The Influence of Avatar Visual Fidelity on Embodiment and User Experience in Virtual Reality

Judith Hartfill, Frederico Bormann, Erik Wolf and Frank Steinicke

July 14, 2025

Abstract

Consumer grade VR now allows embodying full-body avatars to act within the virtual world. Realism has been shown to be an important factor for the plausibility and self attribution of an avatar. However, the role of self-representation, diverting from the actual user's appearance, is not clear yet.

In this work, we compare a generic, a self-configured, and a photo-generated full-body avatar with similar realism level in terms of embodiment and user experience. We conducted a study with thirty participants, that revealed higher ratings of attractiveness, valence, and threat reaction (as an indirect measure for embodiment) in the self-configured condition, compared to the photo-generated avatar. However, direct measures for embodiment remained inconclusive. These results motivate further research in the field of self-configured avatars, considering the feedback about avatar diversity that is still worthy of improvement.

Figure 1: Three representations used by a participant during the study. From left to right: Generic representation with limited choices for gender and ethnicity. Self-configured avatar with instruction to strive for the best representation detached from physical appearance. Photo-generated avatar to match the physical appearance.

1 Introduction

Due to the rise in popularity of virtual reality (VR), there has been a growing interest in the design of virtual avatars and how these differences influence the user experience (UX). At the same time, major improvements have been made to naturalistic input methods that enable the user to use their body movements for direct input. Previous research has marked it as advantageous in terms of user engagement. [95]

Interaction in VR using body movements raises the question of a virtual representation of this body in the virtual world, specifically, the use of a first-person avatar that functions as the digital alter ego in the virtual world. They provide properties to interact with the simulation and extend the body to the virtual world. [89]

Earlier studies have shown various practical benefits of such a self-avatar that matches the human body in terms of morphology, leading to better distance and reachability estimations [72, 73, 74, 25] or the estimation of body weight [94]. Recent research also considered the influence of avatars on the therapeutic potential of VR applications [93, 28]. Considering the UX perspective of immersive VR setups, the existence of avatars improves the illusion of presence in the virtual world as it introduces the visual plausibility of being in the displayed place [78, 82].

Beyond that, self-avatars are not simply viewed as external entities but rather tend to be processed in similar ways as the real body, which has been termed as embodiment [19] or considering the subjective impression related to it the sense of embodiment [49]. It combines multiple factors, including accepting the virtual body as your own body (body ownership) and oneself as the cause of its actions (sense of agency). Both bottom-up and top-down influences contribute to such an illusion, the first one mainly consisting of synchronous visuo-motor stimuli. However, previous studies have shown that the visual appearance of an avatar is an important top-down factor that can decrease the need for visuo-motor synchronicity, even tolerating inconsistencies to create an ownership illusion, which is one important sub-dimension of SoE [65, 31, 89, 42].

Previous work has categorized the influence factors on an avatar's visual appearance or avatar visual fidelity into three dimensions [38, p. 49]: Anthropomorphism describes the level of similarity to a humanoid morphology. Realism refers to the degree of visual detail, from a comic - like representation to a photorealistic one. While the those two categories refer to the avatar's general properties, the third dimension, truthfulness, is defined by the similarity to the user whom the avatar should represent.

Previous studies have already investigated how changes in some dimension of avatar visual fidelity might support the SoE of a user towards an avatar, often focusing on the level of truthfulness [77, 89, 24, 42, 48], reporting a higher SoE for a higher level of truthfulness. Matching gender and ethnicity seem especially important for invoking a SoE and SoO [21].

At the same time, the results for realism remain inconclusive [48, 77, 55].

Photorealistic full-body scans can be time and cost intensive[1]. In addition, they tend to create uncanny [67] and eerie sensations due to their nature of being imperfect visual reproductions [42, 55, 89] and were sometimes even inferior to hand modeled avatars [48]. Even if there are advances to overcome the uncanny valley [44], to explore a solution reducing effort and costs, in this work we use self-configured avatars for use in VR created by the user in a conventional character editor. This approach allows for avatars without the constraint of recreating the user's real-world appearance, which is prone to creating uncanny results because of easily noticeable differences from the actual appearance. This would lead to a higher SoE and better emotional response, especially regarding valence. Approaches that let users configure their appearance have rarely been tested and provided quite limited choices[24].

In this work, we further investigate the role of user-configured avatars compared to photo based and generic avatars in terms of user experience and embodiment.

All avatars are created using the same tool, so they follow the same level of anthropomorphism and realism. For clearer distinction, the words generic, customized or self-configured, and personalized or photo-generated are used throughout this work to refer to these three types. Additionally, this study's setup will use consumer-grade hardware evaluating a cost-efficient full-body tracking solution.

This paper is structured as follows: Section 2 gives an overview of related research on avatar visual fidelity and diversity in avatar appearance. Section 3describes the experimental setup and conduction of the study. Section 4) summarizes the results and and their discussion, including the limitations and implications on future research (Section 5).

2 Related Work

This section gives an overview of related works from the fields of embodiment and avatar visual fidelity

2.1 Avatar Visual Fidelity

The visual impression of an avatar is referred to as visual fidelity and can be divided into three dimensions [38, 64]: (i) anthropomorphism, (ii) realism and (iii) truthfulness.

In this context, anthropomorphism relates to the layout of the virtual body, the amount, position, and DoFs of its limbs, ranging from humanoid to non-humanoid, in terms of structure and motion range. A humanoid avatar makes it easier for the user to integrate the virtual body into the users' body schema. Layouts diverging from the users body schema are possible, as it is adaptable due to homuncular flexibility [95, 54]. It has been shown that it is possible to control different body layouts in VR like prolonged arms [50], tails [86], additional fingers [46], or a third arm [53].

Realism describes the style of how virtual objects and environments, are visualized [66, 41] and can range from abstract and cartoonish to photorealistic. The maximum level of detail is limited by the computing power and display technology. Zell et al. [97] found that the shape is the key dimension for perceived realism, as it directly interacts with anthropomorphism. An extreme example of this would be visualizing a hand with a single square pixel, which would make it hard without context to recognize it at all or use it intuitively. Regarding avatars, Fleming et al. [32] found that improvements in render quality had no significant impact on appeal ratings of a virtual avatar.

The truthfulness of a virtual object or avatar characterizes how close it is designed regarding its real-world counterpart, ranging from a generic representation, like choosing an avatar from a set of pre-defined avatars, to personalized, where the avatar would be fitted according to the user. When confronted with a virtual avatar that resembles the self, users are willing to get closer to such a self-avatar than they would with an avatar that is considered someone else [4]; just like approaching their own reflection in the mirror. Moreover, there is evidence that users are more engaged when using a truthful avatar and also prefer such an avatar over a generic one [60]. Waltemate et al. [89] showed that personalized avatars can increase presence, compared to generic avatars.

2.2 Sense of Embodiment

Sense of Embodiment (SoE) describes the feeling that an object (virtual or real) belongs to or replaces the own body or parts of it. Kilteni et al. [49] define SoE toward a body B as "the sense that emerges when B's properties are processed as if they were the properties of one's own biological body." The state of the own body can be perceived through different sensory channels: Visually, but also through spatial orientation by the vestibular system, proprioception, touch perception as well as the perception of temperature and pain. These channels are not possessed separately in the brain, but are rather integrated to build a reliable perception of the body and environment, known as multisensory integration [85]. The rubber hand illusion, which is based on a tactile and visual impression of a fake limb, was first described by Botvinick et al. [10]. They showed that it is possible to induce a feeling of disembodying the own arm and instead embodying a rubber arm, by providing a synchronous visuo-tactile impression on both the real and the fake arm. SoE over a virtual body is fostered, when the artificial stimuli fit the perceived stimuli, as it makes multisensory integration is easier. It has been shown to be linked to several factors [43], such as the avatar's appearance [3, 87, 88, 63, 61, 62, 89, 30], its control [90, 30], as well as the viewpoint [3, 9, 27, 57, 56, 71, 50]. However, also the locus of control (the degree to which people believe that they have control over the outcome of events or attribute them to external factors) [47, 20], use's preference [35, 14], and emotion [33] as well as vestibular signals [59, 9, 50, 8] and visuo-tactile stimuli [3, 10, 57, 81, 52, 69, 50] play a role in SoE.

Regarding virtual avatars in immersive VR setups, Kilteni et al. [49] suggest SoE is based on three components: (i) Sense of Self-Location (SSL), (ii) Sense of Ownership (SoO), and (iii) Sense of Agency (SoA).

Sense of Self-Location The experience of being inside a body and the perception of the spatial volume of that body is referred to as SSL [49]. Usually, one feels located inside their own physical body, however, out-of body illusions are also a topic of research [8]. SSL is based on visuospatial perspective (first-person perspective (1PP) or third-person perspective (3PP)), vestibular signals, as well as tactile input [49]. Although, we usually experience a 1PP in our body, it is possible to induce SSL from a 3PP [57, 56, 9]. However, the effect of SSL is stronger for first-person experiences. It could be shown, that when a virtual body is exposed to a threat, the physical reaction is stronger for 1PP than for 3PP [70, 71, 3].

Sense of Ownership When a body is attributed to the self, this evokes a sense of ownership (SoO) towards that body or body parts [37]. That body then is perceived to be the source of the sensations. Both top-down and bottom-up processes influence SoO [87, 88]. Regarding bottom-up influences, both spatial and temporal synchrony are important, such as temporal integration of visual and tactile information, as in the rubber hand illusion (RHI) [10, 87]. Furthermore, visuo-proprioceptive synchrony was also found to enhancy SoO [23, 87]. For top-down influences, the similarity of the virtual body in terms of anthropomorphism [3] or configuration [26, 87, 17] has been shown to be critical.

Sense of Agency While SoO can be experienced over a static body, sense of agency (SoA) refers to the sense of having "global motor control, including the subjective experience of action, control, intention, motor selection, and the conscious experience of will" [9]. SoA also depends on visuo-motor synchronicity and has been suggested to be linked to the correlation of the predicted sensory outcome of an action (efference copy) and its actual sensory outcome [49, 79, 91]. This effect appears to be independent of the consequences of the action, but rather limited to the sensory feedback of the action [79].

2.3 Presence, Social Presence and Co-presence

According to Slater, presence refers to the feeling of being in a place despite knowing one is not physically there [82]. The effect is influenced by the technical capabilities of the used system, such as setups that resemble real-world impressions more closely like displays with a large field of view (FOV), a high-resolution, or head tracking, to adjust the virtual field of view accordingly, as in HMDs, but presence can occur when watching a 2D movie. Slater [82] distinguishes two types of presence: (i) place illusion and (ii) plausibility illusion. Place illusion refers to the feeling of being in a specific place, despite knowing where one actually is. It can occur without a virtual body, i.e., in a 3D movie; however, being embodied in a virtual avatar (cf. Section **??**) enhances sensations that resemble the real world, increasing the place illusion [82, 78]. The plausibility illusion describes the feeling that actions in VR are actually happening, even when being aware that the actions are only virtual [82, 83]. Therefore, users' reactions to virtual threats are often similar to their real-world reactions, such as cringing when something unexpected happens.

Avatars can also increase *co-presence*, the sense of being together and becoming "accessible, available and subject to another" [39]. Co-presence consists of two aspects: Perceiving others and the feeling of being perceived by others [39], which usually without communication devices happens face-to-face, and can also be realized with virtual avatars.

Another concept of presence is social presence [80], that focusses on the perception of others and the relationship between individuals. The Media Richness Theory adds to this idea, that social presence depends on the communication media used, increasing with both extension of the covered information and their quality [18]. Moreover, it has been suggested, that intimacy increases with subtle communication cues such as eye contact, physical proximity, and the amount of smiling [2]. In that sense, immersive VR systems are a powerful communication tool, as they not only provide a high quality sensual impression, but some devices can also capture eye gaze and smile. Additionally, avatars make it easy to judge the interpersonal distance.

2.4 Uncanny Valley Effect

Since there is an ongoing effort to increase the human likeness of robots and avatars, one might assume a monotonous increase in the affinity felt for those agents. Nevertheless, as first described by Mori et al. [67], it is a now well-known phenomenon that this affinity drops to a negative level before the look of a healthy human is perfectly recreated (see Fig. 2). This invokes an eerie and spine-tingling sensation. This minimum in affinity has been termed by Mori et al. as the uncanny valley. For example, this sensation may be invoked by a prosthetic hand that imitates the look of a human hand.

This effect can even be increased if movement is incorporated, like in the current study, increasing the peaks but also creating an even deeper valley, e.g., representing the human likeness and affinity felt for zombies. Mori et al. explain the cause of this by a human instinct of self-preservation since the eerie sensation protects us from the danger of corpses, for example. As the human's visual appearance changes after death, decreasing the human likeness, it puts a dead body in the uncanny valley, providing those with a self-preservation advantage that instinctively avoid entities in that region of human likeness.

