Avatar Type, Self-Congruence, and Presence in Virtual Reality

Tianqi Huang
School of Advanced Technology
Xi'an Jiaotong-Liverpool University
Suzhou, China
tianqi.huang17@alumni.xjtlu.edu.cn

Yue Li*
Department of Computing
School of Advanced Technology
Xi'an Jiaotong-Liverpool University
Suzhou, China
yue.li@xjtlu.edu.cn

Hai-Ning Liang
Department of Computing
School of Advanced Technology
Xi'an Jiaotong-Liverpool University
Suzhou, China
haining.liang@xjtlu.edu.cn

Fig. 1. Real-life environments populated with different avatar types. (a) Classroom scene with anime avatars; (b) Gallery scene with human avatars; (c) Café scene with human and anime avatars; (d) Street scene with human, anime, animal, and item avatars; (e) Forest scene with animal avatars.

ABSTRACT

Avatars serve as users' virtual identities and hold a significant role in shaping the user experience within the realm of Virtual Reality (VR). The appearance of individual avatars and the perceived self-congruence within the environment are likely to influence users' perceived presence in VR. In this paper, we present a study that investigates four types of avatars in VR: anime, human, animal, and item. Participants were asked to choose an avatar before entering a virtual environment (classroom, gallery, café, street, and forest) populated with avatars of different types and to evaluate their perceived self-congruence within the environment and the perceived presence. Our study results showed no significant difference in presence when users use different avatars. However, there is a correlation between users' perceived self-congruence and social presence. We discuss the findings and provide suggestions for the future use of avatars in VR.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **Virtual reality**; **Collaborative and social computing**.

KEYWORDS

virtual avatar, virtual reality, presence, social presence, self-congruence

*Corresponding author

CHCHI 2023, November 13–16, 2023, Denpasar, Bali, Indonesia

ACM ISBN 979-8-4007-1645-4/23/11…$15.00
https://doi.org/10.1145/3629606.3629614

ACM Reference Format:
Tianqi Huang, Yue Li, and Hai-Ning Liang. 2023. Avatar Type, Self-Congruence, and Presence in Virtual Reality. In *Chinese CHI 2023 (CHCHI 2023), November 13–16, 2023, Denpasar, Bali, Indonesia.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3629606.3629614

1 INTRODUCTION

In recent years, the world of technology has witnessed the rapid rise of Virtual Reality (VR) and the emergence of the Metaverse. These two interconnected trends have captured the imagination of individuals and industries alike, presenting a dynamic landscape where digital experiences and real-world interactions seamlessly blend. Central to this evolution are virtual avatars, digital representations of users that serve as their presence in the virtual realm [14, 43]. Avatars have gained remarkable significance, acting as the conduits through which people engage with this new digital frontier. Platforms such as Meta Horizon[1] and VRChat[2] showcase the growing importance of virtual avatars in shaping our online interactions and pushing the boundaries of what is possible in this exciting era of digital existence. In addition to the VR social and gaming platforms, online platforms such as Mozilla Hubs[3], Spatial[4], and Gather Town[5] have also become popular in the past few years, with online meetings, conferences, and other remote events taking place in virtual environments, all facilitated by avatars. These platforms offer a versatile way for collaborative gathering and allow customizable environments with cartoonized avatars, lifelike avatars, or pixelated avatars that facilitate users' interactions and affect their feelings, behavior, and performance [2, 12, 27, 28]. Avatars, despite their various forms and appearances, play a vital role in virtual environments, online gaming, and collaborative systems [6, 9]. Understanding the role of virtual avatars is essential to grasp

[1]https://horizon.meta.com
[2]https://hello.vrchat.com
[3]https://hubs.mozilla.com/
[4]https://www.spatial.io
[5]https://www.gather.town

the transformative potential of the Metaverse, as they redefine how we socialize, work, learn, and play, creating opportunities and challenges that demand our attention and exploration.

Within the virtual environment, an avatar is closely related to the sense of presence, namely, the sense of "being there". Lee [22] categorized presence into three categories: spatial presence (telepresence), social presence, and self-presence. Spatial presence refers to the subjective experience of being in a mediated environment or an immersive virtual environment rather than a true physical environment; social presence refers to the "sense of being with another" [3]; and self-presence refers to someone who can experience as the actual self-including body, and emotions even though using the virtual avatar. With the perceived social presence in particular, someone can feel the presence of another in the same environment, whether the other is a real person or a human-like artificial intelligence [7, 31]. Facilitating a high level of presence is closely related to the design of virtual avatars. In particular, avatar realism is essential for the development of collaborative virtual environments in the future. It is one of the main factors that can influence interpersonal interactions and social presence in VR [37]. The visual and behavioral realism of avatars fostered by natural facial expressions, gestures, and body movements were shown to have effects on social presence [34]. These non-verbal cues embodied in virtual avatars enable natural and effective communication in VR, thereby increasing social presence.

Self-congruence occurs when one's perception of one's real self and one's ideal self are identical. For example, the characters that players use in video games can create self-congruence with them, which can influence players' behavior and perception in the game world [18]. Similar to video games, avatars in virtual worlds serve almost the same functions as characters in video games. Self-congruence in this study refers to how consistent users feel with their chosen avatars in different environments.

This study is motivated by the various forms of virtual avatars used in existing social VR platforms, such as human-like, anime, and non-human avatars (e.g., animals and items). We designed a controlled experiment to investigate the effects of different types of avatars in different virtual environments on presence. Five real-life environments were pre-populated with avatars of various types (see Fig. 1). Participants were asked to select an avatar before entering the scene and to rate the self-congruence (i.e., how well their avatars fit the scene) and presence.

Our study explored the impact of different avatar types on users' perceived presence and the relationship between self-congruence and social presence in virtual environments, making the following four main contributions. First, by systematically evaluating various avatar types in different virtual environments, the research showed that the avatar type did not have a significant impact on the perceived sense of presence, and the perceived presence using different avatars in the same virtual environment did not vary significantly. Second, the study unveiled the connection between users' perceived self-congruence of their avatars in the virtual environment and the level of social presence experienced. The findings showed a significant positive correlation, shedding light on the importance of aligning users' perception of avatar characteristics and the situated environment to foster a greater sense of presence. Third, the research provided valuable insights into user preferences in different environments. This information is crucial for developers seeking to create more appealing and user-centered virtual experiences. Finally, the study's inclusion of non-human avatars (animals and items) opened new avenues for exploration, showing how these unconventional avatars may contribute to presence and expand the boundaries of avatar-mediated interactions. Together, these contributions contribute to a richer understanding of the dynamics between avatars, virtual environments, and the perception of presence in the evolving landscape of VR and Metaverse research.

2 RELATED WORK

2.1 Avatar Type and Presence in VR

Previous research has shown that the appearance and type of avatars have an impact on cognition and presence [40]. Steed et al. [36] argued the necessity of a self-avatar in VR, as it affects users' cognitive load. Expanding on this work, Pan and Steed [32] conducted further studies and found that the cognitive load of a user is influenced by the type of avatar. Their study had four conditions: no-body avatar, hands-only avatar, full body avatar, and real person. Each participant took part in one of the four conditions and completed two cognitive tasks. The results showed that the memory task performance was significantly better in the full body condition and the real person condition than in the no-body condition. However, the hands-only condition did not differ significantly from the two extremes. This is consistent with recent literature suggesting that partial body expression can induce the full body illusion [19].

Heidicker et al. [16] designed an experiment to study the influence of avatar appearance on social presence. They designed three avatars with different appearances: an avatar with a full body and regular predefined idle animations; an avatar with a full body and allowed one-to-one mapping of user actions to avatar actions; an avatar with a body consisting of just head and hands and allowed one-to-one mapping of user actions to avatar actions. They asked participants in pairs to complete a collaborative task using the same type of avatar in two different rooms. Their results showed no significant difference among the three avatars on social presence. However, an avatar with mapped condition had a higher presence than that with idle animation, and a full body avatar resulted in higher co-presence and behavioral interdependence than avatars with only the head and hands. Avatar appearance in this study concerned the completeness of human avatars and the behavioral realism associated with the human avatar. Based on these two studies, although there were no significant differences between full-body and half-body avatars in cognitive load, task performance, and social presence, full-body avatars were selected for this work, which had shown relatively higher performance in previous studies.

