Becoming An Animal? Exploring Proteus Effect Based on Human-avatar Hand Gesture Consistency

Tangjun Qu
Shandong University, Shandong, China

Junjie Wang
Shandong University, Shandong, China

Yongjiu Lin
Shandong University, Shandong, China

Juan Liu
Shandong University, Shandong, China

Chao Zhou*
Institute of Software Chinese Acadamy of Sciences, Beijing, China

Baiqiao Zhang
Shandong University, Shandong, China

Kaiyuan Jiang
Shandong University, Shandong, China

Yulong Bian†
Shandong University, Shandong, China

Abstract

Human cognition and behavior can be unconsciously affected by personal avatars in the virtual world, a phenomenon known as the Proteus Effect. When using first-person non-human avatars, the characteristics of virtual hands may also induce relevant cognitive and even behavioral patterns in real human hands. Therefore, evaluating human-avatar gesture consistency may be a potentially effective method for objectively assessing the Proteus Effect when using non-human avatars. To explore this question, we first created human and non-human avatars, including three animals. Then, we constructed a dataset of hand gestures and trained a model to dynamically recognize real-hand gestures that were consistent with corresponding avatar hands. Next, we designed a virtual reality experimental task involving grasping objects with intuitive gestures and performed a 2 (avatar type: human/non-human) * 2 (virtual hand: presence/absence of spontaneous movement) within-subject experiment to examine the effects of avatar characteristics on self-illusion and human-avatar gesture consistency. The results showed that participants performed a significantly larger percentage of gestures that were consistent with their currently used avatars. Additionally, participants did experience self-illusion when using non-human avatars, although the levels were significantly lower than those when using human avatars. Therefore, self-illusion may serve as a perceptual antecedent of the Proteus Effect, even with non-human avatars, inadvertently altering the behavioral gestures of their real hands. In conclusion, detecting human-avatar gesture consistency can help evaluate the Proteus Effect.

Index Terms: Proteus Effect, self-illusion, virtual hands, non-human avatar.

1 Introduction

Avatar serves as a digital representation of the user, allowing them to engage in actions and immerse themselves in virtual spaces. Yee et al. discovered that human perception, attitude and behavior in virtual world would be affected by the characteristics of avatar, which is known as the Proteus Effect [36]. Researchers have verified the Proteus Effect in various virtual activities [17, 9, 11, 6], noting a significant correlation with immersion [30, 24]. Virtual reality (VR) designers purposefully design avatars with specific functions to enhance user interaction and performance in virtual environments [14]. The extent of the Proteus Effect has been regarded as a psychological phenomenon for evaluating user experience in VR [15]. Therefore, assessing this effect not only aids designers in gaining a deeper understanding of user states when utilizing avatars but also assists them in optimizing avatar utilization [14, 15].

While most previous studies on the Proteus Effect primarily focus on human avatars [25], the exploration of non-human avatars remains limited. Despite the increasing use of non-human avatars, their impact has not been thoroughly investigated in the context of the Proteus Effect. Unlike human avatars, non-human avatars possess less self-relevance to users due to substantial differences in body structure and behavioral patterns. Lan et al. have confirmed the persistence of the Proteus Effect even with non-human avatars [16], employing subjective questionnaires to evaluate users' social engagement. Li et al. developed a virtual environment where participants embodied a cat avatar and introduced the psychological concept of self-illusion. Their findings indicate that participants with a heightened sense of self-illusion were more inclined to identify themselves with the virtual role [18]. This may suggest a mechanism by which people associate themselves with the virtual role, producing a cognitive self-illusion and further induces consistent behavioral performance.

Therefore, this paper aims to explore a potential objective method to evaluate the Proteus Effect through unconscious hand gestures and movements of humans, considering the disparities between gestures of human hands and non-human avatars. This serves as the primary research question of the study. Additionally, this research also focuses on examining the potential effects of avatar characteristics on self-illusion and human-avatar gesture consistency, further understanding whether self-illusion may act as a perceptual precursor to the Proteus Effect.

To address these questions, the study initially developed a human avatar alongside three non-human avatars. Subsequently, we constructed a hand gesture dataset that intuitively reflected gestures those were consistent with corresponding avatar hands. A recognition model for human-avatar gesture consistency was then trained. These avatars and the recognition model were integrated into a virtual reality (VR) experimental system. Using this system as a foundation, a user experiment was conducted to analyze the effects of avatar characteristics on self-illusion and human-avatar gesture consistency.

The main contributions can be summarized as follows:

- We introduce an objective and real-time method for evaluating the Proteus Effect in the utilization of non-human avatars, based on detecting human-avatar gesture consistency. Our findings reveal that participants exhibited a significantly higher percentage of gestures consistent with their presently used avatars.
- We discovered that the types of virtual avatars significantly impact users' self-illusion. Utilizing non-human avatars can indeed evoke self-illusion of participants, although its levels were significantly lower than those experienced when using human avatars.

*Corresponding author. E-mail: zhouchao@iscas.ac.cn
†Corresponding author. E-mail: bianyulong@sdu.edu.cn

- We conclude that self-illusion may be a perceptual antecedent of the Proteus Effect (even non-human avatars). The characteristics of virtual avatars influence human self-illusion at a cognitive level, subsequently may affecting hand gestures at the behavioral level.

2 Related Work

2.1 Avatar and the Proteus Effect

In virtual reality, electronic games, and various media technologies, avatars serve as digital representations of humans [21, 29]. Human often form significant relationships with their virtual avatars, and the utilization of avatars heightens individuals' enjoyment of these media technologies [19, 31, 13]. Yee and Bailenson investigated the impact of avatar characteristics on individual behavior within virtual reality environments, a phenomenon known as the Proteus Effect [36, 37].

The Proteus Effect plays a crucial role in VR, as it not only shapes and influences people's cognition and behavior but also reflects their user experience [15]. For instance, when white individuals portray black characters in virtual environments, it can persistently diminish implicit biases among white individuals [3]. Additionally, employing avatars with normal weight has been shown to stimulate children's motivation for physical activity [17]. Kocur proposed that VR designers and researchers could enhance user interaction and performance in virtual environments by crafting characters with specific functions [14]. As research into the Proteus Effect deepens, the influence of avatars on users has extended beyond behavioral aspects to encompass cognitive and emotional levels [9, 11].

Kocur et al. suggested employing the degree of the Proteus Effect in VR as a metric for evaluating User Experience (UX). For example, if users exhibit slower movements while embodying elderly characters in VR, it could be inferred that their experience is better [15].

