International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

The Impact of Gender and Level of Control over Agents' Aesthetics on User Experiences in VR Training

Sonam Adinolf, David Conroy, Peta Wyeth & Leonie Simpson

To cite this article: Sonam Adinolf, David Conroy, Peta Wyeth & Leonie Simpson (20 May 2025): The Impact of Gender and Level of Control over Agents' Aesthetics on User Experiences in VR Training, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2025.2498481

To link to this article: https://doi.org/10.1080/10447318.2025.2498481

Published online: 20 May 2025.

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2025.2498481

The Impact of Gender and Level of Control over Agents' Aesthetics on User Experiences in VR Training

Sonam Adinolf, David Conroy, Peta Wyeth, and Leonie Simpson

Queensland University of Technology, Brisbane, Australia

ABSTRACT

This study explores how user gender, agent gender, and control over agent aesthetics affect experiences in a VR phishing awareness training application. As digital agents become more common in immersive training, understanding their design impact is critical. Participants were assigned to one of three conditions: (1) full customization of the agent's appearance, (2) selection from premade agents, or (3) no choice. We measured outcomes including phishing awareness (immediate and retained), mental effort, enjoyment, presence, and relatedness. Results showed no significant effects of user gender or choice condition on awareness, mental effort, or presence. However, female participants reported greater relatedness with their agent at a three-month follow-up, suggesting potential long-term engagement benefits. Aesthetic control did not significantly influence learning or awareness retention, underscoring the importance of content quality over personalization. These findings contribute to research on gender dynamics and customization in immersive learning environments.

KEYWORDS

Virtual Reality; VR training; customization; embodied agents; gender; user experiences

Virtual Reality (VR) has emerged as a potential medium for training across various domains, enabling individuals to practice skills and apply knowledge in a safe, controlled environment (Radianti et al., 2020). This immersive nature of VR, combined with its ability to provide immediate feedback, makes it an ideal platform for various training needs, particularly in fields where hands-on experience is crucial (Johnston et al., 2018). Recent studies show positive impacts of VR on learning and training outcomes in both educational (Merchant et al., 2014) and professional training settings (Renganayagalu et al., 2021).

A key component of immersive VR training environments is the integration of embodied agents (Kyrlitsias & Michael-Grigoriou, 2022). These digital entities can assume various roles—such as mentors and guides adding depth to the user experiences. The presence of these agents in VR not only enriches the interaction but also plays a significant role in the social aspects of training. By providing immediate feedback and guidance, these agents can simulate social contexts, which are crucial for developing complex skills (Grivokostopoulou et al., 2020; Petersen et al., 2021). The design and aesthetics of these agents are pivotal, influencing learners' cognitive and affective engagement. Recent research show that users desire agents to be more relatable (Zargham et al., 2023). The extent of customization and realism in agents' appearance and behavior potential to impact users' sense of presence, motivation, and cognitive load, thereby affecting training outcomes in VR (K. Kim et al., 2018; Petersen et al., 2021).

Additionally, user gender plays a complex role in interactions within VR training environments. Studies suggest that gender congruence between learners and agents could potentially benefit the learning process by affecting social presence, spatial presence, and learner satisfaction (Makransky et al., 2019; Shang et al., 2020). However, literature on this topic shows mixed results, pointing to a need for further research (Moreno & Flowerday, 2006; Weidner et al., 2023).

This study uses phishing awareness training as the context due to the evolving nature of cyber threats requiring training programs that not only convey information but also instill practical skills and a deep understanding of cyber risks. Traditional training methods often fall short in engaging participants and embedding long-lasting phishing awareness. VR-based training offers an engaging alternative. It can enable learners to experience the consequences of security breaches and practice response strategies in a risk-free environment, thus enhancing both the retention and application of phishing awareness principles (Khan, 2018; Schlecht, 2018).

This study aims to explore the relationship between digital agent customization, user gender, agent gender, and their impact on phishing awareness training within a VR context. It seeks to understand how varying levels of control

CONTACT Sonam Adinolf adinolfd@qut.edu.au School of Computer Science, Faculty of Science, Queensland University of Technology, Brisbane, QLD, Australia 4001

over an agent's aesthetic features and the gender pairing between the user and the agent influence key aspects of the learning experience—namely, mental effort, sense of presence, enjoyment, and the sense of relatedness. The latter, rooted in the psychological need for connection and belonging, is crucial in learning contexts as it can significantly influence motivation and engagement, thereby impacting learning efficacy (Ryan & Deci, 2000). Employing a custom designed VR-based phishing training application, this research adopts a nested control design to examine the effects of agent customization and gender pairing, aiming to contribute to the development of more effective agents and learning experiences in VR training applications.

1. Background

1.1. VR and its affordances for learning and training

VR has been a subject of extensive research for over three decades, exploring its potential to simulate experiences that either mirror real-life or present entirely new realities (Radianti et al., 2020; Slater & Sanchez-Vives, 2016). VR enables learners and trainees to grasp spatial relationships and concepts more intuitively, offering them learning experiences that are contextual, situated, and rich with meaning (Merchant et al., 2014; Radianti et al., 2020).

Central to the VR experience is the *sense of presence*, a psychological state where users feel genuinely "in" the virtual environment despite knowing otherwise (Bowman & McMahan, 2007; Slater, 2009). Research has linked a high sense of presence to improved emotional responses, enjoyment, and even learning outcomes within VR settings (Chirico & Gaggioli, 2019; Grassini et al., 2020; Servotte et al., 2020; Tang et al., 2022; Tussyadiah et al., 2018). However, a systematic literature review spanning 24 years found that while the sense of presence is crucial for VR experiences in education, the outcomes of its association with learning are not uniformly agreed upon in the literature (Krassmann et al., 2019). They highlighted the need for a deeper understanding of how sense of presence can be leveraged for learning purposes in VR.

VR's immersive capabilities, alongside VR's ability to evoke a strong sense of presence and place, underpins its suitability for experiential and constructivist learning methods (Johnston et al., 2018). However, there are conflicting findings regarding the relationship between the enjoyment of VR experiences and knowledge gain, with some research showing no relationship (Makransky et al., 2021) and others indicating a positive relationship (Huang et al., 2022). This highlights the need for further research into the relationship between enjoyment and training outcomes.

Co-presence, or social presence, extends the concept of presence to include the feeling of being with others in the virtual space (Biocca et al., 2003; Bulu, 2012). This sense of shared experience, facilitated by VR's ability to replicate social cues, is crucial for creating engaging and interactive learning environments. It fosters a sense of social connection, or sense of relatedness, which is vital for education and can be significantly enhanced by the presence of virtual agents designed to interact with users in meaningful ways.

1.2. Agents for education and training

Pedagogical agents, whether embedded in multimedia learning environments, intelligent tutoring systems, or educational video games, have shown a modest but significant impact on educational outcomes (Craig et al., 2015; Graesser & McNamara, 2010; Kinzer et al., 2012; Schroeder et al., 2013; Sikström et al., 2022; Wang et al., 2018). By simulating social interactions, agents can provide customizable and scalable training opportunities (Okita et al., 2013; Othlinghaus-Wulhorst & Hoppe, 2020). For instance, role-playing, a training technique for developing interpersonal skills and increasing student satisfaction (Sherman et al., 2020), faces scalability challenges that agents could potentially address (Nikendei et al., 2005; 2007; Sogunro, 2004). The effectiveness of these agents is influenced by a multitude of factors, including their design, voice, and behavior, which can affect user engagement and learning processes (Heidig & Clarebout, 2011; Hoorn, 2015; Veletsianos, 2007).

The concept of customizing agents has been suggested as a promising way to foster social connections with these digital entities. Baylor highlighted the potential of agent customization to serve as effective social models (Baylor, 2009). This aligns with findings from various studies, indicating that customizing avatars or player characters in virtual environments can have a positive impact on outcomes like motivation and learning (Lim & Reeves, 2009; Lin et al., 2017; Türkay & Kinzer, 2016). For instance, Herrera and Bailenson's study in a VR context showed that choosing an avatar's skin color can enhance prosocial behavior towards social issues like homelessness (Herrera & Bailenson, 2021).

Moreover, the choice in customizing on-screen agents has been shown to influence user perceptions significantly. Xiao et al. found that users who chose their agents rated them more favorably in terms of likability, trust, and effectiveness, suggesting that the act of choosing can enhance user engagement and motivation (Xiao et al., 2007). Similarly, Wald et al. demonstrated that customizing chatbots led to increased perceptions of trust and anthropomorphism, indicating that personalized agents are perceived as more competent (Wald et al., 2021).

However, studies focusing on the choice of pedagogical agents have yielded mixed results, especially concerning cognitive and affective outcomes. These inconsistencies may be attributed to differences in learning contexts, learner characteristics, and technological design among the studies. For example, Kim and Wei found no significant impact of choosing among four agents in a math tutorial interface on learning outcomes (Kim & Wei, 2011), while Kinzer et al. observed that the choice of a virtual mentor in a math game led to improved game engagement (Kinzer et al., 2012). Furthermore, Behrend and Thompson noted that giving participants control over their pedagogical agent's appearance, personality, and feedback behavior in web-based training enhanced learning outcomes when multiple choices were provided (Behrend & Thompson, 2012).

While much of the existing research has focused on avatar customization or interface design, there is a gap in the literature concerning agent and avatar customization in VR settings. Weidner et al.'s review highlighted this lack of investigation and called for further research in this area (Weidner et al., 2023). Notably, the impact of choice as a way to provide users sense of agency in VR contexts remains unexplored, and the optimal number of choices to offer is still an open question. The phenomenon of choice overload, well-documented in other domains like shopping (Iyengar & Lepper, 2000), has only recently begun to be explored in virtual environments, with preliminary data suggesting that its threshold may be significantly higher in these settings (Adinolf et al., 2020b).

1.2.1. Effect of gender when interacting with agents in VR

The impact of gender on interactions with virtual agents in VR settings has produced mixed results in the literature, with studies reporting both significant and negligible effects on user experiences and educational outcomes.

Makransky et al. (2019) reported gender-specific preferences and performance in a VR setting, where girls performed better with a young female scientist agent, and boys with a futuristic robotic agent. Boys also reported more social presence with the female agent, a difference not observed among female participants. Contrasting this, a study on non player characters (NPCs) showed that participants of both genders found opposite-gender agents more convincing as mentors (Rogers et al., 2018). In a VR training context, male participants reported finding female agents more attractive while there were no differences on performances outcomes across the gender groups (Shang et al., 2020). In medical and therapeutic VR applications, participants generally preferred female agents due to their perception that female agents were more helpful and professional, thus enhancing user's comfort (Alsharbi & Richards, 2017; Feijóo-García et al., 2021; Mostajeran et al., 2019), and in an educational VR context, a female agent was reported to elicit more rapport from students than a male agent (Novick et al., 2019). Gender-based preferences were also evident in avatar customization within MMOs, where female players derived more enjoyment from this feature than male players, affecting their game interest (Turkay & Adinolf, 2010).