In addition to the comparison between a half-body avatar and a full-body avatar on cognitive load, task performance, and social presence, some studies have suggested important aspects of avatars that affect perceived presence. Freeman et al. [8] conducted a series of interview studies and found that aesthetics, gender [35], ethnicity, and age [1] embodied in virtual avatars affect social interactions in VR. The interviews showed that most participants considered that avatars were an extension of themselves, so they tended to use avatars that were similar to their physical selves, which could give them a much stronger sense of presence. However, some users tended to use non-human avatars or non-human features to enrich their avatar looks. For example, one participant found the use of rabbit ears and rabbit buck teeth socially beneficial as they could make them look cute and friendly. The other participant used a blue bird to represent the self and enjoyed being a completely different species in VR. Participants also reported the use of creative avatars, such as using alien-like creatures. For these users, experiencing a completely different self from real life was one of the main reasons for engaging in social VR. These are related to users' sense of virtual body ownership.

Virtual body ownership refers to the psychological phenomenon in which a person using VR or a similar immersive technology starts to perceive and experience an avatar or digital body as their own [35]. It is a key factor in creating an immersive VR experience, as it enhances users' sense of presence and engagement within the virtual world [4]. Krekhov et al. [20, 21] have studied non-human avatars such as animal avatars. They found that virtual body ownership was also applicable to animals and argued that an illusion of virtual body ownership had a high potential for non-human avatars. In Krekhov et al.'s study, the results showed that although the mapping was different for each animal, participants had no problems controlling the body of non-human structures and performing the tasks. However, it is not clear under what circumstances users would choose to use this kind of non-human avatars and whether it affects perceived presence in VR.

Previous works mainly focused on the inclusion of body parts, such as the head, hands, and the main body, that consist of half-body avatars or full-body avatars. While half-body avatars are easier to configure in technical development and have comparable performance (e.g., [24]), the use of full-body avatars was shown to have better performances in cognitive tasks and have contributed to perceived presence in social environments. Additionally, factors that affect perceived presence, like aesthetics, gender, ethnicity, and age, were also discussed in previous works. These studies have focused on the impact of avatar appearances and most of them have taken the form of human avatars. While some works have explored the use of non-human avatars in VR, few of them were used as virtual self-representations in a social environment. In addition, it is not clear whether users' perceived presence differs when using human-like, anime, and non-human avatars.

2.2 Virtual Avatars and Self-Congruence in VR

The selection and creation of avatars is often one of the first tasks users should do in virtual worlds. Some studies have suggested that people are more likely to choose avatars that are the same gender as themselves or have similar characteristics to themselves [30] and that users are more accepting of avatars that have a similar appearance to the users [41]. Jinsu et al. [33] argued that self-congruence was formed when users first created or chose their avatars. Their experiments showed that the similarity in appearance between users and avatars affected purchase intention in the virtual world, which also influenced their purchase intentions in the real world. Thus, they suggested that a high level of self-congruence is crucial in VR. In addition, the visual similarity of an avatar within an environment together with others also affects the sense of commonality and shared identity [11], which sheds light on the environment-related factors related to perceived self-congruence.

In another study, Unal et al. [38] focused on the congruence of participants' avatars (self-images) and studied their relationship with brand images and purchasing intention. The results showed that participants created avatars that have different personal characteristics from themselves, which are more free and controllable and resemble their ideal selves instead of their real selves. Their study also showed that participants are more attracted to male avatars than female avatars, regardless of their own gender. These previous studies have discussed the self-congruence between users and avatars influenced by different genders and appearances within the field of psychology as well as consumption and marketing. Researchers analyzed the perceived self-congruence between users and avatars and evaluated the purchase intention generated by avatars in the virtual environments. However, how users' perceived self-congruence varies in different environments was not systematically studied in previous works.

In summary, most of the previous studies have focused on different forms of human avatars, such as full-body and half-body avatars, and different appearances of human avatars, such as different clothes, different genders, etc. Yet, there is a growing body of applications that allow different avatar types in social VR experiences. The effects of various avatar types on perceived presence have not been explored in previous work. Also, environments hosting virtual avatars are likely to vary, both physically and socially. Users' choices of avatars and perceived self-congruence within the environments have not been systematically studied, which may yield interesting findings for future design of social VR experiences. This research aims to address these research gaps.

3 SYSTEM IMPLEMENTATION

3.1 Avatars

We configured four different types of avatars in the system: anime avatars, human avatars, animal avatars, and item avatars (see Fig. 2). Before entering the scene, the user can select the avatar by selecting the avatar's profile photo through the ray. There were 4 choices for anime and human avatars (2 females and 2 males) and 2 choices for animal and item avatars. The animal avatars were Judy and Nick from *Zootopia*; the two item avatars were a garbage bin and a camera. Except for item avatars, all other avatars were humanoid using the VRIK plugin[6]. We tested real animal avatars, such as cats and dogs. However, these animal models cannot be rigged to map the movements of the head and four limbs. If they are bound, it will cause deformities in the avatar shape; otherwise, if only the head position is mapped, the avatars will float in the air. Thus, we also opted to implement animal avatars in the humanoid type. Item avatars do not have bones, so they cannot be humanoid. We only mapped it with the head position and fine-tuned the camera position to show a natural viewpoint. Avatars as users' self-representations in VR will move or turn when users push or turn the joysticks on the controllers. The triggers of the two controllers were used to teleport in the virtual environment, and the grips were used to grab virtual objects.

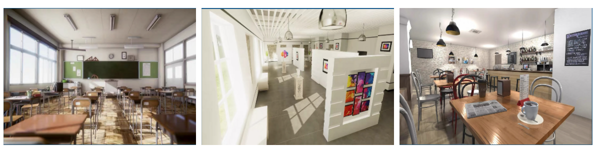
[6]http://www.root-motion.com/finalikdox/html/page16.html

(a) Anime (b) Human (c) Animal (d) Item

Fig. 2. Four types of avatars implemented in the system.

3.2 Environments

Five virtual environments were implemented in the system: a classroom, an art gallery, a café, a street, and a forest. As shown in Fig 3, images of the empty scenes were shown to users when selecting an avatar. Users could not see the type(s) of avatars in the scene until they entered it. In all environments, a mirror was placed at the user's birth point, allowing users to observe their own appearances. Pre-populated avatars in each scene have animations, such as standing, sitting, walking, and chatting. More details about avatars in each environment are shown in Table 1.



(a) Classroom (b) Gallery (c) Café

(d) Street (e) Forest

Fig. 3. Five real-life environments.

3.2.1 Classroom. In the classroom scene, there were only anime type avatars, with student avatars sitting on chairs and a teacher avatar standing at a podium.

3.2.2 Gallery. In the art gallery scene, there were only human avatars, some of which paced back and forth within the scene while others were looking at the paintings.

3.2.3 Café. Within the café scene, there were both anime and human avatars. Some of these avatars were waiting in line to order food, and some were seated and chatting.

3.2.4 Street. The street scene included all four types of avatars. Anime avatars were chatting by the street, human avatars were walking or forming small groups to have conversations, animal avatars, both real ones and humanoid ones, were standing and watching pedestrians, and item avatars like garbage bins were positioned on the roadside.

3.2.5 Forest. The forest scene features an abundance of trees and grasslands, where leaves and grass sway with the breeze. The scene was enriched by the inclusion of diverse real animal models such as horses, cattle, elephants, and rhinoceroses.