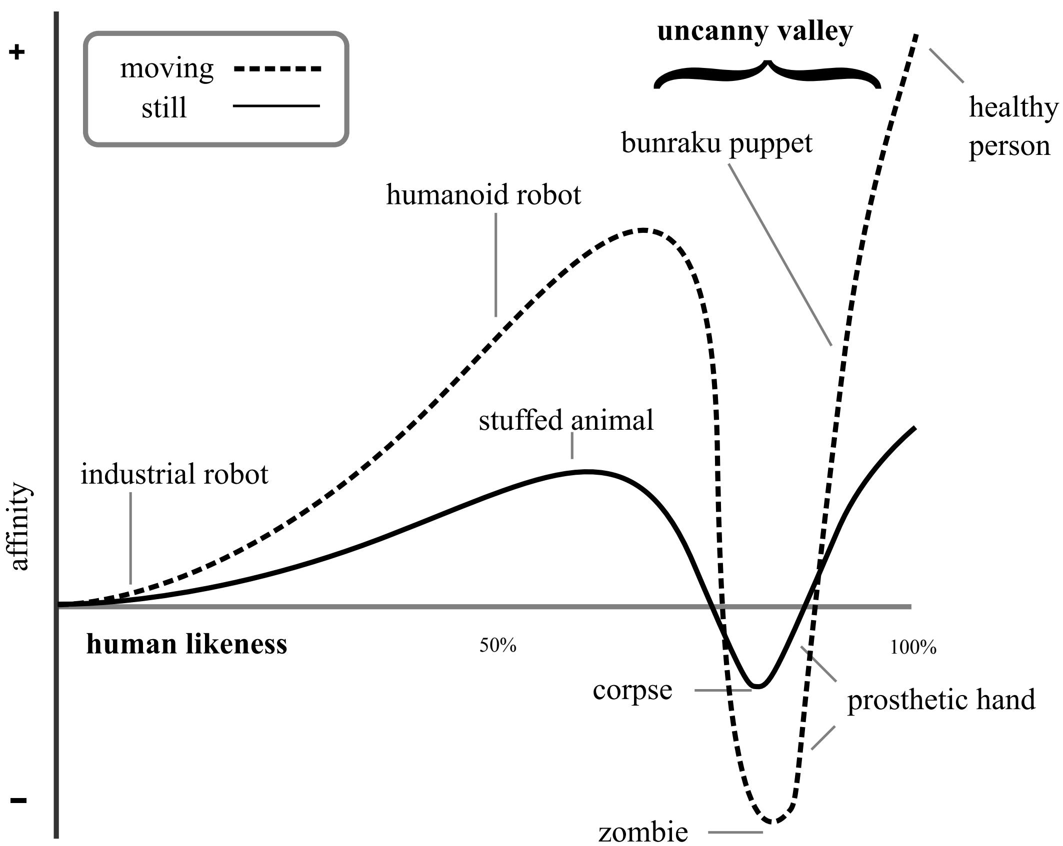

Figure 2: Relationship of human likeness and affinity according to Mori et al. showing a large drop in affinity close to the human likeness of a healthy person. The effect is increased for moving entities. Image by Smurrayinchester distributed under CC BY-SA 3.0 based on the figure in [67].

2.5 Sense of Embodiment for Self-Avatars Depending on Avatar Visual Fidelity

While the sense of embodiment is influenced by multiple factors, in this work we focus on the effects of differing levels of avatar visual fidelity, particularly in full-body representations. For example, Maselli et al. found realism an important top-down factor for an ownership illusion that diminished the need for bottom-up factors like synchronous stimulation [65].

Waltemate et al. [89] have considered this in terms of different levels of personalization. They provided users with three different avatar representations as a within factor, one generic hand-modeled, one non-personalized created with photogrammetry, and one created by a photogrammetry scan of the participant. In the HMD condition, participants were then immersed in a virtual reality room using those avatars from a first-person perspective in front of a virtual mirror with full-body motion tracking and received multiple movement instructions that incentivized them to pay attention to the visual appearance of their bodies. Sense of embodiment and emotional response were assessed with several mid- and post-immersion questionnaires. In accordance with their initial hypothesis, they found ownership, presence, and dominance (according to SAM) to be significantly higher in the personalized scanned condition than in the other two. However, the two generic conditions did not differ in most measures. For that reason, they attributed the effect to the factor of personalization (or truthfulness) as the scanned conditions did not differ in realism or anthropomorphism. Agency did not show any significant effect on personalization. Waltemate et al. did not encounter any uncanny valley effects, but they did not implement any measures for that.

This aligns with the findings of Salagean et al. [77] comparing different levels of personalization and photorealism in a 2×2 setup. The setup used in their experiments was similar to that of [89] with a full-body tracked first-person avatar in front of a mirror and multiple motion instructions. They found a higher sense of ownership in the personalized (scanned) condition with high photorealism compared to all other ones and, in general, a positive effect of truthfulness and realism on the sense of embodiment. As an implicit measure of ownership, they induced a threat by a heavy box falling down near the avatar and measured the skin conductance response, but this did not reveal a significant effect.

Döllinger et al. [24] used three different truthfulness levels with a generic and a personalized photogrammetric avatar as extremes and a customized one with different, prepared body characteristics that allowed users to choose an avatar that matched their physical appearance from 67 representations in total. Introducing the UVI to the questionnaires, they explicitly found photogrammetric personalized avatars to invoke a greater feeling of eeriness, although the sense of ownership was better.

Gorisse et al. continued this research by implementing a setup with a third-person perspective [42]. The conditions they realized used the same level of realism and anthropomorphism but differed in the degree of truthfulness. However, the first two versions used robotic representations without any similarities to the participant that either could or could not contain the human body due to their outer shape. Only the third version depicted a scanned version of the user's head. Similar to the previous results, they found the avatar version with the scanned representation of the user to induce a higher subjective SoO and abstract ones to allow for a stronger disconnection and risky behavior. However, some participants gave feedback hinting at a greater feeling of eeriness for the personalized avatar.

While most studies that investigate the effect of personalization focus solely on photogrammetric reconstructions, Jo et al. [48] added a cartoonish representation created by a 3D artist for each participant and found the sense of ownership to be best compared to a generic and a highly realistic representation from a photo. However, they did not use tracking; instead, they used a Wizard-of-Oz setup that required participants to follow the avatar's movements.

Only very few previous works have examined the effects of customization without scanning, or even with self-configuration in avatar depiction. However, they either focused on rather realistic renderings and constrained the choice to match the actual appearance [24] or did not implement a full-body tracking setup and lacked the comparison with a personalized condition on the same realism level [48]. The results of these studies remain inconclusive in terms of customization versus photorealistic personalization and miss any investigation of giving participants complete freedom over their avatars, only focusing on congruence with their real-world visual appearance.

2.6 Diversity in Avatar Appearance

As previously discussed, the visual appearance of an avatar and its similarity to the user impact the VR experience. Do et al. [21] aimed to investigate the importance of matching main features like gender or ethnicity. They provided users with four different avatars, each either matching or not matching their gender or ethnicity, that the participants embodied in front of a virtual mirror while performing multiple motion tasks. They found that matching ethnicity positively influences the sense of embodiment in general, and matching gender contributes to the sense of ownership.

Despite its necessity, Do et al. [22] found a lack of representation of different ethnic backgrounds in existing avatar libraries and identified the need for such a library focusing on diversity. To tackle this, they averaged publicly available photos as a basis for modeling 42 base avatars of different ethnic groups, each equipped with five different outfits, resulting in a total of 210 models. These avatars were then refined in an iterative interview process and validated via online surveys. The evaluation showed that most avatars were correctly attributed to the intended group. In their future work, they state that additional development is necessary to include more ethnic backgrounds, professions (by outfits), body types, age ranges, and gender since only two genders are currently included in the library.

2.7 Embodied avatars as virtual identity

People's rationale for choosing the visual appearance of their virtual self-avatar can have several motivations [15]. In the context of immersive VR setups with HMDs and full-body tracking, reasons include aesteticism, showing an aspect of their identity, and adapting to or or distinguishing from a social group [34, 7, 58]. Creating an avatar enables people to decide which aspects of their body to include in the avatar or not, such as gender [34] or disability [98]. While dissimilar avatars that hide aspects of the self can be chosen to avoid social exclusion or harassment [7, 34, 98], they can also be used to entirely change the own identity to facilitate communication [7, 34] or to explore an ideal version of the self [58].

3 Study

We conducted a study to investigate the effect of avatar visual fidelity on embodiment and user experience.

3.1 Participants

Thirty people took part in the study. They were offered credit for their participation if applicable to their course of study. Seven identified as female, four as non-binary or genderqueer, and nineteen as male. Six of the participants were left-handed. The mean age was 24.93 ± 3.97 years. 22 had a computer science, six a psychology background. Ten+ people used glasses and two contact lenses for vision correction.

25 participants had used a head-mounted display before. Except for four people who used it almost once a month and one who used it almost daily, all of them used it once every half a year or less. While 13 participants indicated experience with hand tracking technologies before the study, only three mentioned any previous usage of full-body tracking.

3.2 Technical Setup

The study conducted using the Meta Quest 2 and 3 HMDs with a resolution of 1832×1920 or 2064×2208, respectively, and a refresh rate of 120 Hz. The VR application was created using Unity and ran on a PC equipped with an Intel Core i9-10900K CPU with $10 \times 3.70\ GHz$, 32 GB RAM, and an NVIDIA GeForce RTX 3090 running Windows 10.

The participants' motions were tracked with the Sony Mocopi system [84] that uses six sensors with accelerometers and gyroscopes placed at the head, wrists, hip, and ankles. The system estimates the position of the body parts using machine learning. This allows for an easy and cost-efficient motion-tracking solution. However, tracking errors will build up over time, and recalibration is needed to correct them. The sensor data was transferred via Bluetooth to an iPhone 15 Pro and sent to the PC via a Wi-Fi connection, since the system does not allow for a direct connection between sensors and the PC.

As the position and rotation of the character pose outputted by the Mocopi system did not align with the head pose from the HMD that was used for calculating the perspective in virtual reality, a position and rotation offset had to be calculated to avoid the user's view being placed outside the avatar's head. For doing so, the position of the avatar's head was fixed to the current head position calculated by the HMD. To compensate for rotational drift, the rotation offset around the vertical axis was constantly calibrated every five seconds as long as the viewing direction was within a 45 degree angle up or down from the horizontal axis, and the angular velocity did not exceed 5 degrees per second to avoid glitches by edge cases when looking up or down and during fast movements due to differing latencies. The avatar was manually resized according to the body height of participants specified during preparation.

The sliding feet problem (see [93, 92]) would have been apparent here by a user looking down and moving their head, causing their whole body, including their feet, to move too, although their real feet remained in place on the ground. That could have happened as the body's position is fixed to the HMD's position and the latencies differ between the HMD and the other sensors. To compensate for that, a dual approach was used for the feet positioning: As long as the feet were not directly in the field of view of the user, the joint rotations of the legs and feet were directly taken from the Mocopi system's output as this provided the more natural-looking leg configuration and movement in the mirror. But if the feet became directly visible, they were fixed in place, and only the positional and rotational offset from that point on were used from the Mocopi system with the legs moved according to Unity's inverse kinematics.

The Meta Quest Headset tracked the positions of hands and fingers via inside-out cameras and image processing. The avatar's arms were placed according to Unity IK. In case the hands moved outside the tracked area or other tracking loss, the data from the Mocopi system was used as a fallback (similar to [93]) for the hand position smoothly morphing between the potentially differing positions in 0.5 s. The Meta Quest hand tracking data was also used to adapt the avatar's hand size to the participant automatically.

To give the face a lifelike appearance, the HMD microphone captured the user's speech and transformed it into matching lip movements, facilitating the Oculus Lipsync plugin[1].

Within VR, users were placed in a simple room in front of a virtual mirror, implemented with the Magic Mirror Pro Unity package. Instructions and questions were played back via audio and simultaneously displayed at the top of the mirror, similar to [77].

3.3 Procedure

3.3.1 Preparation

After giving informed consent, the participants created their avatars to be used in VR, depending on the condition order. This order of conditions was counterbalanced.

In the generic condition A, they were offered 13 avatars that were created with the ReadyPlayerMe avatar creation tool from images of a selection of the VALID avatars [22] to offer various ethnicities and different genders and avoid a white male centered bias [21], while deliberately limiting their choice to very little options. Participants were instructed to choose an avatar that resembled their visual appearance the best, eliminating the necessity to make assumptions about the person's gender or ethnic background.

In customized condition B, participants were allowed to create their avatar with ReadyPlayerMe, giving them complete freedom over the look of their character. They were given the explicit instruction: "Please create the avatar in the way you would like to be represented in VR. This may coincide with your real-world visual appearance, but it does not have to." However, they were told not to use any assets that impaired their vision or theoretical freedom of movement in any way.