2.2 Behavior Changes Caused by Avatars

Numerous previous studies have demonstrated behavior changes caused by avatars from various perspectives. Regarding the Proteus effect, the most typical behavior change is that the characteristics of the avatar alter the participant's social behavior, and this change occurs with both human and non-human avatars [36, 16]. Moreover, the self-avatar follower effect was proposed to describe the tendency of VR users to follow the movements of their avatars to reduce potential multi-sensory discrepancies between their (real) body cues and the visual feedback [10]. Won et al. discussed homuncular flexibility in VR, which refers to the ability of users to learn to control bodies different from their own by changing the relationship between tracked and rendered motion, and found through two experiments that participants can quickly learn to use a novel avatar and will change their own behavior to effectively complete the task [34]. This seems to provide a subjective reason for behavior changes induced by avatars.

2.3 The Proteus Effect in non-human avatars

Ratan et al. conducted a meta-analysis of studies on the Proteus Effect over the past decade, demonstrating the consistency of the Proteus Effect under human avatar [25]. In recent years, there has been a growing interest in the embodiment of animals in VR experiences. Researchers have utilized animal avatars to investigate their influence on participants' empathy, with the aim of fostering a better understanding of natural environments and promoting animal conservation through these animal avatars [27, 1, 23]. Furthermore, research has explored the impact of providing tactile and auditory feedback within animal avatars on the experience of embodiment [32]. Recently, Jiang et al. proposed a solution called HandAvatar to enrich users' flexible control over non-human bird avatars [12], then used it to overcoming the height of fear by inducing the proteus effect. As for the Proteus Effect in non-human avatars, Oyanagi et al. investigated whether experiencing bird avatars can reduce people's fear of heights [22]. Based on the work of Bian et al. [6], Lan et al. explored the impact of attractiveness of non-human avatars (dogs) on users' social participation, and finally found that participants with higher attractiveness of non-human avatars showed higher social participation, proving that Proteus Effect could still exist in using non-human avatars [16]. However, the Proteus Effect under non-human avatars were not fully explored.

2.4 Measurement of the Proteus Effect

Previous studies have utilized various methods to assess the impact of avatars on users, including implicit behavioral measures, self-report scales, and physiological measures [19]. Yee et al. validated the Proteus Effect using behavioral measures such as interpersonal distance measurements and ultimatum games [2, 8]. Regarding physiological measurements, prior research has explored indicators such as electrocardiography, heart rate, and skin conductance [19, 26]. However, for the Proteus Effect under non-human avatars, existing studies have mainly relied on questionnaire-based measurements, lacking objectivity [22, 16]. Overall, there is currently a shortage of specialized questionnaires for measuring the Proteus Effect, as well as objective and real-time evaluation methods.

2.5 Self-illusion

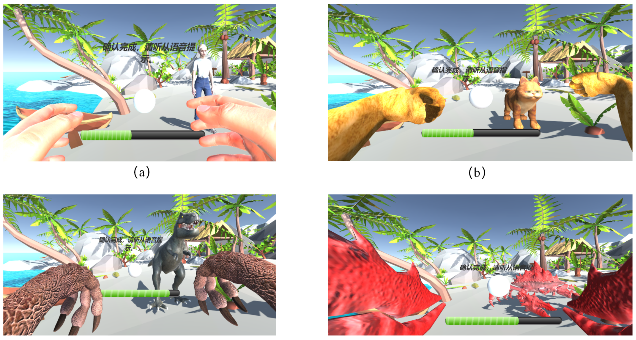
The most widely accepted theoretical basis of the Proteus Effect is the self-perception theory, which posits that individuals may infer their own emotional, cognitive, and other states based on the characteristics of avatars [5]. Self-concept [4] is the individual's belief about himself/herself, including the person's attributes and who and what the self is. In everyday life, self-concept typically doesn't undergo significant changes. However, in virtual reality, users can experience various avatars, leading to self-illusion. The self-illusion is a general illusion about the self in cognition, giving participants a feeling of being associated with a role in the virtual world, even though the person is convinced that the role is not the real self in reality. Li et al. devised a virtual scene in which participants experienced the cat avatar from the first-person perspective and proposed a new psychological component of presence, namely self-illusion [18]. Their study revealed that participants exhibiting heightened levels of self-illusion were more inclined to identify themselves with the virtual character. Consequently, self-illusion is an important aspect of self-perception and may constitute a perceptual outcome linked to the Proteus Effect.

3 Developing Human-Avatar Gesture Consistency Recognition Model

To explore the Proteus Effect under non-human avatars from gesture features, we built a hand gesture dataset and developed a recognition model to detect human-avatar gesture consistency.

3.1 Design of Avatar

The human avatars were designed as young adults (one male and one female), while the non-human avatars was selected based on similarity between their hands with human hands. To comprehensively cover various degrees of similarity to human hands, three typical non-human avatars (animals) were carefully chosen, including cat, dinosaur, and crab. Subsequently, using professional modeling software, we created hand models for both human avatars and non-human avatars, and produced spontaneous hand animations as well as grasping motion animations for each hand model. The full-body of each type of avatar, along with the according virtual hand from the first-person perspective are illustrated in Fig. 1.



(a) (b) (c) (d)

Figure 1: Four kinds of virtual hands from the first-person-view and the full-body of each avatar. According to the similarity with real human hands, they are ranked from high to low as human hands and full-body human model(taking the female as an example) (a), cat claws and full-body cat model(b), dinosaur claws and full-body dinosaur model(c), and crab claws and full-body crab model(d), respectively.

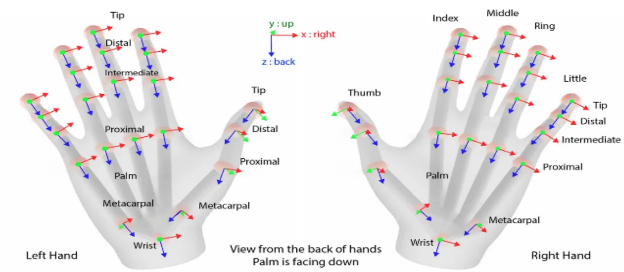


Figure 2: The 22 hand joints extracted in this study.

3.2 Building Human Hand Gesture Dataset Consistent with Avatars

The purpose of building a human gesture dataset consistent with avatars is to comprehensively cover the gesture data information representing each avatar as fully as possible. Specifically, this entails constructing gestures representing human avatars, gestures representing cat avatars, gestures representing dinosaur avatars, and gestures representing crab avatars.

3.2.1 Apparatus and the principle of gesture information collection

The four aforementioned virtual avatars were experienced using the Pico 4 all-in-one VR headset (short as Pico 4). The Pico4 captures and generates hand data through its front-facing quad-eye high-precision environmental tracking camera (VGA 640*480@60HZ). By combining OpenXR hand recognition information with Pico's hand recognition SDK, we ultimately extracted 22 nodes from each hand. During the virtual scene experience, real-time three-dimensional world coordinates of all 22 nodes for each hand were obtained. In order to avoid the influence caused by participants' own positions and rotations during the experience, the wrist node served as the reference point for the entire hand, while the remaining 21 nodes were converted to local coordinates relative to the wrist node to record hand shape information. The hand nodes are depicted in Fig. 2.