3.3 Tasks

Once users selected an avatar for a given environment, they needed to complete a unique task in the environment.

3.3.1 Classroom. In the classroom scene, users were asked to find the teacher and complete the prompted task. They needed first to find the teacher's location and then walk towards the teacher until the dialogue prompt box appeared. The teacher would ask the user to draw a picture on the whiteboard using a marker (see Fig. 4a). Users needed to use the controller to grab the marker (by pressing the grab button on the controller) and move to the whiteboard to complete the drawing tasks.

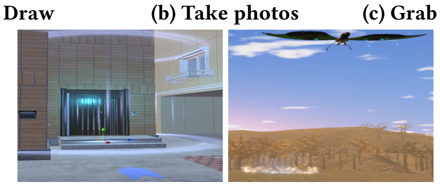
3.3.2 Gallery. In the gallery scene, users were asked to find the camera in the scene and take a picture of their favorite artwork. The camera was near the user's birth point and could be grabbed in hand. The camera included a display that showed an image composition, and the user could take a photo by pressing the trigger button (see Fig. 4b). When the button was pressed, the camera would flash and take a photo.

3.3.3 Café. In the café scene, users were asked to act as a cashier and prepare the drinks for the customer. They needed to talk to the first person in line, who would ask for two drinks. Users then needed to turn around, find the two drinks, and grab and place them on the two glowing circles on the cashier's table (see Fig. 4c), where the customer will pick up the order. There were prompts showing whether the task was performed correctly. The task was completed when both items were placed correctly.

3.3.4 Street. Within the street scene, users needed to talk to the avatar next to them. The avatar would ask the user to walk along the blue arrows on the road and reach the destination (see Fig. 4d).

3.3.5 Forest. Similar to the street scene, in the forest scene, the user needed to find the butterfly in the air and trigger the dialogue box. It would lead the user to reach the destination (see Fig. 4e).

(a) Draw (b) Take photos (c) Grab objects



(d) Street destination (e) Forest destination

Fig. 4. Interaction tasks in each environment.

Table 1. The type and number of pre-populated avatars in the five virtual environments and the task details.

Environment	Anime	Human	Animal	Item	Task
Classroom	9	0	0	0	Draw a picture on the blackboard
Gallery	0	7	0	0	Find the camera and take a picture
Café	5	5	0	0	Serve the correct drinks
Street	6	19	5	1	Follow the arrow and move to the destination
Forest	0	0	9	0	Follow the butterfly and move to the destination

4 RESEARCH DESIGN

4.1 Research Questions and Hypotheses

In this work, we use presence as an inclusive concept that includes both spatial presence and social presence. Our research aims to answer two research questions:

RQ1. *Do avatar types (anime, human, animal, and item) influence users' perceived presence?*
RQ2. *Is there a correlation between users' perceived self-congruence and social presence?*

Based on the related work introduced in Section 2, it was indicated that the appearance and type of avatars have an impact on social presence [8, 16], and the similarity in appearance between users and avatars was positively correlated with self-congruence [33]. Thus, the following hypotheses were proposed and tested:

H1. There is a difference in users' perceived presence when using different avatars.
H2. There is a correlation between users' perceived self-congruence and social presence.

4.2 Controlled Variables: Avatar Type and Environment

4.2.1 Avatar Type. Human-like avatars, such as realistic human and anime avatars, have been widely adopted and studied in previous work. In addition to these two, this research included animal avatars and item avatars to enrich the avatar types. These avatars are seen in recent social applications but are not well understood in the literature.

4.2.2 Environment. We made a total of five virtual environments available to participants: classroom, art gallery, café, street, and forest. The reason for including various environments is to configure different social environments so that we could evaluate perceived self-congruence with different choices of avatars. This also allowed us to understand its relationship with the sense of social presence.

4.3 Measured Variables: Self-Congruence and Presence

4.3.1 Self-Congruence. Self-congruence refers to the appropriateness or compatibility of an avatar within a given virtual environment. It encompasses how well the avatar's appearance, behavior, and characteristics align with the context and setting of the virtual world. This was measured by asking the question: *How well do you think the avatar you chose fits the scene?*

4.3.2 Presence. We evaluated presence from two established questionnaires [15, 39] including six dimensions: spatial presence (SP), overall presence (OP), co-presence (CP), attentional allocation (AA), perceived message understanding (PMU), and perceived behavioral interdependence (PBI). The spatial presence question was a classic measure asking the sense of "being there"; the three degree questions about overall presence asked users about their overall feeling in VR, which has been widely used in the VR community. These were taken from the Slater-Usoh-Steed questionnaire [39]. The remaining four constructs about social presence measures were adapted from the Networked Minds Social Presence Measure [15]. Specifically, co-presence is characterized by the feeling of being in the same space with another person. Attentional allocation addresses the attention the user allocates to and receives from an avatar. Perceived message understanding is the ability of users to understand the message from the avatar. Perceived behavioral interdependence refers to the degree to which user behavior affects and is influenced by the behavior of avatars. Full details are shown in Table 2. Items were rated on a five-point Likert scale (1 = Strongly Disagree / Not at all and 5 = Strongly Agree / Very Much).

4.4 Experimental Design

4.4.1 Setup and System Features. The user study took place in a laboratory environment where the participants had enough space to use a VR device and complete the experiment tasks. The VR device used in the experiment is Meta Quest 2 (see Fig. 5), and the computer was configured with an Intel(R) Core(TM) i9-10900K CPU and the NVIDIA GeForce RTX 2080 Ti GPU. We used Unity (version 2021.3.8f1c1) to develop the VR system, including the setup of five virtual environments and the implementation of the interaction tasks described in Section 3.3 and Table 1.

4.4.2 Experimental Procedure. The study consisted of three parts: 1) briefing, 2) five experimental sessions, and 3) a follow-up interview. At the briefing session, we collected participants' consent to ensure that they understood the purpose of the study and agreed to take part in it. Then, the demographic information of participants was collected, and tutorials were provided for participants to familiarize themselves with the devices and controls. The main study included five experimental sessions, where we followed a Latin Square Design and asked participants to complete tasks in the five virtual environments. In each environment, participants needed to select an avatar they felt like using in the environment and complete the given task. They were invited to evaluate their perceived presence after each experimental session. Upon completing all five sessions, participants were interviewed and asked to

Fig. 5. A participant in the experiment wearing the Meta Quest 2 HMD and two hand-held controllers.

Table 2. Presence questionnaire adapted from [15]. ***Italic items*** **are reverse-coded.**

#	Question
SP	In the computer-generated world, I had a sense of "being there".
OP1	To what extent did you feel you could interact with the person or people you saw/heard?
OP2	To what extent were there times during the experience when the virtual environment was the reality for you?
OP3	During the time of your experience, did you often think to yourself that you were actually in the virtual environment?
CP1	I noticed other avatars.
CP2	Other avatars' presence was obvious to me.
CP3	Other avatars caught my attention.
AA1	*I was easily distracted from avatars when other things were going on.*
AA2	I remained focused on avatars throughout our interaction.
AA3	*Avatars did not receive my full attention.*
PMU1	It was easy to understand avatars.
PMU2	*Understanding avatars was difficult.*
PBI1	My behavior was often in direct response to avatars' behavior.
PBI2	I reciprocated the avatars' actions.
PBI3	My behavior was closely tied to avatars' behavior.

talk about their choice of avatars and their corresponding experience in the five environments. The experience of each scene lasted about 5 minutes, and the whole experiment lasted 30-40 minutes for each participant. Participants were instructed that they could stop the experiment immediately if they felt uncomfortable in the experiment, such as vertigo, headache, nausea, etc. Yet there was no such instance. The study was identified as low-risk research and has been approved by our university's Ethics Committee prior to data collection.