Conversely, Moreno and Flowerday (2006) found that the choice of pedagogical agent was not influenced by the user's gender. Shang's research (Shang et al., 2020) in a VR training context also found no significant performances but found a preference towards female agent though this study included a small sample size. Despite the noted preferences for female agents in certain contexts, a study by Novick and colleagues (2019) found that having a same-gender agent did not significantly affect rapport, and agent gender had no significant effect on recall. Additionally, research on MMOs showed that customization affected both male and female players' engagement similarly (Turkay, 2013).

These contrasting findings highlight the complexity of gender dynamics in VR interactions with agents. While some studies suggest gender influences preferences and outcomes, others report no significant gender-based effects. This inconsistency underlines the need for further exploration in this area to better understand how gender interplays with virtual agent design and user experiences in VR settings. Consequently, a recent review study (Weidner et al., 2023) encouraged further research on the impact of gender on user experiences with agents.

1.3. Phishing user training as the case study for the application

The focus on phishing for the training activity in this study is driven by multiple factors. Firstly, cybersecurity is critically important in our increasingly digital world, underscored by the escalating economic impact and frequency of cybercrime. The FBI's 2021 Internet Crime Report highlighted a significant increase in losses due to cybercrime, amounting to $6.9 billion, which is about $2 billion more than in 2020 (Threatpost, 2021). Phishing remains one of the most prevalent and successful social engineering tactics, relying on deception to manipulate individuals into revealing sensitive information or granting unauthorized access to systems. Despite advancements in security technologies, human susceptibility continues to be a primary factor in phishing attacks, highlighting the necessity of improved user education and awareness training (Khan, 2018; Vishwanath et al., 2018).

Secondly, phishing awareness is a fundamental component of cybersecurity training. Phishing attacks exploit human vulnerabilities through social engineering tactics, making user awareness and education a critical line of defense (Aldawood & Skinner, 2019; Pattinson et al., 2012). Studies show that awareness training significantly reduces phishing susceptibility by enhancing users' ability to recognize deceptive emails and malicious links (Caputo et al., 2014; Jampen et al., 2020; Reinheimer et al., 2020). However, the effectiveness of different training methodologies in fostering phishing awareness remains an area of active research (Hillman et al., 2023).

The universal relevance of cybersecurity knowledge, given the widespread use of networked computers, makes it an ideal topic for training that could benefit a broad audience. Improving phishing awareness through innovative training methods could mitigate the frequency and severity of cyber attacks, providing economic advantages and enhancing digital safety for individuals and organizations (Chen et al., 2023). Traditional Computer-Based Training (CBT) is the most common form of cybersecurity training (Abawajy, 2014), yet it has several drawbacks, including monotony and lack of engagement, which can lead to decreased motivation and feelings of isolation among learners (Cone et al., 2007). In response to these limitations, there's a growing interest in leveraging immersive technologies like Augmented Reality (AR) and VR for cybersecurity awareness training. The unique capabilities of AR and VR, such as interactive feedback and the ability to simulate scenarios in a realistic controlled environment, offer promising advantages over conventional CBT (Alqahtani & Kavakli-Thorne, 2020; Kroll, 2019; Schlecht, 2018).

Although VR has been widely explored for training, its specific application in phishing awareness remains limited. The immersive nature of VR provides a unique opportunity to enhance user engagement and realism, potentially improving training outcomes by simulating real-world phishing attacks in an interactive environment (Veneruso et al., 2020). Some studies have examined VR for security-related training, such as social engineering awareness and password security, but the literature specifically addressing phishing training in VR is scarce (Chandrashekar et al., 2023; Rana & Chicone, 2023). Existing evidence from non-VR studies suggests that interactive and experiential learning approaches, such as gamification and role-playing, improve engagement and retention of cybersecurity knowledge, indicating a strong potential for VR-based phishing awareness training to be effective (Moumouh et al., 2023; Salameh & Loh, 2022)

Individual differences in phishing susceptibility significantly influence training effectiveness, with gender emerging as a notable factor. Research presents mixed findings: some studies indicate that men report higher phishing awareness, while women engage more with interactive learning tools, potentially enhancing their retention of cybersecurity concepts (McGill & Thompson, 2021; Pratama et al., 2023). Women tend to perceive higher phishing risks but often report lower confidence in detecting phishing attempts, whereas men express greater confidence—even when their actual detection ability does not always align with this confidence (Kavvadias & Kotsilieris, 2025; Sun et al., 2016). Differences in information processing styles may also contribute to phishing vulnerability, with men focusing more on message credibility and women being more influenced by urgency cues (Lee et al., 2023). These variations suggest that gender-responsive training approaches could enhance training outcomes (Pratama et al., 2023). While these dynamics have been studied in traditional digital contexts, little research has examined whether they persist in immersive VR-based training. Given VR's potential for engagement and embodied learning, this study explores how gender influences phishing awareness, training effectiveness, and knowledge retention within a gamified VR cybersecurity training environment.

Given the limitations of traditional CBT and the advantages of VR training, this study examines the effectiveness of a gamified, agent-based VR application for phishing awareness.

2. Current study

Based on the literature presented above, this study examines how different levels of control over a digital agent's aesthetic features affect phishing awareness retention in a VR-based training environment. Training effectiveness was measured through participants' self-reported phishing awareness immediately after the training and three months later, assessing changes in their ability to recognize phishing indicators.

In this study, we consider phishing awareness, mental effort, and enjoyment as key aspects of the training experience. Phishing awareness refers to participants' self-reported ability to recognize phishing indicators before and after the training. Enjoyment reflects participants' subjective engagement with the training, measured using the Intrinsic Motivation Inventory (IMI) (see the Methods section for details). Lastly, iven the multi-sensory experiences of VR environments, it is necessary to measure mental effort and its impact on user experiences (Huang et al., 2020; Sari et al., 2024) to understand the cognitive demands imposed on users while engaging with VR-based training. Mental effort provides insight into the cognitive load experienced by participants, which can influence engagement. Higher mental effort may indicate increased difficulty in processing training content, while lower mental effort may suggest greater ease of comprehension and interaction with the VR environment (Paas et al., 1994). More specifically, the study addresses the following aims.

Aim 1) To understand how the interaction between user gender and agent gender affects training experiences and social connection with the agent in the VR experience. The following research questions were asked for this aim.

RQ1a: Are users' mental effort and awareness of phishing affected by the interaction between user gender and agent gender?

RQ1b: Are users' sense of presence, enjoyment, and sense of relatedness impacted by the interaction between user gender and agent gender?

Aim 2) To examine whether the VR intervention had any influence on participants' awareness of phishing which was the context for the VR training.

RQ2: Does participants' reported awareness of phishing vary from immediate (t1) to long-term (t2) assessments?

Aim 3) To understand how different levels of control over agents' aesthetic features affect learner experiences with a training task in a gamified VR based phishing training application. Learner experience is measured in terms of 1) reported awareness of phishing, 2) mental effort, 3) enjoyment. Of these, first was measured both immediately after the experience and based on recollection three months later. The latter two were only measured immediately after the experience. Thus, the associated research questions are:

RQ3a: To what extent does varying the level of control over agents' aesthetic features impact users' *reported awareness of phishing* ?

1. The immediate (t1) effects after the training experience
2. The long-term (t2) effects, change from immediate to three months later

RQ3b: To what extent does varying the level of control over agents' aesthetics features affect users' *mental effort*, *enjoyment*, *sense of presence* and *sense of relatedness* in a VR based phishing awareness training application as measured immediately after the experience?

Given the inconsistent findings on the relationships found between sense of presence and training experience measures in prior research (Krassmann et al., 2019; Servotte et al., 2020), the following research question is asked:

RQ4: What is the relationship between sense of presence in the VR training and key training experience measures of 1) self-reported awareness phishing awereness, 2) mental effort and 3) enjoyment?

3. Methods

In order to achieve the aims listed above, we employed a between-subjects design with nested conditions. Participants were randomly assigned to one of four conditions based on their level of control over agent aesthetics: (1) Choice, (2) Customization, (3a) No-choice: Gender-matched, and (3b) No-choice: Gender-mismatched (see Figure 1). Aim 1 is addressed by the nested co-choice groups; since one group involves gender match, and the other involves gender mismatch. The gender of the embodied agent is determined by their voice: male voice is assigned male gender and female voice assigned female gender. This design allows us to investigate any interactions between participant's gender and that of the agent. Aim 2 is achieved by the repeated measures design, and Aim 3 is addressed by the three different levels of control over the agent's aesthetics -high levels of control in the customization condition, moderate levels in the choice condition, and none in the no-choice conditions.

3.1. Procedure

This study took place in two stages: 1) Experiencing the VR training, with pre-and-post surveys; 2) A three-month follow-up survey (see Figure 2 for the study outline).

Upon arrival at the study site, each participant was provided with study information and an informed consent form, as approved by the university's Ethics Review Board. After participants' granted consent, they were requested to complete a brief pre-survey lasting approximately 2 to 3 min using Qualtrics survey tool. Participants were randomized into treatment groups and were given an ID number using a randomization script in Qualtrics. Subsequently, participants were equipped with a VR headset and handheld controllers, and the application commenced. They immediately entered the ID number in the VR experiences as their "employee training ID number" as it aligned with the narrative of the VR training. This allowed us to match log data from the VR experience with the survey data for participants. After a brief training of the controls, the researcher refrained from offering any unsolicited comments or instructions to the participants during the VR training experience, unless explicitly requested for assistance. Following the experience, participants were administered a post-survey, which took around 10 to 15 min to complete.

Three months after completing the post-survey, participants were invited to complete a follow-up survey via email. Non-respondents received up to four reminder emails. A total of 61 participants (53.5% of the original sample)

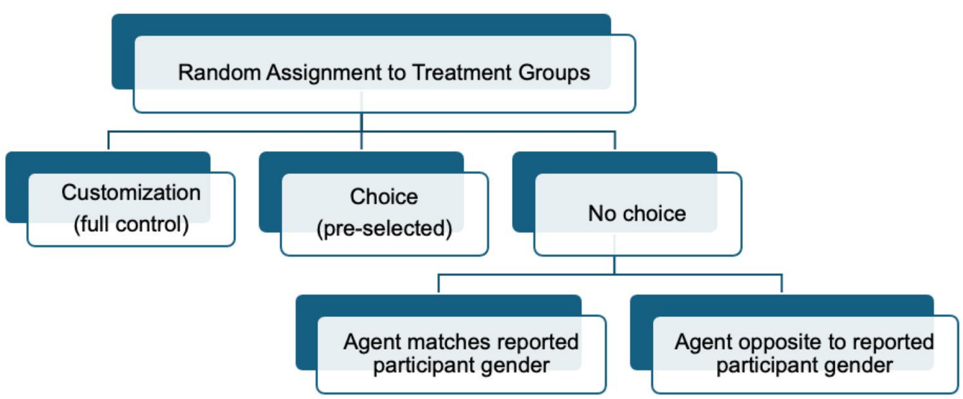

Figure 1. Schematic overview of the nested study design.