3.2.2 Participants

A total of 46 participants(18-21 years old, M=19.46, SD=1.13, 24 males and 22 females) from a local university were recruited to participate in the construction of the gesture dataset. All participants were interested in experiencing VR, and most of them had no prior VR experience. All participants were in good health with normal vision, and none had color blindness or color vision deficiency.

Figure 3: Partial instructions for gesture collection tasks (taking the dinosaur avatar as an example).

3.2.3 Procedure

The order of data collection for different avatars was randomly predetermined. Upon entering the VR scene, they were instructed to experience the four virtual avatars in the predetermined random order. At the start of collecting data, each participant was instructed to perform gestures based on their intuition under each avatar condition. We did not guide/require/force them to perform any specific gestures to represent an avatar, all gestures were made entirely intuitively. For each avatar, five instances of hand movements indicating grasping and five instances of baseline gestures were collected. For the former, participants were instructed to grasp a white ball in front of them to induce a grasping action, and each participant was required to perform one grasping action during each data collection session. For the latter, they were instructed to remain in a purposeless idle state. Tasks were executed separately for the left and right hands, and their data were collected. An example of the prompt text for the gesture information collection is shown in the Fig. 3. We avoid using suggestive words that may interfere with the intuitive gestures made by participants. A progress bar in the participants' field of view indicated the progress of each data collection session, with each session consisting of 40 frames. The data collection system is illustrated in Fig. 1. Consequently, the dataset comprises eight classification labels: human hand (grasping and idle), cat claws (grasping and idle), dinosaur claws (grasping and idle), and crab claws (grasping and idle). Fig. 4 displays examples of gestures made by most participants that were consistent with each virtual avatar. For the human avatar, Fig. 4(a) illustrates the participants' gestures from different angles and with varying degrees of finger bending. For the cat avatar, most participants presented a half-fist or fist gesture. For the dinosaur avatar, participants tended to use two or three fingers to represent the dinosaur. For the crab avatar, participants tended to make various scissor-like gestures. Ultimately, the database contains 1840 gesture state sequences.

3.2.4 Description of gesture sequence

Each gesture state sequence in the dataset comprises 40 frames of gesture states. Each frame's gesture state consists of the local three-dimensional coordinates of the remaining 21 hand nodes, excluding the wrist. Each frame's gesture state is akin to taking a photograph of the hand, and with 40 frames, it creates an animation-like effect. Fig. 5 illustrates the gesture state of a participant captured from the obtained local three-dimensional coordinates within a single frame.

(a) (b) (c) (d)

Figure 4: The most common three gestures of participants consistent with the avatars are: (a) three gestures consistent with the human avatar, (b) three gestures consistent with the cat avatar, (c) three gestures consistent with the dinosaur avatar, and (d) three gestures consistent with the crab avatar.

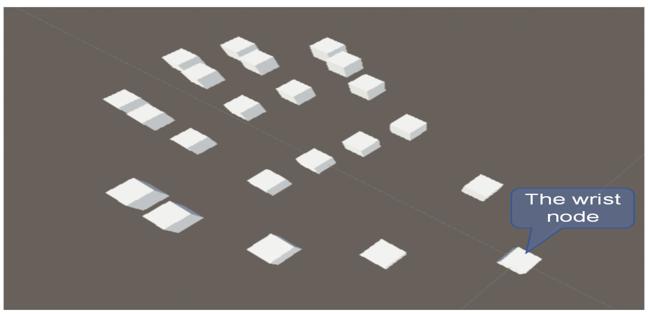


Figure 5: A frame of gesture state reconstructed from the recorded local three-dimensional coordinates of 21 hand nodes (with the wrist serving as the origin of the local coordinate system), the wrist node in the figure is for demonstration purposes only; in actual data collection, the wrist serves as the origin of the local coordinate system.

Each gesture state sequence is associated with one of the eight classifications.

3.3 Training of Human-avatar Gesture Consistency Recognition Model

We first divided the dataset into a predefined training set and a predefined test set containing 185 samples. Then, we trained a human-avatar gesture consistency recognition model by using 80% of the predefined training set and the rest 20% as the validation set. When dealing with classification problems involving temporal data, common models include Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU). We ultimately adopted GRU because it is another improved type of RNN. Similar to LSTM but with a simplified structure, GRU offers faster processing speed while maintaining classification accuracy [35].

Finally, we constructed a deep neural network model based on Gated Recurrent Units (GRU). Initially, the input data is processed through GRU layers, which effectively capture long-term dependencies in time-series data, crucial for modeling sequential data. Subsequently, the data passes through two fully connected hidden layers, aiding the network in learning abstract feature representations of the data. Lastly, the output layer employs a softmax activation function to convert the raw outputs of the network into a probability distribution of categories, facilitating multi-class classification. Additionally, to mitigate overfitting, Dropout layers were added within the model, reducing network complexity by randomly dropping some neurons during training. Overall, our model combines GRU and fully connected layers, demonstrating strong learning and generalization capabilities suitable for classifying our gesture sequence data. Through this model, we can identify the avatar type represented by real human gestures and whether grasping motions have been performed. For the recognition of similar gestures, our model can effectively capture gesture features, and the softmax function can calculate the probability belonging to each category of all the gesture categories (including the similar ones) based on the input data, then we select the one with highest probability as the output. Ultimately, we referred to this model as the human-avatar gesture consistency recognition model. The structure of the human-avatar gesture consistency recognition model is shown in the Fig. 6.

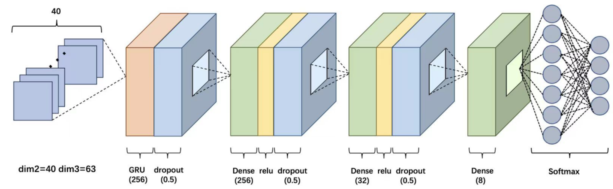


Figure 6: The structure of the human-avatar gesture consistency recognition model.

We used a predefined test set for testing, and the theoretical accuracy of the model's recognition results is 99.5% (see the confusion matrix in Fig. 7). The numbers "0"-"7" on the left and bottom of the confusion matrix represent the data labels, namely cat idle, human idle, dragon idle, crab idle, cat grasping, human grasping, dragon grasping, and crab grasping, respectively.