4.5 Participants

Eighteen participants (8 females and 10 males) voluntarily signed up for the study. Participants were undergraduate and graduate students aged between 20 and 25 ($M = 22.4, SD = 1.3$). We asked participants to rate their familiarity with 3D graphics, virtual avatars, and VR systems on a scale from 1 (Not at all familiar) to 5 (Extremely familiar). Participants in our sample was moderately familiarity with 3D graphics ($M = 3.44, SD = 0.98$), virtual avatar ($M = 3.00, SD = 0.84$), and VR system ($M = 3.28, SD = 1.07$). Concerning their prior experience with VR, two participants reported that they had not used VR before.

5 RESULTS

In this section, we report our findings in response to the research questions. We adopted non-parametric statistical tests, given that the questionnaire data were measured on Likert scales. All analyses were conducted using IBM SPSS Statistics.

5.1 Avatar Types and Presence

Fig. 6 illustrates user responses to subcategories of the presence questionnaire using different types of avatars. We calculated the mean value of the perceived presence of the same avatar type in five environments and conducted a comparative analysis.

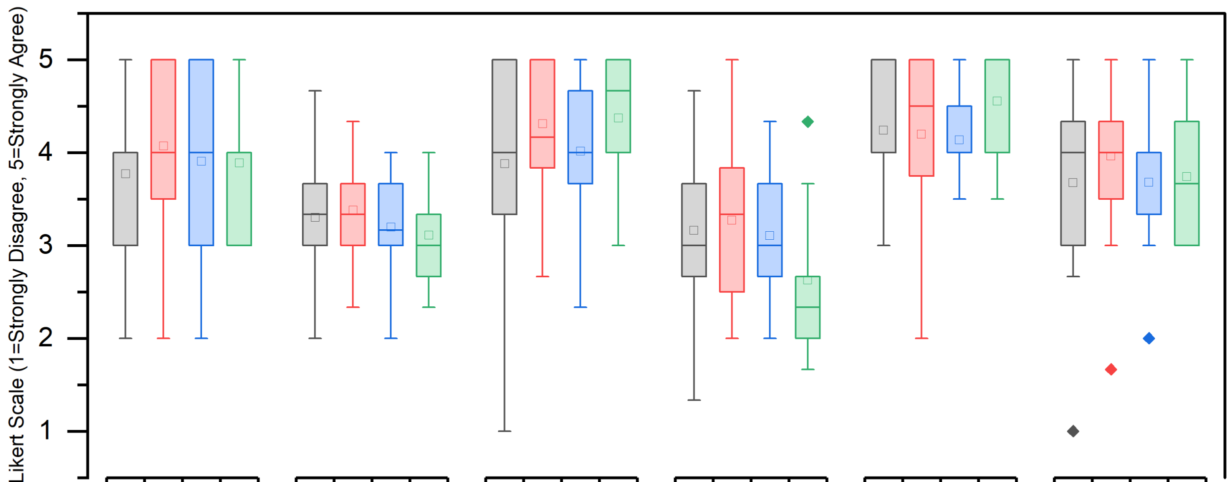
5.1.1 Spatial Presence (SP). . A Kruskal-Wallis test showed no significant difference in spatial presence between different types of avatars, $\chi^2(3) = 2.012, p = 0.57$, with a mean rank spatial presence of 41.15 for anime avatars, 49.96 for human avatars, 46.73 for animal avatars and 43.61 for item avatars.

5.1.2 Overall Presence (OP). . A Kruskal-Wallis test showed no significant difference in overall presence between different types of avatars, $\chi^2(3) = 2.033, p = 0.566$, with a mean rank overall presence of 46.06 for anime avatars, 49.88 for human avatars, 42.43 for animal avatars and 37.44 for item avatars.

5.1.3 Co-Presence (CP). . A Kruskal-Wallis test showed no significant difference in co-presence between different types of avatars, $\chi^2(3) = 3,369, p = 0.338$, with a mean rank co-presence of 40.71 for anime avatars, 50.91 for human avatars, 42.43 for animal avatars and 52.67 for item avatars.

5.1.4 Attentional Allocation (AA). . A Kruskal-Wallis test showed no significant difference in attentional allocation between different types of avatars, $\chi^2(3) = 5.443, p = 0.142$, with a mean rank attentional allocation of 49.79 for anime avatars, 50.32 for human avatars, 35.64 for animal avatars and 39.83 for item avatars.

5.1.5 Perceived Message Understanding (PMU). . A Kruskal-Wallis test showed no significant difference in perceived message understanding between different types of avatars, $\chi^2(3) = 3.171, p = 0.366$, with a mean rank perceived message understanding of 45.26 for anime avatars, 46.63 for human avatars, 39.66 for animal avatars and 57.11 for item avatars.



Avatar	AN	HM	AM	IT	AN	HM	AM	IT
Mean	3.77	4.07	3.91	3.89	3.30	3.38	3.20	3.11
Std.Dev	0.80	0.86	1.02	0.78	0.58	0.55	0.56	0.53
Subcategories	SP				OP			

Avatar	AN	HM	AM	IT	AN	HM	AM	IT
Mean	3.88	4.31	4.02	4.37	3.16	3.27	3.11	2.63
Std.Dev	1.05	0.68	0.79	0.70	0.80	0.80	0.68	0.86
Subcategories	CP				AA			

Avatar	AN	HM	AM	IT	AN	HM	AM	IT
Mean	4.24	4.20	4.14	4.56	3.68	3.96	3.68	3.74
Std.Dev	0.67	0.89	0.47	0.68	0.82	0.75	0.66	0.72
Subcategories	PMU				PBI			

Fig. 6. Boxplots and tables showing the detailed results of the presence questionnaire using different avatars. AN = anime, HM = human, AM = animal, IT = item; SP = spatial presence, OP = overall presence, CP = co-presence, AA = attentional allocation, PMU = perceived message understanding, PBI = perceived behavioral interdependence.

5.1.6 Perceived Behavioral Interdependence (PBI). A Kruskal-Wallis test showed no significant difference in perceived behavioral interdependence between different types of avatars, $\chi^2(3) = 1.972, p = 0.578$, with a mean rank perceived behavioral interdependence of 44.65 for anime avatars, 50.32 for human avatars, 43.77 for animal avatars and 37.67 for item avatars.

5.1.7 Summary. Since there was no significant difference between different types of avatars in all subcategories of presence, we conclude that there was no significant difference in presence when using different types of avatars. Hence, **H1** is not supported.

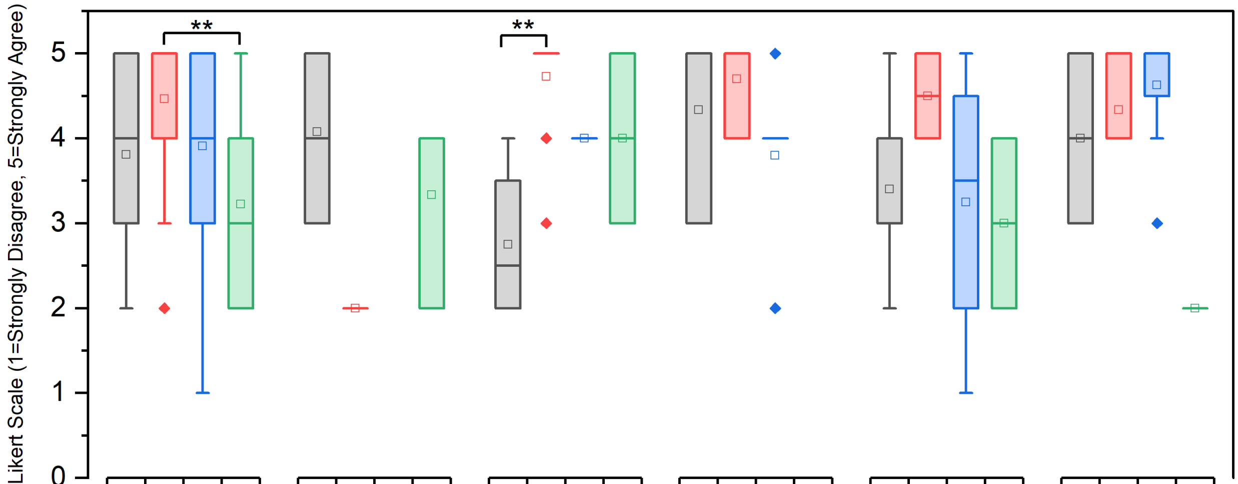
5.2 Self-congruence and Social Presence

5.2.1 Effects of Avatar Types on Self-Congruence. Fig. 7 illustrates the perceived self-congruence using different types of avatars in the five environments. Kruskal-Wallis tests showed a significant difference on users' perceived self-congruence among different types of avatars in different environments, $\chi^2(3) = 11.498, p = 0.009$. A significant difference was found between human avatars (*Mean Rank* = 56.88) and item avatars (*Mean Rank* = 28.11). Within the gallery scene, there was a significant difference on users' perceived self-congruence among different types of avatars, $\chi^2(3) = 9.522, p = 0.023$. A higher level of self-congruence was reported for the human avatars than for anime avatars.

