zhao, Z., Chen, H., and Xi, P. (2015). The Impact of Students And TAs' Participation on Students' Academic Performance in MOOC. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, ASONAM '15, pages 1149–1154, New York, NY, USA. ACM.

Freina, L. and Ott, M. (2015). A literature review on immersive virtual reality in education: state of the art and perspectives.

Freitas, A. and Paredes, J. (2018). Understanding the faculty perspectives influencing their innovative practices in MOOCs/SPOCs: a case study. *International Journal of Educational Technology in Higher Education*, 15(1):5.

Graesser, A. C., Chipman, P., Haynes, B. C., and Olney, A. (2005). AutoTutor: an intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, 48(4):612–618.

Greenwald, S. W., Kulik, A., Kunert, A., Beck, S., Fröhlich, B., Cobb, S., Parsons, S., Newbutt, N., Gouveia, C., Cook, C., Snyder, A., Payne, S., Holland, J., Buessing, S., Fields, G., Corning, W., Lee, V., Xia, L., and Maes, P. (2017). Technology and Applications for Collaborative Learning in Virtual Reality.

Hasegawa, D., Uğurlu, Y., and Sakuta, H. (2014). A human-like embodied agent learning tour guide for e-learning systems. In *2014 IEEE Global Engineering Education Conference (EDUCON)*, pages 50–53.

Hecke, T. V. (2012). Power study of anova versus Kruskal-Wallis test. *Journal of Statistics and Management Systems*, 15(2-3):241–247. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/09720510.2012.10701623.

Heeter, C. (1992). Being There: The Subjective Experience of Presence. *Presence: Teleoperators and Virtual Environments*, 1(2):262–271.

Janse, E. (2002). Time-Compressing Natural and Synthetic Speech. In *Proceedings of the 7th International Conference on Spoken Language Processing*.

Jarosz, A. and Wiley, J. (2014). What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *The Journal of Problem Solving*, 7(1).

Johnson, W. L. and Rickel, J. (1997). Steve: An Animated Pedagogical Agent for Procedural Training in Virtual Environments. *SIGART Bull.*, 8(1-4):16–21.

Joo, Y. J., So, H.-J., and Kim, N. H. (2018). Examination of relationships among students' self-determination, technology acceptance, satisfaction, and continuance intention to use K-MOOCs. *Computers & Education*, 122:260–272.

Kang, S.-H., Watt, J. H., and Ala, S. K. (2008). Communicators' Perceptions of Social Presence as a Function of Avatar Realism in Small Display Mobile Communication Devices. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, pages 147–147. ISSN: 1530-1605.

Kaplan, A. M. and Haenlein, M. (2016). Higher education and the digital revolution: About MOOCs, SPOCs, social media, and the Cookie Monster. *Business Horizons*, 59(4):441–450.

Kauffman, H. (2015). A review of predictive factors of student success in and satisfaction with online learning - ALT Open Access Repository. *Research in Learning Technology*, 23.

King, W. J. and Ohya, J. (1996). The Representation of Agents: Anthropomorphism, Agency, and Intelligence. In *Conference Companion on Human Factors in Computing Systems*, CHI '96, pages 289–290, New York, NY, USA. ACM.

Korallo, L. (2010). *Use of virtual reality environments to improve the learning of historical chronology.* phd, Middlesex University.

Lessick, S. and Kraft, M. (2017). Facing reality: the growth of virtual reality and health sciences libraries. *Journal of the Medical Library Association : JMLA*, 105(4):407–417.

Lester, J. C. and Stone, B. A. (1997). Increasing believability in animated pedagogical agents. In *Proceedings of the first International Conference on Autonmous Agents*, volume 5. Association for Computing Machinery.

Lester, J. C., Stone, B. A., and Stelling, G. D. (1999). Lifelike Pedagogical Agents for Mixed-initiative Problem Solving in Constructivist Learning Environments. *User Modeling and User-Adapted Interaction*, 9(1):1–44.