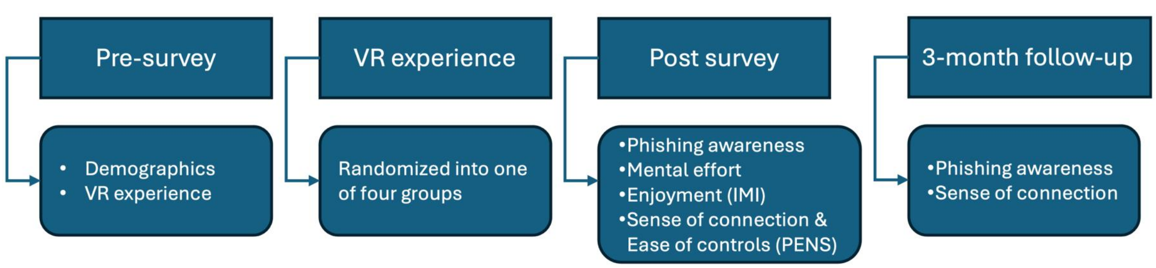

Figure 2. Figure showing the flow of the experiment, including outputs at each step.

Figure 3. A screenshot of the agent customization interface (on the left). The different customizable attributes are listed along the side, with the options for the currently selected attribute in the center and the current state of the robot agent on the right. Customization group is given 10 attributes, with 12 choices in each. On the right, the agents given to the participants in the choice condition. The two female-voiced agents are on the far right, and the two male-voiced agents on the left. The aesthetics of the agents are cartoon-like and aimed to avoid gender stereotypes.

responded to the follow-up, with response rates varying across experimental conditions.

3.1.1. Different treatment groups

The following tables and figures show the numerical differences between the treatment groups (Table 2). Figure 3 shows the customization interface and the pre-constructed agents. When assigning agents in the no choice conditions, the gender of the robot was determined by which voice it was given - male or female. The voice lines were recorded by a male and a female human.

- *Customization:* Participants could customize their agent using the aesthetic choices presented in the application. They could assign male or female voice which determined the gender of the robot agent for the purpose of the analysis.

- *Choice*: Participants could choose an agent out of four pre-constructed ones. Two of these agents had male recorded voice and two had female recorded voice (see Figure 3 for details).

- *No choice, gender matched:* Male identified participants were assigned a male voiced agent from the choice group with grey metallic surface (M2), and female identified participants were assigned a female voiced aged from the group (F1) (see Figure 3 for details).

- *No choice, gender mismatched:* Male identified participants were assigned a female voiced agent (F1) - the same one that was assigned to female identified participants in the previous group; and female identified participants were assigned a male voiced agent (M2) that was assigned to male identified participants in the previous group.

3.2. Materials

The pre-survey consisted of demographic, i.e., age, gender (What gender do you identify with?), country of origin, and VR experience questions. Both the immediate survey and the follow up surveys contained predominantly quantitative scales as detailed below.

3.2.1. Sense of connection and intuitive controls - player experience need satisfaction questionnaire (PENS)

To measure the level of control over agents' aesthetics features on players' social connection with the agent, sense of relatedness items from the PENS questionnaire (Ryan et al., 2006) were utilised due to its common use in similar studies (Karaosmanoglu et al., 2023; Vella et al., 2015). PENS has a five point scale (1- Not at all; 5-To a great extent). The subscale showed good internal consistency (Cronbach's alpha = 0.786). Intuitive Controls subscale of PENS was used to measure users' experiences with the VR controls for the purpose of the equality of the groups (e.g., "When I wanted to do something in the game, it was easy to remember the corresponding control.") The subscale showed good internal consistency (Cronbach's alpha = 0.810).

3.2.2. Mental effort scale

Mental effort was measured using a single item self-rating scale developed by Paas et al. (1994) which is used widely to evaluate users' mental effort in multimedia instructions (Tabbers et al., 2004). The numerical values and labels assigned to the categories ranged from "1 = very, very low mental effort" to "9 = very, very high mental effort."

3.2.3. Presence questionnaire

We measure the "sense of presence" using the revised version (UQO Cyberpsychology Laboratory, 2004) of Witmer and Signer's presence questionnaire (Witmer & Singer, 1998) which is the most commonly used presence scale in prior work (Hein et al., 2018). The questionnaire consists of 24 items on a 7 point Likert scale which cover a range of aspects of presence, including realism, possibility to act, possibility to examine, sound and haptics. Haptics related items (n = 2) were removed as the experience did not involve haptics and those about controllers (n = 2) were removed as PENS scale was used to evaluate users' experiences with the controllers. The remaining 20 items were aggregated into a single measure. The scale showed good internal consistency (Cronbach's alpha = 0.806).

3.2.4. Enjoyment

The Intrinsic Motivation Inventory (IMI) is a multidimensional instrument designed to gauge participants' subjective experiences related to a target activity, particularly their intrinsic motivation (Ryan, 1982). For this study, we only utilized the enjoyment sub-scale as it was one of the key outcomes we were interested in. It consists of seven items, each rated on a 7-point Likert scale (1 = not true at all to 7='very true'). allowing for a nuanced assessment of the participant's enjoyment levels. The subscale showed good internal consistency (Cronbach's alpha = 0.902).

3.2.5. Phishing awareness

We used the phishing awareness scale developed by Vishwanath's (2011) to measure user's day-to-day attention to phishing awareness practices and interest in cybersecurity. This measure was used in both immediately after the experience and in three-month follow-up survey. The items start with "when you receive an email, how much attention do you give to… " They provide sub scales to measure attention including the sender (e.g., "sender's email address"), attention grammar (e.g., "typographical errors") and urgency (e.g., "warnings") with three items in each. The items had 10-point Likert scale. These nine items were averaged to create a single measure for analysis. Internal consistency of the scale is 0.777 for the post-survey and 0.814 for the follow-up survey.

In addition, participants were asked whether they recognized any phishing email in the last three months (Yes or No) and whether they were victims of a phishing email (Yes, No, Maybe). These two questions were asked both in the post-survey and in the three months follow up.

3.2.6. Manipulation check

We conducted a manipulation check to measure participants' sense of control using the following 5-point Likert scale (1 = Not at all to 5 = A lot) item: "How much control did you feel you had over your robot assistant's appearance?." This was asked after the other measures to avoid priming the participants.

Table 1. Different treatment groups showing different levels of control over aesthetic features. Customization group is given 10 attributes, with 12 choices in each.

Customization	Choice	No-Choice	
Voice (1 male/1 female)	4 agents	Gender match	Gender mismatch
Body color			
Hair			
Hair color	Female voice		
Eyes			
Top clothes			
Bottom clothes			
Hat	Male voice		
Facial-features Mustache			

3.3. Virtual reality training application

The VR training application occupies the intersections of numerous disciplines and sub-disciplines. The first stage of this research project consisted of a series of participatory design workshops to obtain expert insight from the relevant fields. This process and the outcomes are outlined in an earlier publication (Adinolf et al., 2020b). The workshops involved experts from Human Computer Interaction (HCI), Human-Robot Interaction (HRI), VR, and Cybersecurity. The design elements from those workshops included overall training scenarios, individual mechanics, pitfalls to avoid (both technical and design), and the role and form of the agent. Once a core design was achieved, the second stage involved a lab and online study to determine the appropriate numbers of attributes and choices to include for the robot agent as there was no prior study to guide these decisions (Adinolf et al., 2020a). That study tested user reactions to a range of pairings of customizable aesthetics for an agent and numbers of choices. These were used to determine an acceptable set of parameters for the customization condition. See Table 1 for the final set of attributes used in this study.

The resulting design was a gamified VR application to help with phishing user training, with a robotic assistant. The robot form was chosen to avoid uncanny valley effect humanoid agents created in previous studies (Petersen et al., 2021). The experience takes place in a virtual mail room. The core mechanic is mail sorting which involves the player picking up "mail" and examining the address and message body for irregularities before sorting it into a series of pneumatic tubes. The participant first looks at mail delivered in tubes, and sorts them based upon an "address" symbol. The participants are also given a demonstration of a "bad" message; one with a symbol that isn't quite correct. Any "bad" mail should be thrown into a furnace. This mechanic aimed to be analogous to examining the sender of an email to determine whether it is legitimate or phishing. As the users progress, the mail becomes slightly more complex. The potential presence of harmful symbols highlighted in the mail body is introduced. This is analogous to examining links in an email to determine their safety. For example, the user must periodically go get coffee from the robot to refill their energy. As the game progresses, two distractor tasks are added in, as the workshops indicated that people are most likely to succumb to phishing attacks when distracted.

A rigorous iterative testing was conducted as VR experts in the workshops had emphasised that bugs and glitches could have an out-sized effect due to the immersion of VR. Overall, 7 play-testers were employed, iterating over 15

Table 2. Gender distribution across treatment groups for main study and follow-up survey. 'M' indicates male and 'F' indicates female participants.

	Customization	Choice	No-Choice: Gender Match	No-Choice: Gender Mismatch
Main Study	22 M, 7 F (n = 29)	23M, 5 F (n = 28)	20M, 10 F (n = 30)	17M, 10 F (n = 27)
Follow-up	8 M, 5 F (n = 13)	9M, 3 F (n = 12)	11M, 9 F (n = 20)	8M, 8 F (n = 16)

major versions of the game. The final game ran on SteamVR, and was compatible with all major headsets. Participants in the lab played the game using the HTC Vive 2. Those in classrooms used an Meta Quest 2 with link cable.

3.3.1. Agent design

Based upon the workshops mentioned above, the agent was designed as a helper and trainer. Prior work suggests that the design of an agent needs to fit with specific narrative roles (Rogers et al., 2018). Thus, the agent's interactions were aimed to be friendly and non-confrontational. So the application begins with the robot agent welcoming the user to their first day at "Steamcorp." The agent then guides the user through the tutorial.

The morphology of the agent is anthropomorphized. It could be categorized as having android properties where it has a complex humanoid design yet is still mechanical in nature (Mobed et al., 2024). The agent does not have legs, instead it rolls on a wheel. The attributes of the agent are cartoon-like to avoid both uncanny valley and gender stereotypes in relation to facial and body structures. Basic skeleton of the robot agent is the same for all agents created and used in the VR experience and aimed to be androgynous and gender neutral to reduce gender stereotypes as suggested by prior work (Koda et al., 2022; Nag & Yalçın, 2020).