5.2.2 Effects of Environments on Social Presence. Fig. 8 illustrates the perceived presence using different types of avatars in the five environments. We calculated the mean value of the four subcategories of social presence measures and conducted a comparative analysis. Kruskal-Wallis tests showed no significant difference in social presence between different types of avatars in classroom ($\chi^2(2) = 2.162, p = 0.339$), gallery ($\chi^2(3) = 4.456, p = 0.216$), café ($\chi^2(2) = 2.169, p = 0.338$), street ($\chi^2(3) = 2.767, p = 0.429$), or forest ($\chi^2(3) = 2.198, p = 0.532$). An overall comparison aggregating results in the five environments showed no significant difference in social presence, $\chi^2(3) = 2.960, p = 0.398$, with a mean rank presence of 41.65 for anime avatars, 52.43 for human avatars, 42.48 for animal avatars and 44.61 for item avatars.

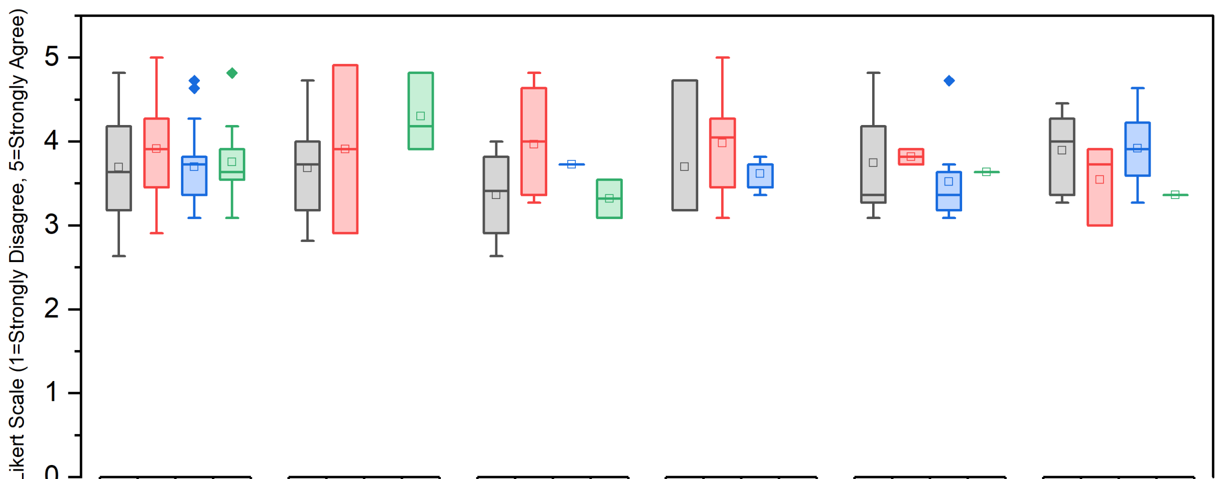
5.2.3 Effects of Avatar Fit on Self-Congruence and Social Presence. Given that the environments have pre-populated avatars, we define a variable *Fit*, where $Fit = 1$ when users' chosen avatar type was included in the scene (e.g., anime avatar in the classroom), otherwise $Fit = 0$ (e.g., anime avatar in the gallery, see Table 1 for more details). Fig. 9 shows the perceived self-congruence and social presence in relation to avatar choices. A Mann-Whitney test showed a significant difference on users' perceived self-congruence between avatar choices ($U = 510, p < 0.001$). Avatars that fit the environment (*Mean Rank* = 54.69) had a significantly higher level of perceived self-congruence compared to avatars that did not fit the environment (*Mean Rank* = 32.92). In addition, a Mann-Whitney test showed a significant difference in attentional allocation between avatars choices ($U = 642.5, p = 0.004$). Avatars that fit the environment (*Mean Rank* = 52.14) had a significantly greater attentional allocation compared to avatars that did not fit the environment (*Mean Rank* = 36.41). There was no statistically significant difference in the other three subcategories of social presence.

5.2.4 Correlation Analysis. A Spearman's correlation analysis showed a statistically significant positive correlation between users' perceived self-congruence and the overall social presence ($r = 0.260, p = 0.013$). Specifically, its correlation with co-presence ($r = 0.130, p =$



Environment	Avatar	Count	Mean	Std.Dev
Overall	AN	31	3.81	1.05
Overall	HM	28	4.46	0.88
Overall	AM	22	3.91	1.23
Overall	IT	9	3.22	1.09
Classroom	AN	13	4.08	0.95
Classroom	HM	2	2.00	0.00
Classroom	AM	0		
Classroom	IT	3	3.33	1.15
Gallery	AN	4	2.75	0.96
Gallery	HM	11	4.73	0.65
Gallery	AM	1	4.00	
Gallery	IT	2	4.00	1.41
Cafe	AN	3	4.33	1.15
Cafe	HM	10	4.70	0.48
Cafe	AM	5	3.80	1.10
Cafe	IT	0		
Street	AN	5	3.40	1.14
Street	HM	2	4.50	0.71
Street	AM	8	3.25	1.49
Street	IT	3	3.00	1.00
Forest	AN	6	4.00	0.89
Forest	HM	3	4.33	0.58
Forest	AM	8	4.63	0.74
Forest	IT	1	2.00	

Fig. 7. Boxplots and tables showing the perceived self-congruence using different avatars in the five environments. AN = anime, HM = human, AM = animal, IT = item.



Environment	Avatar	Count	Mean	Std.Dev
Overall	AN	31	3.70	0.61
Overall	HM	28	3.91	0.58
Overall	AM	22	3.70	0.45
Overall	IT	9	3.76	0.50
Classroom	AN	13	3.69	0.61
Classroom	HM	2	3.91	1.41
Classroom	AM	0		
Classroom	IT	3	4.30	0.47
Gallery	AN	4	3.36	0.59
Gallery	HM	11	3.97	0.57
Gallery	AM	1	3.73	
Gallery	IT	2	3.32	0.32
Cafe	AN	3	3.70	0.89
Cafe	HM	10	3.98	0.55
Cafe	AM	5	3.62	0.20
Cafe	IT	0		
Street	AN	5	3.75	0.73
Street	HM	2	3.82	0.13
Street	AM	8	3.52	0.53
Street	IT	3	3.64	0.00
Forest	AN	6	3.89	0.49
Forest	HM	3	3.55	0.48
Forest	AM	8	3.92	0.45
Forest	IT	1	3.36	

Fig. 8. Boxplots and tables showing the summary results of the social presence questionnaire using different avatars in different environments. AN = anime, HM = human, AM = animal, IT = item.

0.221) and attentional allocation ($r = 0.138, p = 0.194$) was insignificant, but significant correlations were found between self-congruence and perceived message understanding ($r = 0.314, p = 0.003$) and perceived behavioral interdependence ($r = 0.316, p = 0.002$). These results show support for **H2**.