In the photo-generated condition C, a photograph was taken using an iPhone 15 Pro in front of a white background and uploaded to ReadyPlayerMe to create an avatar that, while having the same cartoonish style as the other ones, resembled the participant as closely as possible. After the automatic character creation process, participants were asked if they felt that the visual appearance of the avatar

[1]https://developers.meta.com/horizon/downloads/package/oculus-lipsync-unity/

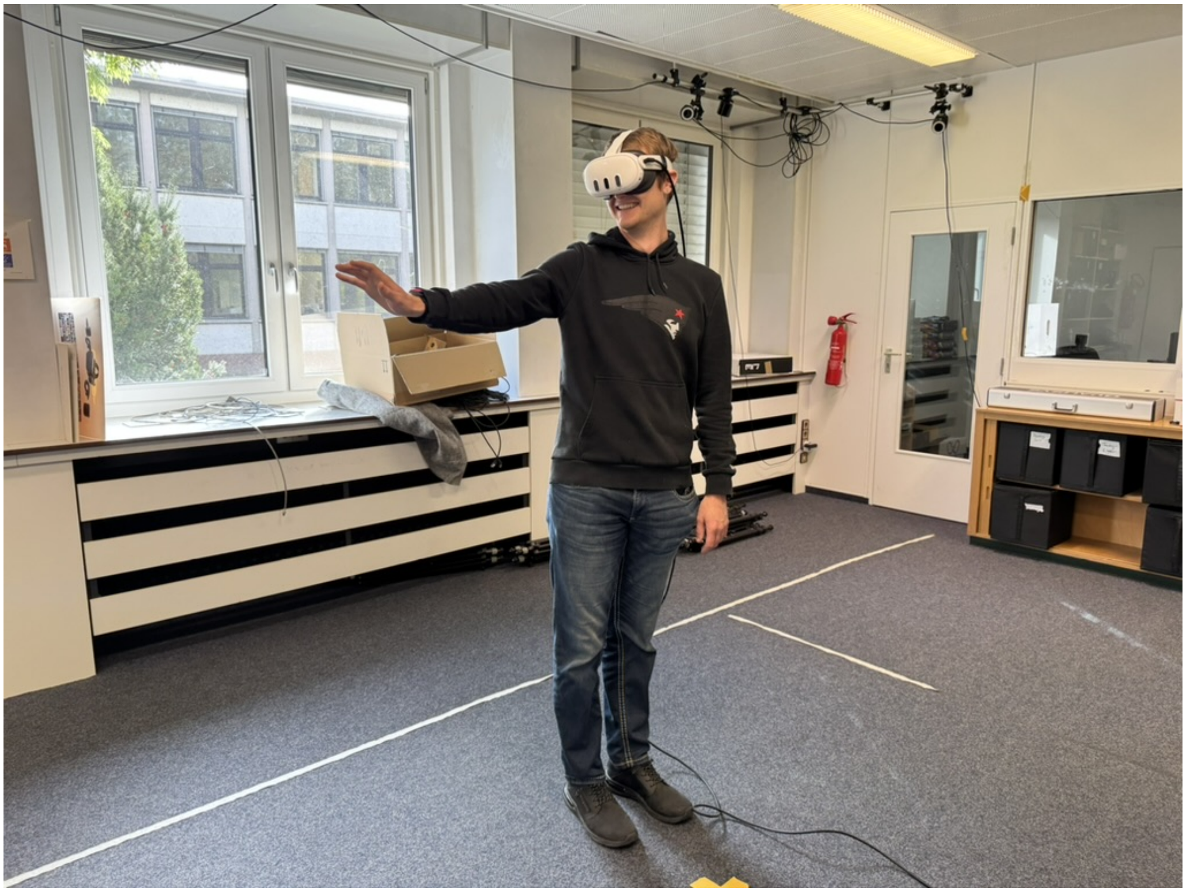

Figure 3: Participant during the VR task performing one of the specified motions

matched their own and offered the possibility to correct features that differed, especially if attributes like gender or hair color were incorrectly inferred from the image.

Following the character creation, the participants filled out a preliminary questionnaire. Then, the participants were equipped with the Mocopi sensors and the VR headset and were guided through the necessary calibration steps. If necessary, this calibration was repeated before each condition to limit the accumulated drift.

3.3.2 VR Exposure

Each condition followed the same structure, only differing in the avatar. Participants started in a virtual room with a mirror initially turned off. They were instructed to stand at a distance of about one meter from the mirror and give verbal answers to the mid-immersion questions, which were immediately transcribed by the experimenter.

After that, the mirror was activated, and a set of movement instructions adapted from Waltemate et al. [89] was played back and simultaneously displayed at the top of the mirror (see Fig. 4). Instruction 3 was adapted as the Mocopi system had problems with the detection of high knees.

1. "Lift your right arm and wave to your mirror image in a relaxed way."
2. "Now wave with your other hand."
3. "Now walk in place."
4. "Now stretch out both arms to the front and perform circular movements."
5. "Now look right, stretch out your right arm to the side, and perform circular movements."
6. "Now look left, stretch out your left arm to the side, and perform circular movements."

Figure 4: Participant's view during a movement task.

During this, participants were repeatedly reminded to alternate between looking at the position of their virtual body and the mirror image following Waltemate et al. [89] to ensure that they perceived the visuo-motor synchrony between their movements and their virtual representation supporting an ownership illusion [40]. In addition, it led to them actively looking at their virtual representation to maximize the time during which they could perceive differences between the conditions. An exception was made for the instructions where participants needed to stretch their arms to the side, as this would have caused their hands to be outside the sensory field of the Meta Quest inside-out cameras (see Fig. 3).

To assess an implicit measure of the sense of ownership, after a randomly selected movement task, a virtual box was dropped left or right next to the participant to induce a potential threat reaction, similar to the procedure in Salagean et al. [77]. The subjective threat impression was captured by a mid-immersion question immediately following the stimulus (see 3.4.1).

When all motion tasks were finished, the participants were confronted with four questions to evaluate their impressions while still immersed in the virtual environment with their avatar (see Section 3.4.1 for details).

At the end of each VR exposure phase, participants were asked to place themselves at a distance from the mirror they felt the most comfortable with, which was then recorded by the experimenter with a button press. By that, the interpersonal distance chosen to the self-avatar in the mirror was recorded (see Section 3.4.6).

After each exposure, participants took off the headset, filled out a post-immersion questionnaire, described in Section 3.4.2ff., and were offered a short break. Finally, they filled in a general questionnaire regarding the experiment, could provide feedback, and ask further questions about the research purpose. The whole session, including the preparation phase, took one to two hours.

Table 1: Selected items of the VEQ+ questionnaire[31] used in this work.

Dimension	Question
Ownership	It felt like the virtual body was my body.
	It felt like the virtual body parts were my body parts.
	The virtual body felt like a human body.
	It felt like the virtual body belonged to me.
Agency	The movements of the virtual body felt like they were my movements.
	I felt like I was controlling the movements of the virtual body.
	I felt like I was causing the movements of the virtual body.
	The movements of the virtual body were in sync with my own movements.
Change	I felt like the form or appearance of my own body had changed.
	I felt like the weight of my own body had changed.
	I felt like the size (height) of my own body had changed.
	I felt like the width of my own body had changed.
Self-Similarity	The appearance of the virtual human's face was similar to mine.
	The overall appearance of the virtual human was similar to me.
	I felt like the virtual human resembled me.
	The appearance of the virtual human reminded me of myself.
Self-Attribution	I felt like the virtual human was me.
	I could identify myself with the virtual human.
	I had the feeling the virtual human was behaving as I would behave.
	I felt like the virtual human had the same attributes as I have.

3.4 Measures

3.4.1 Mid-Immersion Ratings

While still immersed in the virtual environment, participants were presented with five questions or statements in total. All statements had to be answered on a seven-point Likert scale from 1 to 7. The first one was asked immediately after the induced threat by the falling box to rate the statement: "I felt that my own body could be affected by the box." It served as an implicit measurement of the sense of ownership like in [77] since a strong feeling of ownership has been shown to invoke a more intense reaction to a threat [3].

After the movement tasks were completed, all the other questions were asked to picture the feeling after a consistent time of self-observation and synchronous visuo-motor stimuli. Three of the questions were one-item measures for the components of sense of embodiment that also have been used in previous studies during the VR exposure [89, 77]: "To what extent do you have the feeling as if the virtual body is your body?" "To what extent do you have the feeling that the virtual body moves just like you want it to, as if it is obeying your will?" "To what extent do you feel present in the virtual environment right now?" The last question was complemented with an explanation of the concept of presence as the feeling of actually being there in the virtual environment.

The final question addressed the attractiveness of the self-avatar to provide a mid-immersion counterpart to the corresponding post-immersion question (see Section 3.4.5 for details).

3.4.2 Virtual Embodiment Questionnaire

To assess the degree of the sense of embodiment and self-identification, the VEQ+ questionnaire [31] was used after each of the three conditions' VR phase. The Virtual Embodiment Questionnaire (VEQ) it is based on was initially developed by Roth and Latoschik [76] as a standardized and validated way to measure the sense of embodiment in terms of the sense of ownership, sense of agency, and change. Each item has to be rated on a 1 to 7 Likert scale from "strongly disagree" to "strongly agree". The values of each dimension are meant to be averaged to calculate the scores. Fiedler et al. enhanced the questionnaire by including items on self-identification related to visual similarity and personal characteristics [31]. In this work, questions about self-location were excluded, as the virtual body always aligned with the physical body. The selected VEQ+ items used in this study are listed in Table 1.

Table 2: Adjective pairs of the UVI[45]

Dimension	Adjective Pair
Humanness	Inanimate – Living
	Synthetic – Real
	Mechanical Movement – Biological Movement
	Human-Made – Humanlike
	Without Definite Lifespan – Mortal
Eeriness	Dull – Freaky
	Predictable – Eerie
	Plain – Weird
	Ordinary – Supernatural
	Boring – Shocking
	Uninspiring – Spine-tingling
	Predictable – Thrilling
	Bland – Uncanny
	Unemotional – Hair-raising
Attractiveness	Ugly – Beautiful
	Repulsive – Agreeable
	Crude – Stylish
	Messy – Sleek

3.4.3 Uncanny Valley Index

As the human-like avatar representations may be prone to invoke a repelling sensation typical for the uncanny valley, the post-immersion survey contains the Uncanny Valley Index (UVI)[45] questionnaire to monitor this. It consists of 18 adjective pairs in the three main categories humanness, eeriness, and attractiveness, that are answered on a seven-point Likert scale. The adjective pairs are given in Table 2.

3.4.4 Self-Assessment Manikin

To measure the emotional response to the three presented avatars, the Self-Assessment Manikin questionnaire (proposed by [11] and extended by [13]) was used. It consists of three scales – valence, arousal, and dominance – with 9 points and corresponding pictures allowing for a non-verbal response.

Previous research has already compared the emotional reaction to external stimuli in an immersive environment between a high and low ownership condition [36]. The study showed an increase in arousal and dominance as well as valence for positive stimuli in the high ownership condition. This supports the use of emotional response not only as a detached measurement but also as an implicit measure of the sense of embodiment.

3.4.5 Additional Questions

Three additional single questions complemented the previously described and validated questionnaires. The first one asked about the perceived attractiveness of the avatar: "To what extent did your avatar seem attractive to you?" This question captured satisfaction with the avatar's visual appearance. Previous research has shown a positive correlation between attractiveness and the level of truthfulness [41]. However, as many other measures used in this work, it has not been investigated in combination with freely self-configured avatars.

In addition, in this work we specifically investigate the participants' freedom of design and self-expression ("To what extent did you feel that you could express yourself in the virtual world?"). Also, all participants were given a chance to express free-form feedback after each condition concerning the specific avatar and general thoughts on their experience after finishing the last condition.

3.4.6 Mirror Distance

As already referred to in Section 3.3.2, while still immersed in the virtual setup, participants were asked to place themselves at a comfortable distance from the mirror, thereby determining a distance to the mirror image of their avatar. Previous studies have used this measure of interpersonal distance for social behavior in the company of virtual agents: They have found people to maintain a larger distance to agents compared to inanimate objects, but this distance decreased if the self-avatar was more attractive or the agent was presented with the participants face [96, 5, 6]. This may indicate interpersonal distance as an implicit measure of self-identification.

3.5 Design and Hypotheses

A within-subject design with a counterbalanced order of the three conditions with different avatar creation workflows was used, with either a limited choice between generic avatars, a completely customizable, or an avatar generated based on a photo. The level of personalization or customization was the independent variable. As it is apparent from the previous section, several dependent variables are recorded during the process: First, there is the sense of embodiment assessed explicitly both mid- and post-immersion in combination with self-identification as well as implicitly by the threat reaction and distance to the virtual self in the mirror. Due to the differing customization possibilities and different phrasing of the tasks in the condition, the following hypotheses are formulated:

H1 The photo-generated condition invokes the highest self-similarity.