After the initial tutorial, the agent periodically introduces a new activity that the user needs to keep track of. These activities are: drinking coffee served by the robot to replenish energy meter, clearing a conveyor belt to stop boxes from piling up, and periodically cycling a vent fan to clear smoke from the air. The coffee activity was designed to give the participant more direct interaction with the agent. All three activities were also designed to be distractor tasks. To introduce each new task, the robot both speaks and moves to the relevant area and faces towards the user to attract their attention. Once the user completes the goals (or the time limit of 20 min being reached), the robot congratulates the user on their work. They then introduce the score screen as the finale.

3.4. Participants

3.4.1. Main study

The majority of participants were recruited from a prominent Australian university and were subsequently engaged in the study within laboratory or classroom settings. A minority of participants, eight, were sourced through online communities catering to VR enthusiasts, and they were provided with the necessary information and application to partake in the study remotely.

In total, 117 people participated in the study. Among these 82 reported they identified as male, 32 female, and three participants reported as non-binary/third gender (see Table 2 for gender distribution across treatment groups). Due to the latter group having too few participants, their data was removed from analysis. This left 114 participants in for the data analysis. On average, participants were 25.6 years old ($SD = 8.05$), with below average experience in VR ($M = 2.67$, $SD = 1.27$) (see Table 3 for details). No participants reported any simulator sickness during the study and nobody needed to stop the study due to any simulator sickness.

4. Data analysis

Data analysis were performed using SPSS 29 statistical analysis software. We designated an alpha level of 0.05 for significance. We conducted a series of 2×2 ANOVA tests that examined the effect of participant gender and agent gender on outcome variables (RQ1a-RQ1b). We further analysed whether having participant gender matching the agent gender had any impact on outcome variables using One-Way ANOVA tests (RQ1a-RQ1b). In order to do this, we computed a new variable in SPSS. A repeated measures ANOVA was conducted to examine the effects of the treatment conditions on Phishing awareness over time (RQ2). We performed a series of One-Way ANOVA tests to assess the differences in outcome variables across the treatment groups (RQ3a-RQ3b). Assumptions were checked for each analysis. While the normality of the data assumption was violated, ANOVA test are robust against the violation of normality with similar group sizes (Schmider et al., 2010). Research has shown that ANOVA maintains its Type I error rate and statistical power under non-normality when homogeneity of variance is met which was the case in this study (Refinetti, 1996). To analyze RQ4, we conducted Pearson correlation analyses to examine how participants' reported sense of presence in the VR training correlated with (1) phishing awareness (measured immediately after training and in the three-month follow-up), (2) mental effort (assessing

Table 3. Demographics and virtual reality familiarity across treatment groups.

Variable	Group	N	Mean	SD	F	t
Age	Customization	29	23.66	5.90	1.13	0.341
	Choice	28	27.04	10.60		
	No Choice: Gender Match	29	24.90	5.98		
	No Choice: Gender Mismatch	27	26.85	8.20		
VR Familiarity	Customization	29	2.86	1.36	0.774	0.511
	Choice	27	2.67	1.36		
	No Choice: Gender Match	30	2.77	1.22		
	No Choice: Gender Mismatch	27	2.37	1.15		
Hours in VR	Customization	28	1.15	2.21	0.347	0.792
	Choice	28	0.71	1.30		
	No Choice: Gender Match	30	0.77	1.48		
	No Choice: Gender Mismatch	27	0.62	2.03		

cognitive load during training), and (3) enjoyment (evaluating intrinsic motivation and engagement).

4.1. Equality of the groups

We conducted a series of One-Way ANOVA tests to assess the equality of the groups based on the demographic information provided by the participants and their VR.

The tests revealed no statistically significant differences between the groups in terms of their demographics and their experiences with VR. In addition, there were no statistically significant differences between intuitiveness of controls across groups ($p > 0.05$). On average, participants reported VR game as very intuitive (M = 4.09, SD = 0.83). Consequently, we concluded that the groups can be considered equivalent for the purposes of this study.

4.1.1. Manipulation check

A one-way ANOVA test showed that there were significant differences between the groups on how they responded to the manipulation check question F(3,111)=40.33, $p < 0.001$, eta square = 0.52). More specifically customization group reported having the most control (M = 3.97, SD = 1.02), followed by the choice group (M = 3.46, SD = 0.84). Gender matched (M = 1.83, SD = 1.05) and gender mismatched no-choice groups (M = 1.70, SD = 0.91), on average reported little control over the aesthetics of the agent. Therefore, it is reasonable to conclude that the experimental manipulation was successful.

5. Results

5.1. Impact of gender variables on learner' experiences

In the customization group, 13 out of 21 male participants made an agent with a male voice and eight made an agent with a female voice; six out of seven female participants made their agent with a female voice whereas only one made an agent with a male voice.

In the choice group, 19 out of 22 male identified participants chose a male voiced agent and three chose the agent with female voice. Two out of five female identified participants chose a female voiced agent and the rest chose a male voiced agent.

When we combine the choice and customization groups, chi-square test showed that the gender identification of the participants significantly influences the choice of the agent gender (Chi-square statistic = 5.31, $p = 0.021$).

5.1.1. Impact of gender on mental effort and phishing awareness (RQ1a)

We found no statistically significant interaction between the effects of participant gender and agent gender on *mental effort* (F(3, 105) = 2.37, $p = 0.127$). Simple main effects analyses showed that neither participant gender nor agent gender had a statistically significant effect on mental effort ($p > 0.05$). In addition, we found no statistical significance of gender matching on mental effort across four groups (F(3, 105)= 2.32, $p = 0.080$). See Table 4.

A set of two-way ANOVA tests showed no statistically significant interaction between the effects of participant gender and agent gender on *phishing awareness* as measured immediately after the experience (F(3, 105) = 0.014, $p = 0.907$) nor in the follow-up survey (F(3, 53) = 0.044, $p = 0.834$). Simple main effects analyses revealed that neither participant nor agent gender had a statistically significant impact on phishing awareness as measured immediately after the experience and in the follow-up survey ($p > 0.05$). We also found no significant impact of gender matching neither on phishing awareness as measured right after the experience (F(3, 105) = 0.381, $p = 0.767$), nor in the follow up survey (F(3, 53) = 0.449, $p = 0.719$). Please see Table 4 for means and standard deviations.

5.1.2. Impact of gender on sense of presence and relatedness with agent (RQ1b)

We found no statistically significant interaction between the effects of participant gender and agent gender on reported *sense of presence* (F(3, 105) = 0.365, $p = 0.547$). Simple main effects analyses revealed that neither participant nor agent gender had a statistically significant impact on participants' sense of presence either ($p > 0.05$). In addition, we found that gender matching did not have a statistically significant effect on reported sense of presence (F(3, 105) = 0.746, $p = 0.527$). See Table 4 for means ans standard deviations.

A set of two-way ANOVA tests showed no statistically significant interaction between the effects of participant gender and agent gender on *sense of relatedness* as measured immediately after the experience (F(3, 105) = 1.209, $p = 0.274$) and in the follow-up survey (F(3, 105) = 1.02, $p = 0.318$). Simple main effects analyses showed that *agent gender* did not have a statistically significant effect on their

Table 4. Descriptive statistics on the outcome variables based on participant and agent gender.

	Female				Male			
	Female (Agent)		Male (Agent)		Female (Agent)		Male (Agent)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Phishing awareness—t1	7.04	1.91	7.41	0.87	7.18	1.37	7.46	1.80
Phishing awareness—t2	7.56	1.36	7.85	0.91	7.19	1.21	7.12	1.87
Mental effort	5.17	1.72	5.41	1.66	5.23	1.58	4.43	1.77
Enjoyment	5.48	1.04	5.23	1.02	5.53	1.17	5.24	1.22
Sense of presence	5.29	0.63	5.20	0.66	5.38	0.51	5.36	0.59
Sense of relatedness—t1	2.98	0.70	2.98	0.61	3.06	0.56	2.82	0.68
Sense of relatedness—t2	3.33	0.82	2.87	1.14	2.29	0.98	2.39	1.05

sense of relatedness with the agent in neither times ($p > 0.05$).

Simple main effects analyses showed while *participant gender* did not have a statistically significant effect on their sense of relatedness with the agent immediately after the experience ($p > 0.05$), it did in the follow-up survey (F(3, 54) = 6.36, $p = 0.015$). Participants who identified as female (M = 3.13, SD = 1.02) reported higher sense of relatedness with the agent than male participants did (M = 2.36, SD = 1.02).

Gender matching did not have a statistically significant effect on relatedness with the agent as measured immediately after the experience (F(3, 105) = 1.02, $p = 0.388$). But, it did have a statistically significant effect on follow-up sense of relatedness scores (F(3, 54) = 3.17, $p = 0.032$, eta-squared = 0.15). Post-hoc anaysis showed that those who identified as female reported significantly higher sense of relatedness with their female robot agents compared to male identified participants with their male robot agents. See Table 4 for means and standard deviations.

5.3. Impact of VR phishing training on phishing awareness (RQ2)

A paired sample t-tests were utilized to compare the immediate (t1) and follow-up (t2) phishing awareness survey responses across each specific sub-category rather than averaging the metrics. This approach aimed to ascertain whether the intervention prompted participants to scrutinize different aspects of their emails more closely. A statistically significant difference was noted in the category of *Grammar*, with participants demonstrating increased vigilance towards grammatical details in their emails during t2 (t = −2.52, $p = 0.014$, N = 61) (See Table 5). An exact McNemar's tests were conducted to ascertain if there were any statistically significant changes in the proportion of participants who recognized phishing emails and those who fell victim to phishing between the pre-survey and the follow-up survey. The test revealed no statistically significant changes in either category ($p > 0.05$).

Table 5. Two-Sided Paired-Samples t-test statistics comparing phishing awareness (N = 61).

Variable	Mean	SD	t	Two-Sided p
Sender—t1	6.87	2.17	0.23	0.823
Sender—t2	6.77	2.08		
Grammar—t1	7.47	2.06	−2.52	0.014
Grammar—t2	8.04	1.86		
Urgency—t1	7.49	2.42	−0.06	0.956
Urgency—t2	7.43	2.23		

5.4. Impact of level of control over agent's aesthetic features on learner experiences

5.4.1. Impact on reported awareness of phishing (RQ3a)

We found no statistically significant main effects of levels of control over agent's aesthetic features on phishing awareness as measured immediately after the experience (F(3, 110) = 0.379, $p = 0.768$) and three months later (F(3, 56) = 0.981, $p = 0.408$). Descriptive statistics can be seen in Table 6.