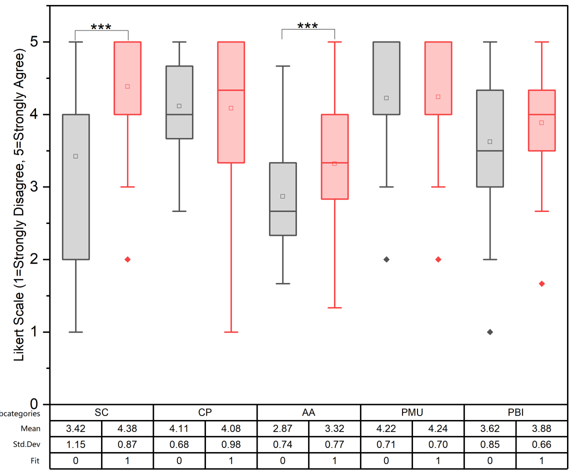


Fig. 9. Boxplots and tables showing users' perceived self-congruence and social presence in relation to avatar choices. SC = self-congruence, CP = co-presence, AA = attentional allocation, PMU = perceived message understanding, PBI = perceived behavioral interdependence.

Table 3. Frequencies of the chosen avatar types in each environment, and users' perceived self-congruence. Avatar types included in the environment are highlighted in bold.

Environment	Avatar type	N	Self-congruence Mean (SD)
Classroom	**Anime**	13	4.08 (0.95)
Classroom	Human	2	2 (0)
Classroom	Animal	0	N/A
Classroom	Item	3	3.33 (1.15)
Gallery	Anime	4	2.75 (0.98)
Gallery	**Human**	11	4.73 (0.65)
Gallery	Animal	1	4 (N/A)
Gallery	Item	2	4 (1.41)
Café	**Anime**	3	4.33 (1.15)
Café	**Human**	10	4.7 (0.48)
Café	Animal	5	3.8 (1.10)
Café	Item	0	N/A
Street	**Anime**	5	3.4 (1.14)
Street	**Human**	2	4.5 (0.7)
Street	**Animal**	8	3.25 (1.48)
Street	**Item**	3	3 (1)
Forest	Anime	6	4 (0.89)
Forest	Human	3	4.33 (0.58)
Forest	**Animal**	8	4.625 (0.74)
Forest	Item	1	2 (N/A)

5.3 User Preferences and Interview Findings

We calculated the frequency of participants' avatar selections and labeled their selections as 1 (selected) and 0 (not selected) to analyze user preferences. Table 3 shows the summary results.

In the classroom scene, a Friedman test showed a statistically significant difference in the choices of avatar types, $\chi^2(3) = 22.444, p < 0.001$. Pair-wise comparisons showed that user preferences were significantly higher for anime avatars (*Mean Rank* = 3.44) than human avatars (*Mean Rank* = 2.22), animal avatars (*Mean Rank* = 2), and item avatars (*Mean Rank* = 2.33). Participants preferred anime avatars and had the highest self-congruence ($M = 4.08, SD = 0.95$) using it in the classroom. In the gallery scene, a Friedman test showed a statistically significant difference in choices of avatar types, $\chi^2(3) = 13.556, p = 0.004$. Pair-wise comparisons showed that user preferences were significantly higher for human avatars (*Mean Rank* = 3.22) than animal avatars (*Mean Rank* = 2.11) and item avatars (*Mean Rank* = 2.22). Participants preferred human avatars and had the highest self-congruence ($M = 4.73, SD = 0.65$) using it in the gallery. In the café scene, a Friedman test showed a statistically significant difference in choices of avatar types, $\chi^2(3) = 11.778, p = 0.008$. A significant difference was found between human avatars (*Mean Rank* = 3.11) and item avatars (*Mean Rank* = 2.00). Participants preferred human avatars and had the highest self-congruence ($M = 4,7, SD = 0.48$) using it in the café. There was no statistically significant difference in choices of avatar types in the scene ($\chi^2(3) = 4.667, p = 0.198$). Similar results were found for the forest scene ($\chi^2(3) = 6.444, p = 0.092$).

Participants' interview comments were analyzed using Theme-Based Content Analysis [29]. The results revealed four factors that may influence their perceived presence and VR experience, including virtual object interactions (N=9), movements in VR (N=7), social interactions with other users (N=4), and visual fidelity (N=2). Detailed findings will be discussed in the next section.

6 DISCUSSION

Our observations in the study showed that participants reported an overall positive experience and a high level of perceived presence in the VR environments. They acknowledged the feeling of actually being in the virtual scenes and being the avatar that they chose as their self-representations. In particular, participants showed surprise and excitement to be an animal and an item, namely forms that are otherwise impossible in the real world. In this section, we discuss our findings on the research questions.

6.1 Avatar Type and Presence

In response to **RQ1**, our results showed no significant effects of avatar types on perceived presence. Thus, **H1** was not supported. This finding is inconsistent with previous work [8], where significant differences were found in social presence among different avatars. However, our results should be interpreted by taking into account the experimental setting. While previous work mainly compared different avatar types within a single environment, our study investigated multiple environments, simulating the actual use of avatars in the social virtual world. It is likely that the diverse environments and the avatars within the environments reduced the direct impact of the avatar type itself. Still, our study simulates real-life settings, and the findings shed light on the diverse design space of virtual avatars in VR and the Metaverse.

In addition to the effects of avatar types, participants commented on other aspects that have influenced their experience in VR during the interview, which can be grouped into four themes. The first one is about virtual object interactions. For example, P5 commented that *"there could be more interactions with the animals (in the forest scene)"* and P16 also expressed a similar feeling, *"I wish I could talk to the butterfly."* Participants also reported on perceived affordances in VR. For example, P18 commented that *"it would be nice if I could save the picture I took (in the gallery) to an album."* Supporting diverse virtual object interactions is likely to positively affect users' sense of presence. Users also commented on the movements in VR. P15 mentioned that *"I want to have more kinds of movements in VR, such as jumping and flying"*, and P14 mentioned that *"I prefer the rotation of views to be continuous but not discrete"*. As investigated in previous works [44], locomotion is one of the most important factors affecting VR experience. Finding appropriate locomotion techniques will contribute to users' perceived presence. Aside from the techniques, designers should also consider user characteristics such as height. For example, P4 mentioned that *"somehow I felt I was too short in the street scene and my feet were nearly under the ground"*. This is a limitation of using full-body avatars without tracking users' feet. Aside from the object interactions and movements, participants also commented on their social interactions. For example, *"the experience will be more realistic if the avatars have voice and I could talk to them"* (P7). Users value behavioral realism and realistic interactions, which are important criteria for improving social presence in VR [13, 25]. Future work may consider incorporating large language models to enrich social interactions, either text or voice dialogues. Finally, two participants suggested increasing the fidelity of the environment and avatars. P13 commented that *"the 3D models (of the environment) could be more realistic"* and P10 mentioned specifically the fidelity of avatar costumes and suggested including cloth simulation for the sleeves. Some recent works have been done to simulate realistic complex costumes, such as the AR try-on system [10]. These factors (object interactions, movements in VR, social interactions, and visual fidelity) were found to be closely related to users' perceived presence and could implicate the future design of immersive social VR experiences.

6.2 Self-Congruence and Social Presence

In response to **RQ2**, we found significant correlations between perceived self-congruence and social presence, which supported **H2**. We found that overall, users perceived the highest self-congruence using human avatars and the least self-congruence using item avatars. Anime and animal avatars were rated in the middle and did not vary significantly from the two extremes. Perceived self-congruence and attentional allocation were found to be significantly higher when users chose an avatar type that was the same as one of the pre-populated avatars. In addition, we found that users tended to choose human avatars (human and anime) over non-human avatars (animal and item). This finding is in line with Freeman and Maloney's findings [8] that users have a high demand to construct self-presentations that closely resemble one's physical self. Nevertheless, for those who chose animal and item avatars, their perceived presence was not significantly worse. This indicates the feasibility of further incorporating non-human avatars in social experiences, such as games [20] and social gatherings [9].