H2 The customizable avatar shows a higher sense of embodiment.

H2.1 The higher sense of embodiment increases reaction to the threatening stimulus.

H2.2 The higher sense of embodiment decreases distance to the virtual self in the mirror.

The second hypothesis is based on the assumption that the participant's ability to customize their avatar freely to match their body image without necessarily showing their real-world appearance reduces the impact of the uncanny valley effect. By avoiding appearances that invoke an eerie sensation, they could choose more attractive features. As the UVI was also captured, this assumption is covered in **H3**.

H3 The uncanny valley reaction is lower in the customizable condition than in the photo-generated.

H4 The customizable is perceived to be more attractive than the other avatars.

As the customizable condition gives more power over their preferred representation to the user, not being constrained by the number of options or their real-world appearance, it is also expected to create better ratings in the questions related to emotional reaction and user experience:

H5 The customizable avatar increases the emotional response, especially valence, compared to the other conditions.

H6 The freedom of choice results in a higher rating of the possibility to express oneself.

4 Results

The analysis of the results was conducted in R. The Shapiro-Wilk test was used to check for normal distribution. If the data was normally distributed, repeated measures ANOVA with optional sphericity correction was used to check for significant differences. In the non-normally distributed case, the Friedman test was performed. In cases with significant differences, a pairwise t-test or a paired Wilcoxon test was conducted depending on the result of the preliminary test. An α of 0.05 was used for all statistical tests.

The questionnaires used to monitor the usability of the system revealed no issues interfering with the validity of the study: The System Usability Scale [12] resulted in a score of 74.2 which can be interpreted as good usability. The NASA TLX [68] score of 16.38 is inconspicuous. The VR Sickness Questionnaire [51] showed significant – according to the Wilcoxon signed-rank test with paired samples – but only little increases from the pre- to the post-questionnaire for eyestrain ($W = 0, p = 0.00601$, pre: $M = 0.1 SD = 0.305$, post: $M = 0.467 SD = 0.629$) and blurred vision ($W = 0, p = 0.0477$, pre: $M = 0.033 SD = 0.183$, post: $M = 0.233 SD = 0.504$).

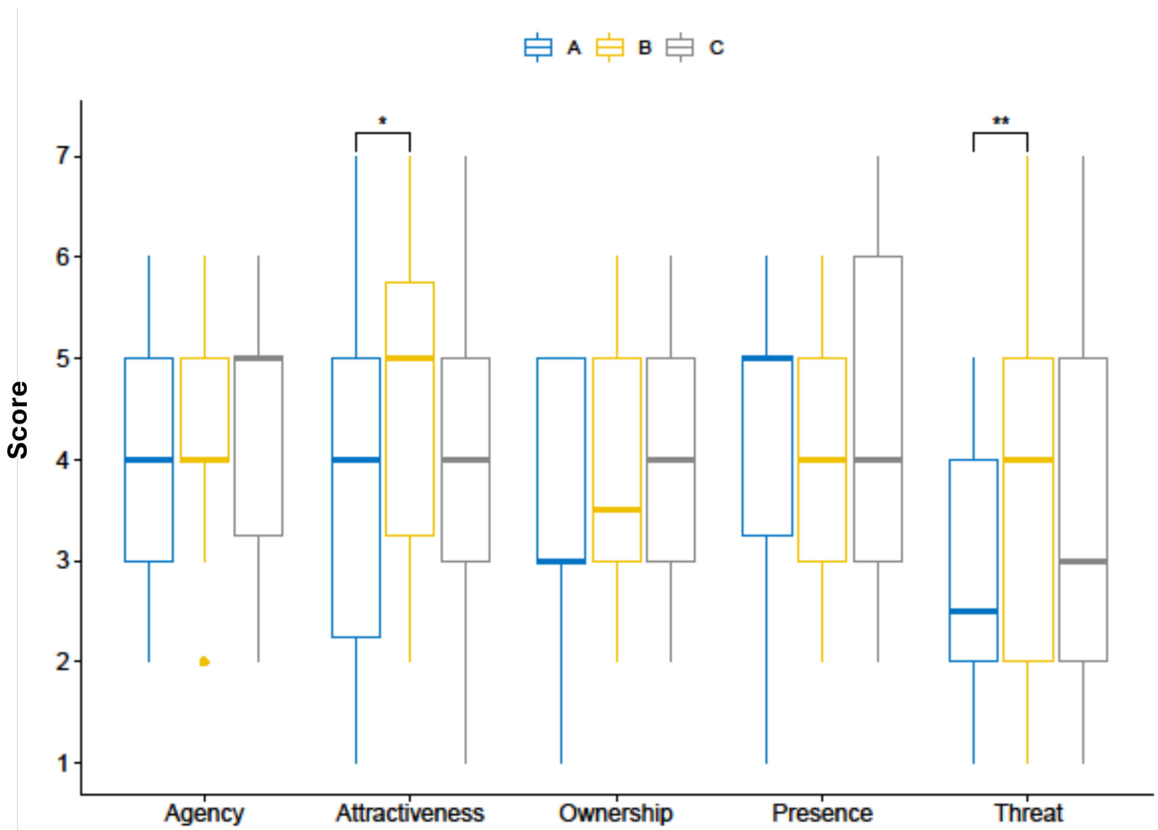

Figure 5: Boxplot of the mid-immersion ratings on agency, attractiveness, ownership, presence, and subjective threat by the falling box on a scale from 1 to 7. A: generic avatar, B: customized avatar, C: photo-generated avatar. $*$ and $**$ indicate significant differences according to pairwise signed-rank Wilcoxon test with Bonferroni adjustment with $p < 0.05$ or $p < 0.01$, respectively.

4.1 Mid-Immersion Ratings

While still immersed in the VR setup, participants were asked several questions regarding their user experience and sense of embodiment as stated in Section 3.4.1. A box plot of the results is given in Fig. 5. As the Shapiro-Wilk test showed a violation of the normal distribution, non-parametric tests were used to analyze the results.

A Friedman test revealed no significant effect of avatar fidelity in agency ($\chi^2 = 1.47, p = 0.481$), ownership ($\chi^2 = 2.70, p = 0.259$), and presence ($\chi^2 = 1.11, p = 0.573$). However, it showed a significant effect of this factor on the rating of attractiveness ($\chi^2 = 8.30, p = 0.0158$) and subjective threat by the falling box ($\chi^2 = 12.9, p = 0.00161$). A Friedman test conducted to investigate the factor of order revealed no significant effect of order for any of the questions.

For further investigation, the Wilcoxon signed-rank test with paired samples and Bonferroni correction was applied as a post-hoc test to the measures with a significant effect. It revealed that participants rated the attractiveness of the avatar created by themselves (Condition B, $M = 4.47, SD = 1.61$) significantly higher than the generic avatar (Condition A, $M = 3.83, SD = 1.62$) ($W = 38.5, p = 0.034$). The effect size was calculated as $r = 0.489$, which can be interpreted as a large effect according to Cohen [16]. In addition, it was found that the subjectively perceived threat by the falling box in Condition B ($M = 3.73, SD = 1.78$) was higher than in Condition A ($M = 2.8, SD = 1.47$) ($W = 4, p = 0.000852$) with an effect size of $r = 0.646$ which can be interpreted as a large effect. All other post-hoc tests conducted on attractiveness and threat did not result in the finding of a significant effect.

4.2 Post-Immersion Ratings

After each of the three VR exposures, participants were asked to complete several surveys to get a detailed view of their emotional response, user experience, and sense of embodiment.

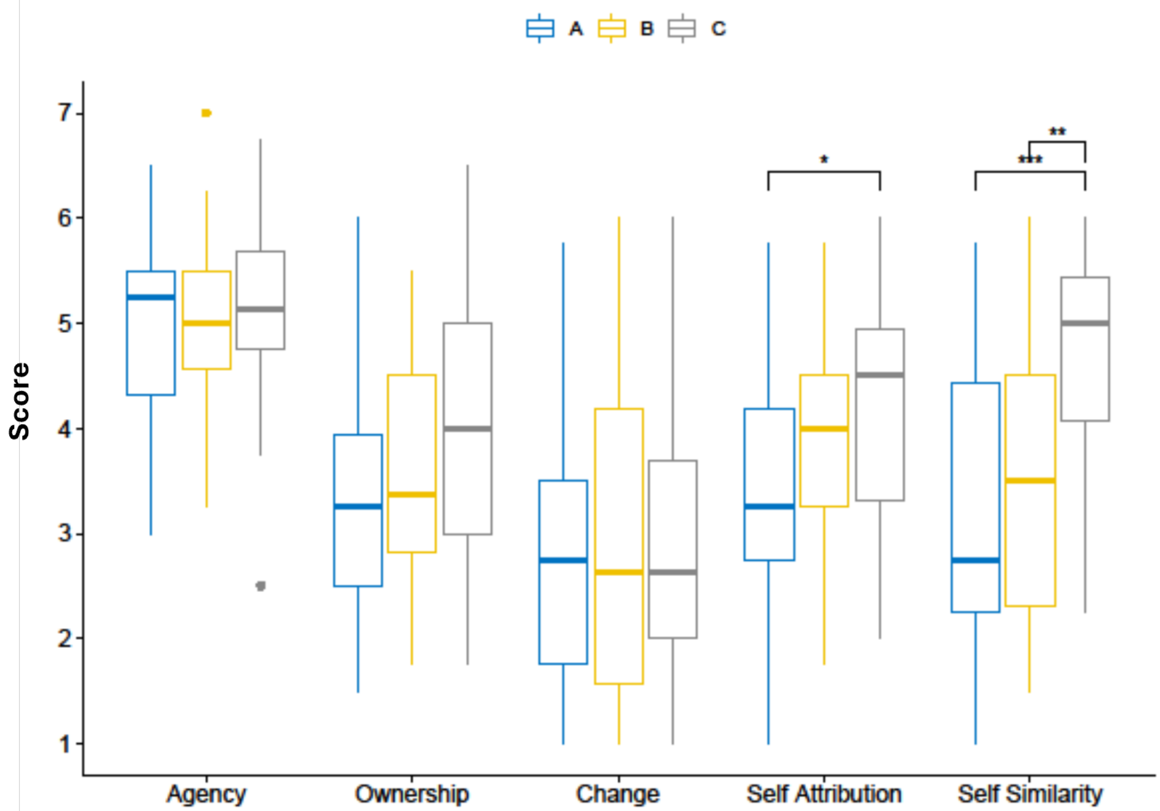

Figure 6: Results of the VEQ+ completed immediately after each VR exposure phase on agency, body ownership, change, self-attribution, and self-similarity. Ratings were given on a scale from 1 to 7. A: generic avatar, B: customized avatar, C: photo-generated avatar. Statistically significant differences were checked by a paired t-test for self-attribution, and a signed-rank Wilcoxon test for self-similarity, both with Bonferroni correction applied. $*, **, ***$ indicate significance levels of $p < 0.05, p < 0.01, p < 0.001$.

4.2.1 Virtual Embodiment Questionnaire

To record the sense of embodiment and the experienced self-similarity as well as self-attribution, the participants had to answer a questionnaire adapted from the VEQ+ [31] as it is described in Section 3.4.2. Answers were aggregated by mean for each participant and the five dimensions: agency, body ownership, change, self-attribution, and self-similarity (see Fig. 6). The Shapiro-Wilk test showed a violation of the normal distribution for questions regarding change and self-similarity in some conditions. The results of these categories were analyzed using the non-parametric Friedman test.

As all other answers followed a normal distribution, repeated measures ANOVA was used. It revealed a significant main effect of avatar visual appearance for self-attribution ($F(1.93, 55.9) = 4.83, p = 0.013$). With an $\eta^2 = 0.093$, the effect size can be interpreted as medium. No significant main effect of avatar visual fidelity was found for the other dimensions, i.e., agency ($F(1.9, 55.2) = 0.473, p = 0.616$) and body ownership ($F(1.94, 56.2) = 2.50, p = 0.093$). A pairwise comparison with a paired t-test with Bonferroni adjustment was conducted as a post-hoc test to investigate the particular differences between the conditions. It showed the generic avatar (Condition A, $M = 3.42, SD = 1.51$) to have a significantly lower score in self-attribution than the avatar generated by photo (Condition C, $M = 4.21, SD = 1.50$) ($t(29) = -2.83, p = 0.025$). A calculation of Cohen's d resulted in $d = -0.517$, which indicates a moderate effect size.