Next we tested whether there was a change over time. Repeated measures ANOVA test showed that there was no statistically significant main effect of neither time (F(3, 59) = 0.290, $p = 0.592$) nor group on reported phishing awareness (F(3, 59) = 0.736, $p = 0.535$). In addition, there was also no statistically significant interaction between group and time (F(3,59) = 1.39, $p = 0.254$) on reported phishing awareness.

5.4.2. Impact on mental effort, enjoyment, sense of presence and relatedness (RQ3b)

One-Way ANOVA tests showed that the different levels of control over agent's aesthetic features did not have a statistically significant impact on participants' mental effort (F(3, 113) = 0.603, $p = 0.549$) (see Table 6). Overall, reported mental effort was 4.82 (SD = 1.77) out of 9 which is about average across all participants. Among the four groups, both customization and the gender matched no-choice group reported below average mental effort in their VR experience.

There was also no statistically significant main effects of group on participants' sense of presence (F(3, 113) = 0.948, $p = 0.420$), enjoyment (F(3, 111) = 0.937, $p = 0.425$) or sense of relatedness in t1(F(3, 111) = 1.01, $p = 0.391$) or in t2 (F(3, 111) = 1.303, $p = 0.282$).

5.5. Relationship between sense of presence and key training experience measures (RQ4)

We ran a set of Pearson correlation between sense of presence and training experience outcomes. We found a statistically significant positive correlation between sense of presence and enjoyment measures, as well as with reported phishing awareness both in t1 and t2. While the correlation between presence and mental effort was negative, it did not reach significance. Please see Table 7 for the correlation table.

Table 6. Means and standard deviations on the outcome variables based on level of control over agent's aesthetics.

	Customization		Choice		Gender match		Gender mismatch	
	M	SD	M	SD	M	SD	M	SD
Phishing awareness—t1	7.33	1.43	7.52	1.57	7.23	2.06	7.06	1.36
Phishing awareness—t2	6.96	1.33	7.27	1.72	7.35	1.79	7.91	1.19
Mental effort	4.72	1.67	5.14	1.92	4.43	1.81	5.00	1.70
Enjoyment	5.64	0.90	5.26	1.21	5.28	1.32	5.20	1.16
Sense of presence	5.44	0.60	5.40	.56	5.30	0.58	5.17	0.60
Sense of relatedness—t1	3.09	0.52	2.92	.84	2.89	0.69	2.83	0.41
Sense of relatedness—t2	2.95	1.01	3.03	1.07	2.45	1.16	2.31	1.04

Table 7. Means, standard deviations, and correlations among the variables.

	Mean	SD	1	2	3	4	5	6
1. Sense of Presence	5.34	0.58						
2. Phishing awareness—t1	7.31	1.61	0.232*					
3. Phishing awareness—t2	7.44	1.51	0.303*	0.605**				
4. Mental effort	4.82	1.77	−0.158	0.041	0.036			
5. Enjoyment	5.32	1.15	0.536**	0.249**	0.335**	0.045		
6. Sense of relatedness—t1	2.93	0.64	0.334**	0.253**	0.159	0.115	0.442**	
7. Sense of relatedness—t2	2.64	1.10	0.022	−0.059	−0.038	0.123	0.112	0.347**

Note: * indicates $p < 0.05$; ** indicates $p < 0.01$.

6. Discussion

The primary objective of this study was to explore the effects of user gender and agent characteristics, along with the level of control users have over an embodied agent's aesthetic features, on learning experiences. These effects were measured both immediately after the VR experience and three months later, specifically within the context of phishing training. This paper posed four aims with associated research questions. The findings are discussed below, under headings for each aim.

6.1. Gender dynamics in VR interactions (RQ1a-RQ1b)

The first aim was to understand how the interaction between user gender and agent gender affects learner experiences and social connection with the agent in the VR training, an area that has produced mixed results in existing literature. We found that female identified participants selected female voiced agents compared to male participants which aligns with prior work on avatar studies where female player created avatars that match their gender identity (Turkay & Adinolf, 2010). Contrary to findings by Makransky et al. (2019), which highlighted gender-specific performance in VR settings, our study found no significant interaction between participant gender and agent gender as determined by the study on mental effort, enjoyment, sense of presence, or phishing awareness. These findings echo those of Moreno and Flowerday (2006) and Shang et al. (2020). However, it is important to recognize that gender may serve as a proxy for other factors that influence VR-based performance, such as prior gaming experience. Gaming familiarity is more common among men and has been shown to impact VR navigation, engagement, and performance (Nenna & Gamberini, 2022). Future studies may consider measuring gaming experience as a potential covariate when investigating gender effects in VR training applications. The absence of significant findings might be also indicative of VR's potential to equalize learning experiences across genders in VR training. This finding might be because of the deliberate design of the robot agent where it had an androgynous structure rather than a gendered one. Thus our findings show support for predictions from prior studies on the impact of androgynous agents for inclusive design (Nag & Yalçın, 2020).

The literature suggests that relatedness is a fundamental human need that influences motivation and engagement (Ryan & Deci, 2000), yet our findings suggest that the impact of relatedness with the embodied agents may be subject to a temporal factor. Although there was no immediate post-VR-experience effect, female participants reported a higher sense of relatedness with their agents after three months. In addition, gender matching did have a significant delayed effect. This delayed effect of gender matching could imply that the establishment of a social connection with an embodied agent, while not immediately apparent, may strengthen over time, which can be especially true for female identified individuals. This is encouraging as prior work showed positive association between social relatedness and learning motivation (Müller & Palekčić, 2005).

One possible explanation for the delayed effects could be the role of reflection and memory reconsolidation (Nader & Einarsson, 2010). Over time, participants may reflect on their experiences, which could alter their perception of relatedness with the agent. This reflective process might enhance their appreciation of the agent's role and the interactions, particularly when these interactions are designed to be socially meaningful. For instance, while immediate measures of relatedness might capture initial reactions to novelty or usability aspects of the VR experience, delayed measures might better capture deeper educational impacts, such as the integration of the VR training into one's professional identity or long-term behavioral changes. This distinction underscores the need for longitudinal approaches in VR research to truly understand how interactions with agents mature and solidify into meaningful educational experiences.

6.2. Influence of VR on phishing awareness (RQ2)

This study investigated the long-term effects of VR training on phishing awareness. Based on prior research, we anticipated that the effects of VR training on phishing awareness might diminish over time. However, our findings suggest that participants retained increased vigilance towards phishing cues, particularly in grammatical errors, even after three months.

This delayed increase in phishing awareness suggests that VR training may have latent effects, where initial exposure sets a foundation for subsequent reflection and knowledge integration. In educational contexts, this indicates that the benefits of VR training may unfold over a longer timeframe than traditional assessments typically capture. Future research should consider incorporating delayed testing intervals to better capture the long-term retention and behavior change effects of VR training.

The enhancement in phishing awareness aligns with the core training mechanic, which emphasizes scrutinizing email content for anomalies—a crucial skill for detecting phishing attempts. This observation is supported by research highlighting the effectiveness of targeted training in boosting phishing awareness, such as the study by Vishwanath et al. (2011). It also adds to the discussion about the enduring impact of VR interventions in educational settings, suggesting that immersive VR environments can effectively influence behavior change, as discussed in Krassmann et al. (2019).

Furthermore, this finding underscores the potential of VR as a vital tool in phishing awareness training, particularly in fostering critical observational skills essential for identifying subtle phishing cues. Future studies should delve deeper into identifying which specific elements of VR training contribute to these durable outcomes and how these elements can be optimized to enhance different facets of phishing awareness.

It is important to acknowledge that we cannot entirely isolate other factors that might have influenced these findings. For instance, participants' phishing awareness might have naturally improved over time, coinciding with our follow-up surveys.

6.3. Effect of control over agent's aesthetics on training experiences (RQ3)

Our research explored the impact of allowing users to customize the aesthetic features of virtual agents on training experiences and retention of phishing awareness. Contrary to expectations, we discovered that such customization does not significantly affect training experiences. This finding is surprising given previous research that has documented positive effects of customization on enhancing prosocial behavior and user engagement (Cuthbert et al., 2019; Herrera & Bailenson, 2021; Wald et al., 2021; Xiao et al., 2007).

The influence of customization may depend on how it engages users. Customization that directly affects learning content—such as choosing problem-solving strategies or paths—might substantially enhance motivation (Cuthbert et al., 2019). In contrast, aesthetic customization, which alters only the superficial features like an agent's appearance, appears to have minimal impact on the deeper cognitive or affective processes.

Additionally, our data showed that phishing awareness increased over time across all participant groups, regardless of the level of customization. This improvement suggests that simply engaging with VR training may yield enduring educational benefits, indicating the critical role of immersive content quality over aesthetic customization in VR training environments.

This phenomenon aligns with media richness theory, which asserts that the effectiveness of a communication medium is rooted in its ability to convey information-rich interactions (Maity et al., 2018). The inherent richness and immersive capability of VR potentially may make the details of customization less influential than the overall quality and relevance of the content presented. Therefore, prioritizing interaction and relevance in VR scenarios might be more impactful than extensive customization options.

This perspective recommends that future VR training programs to focus more on content that deeply engages users' learning processes before features like modifying aesthetic elements, which may have negligible effects on training effectiveness.

6.4. Relationship between sense of presence and key training measures (RQ4)

The relationship between the sense of presence and key training measures was also explored. The study found a positive correlation between sense of presence and measures of enjoyment, as well as reported phishing awareness immediately and three months after the VR experience, and no statistically significant correlation with mental effort was observed. These results are in line with previous research by Chirico and Gaggioli (2019) and Tussyadiah et al. (2018), who highlighted the connection between presence and emotional response. They underscore the importance of a strong sense of presence as a factor for enhancing enjoyment and potentially reinforcing learning retention, even if it does not necessarily impact mental effort. Our results also align with Huang et al. (2020)'s findings, which suggested that the sense of presence could enhance involvement and enjoyment in the learning process without directly influencing educational outcomes.

The positive correlation between sense of presence and phishing awareness shows evidence that the immersive and interactive aspects of VR settings improve attention to and retention of educational content (Krassmann et al., 2019; Tang et al., 2022). By simulating scenarios, VR allows learners to practice skills in a context-rich setting which is often more impactful than traditional learning environments due to the increased attention, relevance and application of the knowledge gained (Krassmann et al., 2019).

Interestingly, the correlation between sense of presence and mental effort was negative, although it did not reach statistical significance. This trend might suggest that a stronger sense of presence could potentially reduce the cognitive load on learners by making interactions more intuitive or natural. However, the lack of significance indicates that further research is needed to explore this relationship. It is possible that the immersive nature of VR could either simplify cognitive processes by providing context-rich cues or complicate them due to the complexity of navigating 3D immersive virtual environments (Makransky & Petersen, 2021).