Lin, L., Parmar, D., Babu, S. V., Leonard, A. E., Daily, S. B., and Jörg, S. (2017). How Character Customization Affects Learning in Computational Thinking. In *Proceedings of the ACM Symposium on Applied Perception*, SAP '17, pages 1:1–1:8, New York, NY, USA. ACM.

Lugrin, J.-L., Wiedemann, M., Bieberstein, D., and Latoschik, M. E. (2015). Influence of avatar realism on stressful situation in VR. In *2015 IEEE Virtual Reality (VR)*, pages 227–228. ISSN: 2375-5334.

Madden, J., Pandita, S., Schuldt, J. P., Kim, B., Won, A. S., and Holmes, N. G. (2020). Ready student one: Exploring the predictors of student learning in virtual reality. *PLOS ONE*, 15(3):e0229788.

Maldonado, H. and Nass, C. (2007). Emotive Characters Can Make Learning More Productive and Enjoyable: It Takes Two to Learn to Tango. *Educational Technology*, 47(1):33–38.

Martha, A. S. D. and Santoso, H. B. (2019). The Design and Impact of the Pedagogical Agent: A Systematic Literature Review. *Journal of Educators Online*, 16(1). Publisher: Journal of Educators Online.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., and Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, 70:29–40.

Misbhauddin, M. (2018). VREdu: A Framework for Interactive Immersive Lectures using Virtual Reality. In *2018 21st Saudi Computer Society National Computer Conference (NCC)*, pages 1–6.

Mori, M. (1970). On the Uncanny Valley. *Energy*, 7(4).

Moro, C., Stromberga, Z., and Stirling, A. (2017). Virtualisation devices for student learning: Comparison between desktop-based (Oculus Rift) and mobile-based (Gear VR) virtual reality in medical and health science education. *Australasian Journal of Educational Technology*, 33(6).

Nemer, D. and O'Neill, J. (2019). Rethinking MOOCs: The Promises for Better Education in India. *IJICTHD*, 11(1):36–50.

Novick, D., Afravi, M., Camacho, A., Rodriguez, A., and Hinojos, L. (2019). Pedagogical-Agent Learning Companions in a Virtual Reality Educational Experience. In Zaphiris, P. and Ioannou, A., editors, *Learning and Collaboration Technologies. Ubiquitous and Virtual Environments for Learning and Collaboration*, Lecture Notes in Computer Science, pages 193–203, Cham. Springer International Publishing.

Nowak, K. L. (2004). The Influence of Anthropomorphism and Agency on Social Judgment in Virtual Environments. *Journal of Computer-Mediated Communication*, 9(2).

Peperkorn, H. M., Diemer, J., and Mühlberger, A. (2015). Temporal dynamics in the relation between presence and fear in virtual reality. *Computers in Human Behavior*, 48:542–547.

Pirker, J., Lesjak, I., Parger, M., and Gütl, C. (2018). An Educational Physics Laboratory in Mobile Versus Room Scale Virtual Reality - A Comparative Study. In Auer, M. E. and Zutin, D. G., editors, *Online Engineering & Internet of Things*, Lecture Notes in Networks and Systems, pages 1029–1043. Springer International Publishing.

Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *Instructional Science*, 23(5):405–431.

Rajeswaran, P., Hung, N., Kesavadas, T., Vozenilek, J., and Kumar, P. (2018). AirwayVR: Learning Endotracheal Intubation in Virtual Reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 669–670.

Rosé, C. P., Carlson, R., Yang, D., Wen, M., Resnick, L., Goldman, P., and Sherer, J. (2014). Social factors that contribute to attrition in MOOCs. In *Proceedings of the first ACM conference on Learning @ scale conference - L@S '14*, pages 197–198, Atlanta, Georgia, USA. ACM Press.