The non-parametric Friedman test was used to analyze self-similarity and change, which showed a significant main effect of avatar visual fidelity for self-similarity ($\chi^2 = 13.6, p = 0.00110$). A signed-rank Wilcoxon test with paired samples and Bonferroni adjustment was conducted to further explore which differences between the samples contributed to this effect. This indicated that self-similarity ratings were significantly higher using the photo-generated avatar (Condition C, $M = 4.68, SD = 1.24$) compared to the generic avatar (Condition A, $M = 3.18, SD = 1.37$, $W = 31, p = 0.000166$) and the self-configured avatar (Condition B, $M = 3.52, SD = 1.51$, $W = 69, p = 0.007$). The effect size was calculated to be $r = 0.741$ comparing Condition C to A and $r = 0.570$ comparing Condition C to B, both of which can be interpreted as large effects.

In all other dimensions, no significant effect of avatar visual fidelity was found. While ANOVA revealed no statistically significant effect for body ownership ($F(1.94, 56.2) = 2.50, p = 0.093$), a visual inspection of the results using a box plot (see Fig. 6) suggested a possible trend in body ownership, which is supported by a Cohen's d $d = 0.301$ for the comparison of Condition C and A indicating a small effect size. However, as stated before, it is not strong enough to be considered statistically meaningful, which is also confirmed by a t-test with Bonferroni adjustment ($t(29) = -1.65, p = 0.33$).

As a visualization of the data (see Fig. 6) might suggest a difference in variance for change and self-similarity, this was checked with Levene's test, revealing no difference in variances in any category.

The preceding tests were repeated with order as the considered factor to check for potential effects of condition order. In some cases, the Shapiro-Wilk test showed a violation of the normality assumption in agency and change. Those dimensions were analyzed using the non-parametric Friedman test. For the other ones, repeated measures ANOVA was used. The ANOVA showed no statistically meaningful effect of order for body ownership ($F(2, 58) = 0.0647, p = 0.9374$), self-similarity ($F(2, 58) = 0.4705, p = 0.627$), or self-attribution ($F(2, 58) = 0.7714, p = 0.4671$). Similarly, no significant main effect of order was found with the Friedman test for agency ($\chi^2 = 2.25, p = 0.325$) and change ($\chi^2 = 4.93, p = 0.0849$).

4.2.2 UVI Questionnaire

To assess the impact of the uncanny valley effect, participants filled out the Uncanny Valley Index questionnaire [45] already described in 3.4.3. The results were aggregated by mean for each question dimension: Attractiveness, eeriness, and humanness (see Fig. 7). A preliminary Shapiro-Wilk test showed a violation of the normal distribution assumption in all three dimensions, so non-parametric tests were used for analysis.

A Friedman test showed a significant main effect of avatar visual fidelity for eeriness ($\chi^2 = 13.8, p = 0.00101$) and humanness ($\chi^2 = 6.39, p = 0.0410$) while attractiveness revealed no clear effect ($\chi^2 = 4.88, p = 0.0870$). To further investigate these effects, the signed-rank Wilcoxon test with Bonferroni correction was used as a post-hoc test to check for significant pairwise differences. Considering the dimension of eeriness, it was found that the ratings were significantly higher for the self-configured avatar (Condition B, $M = 3.79, SD = 1.06$, $W = 32.5, p = 0.000324$) and the photo-generated avatar (Condition C, $M = 3.28, SD = 0.74$, $W = 78, p = 0.008$) compared to the generic avatar (Condition A, $M = 2.84, SD = 0.688$). The effect size was calculated as $r = 0.701$ for comparing Conditions A and B and $r = 0.552$ for comparing A and C, both of which can be interpreted as a large effect. However, the pairwise comparisons of the conditions regarding the humanness ratings did not reach statistical significance, which might indicate that the sample size was too small to detect a specific difference between any two conditions. A visual inspection of the attractiveness results might suggest further differences that are supported by a calculation of the effect size ($r = 0.463$ for Conditions A and B, $r = 0.334$ for Conditions B and C). Nevertheless, it has to be stated that as the differences did not reach the significance level ($\alpha = 0.05$), so they cannot be considered statistically meaningful, and further research is necessary to check this.

4.2.3 Self-Assessment Manikin

To investigate the emotional response to the VR exposure depending on the avatar visual fidelity, participants completed the Self-Assessment Manikin questionnaire [11] described in 3.4.4 by selecting on a scale from 1 to 9, which picture mostly resembles their emotional reaction to the preceding experience (for a visualization of the results see Fig. 8). The Shapiro-Wilk test showed several violations of the normality assumptions through all questions and conditions, so non-parametric tests were used to check for statistical significance.

A Friedman test indicated a significant main effect for valence ($\chi^2 = 6.44, p = 0.0400$) while arousal ($\chi^2 = 3.27, p = 0.195$) and dominance ($\chi^2 = 4.63, p = 0.0990$) did not reach significance. As post-hoc tests, a pairwise comparison with the signed-rank Wilcoxon test revealed that the self-configured avatar (Condition B, $M = 6.33, SD = 1.40$) was significantly higher than the generic avatar (Condition A,

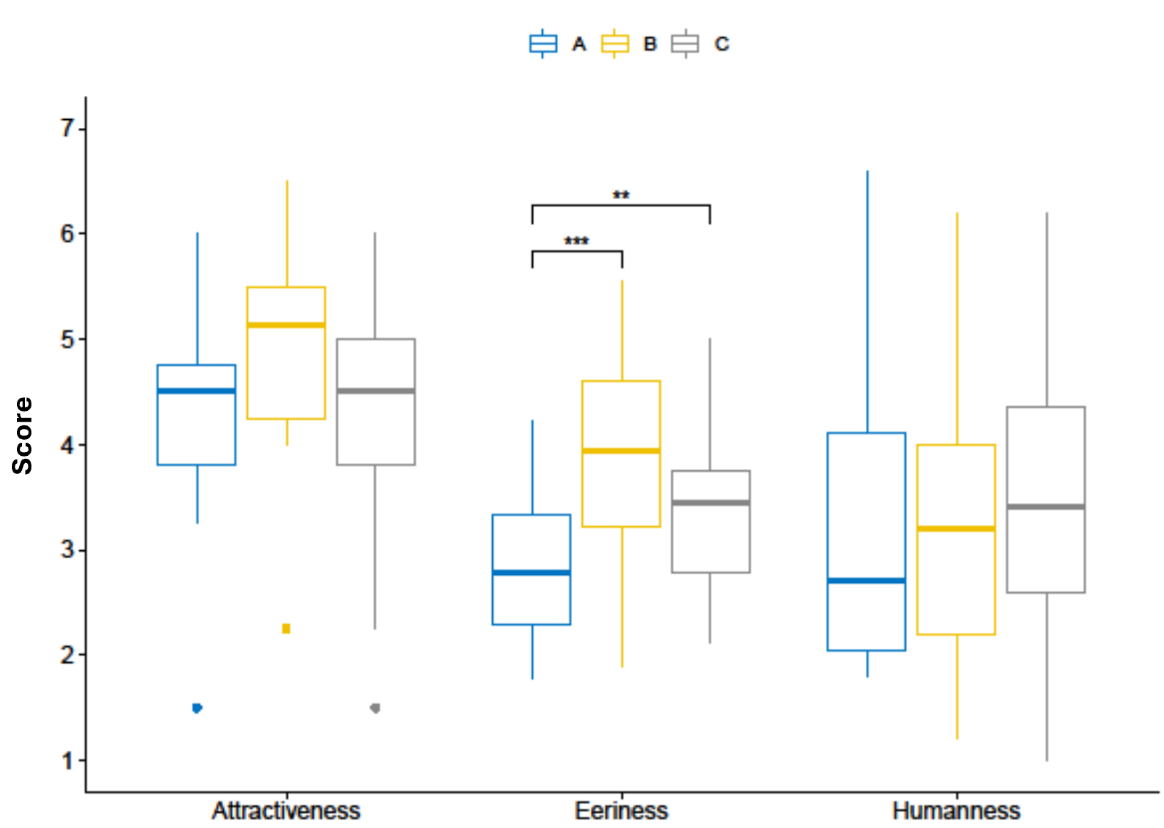

Figure 7: Results of the UVI questionnaire aggregated by the three main dimensions attractiveness, eeriness, and humanness on a scale from 1 to 7. A: generic avatar, B: customized avatar, C: photo-generated avatar. Friedman tests with signed-rank Wilcoxon tests as post-hoc comparisons were performed to check for significant effects. ** and *** denote significant differences according to the Wilcoxon test with $p < 0.01$ and $p < 0.001$, respectively.

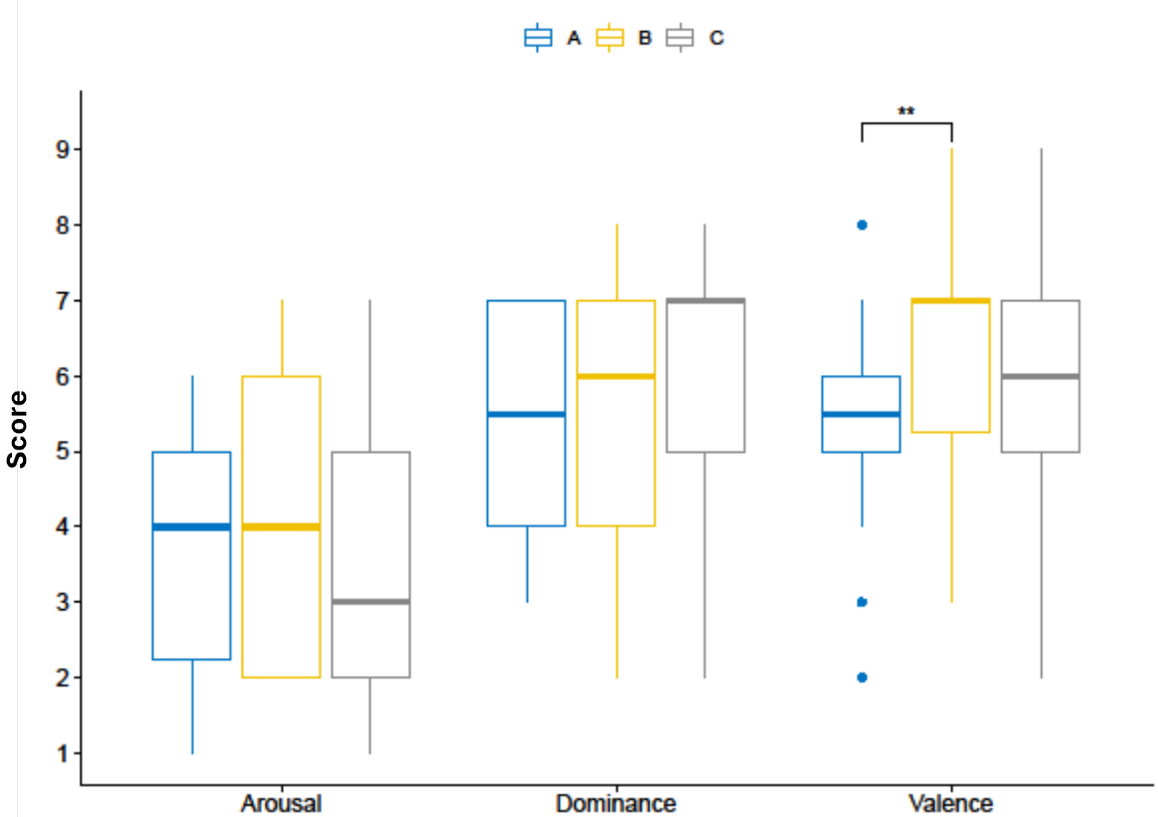

Figure 8: Results of the Self-Assessment Manikin questionnaire. Participants rated their emotional response to the embodied avatar immediately following the VR exposure phase on a scale from 1 to 9. A: generic avatar, B: customized avatar, C: photo-generated avatar. $**$ indicates a significant difference according to the signed-rank Wilcoxon test ($p < 0.01$).

$M = 5.33, SD = 1.35$, $W = 36, p = 0.005$) with an effect size of $r = 0.536$ which suggests a large effect according to Cohen [16]. All other comparisons did not show a statistically meaningful effect.

As the visually different distributions (see Figure 8) might suggest a difference of variances, especially for valence, Levene's test was conducted, revealing no significant differences.

4.2.4 Additional Questions

In addition to the established questionnaires, some questions were added as stated in Section 3.4.5. The first added question asked which degree participants felt they could express themselves in the virtual environment. As the Shapiro-Wilk test indicated a violation of the normal distribution, the non-parametric Friedman test was used to check for a significant main effect of the condition, but no such effect was found ($\chi^2 = 2.34, p = 0.310$).

To complement the mid-immersion question asking explicitly for the attractiveness of the avatar, the same question was added to the post-immersion questionnaire since all other questions asked during the VR exposure were taken from established questionnaires already part of the post-exposure surveys (see Section 3.4.5). A box plot of the result is given in Fig. 9. Because of a violation of the normal distribution shown by the Shapiro-Wilk test in some conditions, the Friedman test was used to analyze the results. It revealed a significant main effect of condition ($\chi^2 = 9.38, p = 0.00917$).