6.5. Implications for future VR training environments

The findings of this study have implications for the design of VR training environments and agent-based training applications. While customization and gender factors may not significantly influence training outcomes, the sustained sense of relatedness with agents reported by female participants suggests that VR interventions could play a role in fostering long-term interest for female identified users on the subject matters.

Understanding that interactions with embodied VR agents might have delayed effects has significant implications for the design of VR experiences. Designers should consider incorporating features that support ongoing engagement and reflection. For example, implementing reflective prompts at the end of sessions could encourage users to think about what they learned and how they felt about the interactions. Additionally, follow-up sessions that revisit the content after some time could help reinforce learning and strengthen the sense of connection with the virtual agent. Designers could also use adaptive algorithms that tailor follow-up activities based on initial user responses, thereby maintaining relevance and engagement. These design considerations can leverage the potential for long-term engagement and transform initial virtual interactions into lasting educational relationships.

The significant correlations between sense of presence and enjoyment, and phishing awareness imply that designing VR experiences that not only foster a strong sense of presence but also actively contribute to enjoyment and learning efficacy is important. Ensuring high-quality, immersive experiences can enhance the educational value of VR, making it a powerful tool in areas like phishing awareness training. Designers may focus on the elements that increase presence, such as consistent graphics, responsive controls, embodied interactions, physiological feedback and scenario realism, to maximize the educational impact of VR training programs (Cho, 2022; Riches et al., 2019; Souza et al., 2018).

6.6. Limitations

This study had multiple limitations which pave the way for future research. The relatively homogeneous participant pool with low numbers of female identified individuals points to the need for studies with more diverse demographics to better understand the universal applicability of VR training. In addition, the gender of agents was solely determined based on voice, potentially leading to a mismatch between the researcher's categorization and participants' perception. Future studies should include a broader range of gender identities and consider participants' perceptions of agent gender, which could influence training effectiveness. In addition, future research may explore the role of linguistic diversity, including language background and accent variations, in VR-based training.

Another limitation of this study is the dropout rate between the immediate post-survey and the three-month follow-up, with only 61 out of 114 participants (53.5%) completing the follow-up assessment. While this retention rate is within the expected range for longitudinal studies (Abdullah et al., 2021), it may have affected the statistical power of long-term phishing awareness findings. Future studies should consider strategies to improve retention, such as incentives or alternative follow-up methods, to strengthen the reliability of delayed outcome measures.

The study was limited to aesthetic customization of agents. Prior research has indicated that functional customization—allowing users to alter or control aspects of the agent's behavior or role—could also significantly impact user experiences. Expanding customization options in future studies could yield different results, particularly in terms of engagement and motivation. Additionally, longitudinal studies that monitor VR engagement over extended periods could yield deeper insights into the lasting effects of VR training on knowledge retention and skill development.

7. Conclusion

This study investigated how control over a digital agent's aesthetics and the interplay of user and agent gender influence learning experiences in a virtual reality (VR) environment, specifically within the context of phishing training. Through nested-control experimental design and analysis, the findings revealed several critical insights into the efficacy and dynamics of VR educational interventions.

Findings shed light on the role of user and agent gender dynamics, impact of the level of control over the aesthetics of embodied agents, and sense of presence in shaping user experiences in a VR training setting. The revelation that customization does not significantly influence immediate training outcomes but may affect long-term behavior change suggests that VR experience designers might prioritize different elements for different training objectives.

The study also shows evidence on the importance of relatedness in behavior change in the long term, aligning with the theoretical frameworks proposed by Ryan and Deci (2000) and Biocca et al. (2003). Additionally, the results on relatedness and sense of presence offer new insights into the psychological impacts of VR learning environments, highlighting the importance of these factors in the design of future VR training experiences, paving the way for more inclusive and effective educational technologies. As VR technology continues to evolve, further research is needed to understand how these social aspects of learning can be effectively integrated into VR training programs to support user motivation and educational engagement.

Author contributions

CRediT: **Sonam Adinolf**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing; **David Conroy**: Data curation, Software, Writing – original draft; **Peta Wyeth**: Conceptualization, Writing – review & editing; **Leonie Simpson**: Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Sonam Adinolf http://orcid.org/0009-0007-8131-5731
David Conroy http://orcid.org/0000-0002-1672-5941
Peta Wyeth http://orcid.org/0000-0003-1867-9536
Leonie Simpson http://orcid.org/0000-0001-8434-9741

References

Abawajy, J. (2014). User preference of cyber security awareness delivery methods. *Behaviour & Information Technology*, *33*(3), 237–248. https://doi.org/10.1080/0144929X.2012.708787

Abdullah, N., Kamaruddin, M. A., Goh, Y.-X., Othman, R., Dauni, A., Jalal, N. A., Yusuf, N. A. M., Kamat, S. A., Basri, N. H., & Jamal, R. (2021). Participants attrition in a longitudinal study: The malaysian cohort study experience. *International Journal of Environmental Research and Public Health*, *18*(14), 7216. https://doi.org/10.3390/ijerph18147216

Adinolf, S., Wyeth, P., Brown, R., & Harman, J. (2020a). *My little robot: User preferences in game agent customization* [Paper presentation]. Proceedings of the Annual Symposium on Computer-Human Interaction in Play (pp. 461–471). https://doi.org/10.1145/3410404.3414241

Adinolf, S., Wyeth, P., Brown, R., & Simpson, L. (2020b). *Near and dear: Designing relatable vr agents for training games* [Paper presentation]. Proceedings of the 32nd Australian Conference on Human-Computer Interaction (pp. 413–425). https://doi.org/10.1145/3441000.3441007

Aldawood, H., & Skinner, G. (2019). Reviewing cyber security social engineering training and awareness programs—pitfalls and ongoing issues. *Future Internet*, *11*(3), 73. https://doi.org/10.3390/fi11030073

Alqahtani, H., & Kavakli-Thorne, M. (2020). Design and evaluation of an augmented reality game for cybersecurity awareness (cybar). *Information*, *11*(2), 121. https://doi.org/10.3390/info11020121

Alsharbi, B., & Richards, D. (2017). *Using virtual reality technology to improve reality for young people with chronic health conditions* [Paper presentation]. Proceedings of the 9th International Conference on Computer and Automation Engineering (pp. 11–15). https://doi.org/10.1145/3057039.3057080

Baylor, A. L. (2009). Promoting motivation with virtual agents and avatars: Role of visual presence and appearance. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1535), 3559–3565. https://doi.org/10.1098/rstb.2009.0148

Behrend, T. S., & Thompson, L. F. (2012). Using animated agents in learner-controlled training: The effects of design control. *International Journal of Training and Development*, *16*(4), 263–283. https://doi.org/10.1111/j.1468-2419.2012.00413.x

Biocca, F., Harms, C., & Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators and Virtual Environments*, *12*(5), 456–480. https://doi.org/10.1162/105474603322761270

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

Bulu, S. T. (2012). Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers & Education*, *58*(1), 154–161. https://doi.org/10.1016/j.compedu.2011.08.024

Caputo, D. D., Pfleeger, S. L., Freeman, J. D., & Johnson, M. E. (2014). Going spear phishing: Exploring embedded training and awareness. *IEEE Security & Privacy*, *12*(1), 28–38. https://doi.org/10.1109/MSP.2013.106

Chandrashekar, N., King, K., Gračanin, D., & Azab, M. (2023). *Design & development of virtual reality empowered cyber-security training testbed for iot systems* [Paper presentation]. 2023 3rd Intelligent Cybersecurity Conference (ICSC) (pp. 86–94). https://doi.org/10.1109/ICSC60084.2023.10349976

Chen, S., Hao, M., Ding, F., Jiang, D., Dong, J., Zhang, S., Guo, Q., & Gao, C. (2023). Exploring the global geography of cybercrime and its driving forces. *Humanities & Social Sciences Communications*, *10*(1), 71. https://doi.org/10.1057/s41599-023-01560-x

Chirico, A., & Gaggioli, A. (2019). When virtual feels real: Comparing emotional responses and presence in virtual and natural environments. *Cyberpsychology, Behavior and Social Networking*, *22*(3), 220–226. https://doi.org/10.1089/cyber.2018.0393

Cho, Y. (2022). *Virtual embodiment for enhancing sense of presence in virtual reality* [Paper presentation]. Proceedings 9th International Conference on Kansei Engineering and Emotion Research. KEER2022. http://dx.doi.org/10.5821/conference-9788419184849.70

Cone, B. D., Irvine, C. E., Thompson, M. F., & Nguyen, T. D. (2007). A video game for cyber security training and awareness. *Computers & Security*, *26*(1), 63–72. https://doi.org/10.1016/j.cose.2006.10.005

Craig, S. D., Twyford, J., Irigoyen, N., & Zipp, S. A. (2015). A test of spatial contiguity for virtual human's gestures in multimedia learning environments. *Journal of Educational Computing Research*, *53*(1), 3–14. https://doi.org/10.1177/0735633115585927

Cuthbert, R., Turkay, S., & Brown, R. (2019). *The effects of customisation on player experiences and motivation in a virtual reality game* [Paper presentation]. Proceedings of the 31st Australian Conference on Human-Computer-Interaction (pp. 221–232). https://doi.org/10.1145/3369457.3369475

Feijóo-García, P. G., Zalake, M., de Siqueira, A. G., Lok, B., & Hamza-Lup, F. (2021). *Effects of virtual humans' gender and spoken accent on users' perceptions of expertise in mental wellness conversations* [Paper presentation]. Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents (pp. 68–75). https://doi.org/10.1145/3472306.3478367

Graesser, A., & McNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist*, *45*(4), 234–244. https://doi.org/10.1080/00461520.2010.515933

Grassini, S., Laumann, K., & Rasmussen Skogstad, M. (2020). The use of virtual reality alone does not promote training performance (but sense of presence does). *Frontiers in Psychology*, *11*, 1743. https://doi.org/10.3389/fpsyg.2020.01743

Grivokostopoulou, F., Kovas, K., & Perikos, I. (2020). The effectiveness of embodied pedagogical agents and their impact on students learning in virtual worlds. *Applied Sciences*, *10*(5), 1739. https://doi.org/10.3390/app10051739

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review*, *6*(1), 27–54. https://doi.org/10.1016/j.edurev.2010.07.004

Hein, D., Mai, C., & Hußmann, H. (2018). *The usage of presence measurements in research: A review* [Paper presentation]. Proceedings of the International Society for Presence Research Annual Conference (Presence) (pp. 21–22). The International Society for Presence Research Prague. https://www.mmi.informatik.uni-muenchen.de/pubdb/publications/pub/mai2018Presence/mai2018Presence.pdf