Rosenberg-Kima, R. B., Baylor, A. L., Plant, E. A., and Doerr, C. E. (2007). The Importance of Interface Agent Visual Presence: Voice Alone Is Less Effective in Impacting Young Women's Attitudes Toward Engineering. In de Kort, Y., IJsselsteijn, W., Midden, C., Eggen, B., and Fogg, B. J., editors, *Persuasive Technology*, Lecture Notes in Computer Science, pages 214–222. Springer Berlin Heidelberg.

Roussou, M., Oliver, M., and Slater, M. (2006). The virtual playground: an educational virtual reality environment for evaluating interactivity and conceptual learning. *Virtual Reality*, 10(3):227–240.

Scaife, M. and Rogers, Y. (2001). Informing the design of a virtual environment to support learning in children. *International Journal of Human-Computer Studies*, 55(2):115–143.

Schönwetter, D. J., Clifton, R. A., and Perry, R. P. (2002). Content Familiarity: Differential Impact of Effective Teaching on Student Achievement Outcomes. *Research in Higher Education*, 43(6):625–655.

Schrader, C. and Bastiaens, T. (2012). Learning in educational computer games for novices: The impact of support provision types on virtual presence, cognitive load, and learning outcomes. *The International Review of Research in Open and Distributed Learning*, 13(3):206–227.

Schut, M. J., Stoep, N. V. d., Fabius, J. H., and Stigchel, S. V. d. (2018). Feature integration is unaffected by saccade landing point, even when saccades land outside of the range of regular oculomotor variance. *Journal of Vision*, 18(7):6–6.

Shin, D. (2018). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior*, 78:64–73.

Short, J., Williams, E., and Christie, B. (1976). *The social psychology of telecommunications*. John Wiley & Sons.

Sitzmann, T. (2011). A Meta-analytic Examination of the Instructional Effectiveness of Computer-based Simulation Games. *Personnel Psychology*, 64(2):489–528.

Sneddon, J., Barlow, G., Bradley, S., Brink, A., Chandy, S. J., and Nathwani, D. (2018). Development and impact of a massive open online course (MOOC) for antimicrobial stewardship. *Journal of Antimicrobial Chemotherapy*, 73(4):1091–1097.

Soliman, M. and Guetl, C. (2010). Intelligent Pedagogical Agents in immersive virtual learning environments: A review. In *The 33rd International Convention MIPRO*, pages 827–832.

Sträfling, N., Fleischer, I., Polzer, C., Leutner, D., and Krämer, N. C. (2010). Teaching Learning Strategies with a Pedagogical Agent. *Journal of Media Psychology*, 22(2):73–83.

Tessier, M.-H., Gingras, C., Robitaille, N., and Jackson, P. L. (2019). Toward dynamic pain expressions in avatars: Perceived realism and pain level of different action unit orders. *Computers in Human Behavior*, 96:95–109.

Tsaramirsis, G., Buhari, S. M., Al-Shammari, K. O., Ghazi, S., Nazmudeen, M. S., and Tsaramirsis, K. (2016). Towards simulation of the classroom learning experience: Virtual reality approach. In *2016*

668     *3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages
669     1343–1346.

670 Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F.,
671     Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Šmíra,
672     M., Epskamp, S., Etz, A., Matzke, D., de Jong, T., van den Bergh, D., Sarafoglou, A., Steingroever, H.,
673     Derks, K., Rouder, J. N., and Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example
674     applications with JASP. *Psychonomic Bulletin & Review*, 25(1):58–76.

675 Wise, A., Chang, J., Duffy, T., and Del Valle, R. (2004). The Effects of Teacher Social Presence on Student
676     Satisfaction, Engagement, and Learning. *Journal of Educational Computing Research*, 31(3):247–271.

677 Witmer, B. G. and Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence
678     Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240.

679 Yang, D., Sinha, T., Adamson, D., and Penstein Rose, C. (2013). Turn on, tune in, drop out: Anticipating
680     student dropouts in massive open online courses. In *Proceedings of the 2013 NIPS Data-driven
681     education workshop*, volume 11.