For further investigation of this effect, the signed-rank Wilcoxon test with Bonferroni correction was applied to check for pairwise differences: The attractiveness ratings for the self-configured avatar (Condition B, $M = 4.57, SD = 1.70$) appeared to be significantly higher than the generic (Condition A, $M = 3.67, SD = 1.49$, $W = 48, p = 0.009, r = 0.549$) and the photo-generated avatar (Condition C, $M = 3.93, SD = 1.60$, $W = 168, p = 0.047, r = 0.399$). The effect size can be interpreted as large and moderate, respectively.

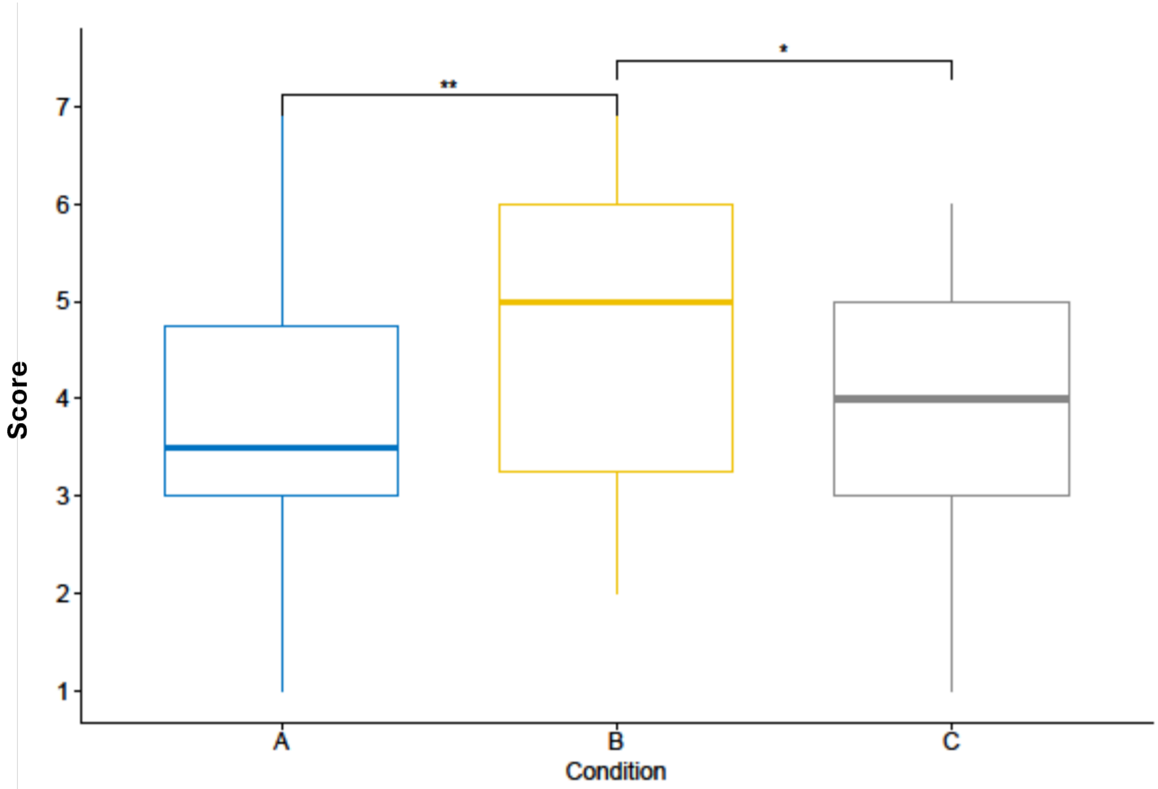

Figure 9: Results of the post-immersion attractiveness rating on a scale from 1 to 7. The self-configured avatar appeared to get a higher rating than both other conditions. A: generic avatar, B: customized avatar, C: photo-generated avatar. $*$ and $**$ denote a significant difference according to the signed-rank Wilcoxon with Bonferroni adjustment with $p < 0.05$ or $p < 0.01$, respectively.

4.3 Mirror Distance

At the end of the exposure, participants performed a positioning task in front of the mirror described in Section 3.4.6 as a potential implicit measurement of attitude to the avatar. The z-axis values of the recorded coordinates were evaluated for a significant effect of condition. As the preliminary Shapiro-Wilk test did not show a violation of the normal distribution assumption, repeated measures ANOVA was used. This indicated no significant main effect of avatar visual fidelity on the distance participants took to the mirror ($F(1.65, 47.92) = 0.73, p = 0.462$).

4.4 Qualitative Data

After each VR exposure phase and at the end of the study, participants were provided the opportunity to give free-form feedback. Concerning general problems, several participants criticized the lower body's tracking as imprecise, having latency, or leading to sliding or hovering feet. Some also noted problems with the layout and positioning of the arms, especially when right next to the body, as the hands sometimes clipped into the body or at least the outer clothing due to imprecise tracking ("My hands [stuck] in my avatar, so there were times when I got the feeling of losing control/the avatar not being me, although sometimes it felt like it was me."). Some tried intuitively to overcome the imperfections by taking poses that matched the avatar they viewed in the mirror. When the arms were stretched forward, some body types and clothing led to impressions of unnatural body configurations ("[like] an inflated rubber suite"). Despite the automatic resizing of the avatar's hands, some participants mentioned their hands being bigger than their real hands. However, most of the participants emphasized the good tracking of their hands and fingers, even being surprised by it ("[When] looking through [the] space between headset and nose, the virtual arm and physical arm seemed to match pretty well").

Concerning the avatar, many participants criticized the limited choice of only two genders, stereotypical body types, everyday clothing, and missing diversity in facial appearance. People who did not identify as cisgender missed the opportunity to either not specify a (binary) gender at all ("[One] should be able to create an avatar without having to specify gender and just straight up choose a body type") or at least configure the body type more detailed to fit the desired appearance ("[The] character creator also only allows 4 different body types to be chosen per gender and even though I am feminine, my body does not reflect this, so I was unable to even create an avatar that matches my real body in certain aspects because the editor did not allow me to choose [...] another body shape than an hourglass figure", "[Those] avatars are quite gendered, especially the female ones").

The avatar-specific feedback did not reveal clear preferences beyond personal likes. However, it did show some valuable insight into potential reasons for the quantitative results. If specifically referenced, the generic avatar was mainly described as not looking like oneself – disturbing identification with it – or even bland or boring and less fun. In contrast, one participant also mentioned the lack of similarity as a potential advantage: "It's so generic that you notice less the differences between you and the avatar and can project more onto it."

The feedback on the other two avatars was also mixed. While the customizable one was most often associated with fun and the pleasure of being able to create it according to one's wishes, some people felt the visual differences in the appearances from their own body that they deliberately chose prevented a stronger self-identification: "I think I could identify less with my own created character because he has a very different appearance than me." "I was surprised to find myself not liking the character I made to look "cool" or finding her more difficult to navigate, although I was looking forward to her." Still, there were participants who explicitly preferred it over the two other options.

In the photo-generated condition, some participants mentioned that they liked it the most as it represented them the most or even referred to a depressing feeling with the other avatars coming from a visual mismatch between the avatar and themselves in the other conditions. Interestingly, it was also the only condition leading to people stating a repelling feeling to their virtual face or being bothered by the close but imperfect reproduction of their appearance: "The character that was supposed to look like me was weird to look at because you wanted to eradicate the differences."

5 Discussion

In the following, the results from the previous section will be checked for their meaningfulness regarding the hypotheses and how these interpretations align with previous research in the field.

The absence of ordering effects indicated by the statistical tests allows for assigning the following effects only to the differences between the avatars.

First, the higher self-similarity found for the photo-generated personalized avatar compared to the two other variants proves hypothesis **H1** and validates the desired distinction between the conditions, i.e., participants did not just create avatars that resembled themselves that, and therefore, did not differ enough to draw meaningful conclusions from the other questionnaires. The same applies to the dimension of self-attribution. However, there is no significant difference to the customized avatar, which might be related to users configuring avatars that differed from their visual appearance without interfering too much with self-attribution. However, it does not support the assumption that users given complete freedom over the visuals of the avatar create a character that might even improve their self-attribution. Reports of that, e.g., creating a "cool" character that they could not identify with, were also found in the qualitative feedback about the customized avatar. This raises the question of the importance of the prompt given to the user for creating the avatar and the surrounding procedure. Future research should further evaluate different versions of it to explore their effect on self-attribution.

Considering the core dimensions of the sense of embodiment, neither mid- nor post-immersion questionnaires showed a clear difference. In particular, agency received high ratings post-immersion throughout all conditions, with a low spread. Despite the criticism of the technical difficulties (see Section 4.4), this implies that the setup was, in general, suitable for creating the desired illusion. The missing difference in the scores between the conditions is in line with previous findings manipulating the truthfulness of avatars [89, 77, 24] explained by agency mainly being related to the coherence of visuo-motor cues that were not manipulated here. The missing differences for presence [77] and change [24] also align with former findings, as change is more important in setups that deliberately manipulate the avatar to differ from the user, which is not the case here.

While visual inspection also does not hint at a difference between the conditions for those dimensions, it suggests a trend for the sense of ownership in the post-immersion VEQ questionnaire, even if it is not statistically significant. The small effect size indicated by Cohen's d can be interpreted so that a potential effect was too small to be detected with the present sample size. This assumption is well-founded on previous studies [77, 89, 24] showing a greater sense of ownership related to higher truthfulness that would at least explain a greater body ownership for the personalized avatar compared to the generic one. However, the statistical insignificance of the results prevents us from making a clear statement on hypothesis **H2**.

In terms of indirect measures, the reaction to the induced threat showed a greater reaction for the customized avatar compared to the generic, which was not achieved by the personalized providing partial support for **H2.1** and implicitly hints at higher ownership for the customized avatar [3, 77]. Due to the lack of differences in the explicit ownership ratings, the final interpretation is unclear, though. The absence of a significant main effect on the distance taken to the mirror might be explained by missing motivation to move as close as acceptable by social standards. In previous studies, this has been done by memorization task for visual features [5] or explicitly asking participants to move closer [96] so a refinement of the task is necessary to investigate the hypothesis **H2.2**.

Aligning with earlier studies [42, 55], the UVI showed a higher eeriness rating for the personalized avatar than the generic. A reason might be due to the generic being easier to dissociate from (see Section 4.4). However, since the same is true for the customized avatar compared to the generic and there is no meaningful difference between the customized and the personalized condition, it could be interpreted as a general effect from a stronger emotional reaction to a non-generic avatar, and the results do not support **H3**.

Although higher eeriness ratings and the expectation of a better sense of embodiment may seem contradictory, earlier studies have observed exactly these counterintuitive behaviors [42, 24]. In their conclusion, Döllinger et al. also called for further investigation of the "complex interplay between personalization, VBO, [and] eeriness". Our results leave this to be at least partially inconclusive and therefore support the necessity for further research, especially since no statistically significant difference was found between the customized and the personalized condition.

In other regards, participants seemed to be satisfied with their customized avatar, showing higher valence ratings and partially confirming **H5** for the comparison with the generic avatar, which might be interpreted as them enjoying the ability to customize according to their wishes. The absence of differences in the other dimensions of the Self-Assessment Manikin probably results from very few other stimuli creating an emotional reaction. At the same time, valence could be invoked by just the pleasure of using a preferred avatar. The effect of the personalized avatar is unclear, though.

The satisfaction with the self-configured customized avatar also shows in the high attractiveness score in mid- and post-immersion questions confirming **H4**, which thereby dampens the negative effects that one otherwise might deduce from the eeriness ratings. As previously mentioned, earlier studies have shown a positive correlation between attractiveness and truthfulness. However, these results show that giving people the power to decide on their own appearance can even improve this.

Finally, the lack of a main effect on the question targeted at the possibility of expressing oneself might be due to the question's phrasing being too generic to capture a specific effect, such that it may be insufficient to make a statement on **H6**.

5.1 Limitations and Future Work

As mentioned in previous sections, this work faces some limitations in tracking and avatar creation that should be considered and improved in future research. However, they do not significantly undermine the conclusions drawn from the results.

First, a lack of tracking fidelity or imprecisions that some participants noticed in the lower body and arms when they left the area that was captured by the visual tracking is generally considered to harm the sense of embodiment [29] and especially the sense of agency [49]. However, the sense of agency was still rated high in all conditions, implicating the capability of the users to adapt to this situation, maintaining a strong illusion described by Won et al. as some kind of homuncular flexibility [95]. In addition, these shortcomings were consistent throughout the study and, therefore, did not impact the comparison between the avatars.