Before the formal experiment began, to assess the practical application effect of the human-avatar gesture consistency model, we built a simple scene using Unity3D. We recruited 10 college students (aged ranging from 19 to 22, *M*=20.5, *SD*=1.08; height ranging from 155 cm to 184 cm, *M*=170.8, *SD*=8.702; weight ranging from 45 kg to 85 kg, *M*=62.9, *SD*=14.97739) to participate in the practical effectiveness testing of the human-avatar gesture consistency recognition model. Among them, there were 5 males and 5 females. Participants were briefed on the gestures in the gesture database before the test. During the test, they sequentially performed hand gestures for each avatar (including idle and grabbing actions). The recognition results and accuracy were computed in real-time. When participants switch gestures, the model's recognition may experience brief errors; however, once the participants' gestures stabilize, the model's recognition results become accurate. Ultimately, the actual usability of the human-avatar gesture consistency recognition model was lower than the accuracy on the test set, reaching 96.4%. The varying body types of participants also validated the model's robustness.

4 User Experiment

To examine the Proteus Effect under different characteristics of avatars, and verify the effectiveness of our proposed method for assessing the effect by detecting human-avatar gesture consistency, we performed an user experiment.

Proteus Effect may be triggered by using non-human avatars. According to our previous analysis that self-illusion may be a perceptual antecedent of the Proteus Effect (even non-human avatars). The virtual avatar might unconsciously influence human's self-perception at cognitive level, then subsequently affect hand gesture at behavioral level. Thus self-illusion and human-avatar gesture consistency can be expected as potential indicators to reflect the

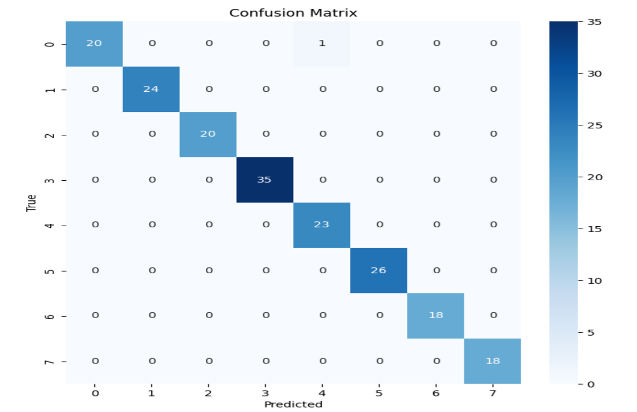


Figure 7: The Confusion Matrix. Each row represents the actual classification of the input, and each column represents the model's identified classification based on the input. The diagonal represents the count of correct predictions.

Proteus Effect.

The follower effect reveals that the self-motion characteristics of avatars can trigger behavioral imitation of VR users. Based on embodied cognition theory, body states may unconsciously influence mental functions such as perception and reasoning[20]. We hypothesize that the follower effect, triggered by the spontaneous animation of the avatar's hands, might further strengthen human's sense of embodied toward the avatar and then promote the occurrence of the Proteus Effect.

According to the above analyse, we proposed the following hypotheses:

- Hypothesis (H1): Different avatar type may significantly affect self-illusion.
- Hypothesis (H2): Spontaneous animation of avatar's virtual hands may significantly affect self-illusion.
- hypothesis (H3): Different avatar type may significantly affect human-avatar gesture consistency.
- hypothesis (H4): Spontaneous animation of avatar's virtual hands may significantly affect human-avatar gesture consistency.

4.1 Participants

A total of 69 college students voluntarily participated in this experiment (17-26 years old, M=19.71, SD=1.895, 53 males and 16 females). They were divided into three groups according to their preferences for virtual animal avatars (cat avatar, crab avatar and dinosaur avatar). Each group consisting of 23 individuals.

4.2 Experimental Design

We performed a 2 (avatar type: human avatar/non-human avatar) * 2 (spontaneous animation of virtual hand: presence/absence) within-subject experimental design. Both avatar type and spontaneous animation of virtual hand were the within-subject factors. The human-avatar gesture consistency and perception of self-illusion were dependent variables. For conditions of the avatar type, human and non-human avatars were same as those designed in Sec. 3.1, each participant needed to experience the human avatar and one non-human avatar that he/she preferred. For conditions of virtual hand, the absence of spontaneous animation means that the avatar hand cannot perform spontaneous motion at all, while the presence of spontaneous animation means that the virtual hand can spontaneously perform some natural, slight, and task independent motions namely excluding hands or forearm movements like reaching out at a certain frequency, Fig. 8 shows the key frames of the spontaneous hand animations for each avatar.

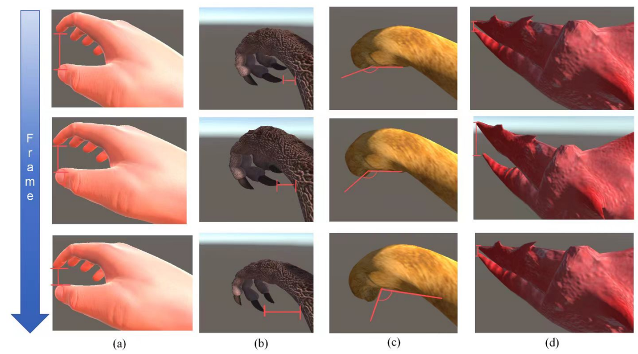


Figure 8: The key frames of the spontaneous hand animations for each avatar. The red outlines in the figure are added as auxiliary lines to help show the difference among key frames of hand animations.

Based on within-subject variables, this experiment comprises four conditions:

- Human avatar with spontaneous animation of virtual hand
- Human avatar without spontaneous animation of virtual hand
- Non-human avatar with spontaneous animation of virtual hand
- Non-human avatar without spontaneous animation of virtual hand

The experience order of the experimental conditions was counterbalanced with the method of Latin square to eliminate sequence effects.

4.3 Experimental Tasks and Procedure

We designed the different avatars and the human-avatar hand gesture consistency recognition model into a VR experimental system, based on which we performed the user experiment.

Before the experiment begins, the researchers will inform the participants about some introductory information related to the experiment, part of which is "Next, you will experience an object-sorting game and you can grasp objects by making natural motion with a single hand based on your intuition...". Before the formal task, participants first entered a practice scene to familiarize with the basic ways to operation (Fig. 9 a).

Secondly, participants entered the virtual environment and selected the corresponding avatar (Fig. 9 b). During the avatar selection phase, participants were free to view each avatar. Fig. 9 b shows a participant viewing the cat avatar. The avatar selection phase was completed when the participant touched the block labeled "Confirm Selection".

Thirdly, before the formal task, the participants engaged in a virtual environment and needed to go through a section of the road to reach the final place of formal task. When passing through this section of the road, they could pick up coconuts on the ground (Fig. 9 c). This process is unrelated to formal tasks, mainly for the purpose of making the participants fully familiar with the virtual world and virtual avatar.