Herrera, F., & Bailenson, J. N. (2021). Virtual reality perspective-taking at scale: Effect of avatar representation, choice, and head movement on prosocial behaviors. *New Media & Society*, *23*(8), 2189–2209. https://doi.org/10.1177/1461444821993121

Hillman, D., Harel, Y., & Toch, E. (2023). Evaluating organizational phishing awareness training on an enterprise scale. *Computers & Security*, *132*, 103364. https://doi.org/10.1016/j.cose.2023.103364

Hoorn, J. F. (2015). Psychological aspects of technology interacting with humans. In *The Handbook of the Psychology of Communication Technology* (Chapter 8, pp. 176). https://doi.org/10.1002/9781118426456.ch8

Huang, C. L., Luo, Y. F., Yang, S. C., Lu, C. M., & Chen, A.-S. (2020). Influence of students' learning style, sense of presence, and cognitive load on learning outcomes in an immersive virtual reality learning environment. *Journal of Educational Computing Research*, *58*(3), 596–615. https://doi.org/10.1177/0735633119867422

Huang, W., Roscoe, R. D., Craig, S. D., & Johnson-Glenberg, M. C. (2022). Extending the cognitive-affective theory of learning with media in virtual reality learning: A structural equation modeling approach. *Journal of Educational Computing Research*, *60*(4), 807–842. https://doi.org/10.1177/07356331211053630

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*(6), 995–1006. https://doi.org/10.1037/0022-3514.79.6.995

Jampen, D., Gür, G., Sutter, T., & Tellenbach, B. (2020). Don't click: Towards an effective anti-phishing training. a comparative literature review. *Human-Centric Computing and Information Sciences*, *10*(1), 33. https://doi.org/10.1186/s13673-020-00237-7

Johnston, E., Olivas, G., Steele, P., Smith, C., & Bailey, L. (2018). Exploring pedagogical foundations of existing virtual reality educational applications: A content analysis study. *Journal of Educational Technology Systems*, *46*(4), 414–439. https://doi.org/10.1177/0047239517745560

Karaosmanoglu, S., Schmolzi, T., & Steinicke, F. (2023). *Playing with friends or strangers? the effects of familiarity between players in an asymmetric multiplayer virtual reality game* [Paper presentation]. Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play, CHI PLAY Companion '23 (pp. 76–82). Association for Computing Machinery. https://doi.org/10.1145/3573382.3616079

Kavvadias, A., & Kotsilieris, T. (2025). Understanding the role of demographic and psychological factors in users' susceptibility to phishing emails: A review. *Applied Sciences*, *15*(4), 2236. https://doi.org/10.3390/app15042236

Khan, M. (2018). *Anti-phishing training vs. does security awareness training work?*. https://securityboulevard.com/2018/08/anti-phishing-training-vs-software-does-security-awareness-training-work/

Kim, K., Boelling, L., Haesler, S., Bailenson, J., Bruder, G., & Welch, G. F. (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* [Paper presentation]. 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 105–114). https://doi.org/10.1109/ISMAR.2018.00039

Kim, Y., & Wei, Q. (2011). The impact of learner attributes and learner choice in an agent-based environment. *Computers & Education*, *56*(2), 505–514. https://doi.org/10.1016/j.compedu.2010.09.016

Kinzer, C. K., Hoffman, D. L., Türkay, S., Günba, N., Chantes, P., Dvorkin, T., & Chaiwinij, A. (2012). The impact of choice and feedback on learning, motivation, and performance in an educational video game (pp. 8). https://www.academia.edu/download/82889429/Kinzer_20et_20al_202012_20Choice_20and_20Feedback.pdf

Koda, T., Tsuji, S., & Takase, M. (2022). *Measuring subconscious gender biases against male and female virtual agents in Japan* [Paper presentation]. Proceedings of the 10th International Conference on Human-Agent Interaction, HAI '22 (pp. 275–277). Association for Computing Machinery. https://doi.org/10.1145/3527188.3563909

Krassmann, A., Melo, M., Pinto, D., Peixoto, B., Bessa, M., & Bercht, M. (2019). What is the relationship between the sense of presence and learning in virtual reality? a 24-year systematic literature review. *PRESENCE: Virtual and Augmented Reality*, *28*, 247–265. https://doi.org/10.1162/pres_a_00350

Kroll, S. T. (2019). *Is virtual reality the next generation in security awareness training?* https://cybersecurityventures.com/is-virtual-reality-the-next-generation-in-security-awareness-training/

Kyrlitsias, C., & Michael-Grigoriou, D. (2022). Social interaction with agents and avatars in immersive virtual environments: A survey. *Frontiers in Virtual Reality*, *2*, 786665. https://doi.org/10.3389/frvir.2021.786665

Lee, Y. Y., Gan, C. L., & Liew, T. W. (2023). Susceptibility to instant messaging phishing attacks: Does systematic information processing differ between genders? *Crime Prevention and Community Safety*, *25*(2), 179–203. https://doi.org/10.1057/s41300-023-00176-2

Lim, S., & Reeves, B. (2009). Being in the game: Effects of avatar choice and point of view on psychophysiological responses during play. *Media Psychology*, *12*(4), 348–370. https://doi.org/10.1080/15213260903287242

Lin, L., Parmar, D., Babu, S. V., Leonard, A. E., Daily, S. B., & Jörg, S. (2017). How character customization affects learning in computational thinking. In *Proceedings of the ACM Symposium on Applied Perception, SAP '17* (pp. 1:1–1:8). ACM. https://doi.org/10.1145/3119881.3119884

Maity, M., Dass, M., & Kumar, P. (2018). The impact of media richness on consumer information search and choice. *Journal of Business Research*, *87*, 36–45. https://doi.org/10.1016/j.jbusres.2018.02.003

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (camil): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, *33*(3), 937–958. https://doi.org/10.1007/s10648-020-09586-2

Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2021). 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*, *113*(4), 719–735. https://doi.org/10.1037/edu0000473

Makransky, G., Wismer, P., & Mayer, R. E. (2019). A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *Journal of Computer Assisted Learning*, *35*(3), 349–358. https://doi.org/10.1111/jcal.12335

McGill, T., & Thompson, N. (2021). Exploring potential gender differences in information security and privacy. *Information & Computer Security*, *29*(5), 850–865. https://doi.org/10.1108/ICS-07-2020-0125

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & 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. https://doi.org/10.1016/j.compedu.2013.07.033

Mobed, D. A. O., Wodehouse, A., & Maier, A. (2024). The aesthetics of robot design: Towards a classification of morphologies. *Proceedings of the Design Society*, *4*, 2413–2422. https://doi.org/10.1017/pds.2024.244

Moreno, R., & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology*, *31*(2), 186–207. https://doi.org/10.1016/j.cedpsych.2005.05.002

Mostajeran, F., Katzakis, N., Ariza, O., Freiwald, J. P., & Steinicke, F. (2019). *Welcoming a holographic virtual coach for balance training at home: Two focus groups with older adults* [Paper presentation].2019 IEEE conference on virtual reality and 3D User Interfaces (VR) (pp. 1465–1470). IEEE. https://doi.org/10.1109/VR.2019.8797813

Moumouh, C., Chkouri, M. Y., & Fernández-Alemán, J. L. (2023). *Cybersecurity awareness through serious games: A systematic literature review* [Paper presentation]. International Conference on Networking, Intelligent Systems and Security (pp. 190–199). Springer. https://doi.org/10.1007/978-3-031-15191-0_18

Müller, F. H., & Palekčić, M. (2005). Continuity of motivation in higher education: A three-year follow-up study. *Review of Psychology*, *12*(1), 31–43.

Nader, K., & Einarsson, E. Ö. (2010). Memory reconsolidation: An update. *Annals of the New York Academy of Sciences*, *1191*(1), 27–41. https://doi.org/10.1111/j.1749-6632.2010.05443.x

Nag, P., & Yalçın, Ö. N. (2020). *Gender stereotypes in virtual agents* [Paper presentation]. Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents (pp. 1–8). https://doi.org/10.1145/3383652.3423876

Nenna, F., & Gamberini, L. (2022). *The influence of gaming experience, gender and other individual factors on robot teleoperations in vr* [Paper presentation]. 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 945–949). IEEE. https://doi.org/10.1109/HRI53351.2022.9889669

Nikendei, C., Kraus, B., Schrauth, M., Weyrich, P., Zipfel, S., Herzog, W., & Jünger, J. (2007). Integration of role-playing into technical skills training: A randomized controlled trial. *Medical Teacher*, *29*(9), 956–960. https://doi.org/10.1080/01421590701601543

Nikendei, C., Zeuch, A., Dieckmann, P., Roth, C., Schäfer, S., Völkl, M., Schellberg, D., Herzog, W., & Jünger, J. (2005). Role-playing for more realistic technical skills training. *Medical Teacher*, *27*(2), 122–126. https://doi.org/10.1080/01421590400019484

Novick, D., Afravi, M., Camacho, A., Rodriguez, A., & Hinojos, L. (2019). Pedagogical-agent learning companions in a virtual reality educational experience. In *Learning and Collaboration Technologies. Ubiquitous and Virtual Environments for Learning and Collaboration: 6th International Conference, LCT 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part II 21* (pp. 193–203). Springer. https://doi.org/10.1007/978-3-030-21817-1_15

Okita, S. Y., Turkay, S., Kim, M., & Murai, Y. (2013). Learning by teaching with virtual peers and the effects of technological design choices on learning. *Computers & Education*, *63*, 176–196. https://doi.org/10.1016/j.compedu.2012.12.005

Othlinghaus-Wulhorst, J., & Hoppe, H. U. (2020). A technical and conceptual framework for serious role-playing games in the area of social skill training. *Frontiers in Computer Science*, *2*, 28. https://doi.org/10.3389/fcomp.2020.00028

Paas, F. G., Van Merriënboer, J. J., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, *79*(1 Pt 2), 419–430. https://doi.org/10.2466/pms.1994.79.1.419

Pattinson, M., Jerram, C., Parsons, K., McCormac, A., & Butavicius, M. (2012). Why do some people manage phishing emails better than others? *Information Management & Computer Security*, *20*(1), 18–28. https://doi.org/10.1108/09685221211219173

Petersen, G. B., Mottelson, A., & Makransky, G. (2021). *Pedagogical agents in educational vr: An in the wild study* [Paper presentation]. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (pp. 1–12). https://doi.org/10.1145/3411764.3445760

Pratama, A. R., Vadila, N., & Firmansyah, F. M. (2023). *Exposing generational and gender gap in phishing awareness among young adults: A survey experiment volume* [Paper presentation]. AIP Conference Proceedings, 2508. AIP Publishing. https://doi.org/10.1063/5.0114868

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778. https://doi.org/10.1063/5.0114868