Second, the visual style of the avatars has been criticized by participants for representing stereotypic body images with vastly differing body shapes depending on the choice of gender in the editor. While the study was deliberately designed to use a cartoonish avatar appearance, this limits the freedom of users to configure an avatar to their liking or their actual physical appearance. The importance of this becomes even more apparent if you consider non-cisgender participants who have been constrained to choose a binary gender for all of their avatars, which might be incongruous with their identity and even more problematic with very gendered shapes of the avatars. While current research emphasizes the importance of gender-matched avatars [21], avatar libraries are often missing non-binary representations, and to the knowledge of the author, there is currently no research considering how the best avatar representations can be achieved for trans* and non-binary people who could largely benefit from a self-determined visual representation in VR. After much progress has been made in the ethnic diversity of avatars recently, much more research is needed on gender diversity.

As the hypotheses of this study largely assumed the potential of higher satisfaction with a self-configured avatar that does not necessarily look like the actual physical appearance of the user, there have been several questions regarding the SoE, emotional response, and attractiveness that could capture the satisfaction with the avatar. However, the questionnaires lack measures to capture the self-esteem and satisfaction of the participants with the actual look of their body that could justify interpersonal differences in the potential of a self-configured avatar, as those factors were not considered in time. Gorisse et al., e.g., have found that people with lower self-esteem prefer more abstract representations. In comparison, when asked, people with higher self-esteem chose representations more similar to them [41]. Future studies should include adequate measures for that, like the Rosenberg self-esteem scale [75], and explore other potential factors for interpersonal differences.

Finally, there was no compensation for the time participants spent longer with the self-configured avatar due to the necessary configuration time. However, it is only to be expected that this contributed to a change in the sense of embodiment to a limited extent, as the decisive elements for a corresponding illusion, such as perspective or synchronous visuo-motor stimuli, were missing.

6 Conclusion

In this work, we used a simple, mainly IMU-based approach to full-body tracking in VR to compare three avatar representations with a varying degree of truthfulness. The motion tracking setup was cost-efficient and generally suitable for this application. While it produced noticeably imprecise tracking results, especially in the lower body area, the general illusion, as well as the additional sensor input and heuristics used as countermeasures, appeared to be effective, resulting in a good sense of agency.

While the intended differences in truthfulness of the generic (gender- and ethnicity-matched), the self-configured, and the personalized photo-generated avatar, were proven by questions for self-similarity, the explicit questionnaires for the sense of embodiment remained largely inconclusive and provided only insignificant trends hinting at a positive contribution of truthfulness to ownership that would be in line with previous research.

The expected positive effect of a self-configured avatar because of the potential of representation without the limit of physical appearance and intuitive avoidance of uncanny features could not be clearly shown as both the self-configured and the photo-generated avatar created a more eerie sensation than the generic one. However, the indirect measure of threat reaction and the emotional reaction showed the advantages of the customizable representation. Those benefits and the creation pipeline requiring significantly less hardware than photogrammetric setups used in other studies definitely motivate further investigation of the potential of self-configured avatars for VR applications.

To investigate individual differences, future research should include questionnaires assessing the self-esteem of participants, since this could explain differences in the preference for visual similarity to the physical self [41]. In addition, the study revealed a large gap in the diversity of gender representation in common avatar libraries that should be considered in future development.

7 Acknowledgements

The authors of this paper received funding from the German Research Foundation (DFG) and the National Science Foundation of China in the project Crossmodal Learning (CML), TRR 169/C8, the European Union's Horizon Europe research and innovation program under grant agreement No. 101135025, PRESENCE project, and Germany's Ministry of Education and Research (BMBF), grant No 16SV8878, HIVAM project. Early drafts of this paper were helped by OpenAI's GPT-4o model for improving writing clarity.

References

[1] Achenbach, J., Waltemate, T., Latoschik, M. E., and Botsch, M. Fast generation of realistic virtual humans. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (Gothenburg Sweden, Nov. 2017), ACM, pp. 1–10.

[2] Argyle, M., and Dean, J. Eye-contact, distance and affiliation. *Sociometry* (1965), 289–304.

[3] Armel, K. C., and Ramachandran, V. S. Projecting sensations to external objects: evidence from skin conductance response. *Proceedings of the Royal Society of London. Series B: Biological Sciences 270*, 1523 (July 2003), 1499–1506.

[4] Bailenson, J. N., Beall, A. C., Blascovich, J., Raimundo, M., and Weisbuch, M. Intelligent agents who wear your face: Users' reactions to the virtual self. In *International workshop on intelligent virtual agents* (2001), Springer, pp. 86–99.

[5] Bailenson, J. N., Beall, A. C., Blascovich, J., Raimundo, M., and Weisbuch, M. Intelligent Agents Who Wear Your Face: Users' Reactions to the Virtual Self. In *Intelligent Virtual Agents*, G. Goos, J. Hartmanis, J. Van Leeuwen, A. De Antonio, R. Aylett, and D. Ballin, Eds., vol. 2190. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001, pp. 86–99. Series Title: Lecture Notes in Computer Science.

[6] Bailenson, J. N., Blascovich, J., and Guadagno, R. E. Self-Representations in Immersive Virtual Environments[1]. *Journal of Applied Social Psychology 38*, 11 (Nov. 2008), 2673–2690.

[7] Baker, S., Waycott, J., Carrasco, R., Kelly, R. M., Jones, A. J., Lilley, J., Dow, B., Batchelor, F., Hoang, T., and Vetere, F. Avatar-mediated communication in social vr: An in-depth exploration of older adult interaction in an emerging communication platform. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2021), CHI '21, Association for Computing Machinery, p. 1–13.

[8] Blanke, O., Landis, T., Spinelli, L., and Seeck, M. Out-of-body experience and autoscopy of neurological origin. *Brain 127*, 2 (2004), 243–258.

[9] Blanke, O., and Metzinger, T. Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences 13*, 1 (Jan. 2009), 7–13.

[10] Botvinick, M., and Cohen, J. Rubber hands 'feel' touch that eyes see. 756–756. Number: 6669 Publisher: Nature Publishing Group.

[11] Bradley, M. M., and Lang, P. J. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry 25*, 1 (Mar. 1994), 49–59.

[12] Brooke, J. SUS: A quick and dirty usability scale. *Usability Evaluation in Industry* (1996).

[13] Bruun, A., Law, E. L.-C., Nielsen, T. D., and Heintz, M. Do You Feel the Same? On the Robustness of Cued-Recall Debriefing for User Experience Evaluation. *ACM Transactions on Computer-Human Interaction 28*, 4 (Aug. 2021), 1–45.

[14] Burin, D., Pignolo, C., Ales, F., Giromini, L., Pyasik, M., Ghirardello, D., Zennaro, A., Angilletta, M., Castellino, L., and Pia, L. Relationships between personality features and the rubber hand illusion: an exploratory study. *Frontiers in psychology 10* (2019), 2762.

[15] Cheymol, A., Fribourg, R., Lécuyer, A., Normand, J.-M., and Argelaguet, F. Beyond my real body: Characterization, impacts, applications and perspectives of "dissimilar" avatars in virtual reality. *IEEE Transactions on Visualization and Computer Graphics 29*, 11 (Nov. 2023), 4426–4437.

[16] Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed ed. Taylor and Francis, Hoboken, 1988.

[17] Costantini, M., and Haggard, P. The rubber hand illusion: sensitivity and reference frame for body ownership. *Consciousness and cognition 16*, 2 (2007), 229–240.

[18] Daft, R. L., and Lengel, R. H. Organizational information requirements, media richness and structural design. *Management science 32*, 5 (1986), 554–571.

[19] De Vignemont, F. Embodiment, ownership and disownership. *Consciousness and Cognition 20*, 1 (Mar. 2011), 82–93.

[20] Dewez, D., Fribourg, R., Argelaguet, F., Hoyet, L., Mestre, D., Slater, M., and Lécuyer, A. Influence of personality traits and body awareness on the sense of embodiment in virtual reality. In *2019 IEEE international symposium on mixed and augmented reality (ISMAR)* (2019), IEEE, pp. 123–134.

[21] Do, T. D., Protko, C. I., and McMahan, R. P. Stepping into the Right Shoes: The Effects of User-Matched Avatar Ethnicity and Gender on Sense of Embodiment in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics 30*, 5 (May 2024), 2434–2443.

[22] Do, T. D., Zelenty, S., Gonzalez-Franco, M., and McMahan, R. P. VALID: A perceptually validated Virtual Avatar Library for Inclusion and Diversity. Publisher: arXiv Version Number: 2.

[23] Dummer, T., Picot-Annand, A., Neal, T., and Moore, C. Movement and the rubber hand illusion. *Perception 38*, 2 (Feb. 2009), 271–280.

[24] Döllinger, N., Beck, M., Wolf, E., Mal, D., Botsch, M., Latoschik, M. E., and Wienrich, C. "If It's Not Me It Doesn't Make a Difference" - The Impact of Avatar Personalization on user Experience and Body Awareness in Virtual Reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Sydney, Australia, Oct. 2023), IEEE, pp. 483–492.

[25] Ebrahimi, E., Robb, A., Hartman, L. S., Pagano, C. C., and Babu, S. V. Effects of anthropomorphic fidelity of self-avatars on reach boundary estimation in immersive virtual environments. In *Proceedings of the 15th ACM Symposium on Applied Perception* (Vancouver British Columbia Canada, Aug. 2018), ACM, pp. 1–8.

[26] Ehrsson, H. H., Spence, C., and Passingham, R. E. That's my hand! activity in premotor cortex reflects feeling of ownership of a limb. *Science 305*, 5685 (2004), 875–877.

[27] Ehrsson, H. H., Wiech, K., Weiskopf, N., Dolan, R. J., and Passingham, R. E. Threatening a rubber hand that you feel is yours elicits a cortical anxiety response. *Proceedings of the National Academy of Sciences 104*, 23 (June 2007), 9828–9833.

[28] Elliott, T. C., Henry, J. D., and Baghaei, N. Designing Humanoid Avatars in Individualised Virtual Reality for Mental Health Applications. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (Sydney, Australia, Oct. 2023), IEEE, pp. 315–321.

[29] Eubanks, J. C., Moore, A. G., Fishwick, P. A., and McMahan, R. P. The Effects of Body Tracking Fidelity on Embodiment of an Inverse-Kinematic Avatar for Male Participants. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Porto de Galinhas, Brazil, Nov. 2020), IEEE, pp. 54–63.

[30] Eubanks, J. C., Moore, A. G., Fishwick, P. A., and McMahan, R. P. A preliminary embodiment short questionnaire. *Frontiers in Virtual Reality 2* (Apr. 2021), 647896.

[31] Fiedler, M. L., Wolf, E., Döllinger, N., Botsch, M., Latoschik, M. E., and Wienrich, C. Embodiment and Personalization for Self-Identification with Virtual Humans. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (Shanghai, China, Mar. 2023), IEEE, pp. 799–800.

[32] Fleming, R., Mohler, B. J., Romero, J., Black, M. J., and Breidt, M. Appealing avatars from 3d body scans: Perceptual effects of stylization. In *Computer Vision, Imaging and Computer Graphics Theory and Applications: 11th International Joint Conference, VISIGRAPP 2016, Rome, Italy, February 27–29, 2016, Revised Selected Papers 11* (2017), Springer, pp. 175–196.

[33] Freeling, B., Lécuyer, F., and Capobianco, A. Petting a cat helps you incarnate the avatar: Influence of the emotions over embodiment in vr. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2022), IEEE, pp. 141–149.

[34] Freeman, G., and Maloney, D. Body, avatar, and me: The presentation and perception of self in social virtual reality. *Proc. ACM Hum.-Comput. Interact. 4*, CSCW3 (Jan. 2021), 239:1–239:27.

[35] Fribourg, R., Argelaguet, F., Lecuyer, A., and Hoyet, L. Avatar and Sense of Embodiment: Studying the Relative Preference Between Appearance, Control and Point of View. *IEEE Transactions on Visualization and Computer Graphics 26*, 5 (May 2020), 2062–2072.

[36] Gall, D., Roth, D., Stauffert, J.-P., Zarges, J., and Latoschik, M. E. Embodiment in Virtual Reality Intensifies Emotional Responses to Virtual Stimuli. *Frontiers in Psychology 12* (Sept. 2021), 674179.

[37] Gallagher, S. Philosophical conceptions of the self: implications for cognitive science. *Trends in cognitive sciences 4*, 1 (2000), 14–21.

[38] Garau, M. *The impact of avatar fidelity on social interaction in virtual environments.* Ph.D., University of London, University College London (United Kingdom), England, 2003. ISBN: 978-1-339-30734-3 Publication Title: PQDT - Global U643695.

[39] Goffman, E. *Behavior in public places.* Simon and Schuster, 1963.