Then, they performed the formal task. The formal task required participants to grasp target objects with intuitively natural gestures.

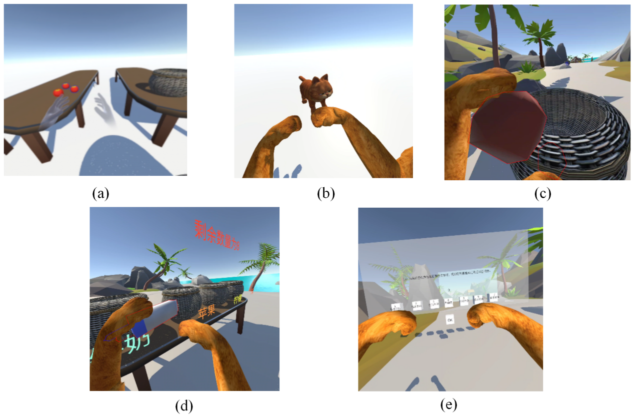


Figure 9: Experimental tasks in each stage, taking the cat avatar as an example. (a) operation familiarization; (b) virtual avatar selection; (c) immersion of virtual avatars; (d) official item sorting; (e) questionnaire filling.

Participants don't need to grip objects with both hands simultaneously, but use just one hand. There were 10 target objects from four categories that were randomly generated and scattered on a virtual table. Four baskets with corresponding category labels were placed on another table. Participants needed to grab up the objects from the table and put them into the right basket one by one (Fig. 9 d). To avoid them using only dominant hand during the task, the number of tasks for each hand (5 times each for the left and right hands) was equal and randomly assigned. They experienced four rounds of task under each experimental conditions in a predetermined order.

After completing each round of tasks, they were immediately complete a brief questionnaire (Fig. 9 e) and take appropriate breaks. The human-avatar gesture consistency recognition model recognized and recorded the gesture classification results throughout the entire task process. Finally, each participants received a present as the reward.

To avoid the VR sickness interfering with the participants' experience, we adopted the real walking method to achieve roaming to reduce Sensory conflict, and we would stop the game and delete the data if the participants suffered from VR sickness or other discomforts during the game. Finally, none of participants faced VR sickness or discomfort in the experiment. The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Human Research Ethics Committee of a local Hospital. When this study involved human participants, informed consent was received from each individual.

4.4 Experimental Apparatus and Interaction Methods

Apparatus. The hardware equipment for the user experiment was still the Pico 4 which supports 6 DOF spatial positioning, allowing participants' real-world movements to be synchronized within the virtual environment.

Interaction method. Throughout the experiment, participants no longer used the controllers of the Pico 4 to interact, but used their nature hand gestures. The participant's hand and the avatar's hand move and are positioned in sync throughout the experiment, which means the virtual hand's movement is controlled by the real hand.

The grabbing operation of targets utilized the human-avatar gesture recognition model described in Sec. 3.3. During the participants' experience, the Pico 4 continuously sent the collected 40-frame hand joint information of the participant into the recognition model. The gesture recognition model predicted the classification results based on the input and sends them to the Pico 4 via computer network. Upon receiving the classification results of the participants' gestures, the experimental program running on Pico 4 determined whether the participant was performing a grabbing action. The classification results were recorded in the Pico 4. To successfully grasp objects, participants don't have to perform specific animal-like gestures or specific "right" gestures. As described in Sec. 3.2, our gesture set includes almost all grasp-related actions for the human hands and those non-human avatars. If participants perform any grasping action and touch an object with their hand, the object will be grasped.

4.5 Measurements

We evaluated human-avatar gesture consistency, self-illusion as dependent variables.

Human-avatar gesture consistency. According to the interaction method described in Sec. 4.4, the human-avatar gesture consistency recognition model recognition results obtained based on the participants' real hand gestures will ultimately be recorded in text format (with each round of item sorting process corresponding to a record file). The human-avatar gestures consistency is calculated in each round of item sorting record file by dividing the number of avatar categories experienced by the participant by the total number of recognition in the record (calculated separately for left and right hands). The main reasons for calculating the left and right hands separately are that our gesture consistency recognition model and interaction methods are based on a single hand, as well as the experimental task was distinctly designed for left and right hands.

For human avatars: In the record file of the item sorting process experienced by participants with human avatar, the percentage is obtained by dividing the number of instances of human hand idle or human hand grasping by the total number of recognition instances in the record.

For non-human avatars: In the record file of the item sorting process experienced by participants with non-human avatars, the percentage is obtained by dividing the number of instances of non-human hand idle or non-human hand grasping experienced by the participant by the total number of recognition instances in the record. For example, when a participant experiences a non-human avatar as a dinosaur avatar, it only includes dragon idle and dragon grasping.

Self-illusion. We adopted the Self-illusion Questionnaire developed in [18]. According the questionnaire, three questions were designed focused on acquisition of behavior patterns (Q1: To what extent do you think you acquire the behavior pattern of the animal you acted?), subjective evaluation (Q2: To what extent do you feel you really come to be the animal?) and spontaneous behavioral intent (Q3: To what extent will you behave like the animal rather than human). Due to that the Q3 focuses on future intentions but this study focuses more on current experiences, thus the Q3 was removed and only Q1 and Q2 were used. The items are scored on the Likert 7-point scale from 1 (very slightly or not at all) to 7 (extremely).

4.6 Results

Kolmogorov–Smirnov tests (K-S tests) were first performed for all dependent variables, and the results supported the normality assumption ($ps > 0.05$).

4.6.1 Effects on self-illusion

A repeated measurements analysis of variance was performed with the feeling of self-illusion as the dependent variable.

Results showed that there was significant main effect of avatar type on feeling of self-illusion (Fig. 10). Specifically, human avatar led to significantly higher level of self-illusion [$F_{(1,66)}$=29.781, p<0.001, η_p^2=0.311]. These results supported H1.

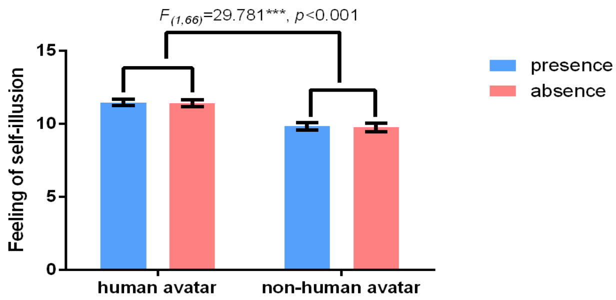


Figure 10: The analysis result of feeling of self-illusion.