One important takeaway from our study is that the avatar type alone does not significantly influence perceived presence, but its similarity with other avatars in the environment is very likely to influence perceived self-congruence, which was shown to be an important indicator of perceived social presence. It is particularly true that if an environment has a consistent type of avatar (such as a classroom and gallery in our study). The perceived self-congruence was the highest when using an avatar type consistent with those in the scene, and using avatars of a different type will negatively influence the perceived self-congruence and social presence. In addition, we also observed similar findings as reported in [25] that users expected a high degree of social interactivity with avatars.

6.3 Limitations and Future Work

Our study has some limitations. First, we provided a limited number of avatars for participants to choose from, instead of letting them customize their own avatars. While we considered this option, we tried to simplify the experimental design and avoid introducing new variations, given that the key variable we investigated was avatar type but not avatar appearance. It is also worth noting that the animal avatars used in our study are humanoid, which differs from the real animal avatars used in [21] and [20]. Second, users' interactions in the system were with computer agents via dialogues but not with real people. It ensures consistent experiment settings but lacks fidelity in communication. Perhaps the same experiment can be duplicated in an online communication platform such as VRChat, and investigate if interacting with real avatars yields the same results. Third, the system only implemented interactions related to tasks in the environments. Participants suggested additional features, such as voice chats and dialogues, which are likely to further improve user experience and social interactions in VR. In addition, our study allowed users to choose from different avatars, thus resulting in different sample sizes in conditions. While the sample sizes for anime (N=31), human (N=28), and animal (N=22) avatars were reasonable, the sample size for items was relatively small (N=9). This limitation in the statistical analysis should be noted. Fourth, previous work on presence has different focuses on spatial presence and social presence. Aside from the measures we used, there are other measures of presence (e.g., [3, 17, 23, 26, 42]) and different scales (5-Likert and 7-Likert) were used in previous work. This may hinder the direct comparability with other works. A unified measure of presence is needed in future work to facilitate more meaningful comparisons. Fifth, the equipment used in this study was limited. We used a standard VR headset and controllers (Meta Quest 2) without any other motion capture device. Thus, only the head and hands can be identified and tracked. The walking was simulated in animations. Future work could consider adopting motion capture technology, such as inverse kinematics solvers, to increase fidelity in movements and better map user postures to the avatar skeletons [5]. Finally, our participants were university students recruited from a local campus. It is not representative of the general population, and it is necessary to increase the sample size and broaden the demographic diversity. As previous work has shown that animal avatars might be more appealing to children [21], it would be interesting to expand the work to more diverse user groups, such as children and the elderly.

7 CONCLUSION

In this study, an experiment with four avatar types and five virtual environments was conducted to investigate the relationship between avatar types, perceived self-congruence, and perceived presence. We did not find significant differences in perceived presence when using different avatars but found a significant correlation between users' perceived level of self-congruence and social presence. Specifically, our analysis showed that participants perceived significantly higher levels of self-congruence in environments containing avatars of the same type as their self-representations. Moreover, avatars that fit the environment were able to significantly increase users' attentional allocation. In terms of user preference for avatar choices, our results showed that participants preferred anime avatars in the classroom, and human avatars in the gallery and the café scenes. The preferred avatar types are also the main avatar types in the scene, indicating that users tend to use an avatar type similar to others', which is likely to contribute to higher levels of perceived self-congruence and social presence. Finally, we present our qualitative analysis findings and highlight that virtual object interactions, movements in VR, social interactions, and visual fidelity were identified by participants as important factors that have impacts on perceived presence.

ACKNOWLEDGMENTS

The authors would like to thank all participants for their time and support. This work is supported by the National Natural Science Foundation of China (62207022), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (22KJB520038), and Xi'an Jiaotong-Liverpool University (RDF-20-02-47).

REFERENCES

[1] Domna Banakou, Raphaela Groten, and Mel Slater. 2013. Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences* 110, 31 (2013), 12846–12851. https://doi.org/10.1073/pnas.1306779110

[2] Alex Barrett, Austin Pack, Diego Monteiro, and Hai-Ning Liang. 2023. Exploring the influence of audience familiarity on speaker anxiety and performance in virtual reality and real-life presentation contexts. *Behaviour & Information Technology* 0, 0 (2023), 1–13. https://doi.org/10.1080/0144929X.2023.2186145

[3] Frank Biocca, Chad Harms, and Judee K. Burgoon. 2003. Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence: Teleoperators and Virtual Environments* 12, 5 (2003), 456–480. https://doi.org/10.1162/105474603322761270

[4] Matthew Botvinick and Jonathan Cohen. 1998. Rubber hands 'feel' touch that eyes see. *Nature* 391, 6669 (01 Feb 1998), 756–756. https://doi.org/10.1038/35784

[5] Samuel R Buss. 2004. Introduction to inverse kinematics with jacobian transpose, pseudoinverse and damped least squares methods. *IEEE Journal of Robotics and Automation* 17, 1-19 (2004), 16.

[6] Lei Chen, Hai-Ning Liang, Feiyu Lu, Jialin Wang, Wenjun Chen, and Yong Yue. 2021. Effect of Collaboration Mode and Position Arrangement on Immersive Analytics Tasks in Virtual Reality: A Pilot Study. *Applied Sciences* 11, 21 (2021). https://doi.org/10.3390/app112110473

[7] Cordelia Erickson-Davis, Tanya M. Luhrmann, Lianne M. Kurina, Kara Weisman, Naomi Cornman, Anna Corwin, and Jeremy Bailenson. 2021. The sense of presence: lessons from virtual reality. *Religion, Brain & Behavior* 11, 3 (2021), 335–351. https://doi.org/10.1080/2153599X.2021.1953573

[8] Guo Freeman and Divine Maloney. 2021. Body, Avatar, and Me: The Presentation and Perception of Self in Social Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 239 (jan 2021), 27 pages. https://doi.org/10.1145/3432938

[9] Guo Freeman, Samaneh Zamanifard, Divine Maloney, and Alexandra Adkins. 2020. My Body, My Avatar: How People Perceive Their Avatars in Social Virtual Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382923

[10] Yukun Fu and Yue Li. 2023. Hanfu AR: Digital Twins of Traditional Chinese Costumes for Augmented Reality Try-On Systems. In *2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)*. 1465–1470. https://doi.org/10.1109/COMPSAC57700.2023.00225

[11] Samuel L Gaertner, John F Dovidio, Jason A Nier, Brenda S Banker, Christine M Ward, Melissa Houlette, and Stephenie Loux. 2000. The Common Ingroup Identity Model for reducing intergroup bias: Progress and challenges. (2000).

[12] Zixuan Guo, Wenge Xu, Jialin Zhang, Hongyu Wang, Cheng-Hung Lo, and Hai-Ning Liang. 2023. Who's Watching Me?: Exploring the Impact of Audience Familiarity on Player Performance, Experience, and Exertion in Virtual Reality Exergames. In *2023 International Symposium on Mixed and Augmented Reality (ISMAR)*.

[13] Wen Hai, Nisha Jain, Andrzej Wydra, Nadia Magnenat Thalmann, and Daniel Thalmann. 2018. Increasing the Feeling of Social Presence by Incorporating Realistic Interactions in Multi-Party VR. In *Proceedings of the 31st International Conference on Computer Animation and Social Agents* (Beijing, China) *(CASA 2018)*. Association for Computing Machinery, New York, NY, USA, 7–10. https://doi.org/10.1145/3205326.3205345

[14] Ilona Halim, Lehan Stemmet, Sylvia Hach, Richard Porter, Hai-Ning Liang, Atiyeh Vaezipour, Julie D Henry, and Nilufar Baghaei. 2023. Individualized Virtual Reality for Increasing Self-Compassion: Evaluation Study. *JMIR Mental Health* 10 (2 Oct 2023), e47617. https://doi.org/10.2196/47617

[15] Chad Harms and Frank Biocca. 2004. Internal consistency and reliability of the networked minds measure of social presence. In *Seventh annual international workshop: Presence*, Vol. 2004. Universidad Politecnica de Valencia Valencia, Spain.