[40] Gonzalez-Franco, M., Perez-Marcos, D., Spanlang, B., and Slater, M. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *2010 IEEE Virtual Reality Conference (VR)* (Waltham, MA, Mar. 2010), IEEE, pp. 111–114.

[41] Gorisse, G., Christmann, O., Houzangbe, S., and Richir, S. From robot to virtual doppelganger: impact of avatar visual fidelity and self-esteem on perceived attractiveness. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces* (Castiglione della Pescaia Grosseto Italy, May 2018), ACM, pp. 1–5.

[42] Gorisse, G., Christmann, O., Houzangbe, S., and Richir, S. From Robot to Virtual Doppelganger: Impact of Visual Fidelity of Avatars Controlled in Third-Person Perspective on Embodiment and Behavior in Immersive Virtual Environments. *Frontiers in Robotics and AI 6* (Feb. 2019), 8.

[43] Guy, M., Normand, J.-M., Jeunet-Kelway, C., and Moreau, G. The sense of embodiment in virtual reality and its assessment methods. *Frontiers in Virtual Reality 4* (Dec. 2023).

[44] Higgins, D., Egan, D., Fribourg, R., Cowan, B., and McDonnell, R. Ascending from the valley: Can state-of-the-art photorealism avoid the uncanny? In *ACM Symposium on Applied Perception 2021* (Virtual Event France, Sept. 2021), ACM, pp. 1–5.

[45] Ho, C.-C., and MacDorman, K. F. Measuring the Uncanny Valley Effect: Refinements to Indices for Perceived Humanness, Attractiveness, and Eeriness. *International Journal of Social Robotics 9*, 1 (Jan. 2017), 129–139.

[46] Hoyet, L., Argelaguet, F., Nicole, C., and Lécuyer, A. “wow! i have six fingers!”: Would you accept structural changes of your hand in vr? *Frontiers in Robotics and AI 3* (2016), 27.

[47] Jeunet, C., Albert, L., Argelaguet, F., and Lécuyer, A. “do you feel in control?”: towards novel approaches to characterise, manipulate and measure the sense of agency in virtual environments. *IEEE transactions on visualization and computer graphics 24*, 4 (2018), 1486–1495.

[48] Jo, D., Kim, K., Welch, G. F., Jeon, W., Kim, Y., Kim, K.-H., and Kim, G. J. The impact of avatar-owner visual similarity on body ownership in immersive virtual reality. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (Gothenburg Sweden, Nov. 2017), ACM, pp. 1–2.

[49] Kilteni, K., Groten, R., and Slater, M. The sense of embodiment in virtual reality. 373–387.

[50] Kilteni, K., Normand, J.-M., Sanchez-Vives, M. V., and Slater, M. Extending Body Space in Immersive Virtual Reality: A Very Long Arm Illusion. *PLoS ONE 7*, 7 (July 2012), e40867.

[51] Kim, H. K., Park, J., Choi, Y., and Choe, M. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied Ergonomics 69* (May 2018), 66–73.

[52] Kocur, M., Kalus, A., Bogon, J., Henze, N., Wolff, C., and Schwind, V. The rubber hand illusion in virtual reality and the real world-comparable but different. In *Proceedings of the 28th ACM symposium on virtual reality software and technology* (2022), pp. 1–12.

[53] Laha, B., Bailenson, J. N., Won, A. S., and Bailey, J. O. Evaluating control schemes for the third arm of an avatar. *Presence: Teleoperators and Virtual Environments 25*, 2 (2016), 129–147.

[54] Lanier, J. Homuncular flexibility, 2006. https://www.edge.org/response-detail/11182, Accessed 11.11.2024.

[55] Latoschik, M. E., Roth, D., Gall, D., Achenbach, J., Waltemate, T., and Botsch, M. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (Gothenburg Sweden, Nov. 2017), ACM, pp. 1–10.

[56] Lenggenhager, B., Mouthon, M., and Blanke, O. Spatial aspects of bodily self-consciousness. *Consciousness and Cognition 18*, 1 (Mar. 2009), 110–117.

[57] Lenggenhager, B., Tadi, T., Metzinger, T., and Blanke, O. Video ergo sum: manipulating bodily self-consciousness. *Science 317*, 5841 (2007), 1096–1099.

[58] Lin, H., and Wang, H. Avatar creation in virtual worlds: Behaviors and motivations. 213–218.

[59] Lopez, C., Halje, P., and Blanke, O. Body ownership and embodiment: vestibular and multisensory mechanisms. *Neurophysiologie Clinique/Clinical Neurophysiology 38*, 3 (2008), 149–161.

[60] Lucas, G., Szablowski, E., Gratch, J., Feng, A., Huang, T., Boberg, J., and Shapiro, A. The effect of operating a virtual doppleganger in a 3d simulation. In *Proceedings of the 9th International Conference on Motion in Games* (2016), pp. 167–174.

[61] Lugrin, J.-L., Landeck, M., and Latoschik, M. E. Avatar embodiment realism and virtual fitness training. In *2015 IEEE Virtual Reality (VR)* (Mar. 2015), IEEE, p. 225–226.

[62] Lugrin, J.-L., Latt, J., and Latoschik, M. E. Anthropomorphism and Illusion of Virtual Body Ownership. *ICAT-EGVE 2015 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments* (2015), 8 pages. Artwork Size: 8 pages ISBN: 9783905674842 Publisher: [object Object].

[63] Ma, K., and Hommel, B. Body-ownership for actively operated non-corporeal objects. *Consciousness and Cognition 36* (Nov. 2015), 75–86.

[64] Mansour, S., El-Said, M., Rude-Parkins, C., and Nandigam, J. The interactive effect of avatar visual fidelity and behavioral fidelity in the collaborative virtual reality environment on the perception of social interaction. *WSEAS Transactions on Communications 5*, 8 (2006), 1501–1509.

[65] Maselli, A., and Slater, M. The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience 7* (2013).

[66] Mcmahan, A. Immersion, engagement, and presence: A method for analyzing 3-d video games. *The Video Game Theory Reader* (01 2003), 67–86.

[67] Mori, M., MacDorman, K., and Kageki, N. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine 19*, 2 (June 2012), 98–100.

[68] NASA. Nasa Task Load Index (TLX) v. 1.0 Manual, 1986. https://humansystems.arc.nasa.gov/groups/TLX/downloads/TLX.pdf, Accessed 20.01.2025.

[69] Normand, J.-M., Giannopoulos, E., Spanlang, B., and Slater, M. Multisensory stimulation can induce an illusion of larger belly size in immersive virtual reality. *PloS one 6*, 1 (2011), e16128.

[70] Petkova, V. I., and Ehrsson, H. H. If i were you: Perceptual illusion of body swapping. *PLOS ONE 3*, 12 (Dec. 2008), e3832.

[71] Petkova, V. I., Khoshnevis, M., and Ehrsson, H. H. The perspective matters! multisensory integration in ego-centric reference frames determines full-body ownership. *Frontiers in psychology 2* (2011), 35.

[72] Phillips, L., Ries, B., Kaedi, M., and Interrante, V. Avatar self-embodiment enhances distance perception accuracy in non-photorealistic immersive virtual environments. In *2010 IEEE Virtual Reality Conference (VR)* (Boston, MA, USA, Mar. 2010), IEEE, pp. 115–1148.

[73] Ries, B., Interrante, V., Kaeding, M., and Anderson, L. The effect of self-embodiment on distance perception in immersive virtual environments. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology* (Bordeaux France, Oct. 2008), ACM, pp. 167–170.

[74] Ries, B., Interrante, V., Kaeding, M., and Phillips, L. Analyzing the effect of a virtual avatar's geometric and motion fidelity on ego-centric spatial perception in immersive virtual environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology* (Kyoto Japan, Nov. 2009), ACM, pp. 59–66.

[75] Rosenberg, M. The Measurement of Self-Esteem. In *Society and the Adolescent Self-Image*. Princeton University Press, 1965, pp. 16–36.

[76] Roth, D., and Latoschik, M. E. Construction of the Virtual Embodiment Questionnaire (VEQ). *IEEE Transactions on Visualization and Computer Graphics 26*, 12 (Dec. 2020), 3546–3556.

[77] Salagean, A., Crellin, E., Parsons, M., Cosker, D., and Stanton Fraser, D. Meeting Your Virtual Twin: Effects of Photorealism and Personalization on Embodiment, Self-Identification and Perception of Self-Avatars in Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg Germany, Apr. 2023), ACM, pp. 1–16.

[78] Sanchez-Vives, M. V., and Slater, M. From presence to consciousness through virtual reality. *Nature Reviews Neuroscience 6*, 4 (Apr. 2005), 332–339.

[79] Sato, A., and Yasuda, A. Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition 94*, 3 (2005), 241–255.

[80] Short, J., Williams, E., and Christie, B. The social psychology of telecommunications. *(No Title)* (1976).

[81] Slater, M. Inducing illusory ownership of a virtual body. *Frontiers in Neuroscience 3*, 2 (Sept. 2009), 214–220.

[82] Slater, M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences 364*, 1535 (Dec. 2009), 3549–3557.

[83] Slater, M., and Sanchez-Vives, M. V. Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI 3* (2016), 236866.

[84] Sony Corporation. About mocopi, 2024. https://www.sony.net/Products/mocopi-dev/en/documents/Home/Aboutmocopi.html, Accessed 26.11.2024.

[85] Stein, B. E., Stanford, T. R., and Rowland, B. A. The neural basis of multisensory integration in the midbrain: its organization and maturation. *Hearing research 258*, 1-2 (2009), 4–15.

[86] Steptoe, W., Steed, A., and Slater, M. Human Tails: Ownership and Control of Extended Humanoid Avatars. *IEEE Transactions on Visualization and Computer Graphics 19*, 4 (Apr. 2013), 583–590.

[87] Tsakiris, M., and Haggard, P. The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of experimental psychology: Human perception and performance 31*, 1 (2005), 80.

[88] Tsakiris, M., Longo, M. R., and Haggard, P. Having a body versus moving your body: Neural signatures of agency and body-ownership. *Neuropsychologia 48*, 9 (July 2010), 2740–2749.

[89] Waltemate, T., Gall, D., Roth, D., Botsch, M., and Latoschik, M. E. The Impact of Avatar Personalization and Immersion on Virtual Body Ownership, Presence, and Emotional Response. *IEEE Transactions on Visualization and Computer Graphics 24*, 4 (Apr. 2018), 1643–1652.

[90] Waltemate, T., Senna, I., Hülsmann, F., Rohde, M., Kopp, S., Ernst, M., and Botsch, M. The impact of latency on perceptual judgments and motor performance in closed-loop interaction in virtual reality. In *Proceedings of the 22nd ACM conference on virtual reality software and technology* (2016), pp. 27–35.

[91] Wegner, D. M., Sparrow, B., and Winerman, L. Vicarious Agency: Experiencing Control Over the Movements of Others. *Journal of Personality and Social Psychology 86*, 6 (2004), 838–848.

[92] Wolf, E., Döllinger, N., Mal, D., Wenninger, S., Bartl, A., Botsch, M., Latoschik, M. E., and Wienrich, C. Does distance matter? Embodiment and perception of personalized avatars in relation to the self-observation distance in virtual reality. *Frontiers in Virtual Reality 3* (Dec. 2022), 1031093.

[93] Wolf, E., Fiedler, M. L., Dollinger, N., Wienrich, C., and Latoschik, M. E. Exploring Presence, Avatar Embodiment, and Body Perception with a Holographic Augmented Reality Mirror. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Christchurch, New Zealand, Mar. 2022), IEEE, pp. 350–359.

[94] Wolf, E., Merdan, N., Dolinger, N., Mal, D., Wienrich, C., Botsch, M., and Latoschik, M. E. The Embodiment of Photorealistic Avatars Influences Female Body Weight Perception in Virtual Reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)* (Lisboa, Portugal, Mar. 2021), IEEE, pp. 65–74.

[95] Won, A. S., Bailenson, J., Lee, J., and Lanier, J. Homuncular Flexibility in Virtual Reality. *Journal of Computer-Mediated Communication 20*, 3 (May 2015), 241–259.

[96] Yee, N., and Bailenson, J. The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human Communication Research 33*, 3 (July 2007), 271–290.

[97] Zell, E., Aliaga, C., Jarabo, A., Zibrek, K., Gutierrez, D., McDonnell, R., and Botsch, M. To stylize or not to stylize?: the effect of shape and material stylization on the perception of computer-generated faces. *ACM Transactions on Graphics 34*, 6 (Nov. 2015), 1–12.

[98] Zhang, K., Deldari, E., Lu, Z., Yao, Y., and Zhao, Y. "it's just part of me:" understanding avatar diversity and self-presentation of people with disabilities in social virtual reality. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA, Oct. 2022), ASSETS '22, Association for Computing Machinery, p. 1–16.