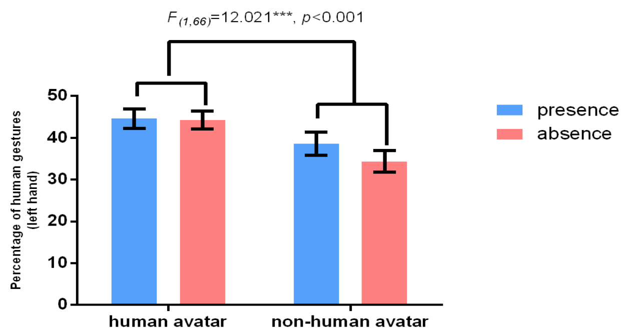


Figure 11: The analysis results of the percentage of human avatar gestures (left hand).

However, the main effect of spontaneous animation of virtual hand(presence/absence) and the interaction effect of avatar type*spontaneous animation of virtual hand on feeling of self-illusion were also not significant (ps>0.05), which did not support H2.

4.6.2 Results on human gestures consistent with the avatar

A series of repeated measurements analysis of variance were performed with the percentage of gestures consistent with the avatar. The data on the left and right hands of the participants were analyzed separately.

There were widely significant main effects of avatar type on percentage relevant measures (Fig. 11, Fig. 12, Fig. 13, Fig. 14). Specifically, human avatar led to significantly higher percentage of human hand gesture on left [$F_{(1,66)}$=12.021, p<0.001, η_p^2=0.154] and right hand [$F_{(1,66)}$=15.089, p<0.001, η_p^2=0.186]. In contrast, non-human avatar led to significantly higher percentage of corresponding non-human gesture on left[$F_{(1,66)}$=18.154, p<0.001, η_p^2=0.216] and right hand [$F_{(1,66)}$=17.392, p<0.001, η_p^2=0.209].

These results supported H3, and demonstrate that different avatar types significantly influence human-avatar gesture consistency.

5 Discussion

This study focused more on the Proteus Effect in using non-human avatars, and first explored an objective method to detect Proteus Effect from the humans' unconscious hand gesture and movement. Based on the method, this study performed an user study and examined the effects of avatar characteristics on self-illusion and consistency of human-avatar gestures. Some important results are discussed below.

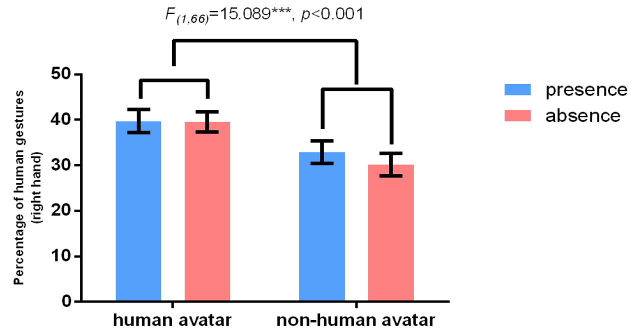


Figure 12: The analysis results of the percentage of human avatar gestures (right hand).

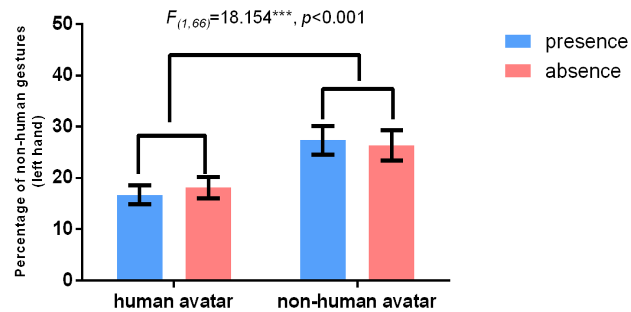


Figure 13: The analysis results of the percentage of non-human avatar gestures (left hand).

5.1 Evaluating Proteus Effect Based on Human- avatar Gesture Consistency

This study focuses on a novel measurement method to evaluate Proteus Effect in using non-human avatars based on human-avatar gesture consistency. As far as we known, we constructed the first gesture dataset that intuitively reflected real hand movements consistent with corresponding non-human avatar hands. Based on the dataset, we trained a recognition model of human-avatar gesture consistency. Then, we validated its effectiveness in evaluating the Proteus Effect in VR when using non-human avatars. The results showed that avatar types had significant effects on these measures, and participants showed gestures consistent with the avatar hands. Proteus Effect indeed occurred during using non-human avatars, and our recognition model can help evaluate the effect by detecting human- virtual avatar gesture consistency.

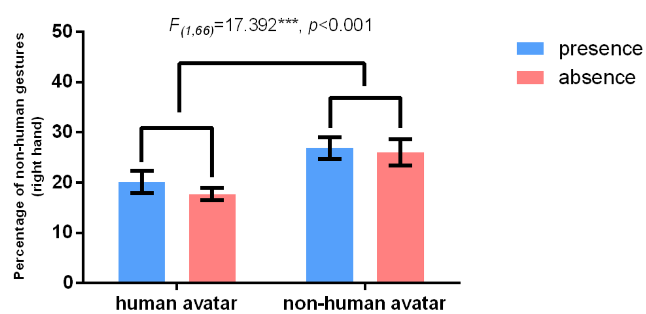


Figure 14: The analysis results of the percentage of non-human avatar gestures (right hand).

Therefore, based on human-avatar gesture consistency recognition model, the user behavior changes can be detected objectively and in real time, providing a new measurement method for exploring the behavioral impact of Proteus Effect.

5.2 The Impact of Different Avatar Types on Self-illusion and the Human-avatar Gestures Consistency

Regarding self-illusion, based on the results in Sec. 4.6.1, we can observe significant differences in the levels of self-illusion between human avatars and non-human avatars. The proposer of self-illusion have pointed out its higher propensity in virtual environments, demonstrating in their research that participants embodying a cat (one of the non-human avatars in this experiment) in a virtual environment experienced self-illusion [18]. The self-illusion scores of participants regarding non-human avatars in this experiment reflect that non-human avatars to some extent induce self-illusion in participants, consistent with previous research. Compared to non-human avatars, human avatars exhibit higher levels of self-illusion. The proposer of self-illusion indicate that possessing a higher level of self-illusion signifies better connection with virtual avatars [18]. Therefore, we can conclude that using human avatars in virtual reality indeed fosters better connection.Perhaps this also explains why the use of human avatars in the past has been better for observing the Proteus Effect.

Regarding the consistency of gestures between humans and avatars, based on the results in Sec. 4.6.2, we can observe significant differences in human and non-human gestures under human and non-human avatars. Specifically, when participants experience human avatars, the percentage of human gestures is higher than non-human gestures, while when participants experience non-human avatars, the percentage of non-human gestures is higher than human gestures. From the changes in gestures, we can infer that the virtual avatars experienced by participants influence their behavioral aspects, consistent with previous research about the Proteus Effect. In previous studies on the Proteus Effect with human avatars, the avatar's influence on participant behavior includes social behavior [36], body movement speed [28], etc., in this study represented by gestures of the human category. In previous studies on the Proteus Effect with non-human avatars, the influence of avatars on participant social behavior was explored [16], while in this study, it is represented by gestures consistent with the non-human avatars experienced by the participants.