[16] Paul Heidicker, Eike Langbehn, and Frank Steinicke. 2017. Influence of avatar appearance on presence in social VR. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 233–234. https://doi.org/10.1109/3DUI.2017.7893357

[17] Wijnand Ijsselsteijn and Giuseppe Riva. 2003. Being There : The experience of presence in mediated environments. *Being There: Concepts, effects and measurement of user presence in synthetic environments* (2003), 14. https://doi.org/citeulike-article-id:4444927

[18] Dong Woo Ko and Jihye Park. 2021. I am you, you are me: game character congruence with the ideal self. *Internet Research* 31, 2 (2021), 613–634.

[19] Ryota Kondo, Maki Sugimoto, Kouta Minamizawa, Takayuki Hoshi, Masahiko Inami, and Michiteru Kitazaki. 2018. Illusory body ownership of an invisible body interpolated between virtual hands and feet via visual-motor synchronicity. *Scientific Reports* 8, 1 (15 May 2018), 7541. https://doi.org/10.1038/s41598-018-25951-2

[20] Andrey Krekhov, Sebastian Cmentowski, Katharina Emmerich, and Jens Krüger. 2019. Beyond Human: Animals as an Escape from Stereotype Avatars in Virtual Reality Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 439–451. https://doi.org/10.1145/3311350.3347172

[21] Andrey Krekhov, Sebastian Cmentowski, and Jens Krüger. 2018. VR Animals: Surreal Body Ownership in Virtual Reality Games. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (Melbourne, VIC, Australia) *(CHI PLAY '18 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 503–511. https://doi.org/10.1145/3270316.3271531

[22] Kwan Min Lee. 2004. Presence, explicated. *Communication theory* 14, 1 (2004), 27–50.

[23] Jane Lessiter, Jonathan Freeman, Edmund Keogh, and Jules Davidoff. 2001. A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence* 10, 3 (2001), 282–297. https://doi.org/10.1162/105474601300343612

[24] Yue Li, Eugene Ch'ng, Sue Cobb, and Simon See. 2019. Presence and Communication in Hybrid Virtual and Augmented Reality Environments. *Presence: Teleoperators and Virtual Environments* 28 (01 2019), 29–52. https://doi.org/10.1162/pres_a_00340

[25] Yue Li, Eugene Ch'ng, and Sue Cobb. 2023. Factors Influencing Engagement in Hybrid Virtual and Augmented Reality. *ACM Trans. Comput.-Hum. Interact.* 30, 4, Article 65 (sep 2023), 27 pages. https://doi.org/10.1145/3589952

[26] Matthew Lombard and Theresa Ditton. 1997. At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication* 3, 2 (1997), 0. https://doi.org/10.1111/j.1083-6101.1997.tb00072.x

[27] Diego Monteiro, Hai-Ning Liang, Jialin Wang, Luhan Wang, Xian Wang, and Yong Yue. 2018. Evaluating the Effects of a Cartoon-Like Character with Emotions on Users' Behaviour within Virtual Reality Environments. In *2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. 229–236. https://doi.org/10.1109/AIVR.2018.00053

[28] Diego Monteiro, Airong Wang, Luhan Wang, Hongji Li, Alex Barrett, Austin Pack, and Hai-Ning Liang. 2023. Effects of audience familiarity on anxiety in a virtual reality public speaking training tool. *Universal Access in the Information Society* (2023). https://doi.org/10.1007/s10209-023-00985-0

[29] Helen Neale and Sarah Nichols. 2001. Theme-based content analysis: a flexible method for virtual environment evaluation. *International Journal of Human-Computer Studies* 55, 2 (2001), 167–189. https://doi.org/10.1006/ijhc.2001.0475

[30] Kristine L. Nowak and Christian Rauh. 2017. The Influence of the Avatar on Online Perceptions of Anthropomorphism, Androgyny, Credibility, Homophily, and Attraction. *Journal of Computer-Mediated Communication* 11, 1 (07 2017), 153–178. https://doi.org/10.1111/j.1083-6101.2006.tb00308.x

[31] Catherine S Oh, Jeremy N Bailenson, and Gregory F Welch. 2018. A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI* 5 (2018), 409295.

[32] Ye Pan and Anthony Steed. 2019. Avatar Type Affects Performance of Cognitive Tasks in Virtual Reality. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 6, 4 pages. https://doi.org/10.1145/3359996.3364270

[33] Jinsu Park and Naeun (Lauren) Kim. 2023. Examining self-congruence between user and avatar in purchasing behavior from the metaverse to the real world. *Journal of Global Fashion Marketing* 0, 0 (2023), 1–16. https://doi.org/10.1080/20932685.2023.2180768

[34] Daniel Roth, Jean-Luc Lugrin, Dmitri Galakhov, Arvid Hofmann, Gary Bente, Marc Erich Latoschik, and Arnulph Fuhrmann. 2016. Avatar realism and social interaction quality in virtual reality. In *2016 IEEE virtual reality (VR)*. IEEE, 277–278.

[35] Mel Slater, Bernhard Spanlang, Maria V. Sanchez-Vives, and Olaf Blanke. 2010. First Person Experience of Body Transfer in Virtual Reality. *PLOS ONE* 5, 5 (05 2010), 1–9. https://doi.org/10.1371/journal.pone.0010564

[36] Anthony Steed, Ye Pan, Fiona Zisch, and William Steptoe. 2016. The impact of a self-avatar on cognitive load in immersive virtual reality. In *2016 IEEE Virtual Reality (VR)*. 67–76. https://doi.org/10.1109/VR.2016.7504689

[37] Anthony Steed and Ralph Schroeder. 2015. *Collaboration in Immersive and Non-immersive Virtual Environments*. Springer International Publishing, Cham, 263–282. https://doi.org/10.1007/978-3-319-10190-3_11

[38] Sevtap Unal, Tevfik Dalgic, and Ezgi Akar. 2018. How avatars help enhancing self-image congruence. *International Journal of Internet Marketing and Advertising* 12, 4 (2018), 374–395.

[39] Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. 2000. Using Presence Questionnaires in Reality. *Presence: Teleoperators and Virtual Environments* 9 (04 2000). https://doi.org/10.1162/105474600566989

[40] Liu Wang, Mengjie Huang, Rui Yang, Hai-Ning Liang, Ji Han, and Ying Sun. 2023. Survey of Movement Reproduction in Immersive Virtual Rehabilitation. *IEEE Transactions on Visualization and Computer Graphics* 29, 4 (2023), 2184–2202. https://doi.org/10.1109/TVCG.2022.3142198

[41] Kevin D. Williams. 2010. The Effects of Homophily, Identification, and Violent Video Games on Players. *Mass Communication and Society* 14, 1 (2010), 3–24. https://doi.org/10.1080/15205430903359701

[42] Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* (aug 1998), 225–240. http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2233447

[43] Jingjing Zhang, Mengjie Huang, Rui Yang, Yiqi Wang, Xiaohang Tang, Ji Han, and Hai-Ning Liang. 2023. Understanding the effects of hand design on embodiment in virtual reality. *AI EDAM* 37 (2023), e10. https://doi.org/10.1017/S0890060423000045

[44] Ziyue Zhao, Yue Li, Lingyun Yu, and Hai-Ning Liang. 2023. TeleSteer: Combining Discrete and Continuous Locomotion Techniques in Virtual Reality. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 755–756. https://doi.org/10.1109/VRW58643.2023.00220