Rana, S., & Chicone, R. G. (2023). The infleunce and gender and acceptance of vr cybersecurity training platforms. *Issues in Information Systems*, *24*(1), 93–100. https://doi.org/10.48009/1_iis_2023_108

Refinetti, R. (1996). Demonstrating the consequences of violations of assumptions in between-subjects analysis of variance. *Teaching of Psychology*, *23*(1), 51–54. https://doi.org/10.1207/s15328023top2301_14

Reinheimer, B., Aldag, L., Mayer, P., Mossano, M., Duezguen, R., Lofthouse, B., Von Landesberger, T., & Volkamer, M. (2020). *An investigation of phishing awareness and education over time: When and how to best remind users* [Paper presentation]. Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020) (pp. 259–284). https://www.usenix.org/conference/soups2020/presentation/reinheimer

Renganayagalu, S. K., Mallam, S. C., & Nazir, S. (2021). Effectiveness of vr head mounted displays in professional training: A systematic review. *Technology, Knowledge and Learning*, *26*(4), 999–1041. https://doi.org/10.1007/s10758-020-09489-9

Riches, S., Elghany, S., Garety, P., Rus-Calafell, M., & Valmaggia, L. (2019). Factors affecting sense of presence in a virtual reality social environment: A qualitative study. *Cyberpsychology, Behavior and Social Networking*, *22*(4), 288–292. https://doi.org/10.1089/cyber.2018.0128

Rogers, K., Aufheimer, M., Weber, M., & Nacke, L. E. (2018). Exploring the Role of Non-Player Characters and Gender in Player Identification. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts - CHI PLAY '18 Extended Abstracts*. (pp. 271–283). ACM Press. https://doi.org/10.1145/3270316.3273041

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, *43*(3), 450–461. https://doi.org/10.1037/0022-3514.43.3.450

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

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, *30*(4), 344–360. https://doi.org/10.1007/s11031-006-9051-8

Salameh, R., & Loh, C. S. (2022). Engagement and players' intended behaviors in a cybersecurity serious game. *International Journal of Gaming and Computer-Mediated Simulations*, *14*(1), 1–21. https://doi.org/10.4018/IJGCMS.313185

Sari, R. C., Pranesti, A., Solikhatun, I., Nurbaiti, N., & Yuniarti, N. (2024). Cognitive overload in immersive virtual reality in education: More presence but less learnt? *Education and Information Technologies*, *29*(10), 12887–12909. https://doi.org/10.1007/s10639-023-12379-z

Schlecht, R. (2018). *Augmented reality could help solve the cybersecurity talent gap*.

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust. *Methodology*, *6*(4), 147–151. https://doi.org/10.1027/1614-2241/a000016

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? a meta-analytic review. *Journal of Educational Computing Research*, *49*(1), 1–39. https://doi.org/10.2190/EC.49.1.a

Servotte, J., Goosse, M., Campbell, S., Dardenne, N., Pilote, B., Simoneau, I. L., Guillaume, M., Bragard, I., & Ghuysen, A. (2020). Virtual reality experience: Immersion, sense of presence, and cybersickness. *Clinical Simulation in Nursing*, *38*, 35–43. https://doi.org/10.1016/j.ecns.2019.09.006

Shang, X., Kallmann, M., & Arif, A. (2020). Effects of virtual agent gender on user performance and preference in a vr training program. In K. Arai & R. Bhatia (Eds.), *Advances in information and communication. FICC 2019. Lecture Notes in Networks and Systems* (Vol. 69). Springer, Cham. https://doi.org/10.1007/978-3-030-12388-8_34

Sherman, G. D., Türkay, S., Moulton, S. T., Friedman, M. C., Darani, N., Daly, B., & Kayden, S. (2020). The generalized sense of power is a psychological resource: Evidence from a disaster response field training exercise. *European Journal of Social Psychology*, *50*(4), 733–748. https://doi.org/10.1002/ejsp.2644

Sikström, P., Valentini, C., Sivunen, A., & Kärkkäinen, T. (2022). How pedagogical agents communicate with students: A two-phase systematic review. *Computers & Education*, *188*, 104564. https://doi.org/10.1016/j.compedu.2022.104564

Slater, M. (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*(1535), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

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

Sogunro, O. A. (2004). Efficacy of role-playing pedagogy in training leaders: Some reflections. *The Journal of Management Development; Bradford*, *23*(3/4), 355–371. https://doi.org/10.1108/02621710410529802

Souza, V., Nedel, L., Kopper, R., Maciel, A., & Tagliaro, L. (2018). *The effects of physiologically-adaptive virtual environment on user's sense of presence* [Paper presentation]. 2018 20th Symposium on Virtual and Augmented Reality (SVR) (pp. 133–142). https://doi.org/10.1109/SVR.2018.00029

Sun, J. C.-Y., Yu, S.-J., Lin, S. S., & Tseng, S.-S. (2016). The mediating effect of anti-phishing self-efficacy between college students' internet self-efficacy and anti-phishing behavior and gender difference. *Computers in Human Behavior*, *59*, 249–257. https://doi.org/10.1016/j.chb.2016.02.004

Tabbers, H. K., Martens, R. L., & Van Merriënboer, J. J. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *The British Journal of Educational Psychology*, *74*(Pt 1), 71–81. https://doi.org/10.1348/000709904322848824

Tang, Q., Wang, Y., Liu, H., Liu, Q., & Jiang, S. (2022). Experiencing an art education program through immersive virtual reality or ipad: Examining the mediating effects of sense of presence and extraneous cognitive load on enjoyment, attention, and retention. *Frontiers in Psychology*, *13*, 957037. https://doi.org/10.3389/fpsyg.2022.957037

Threatpost (2021). Cybercrime gangs get more sophisticated. *Threatpost*. https://threatpost.com/cybercrime-more-sophisticated/179676/ (visited on 2021-09-15).

Turkay, S. (2013). *The effects of customization on player experiences in an extended online social game: A mixed method study* [Ed.D.].

Teachers College, Columbia University. https://www.proquest.com/dissertations-theses/effects-customization-on-player-experiences/docview/1430900227/se-2

Turkay, S., & Adinolf, S. (2010). Free to be me: A survey study on customization with World of Warcraft and City Of Heroes/Villains players. *Procedia - Social and Behavioral Sciences*, *2*(2), 1840–1845. https://doi.org/10.1016/j.sbspro.2010.03.995

Türkay, S., & Kinzer, C. K. (2016). The relationship between avatar-based customization, player identification, and motivation. In *Transforming Gaming and Computer Simulation Technologies across Industries* (pp. 48–79). IGI Global. https://doi.org/10.4018/978-1-5225-1817-4.ch003

Tussyadiah, I. P., Wang, D., Jung, T. H., & Tom Dieck, M. C. (2018). Virtual reality, presence, and attitude change: Empirical evidence from tourism. *Tourism Management*, *66*, 140–154. https://doi.org/10.1016/j.tourman.2017.12.003

UQO Cyberpsychology Laboratory (2004). *Traduction du presence questionnaire de witmer & singer*. Technical report. Université du Québec en Outaouais.

Veletsianos, G. (2007). Cognitive and Affective Benefits of an Animated Pedagogical Agent: Considering Contextual Relevance and Aesthetics. *Journal of Educational Computing Research*, *36*(4), 373–377. https://doi.org/10.2190/T543-742X-033L-9877

Vella, K., Johnson, D., & Hides, L. (2015). *Playing alone, playing with others: Differences in player experience and indicators of wellbeing* [Paper presentation]. Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (pp. 3–12). https://doi.org/10.1145/2793107.2793118

Veneruso, S. V., Ferro, L. S., Marrella, A., Mecella, M., & Catarci, T. (2020). *Cybervr: An interactive learning experience in virtual reality for cybersecurity related issues* [Paper presentation]. Proceedings of the 2020 International Conference on Advanced Visual Interfaces (pp. 1–8). https://doi.org/10.1145/3399715.3399860

Vishwanath, A., Harrison, B., & Ng, Y. J. (2018). Suspicion, cognition, and automaticity model of phishing susceptibility. *Communication Research*, *45*(8), 1146–1166. https://doi.org/10.1177/0093650215627483

Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, H. R. (2011). Why do people get phished? testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems*, *51*(3), 576–586. https://doi.org/10.1016/j.dss.2011.03.002

Wald, R., Heijselaar, E., & Bosse, T. (2021). *Make your own: The potential of chatbot customization for the development of user trust* [Paper presentation]. Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization, UMAP '21 (pp. 382–387). Association for Computing Machinery. https://doi.org/10.1145/3450614.3463600

Wang, F., Li, W., Mayer, R. E., & Liu, H. (2018). Animated pedagogical agents as aids in multimedia learning: Effects on eye-fixations during learning and learning outcomes. *Journal of Educational Psychology*, *110*(2), 250–268. https://doi.org/10.1037/edu0000221

Weidner, F., Boettcher, G., Arboleda, S. A., Diao, C., Sinani, L., Kunert, C., Gerhardt, C., Broll, W., & Raake, A. (2023). A systematic review on the visualization of avatars and agents in ar & vr displayed using head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, *29*(5), 2596–2606. https://doi.org/10.1109/TVCG.2023.3247072

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

Xiao, J., Stasko, J., & Catrambone, R. (2007). *The role of choice and customization on users' interaction with embodied conversational agents: Effects on perception and performance* [paper presentation]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1293–1302). https://doi.org/10.1145/1240624.1240820

Zargham, N., Alexandrovsky, D., Mildner, T., Porzel, R., & Malaka, R. (2023). *"let's face it": Investigating user preferences for virtual humanoid home assistants* [paper presentation]. Proceedings of the 11th International Conference on Human-Agent Interaction (pp. 246–256). https://doi.org/10.1145/3623809.3623821

About the authors

Sonam Adinolf has received his PhD in HCI from QUT. His research includes customization in digital environments, social dynamics surrounding trading card games, and toxicity in online games. His current research seeks to engage digital agents in virtual reality to help improve social and psychological outcomes for users.

David Conroy is a Lecturer, Researcher and Virtual Environment Developer primarily focused on Game AI, Virtual Reality Development and Game Design. David has an interest in the articulation and expression of nuanced gameplay mechanics and their implementation, including educating students on such principles at a tertiary education institutional level.

Peta Wyeth is the Dean of the Faculty Engineering and IT at the University of Technology Sydney. She is internationally recognised in the field of Human-Computer Interaction, employing techniques from the fields of interaction and experience design, and computer science in the design of educational and entertainment technology.

Leonie Simpson is an Associate Professor in the School of Computer Science, at QUT. She is a cryptographer with a passion for information security, particularly the application of cryptography in control measures to provide confidentiality, integrity assurance and authentication, and the management of cryptographic keys.