5.3 The Impact of Spontaneous Animation of Avatar's Hand on Self-illusion and the Human-avatar Gestures Consistency

Based on the results from Sec. 4.6.1, Sec. 4.6.2, unfortunately, we did not find significant effects of spontaneous hand animation of virtual avatars on self-illusion and the human-avatar gesture consistency. That is to say, the results do not support our H2 and H4. The self-avatar follower effect reveals that the influence of the virtual self-body movements on one's actual actions [7]. During the course of this experiment, some participants inquired about the differences in their experiences between two sessions under human avatars, indicating that some participants did not observe significant self-motion of virtual avatars due to the subtle spontaneous finger movements in the hand animation of the human avatars we created. Previous studies on the self-avatar follower effect have mainly focused on human avatars, and research on self-motion of non-human avatars is still lacking. A previous study on the self-avatar follower effect discussed the activation of mirror neuron system [10]. This implies a potential direction for future research to explore the potential effect spontaneous hand animation of virtual avatars. Whether MNS will be activated and the degree of activation may be the key conditions determining the effect of hand animation on self illusion and behavioral aspects.

5.4 The Self-illusion and the Proteus Effect

Based on the results from Sec. 4.6.1 and Sec. 4.6.2, as well as the preceding discussion, it is evident that different types of avatars (human or non-human) significantly influence the level of self-illusion and the consistency of gestures between humans and virtual avatars. Self-illusion refers to participants' self-perception of virtual avatars at the cognitive level in VR. Self-illusion can serve as one of the cognitive-level conditions for the Proteus Effect, and the consistency of gestures between humans and virtual avatars may serve as one of the behavioral indicator of Proteus Effect. In our future research, we can deeply examine the relationship between self-illusion and the Proteus Effect by further manipulating different conditions of self-illusion.

6 Limitations and Future Work

There are still some limitations in this paper, which needs to be solved in future work.

First, this study preliminarily examined the effects of different avatar types (human avatar and non-human avatar), without examining the effects of different avatar characteristics within non-human avatars. In future work, we will further focus on these question and explore wide types,features of different non-human avatars.

Second, the indicators of human-avatar gesture consistency is relatively simplistic, primarily consisting of percentage. In future work, we will further explore indicators that can more fully and accurately reflect human-avatar gesture consistency. Vatavu et al. developed new methods for analyzing the agreement rate of user-elicited data [33], which may help better explore the gestures relevant to each avatar.

Third, the effect of spontaneous hand animation of avatar were not found in this study, which did not support the follower effect hypothesis. Possible reasons may be that the spontaneous features of virtual hand are not prominent enough or the measurement indicators are not sensitive enough. In future work, we will design a wider variety of spontaneous features of virtual hand, and integrate relevant physiological indicators (such as ECG, RSP, etc.) to further examine this question.

7 Conclusions

When using first-perspective non-human avatars, the characteristics of virtual hands may also induce relevant cognitive and even behavioral patterns in real human hands. This study systematically explored the potential of objectively evaluate Proteus Effect in using non-human avatars through evaluating human-avatar gesture consistency. Finally, we propose an objective and real-time method for assessing Proteus Effect in using non-human avatars based on detecting the human-avatar gesture consistency, and found that participants performed significantly larger percentage of gestures that were consistent with their currently used avatars. We found the types of virtual avatar significantly influence users' self-illusion. Using non-human avatars can induce experience self-illusion of participants, although its levels were significantly lower than those in using human avatars. It is concluded that self-illusion may be a perceptual antecedent of the Proteus Effect (even non-human avatars).

Acknowledgments

This work was supported in part by the Ministry of Education of Humanities and Social Science Project (Grant No.22C10422009), Shandong Province Youth Entrepreneurship Technology Support Program for Higher Education Institutions (Grant No. 2022KJN028), and the Young Scholars Program of Shandong University, Weihai (Grant No.20820211005).

References

[1] S. J. G. Ahn, J. Bostick, E. Ogle, K. L. Nowak, K. T. McGillicuddy, and J. N. Bailenson. Experiencing Nature: Embodying Animals in Immersive Virtual Environments Increases Inclusion of Nature in Self and Involvement with Nature. *Journal of Computer-Mediated Communication*, 21(6):399–419, 09 2016. doi: 10.1111/jcc4.12173 2

[2] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis. Interpersonal distance in immersive virtual environments. *Personality and social psychology bulletin*, 29(7):819–833, 2003. doi: 10.1177/0146167203029007002 2

[3] D. Banakou, P. D. Hanumanthu, and M. Slater. Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. *Frontiers in human neuroscience*, 10:226766, 2016. doi: 10.3389/fnhum.2016.00601 2

[4] R. F. Baumeister. *The self in social psychology*. Psychology Press, 1999. 2

[5] D. J. Bem. Self-perception theory. In *Advances in experimental social psychology*, vol. 6, pp. 1–62. Elsevier, 1972. doi: 10.1016/S0065-2601(08)60024-6 2

[6] Y. Bian, C. Zhou, Y. Tian, P. Wang, and F. Gao. The proteus effect: Influence of avatar appearance on social interaction in virtual environments. In *HCI International 2015-Posters' Extended Abstracts: International Conference, HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015. Proceedings, Part II 17*, pp. 78–83. Springer, 2015. doi: 10.1007/978-3-319-21383-5_13 1, 2

[7] L. Boban, L. Strauss, H. Decroix, B. Herbelin, and R. Boulic. Unintentional synchronization with self-avatar for upper-and lower-body movements. *Frontiers in Virtual Reality*, 4:1073549, 2023. doi: 10.3389/frvir.2023.1073549 8

[8] R. Forsythe, J. L. Horowitz, N. E. Savin, and M. Sefton. Fairness in simple bargaining experiments. *Games and Economic behavior*, 6(3):347–369, 1994. doi: 10.1006/game.1994.1021 2

[9] J. Fox, J. N. Bailenson, and L. Tricase. The embodiment of sexualized virtual selves: The proteus effect and experiences of self-objectification via avatars. *Computers in Human Behavior*, 29(3):930–938, 2013. doi: 10.1016/j.chb.2012.12.027 1, 2

[10] M. Gonzalez-Franco, B. Cohn, E. Ofek, D. Burin, and A. Maselli. The self-avatar follower effect in virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 18–25, 2020. doi: 10.1109/VR46266.2020.00019 2, 8

[11] V. Groom, J. N. Bailenson, and C. Nass. The influence of racial embodiment on racial bias in immersive virtual environments. *Social Influence*, 4(3):231–248, 2009. doi: 10.1080/15534510802643750 1, 2

[12] Y. Jiang, Z. Li, M. He, D. Lindlbauer, and Y. Yan. Handavatar: Embodying non-humanoid virtual avatars through hands. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3581027 2

[13] K. Kim, M. G. Schmierbach, M.-Y. Chung, J. D. Fraustino, F. Dardis, L. Ahern, et al. Is it a sense of autonomy, control, or attachment? exploring the effects of in-game customization on game enjoyment. *Computers in Human Behavior*, 48:695–705, 2015. doi: 10.1016/j.chb.2015.02.011 2

[14] M. Kocur. *Utilizing the Proteus Effect to Improve Performance Using Avatars in Virtual Reality*. PhD thesis, 2022. 1, 2

[15] M. Kocur, N. Henze, and V. Schwind. The extent of the proteus effect as a behavioral measure for assessing user experience in virtual reality. In *CHI 2021 - Workshop on Evaluating User Experiences in Mixed Reality.*, pp. 1–3, 2021. doi: 10.5283/epub.45543 1, 2

[16] X. Lan and Z. M. van Berlo. Can non-humanlike avatars induce the proteus effect? the roles of avatar identification and embodiment in influencing social participation. *Computers in Human Behavior: Artificial Humans*, 1(2):100020, 2023. doi: 10.1016/j.chbah.2023.100020 1, 2, 8

[17] B. J. Li, M. O. Lwin, and Y. Jung. Wii, myself, and size: The influence of proteus effect and stereotype threat on overweight children's exercise motivation and behavior in exergames. *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, 3(1):40–48, 2014. doi: 10.1089/g4h.2013.0081 1, 2

[18] S. Li, X. Gu, K. Yi, Y. Yang, G. Wang, and D. Manocha. Self-illusion: A study on cognition of role-playing in immersive virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 28(8):3035–3049, 2022. doi: 10.1109/TVCG.2020.3044563 1, 2, 6, 8

[19] Y. Liu. The proteus effect: Overview, reflection, and recommendations. *Games and Culture*, p. 15554120231202175, 2023. doi: 10.1177/15554120231202175 2

[20] M. G. Maggio, D. Piazzitta, A. Andaloro, D. Latella, F. Sciarrone, C. Casella, A. Naro, A. Manuli, and R. S. Calabrò. Embodied cognition in neurodegenerative disorders: What do we know so far? a narrative review focusing on the mirror neuron system and clinical applications. *Journal of Clinical Neuroscience*, 98:66–72, 2022. doi: 10.1016/j.jocn.2022.01.028 5

[21] K. L. Nowak and C. Rauh. The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction. *Journal of Computer-Mediated Communication*, 11(1):153–178, 2005. doi: 10.1111/j.1083-6101.2006.tb00308.x 2

[22] A. Oyanagi and R. Ohmura. Transformation to a bird: Overcoming the height of fear by inducing the proteus effect of the bird avatar. In *Proceedings of the 2nd international Conference on Image and graphics processing*, pp. 145–149, 2019. doi: 10.1145/3313950.3313976 2

[23] D. Pimentel and S. Kalyanaraman. The effects of embodying wildlife in virtual reality on conservation behaviors. *Scientific Reports*, 12(1):6439, 2022. doi: 10.1038/s41598-022-10268-y 2

[24] J. Quick. *Effects of avatar appearance on user perception and behavior: Role of labels and cognitive mediation in the proteus effect*. PhD thesis, [Honolulu]:[University of Hawaii at Manoa],[December 2016], 2016. 1

[25] R. Ratan, D. Beyea, B. J. Li, and L. Graciano. Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: a meta-analysis of the proteus effect. *Media Psychology*, 23(5):651–675, 2020. doi: 10.1080/15213269.2019.1623698 1, 2

[26] R. A. Ratan and M. Dawson. When mii is me: A psychophysiological examination of avatar self-relevance. *Communication Research*, 43(8):1065–1093, 2016. doi: 10.1177/0093650215570652 2

[27] A. S. Rativa, M. Postma, and M. Van Zaanen. Can virtual reality act as an affective machine? the wild animal embodiment experience and the importance of appearance. In *MIT LINC 2019 Conference: Massachusetts Institute of Technology*, pp. 214–223. EPiC Series in Education Science, 2020. 2

[28] R. Reinhard, K. G. Shah, C. A. Faust-Christmann, and T. Lachmann. Acting your avatar's age: effects of virtual reality avatar embodiment on real life walking speed. *Media Psychology*, 23(2):293–315, 2020. doi: 10.1080/15213269.2019.1598435 8

[29] R. Schroeder. *The social life of avatars: Presence and interaction in shared virtual environments*. Springer Science & Business Media, 2001. 2

[30] V. Stavropoulos, J. Rennie, M. Morcos, R. Gomez, and M. D. Griffiths. Understanding the relationship between the proteus effect, immersion, and gender among world of warcraft players: An empirical survey study. *Behaviour & Information Technology*, 40(8):821–836, 2021. doi: 10.1080/0144929X.2020.1729240 1

[31] S. Trepte and L. Reinecke. Avatar creation and video game enjoyment. *Journal of Media Psychology*, 2010. doi: 10.1027/1864-1105/a000022 2

[32] M. F. Vargas, R. Fribourg, E. Bates, and R. McDonnell. Now i wanna be a dog: Exploring the impact of audio and tactile feedback on animal embodiment. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 912–921, 2023. doi: 10.1109/ISMAR59233.2023.00107 2

[33] R.-D. Vatavu and J. O. Wobbrock. Formalizing agreement analysis for elicitation studies: New measures, significance test, and toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, p. 1325–1334. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2702123.2702223 8

[34] A. S. Won, J. Bailenson, J. Lee, and J. Lanier. Homuncular Flexibility

in Virtual Reality. *Journal of Computer-Mediated Communication*, 20(3):241–259, 01 2015. doi: 10.1111/jcc4.12107 2

[35] S. Yang, X. Yu, and Y. Zhou. Lstm and gru neural network performance comparison study: Taking yelp review dataset as an example. In *2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI)*, pp. 98–101, 2020. doi: 10.1109/IWECAI50956.2020.00027 4

[36] N. Yee and J. Bailenson. The proteus effect: The effect of transformed self-representation on behavior. *Human communication research*, 33(3):271–290, 2007. doi: 10.1111/j.1468-2958.2007.00299.x 1, 2, 8

[37] N. Yee, J. N. Bailenson, and N. Ducheneaut. The proteus effect: Implications of transformed digital self-representation on online and offline behavior. *Communication Research*, 36(2):285–312, 2009. doi: 10.1177/0093650208330254 2