Am I Still Me? Visual Congruence Across Reality–Virtuality and Avatar Appearance in Shaping Self-Perception and Behavior

Marie Luisa Fiedler[*], Christian Merz[*], Lukas Schach, Jonathan Tschanter,
Mario Botsch, Carolin Wienrich, and Marc Erich Latoschik

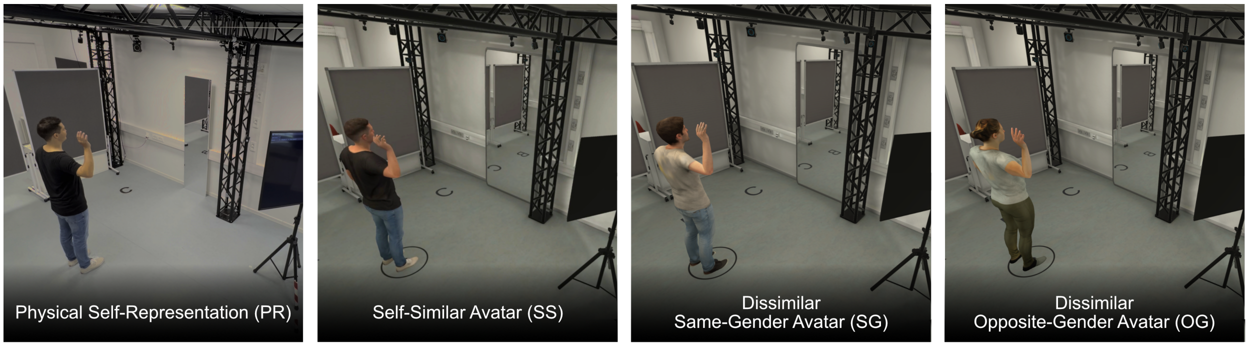


Fig. 1: The figure shows all experimental conditions from an observer's perspective, arranged left to right as follows: (1) physical reality condition while observing their physical self-representation, followed by three VR conditions using a digital twin of the physical lab while embodying a photorealistic (2) self-similar avatar, (3) dissimilar same-gender avatar, (4) dissimilar opposite-gender avatar.

Abstract— This paper presents the first systematic investigation of how congruence in visual self-representation influences self-perception and behavior. We span a continuum from the physical self through avatars with graded self-similarity to clearly dissimilar avatars in virtual reality (VR). In a 1x4 within-user study, participants completed movement and quiz tasks in either physical reality or a digital twin environment in VR, where they embodied one of three avatars: a photorealistic self-similar avatar, a dissimilar same-gender avatar, or a dissimilar opposite-gender avatar. Subjective measures included presence, sense of embodiment, self-identification, and perceived change, and were complemented by an objective movement metric of behavioral change. Compared to physical reality, VR, even with a self-similar avatar, produced lower presence, a weaker sense of embodiment, and reduced self-identification, revealing a persistent gap in visual congruence. Within VR, self-similar avatars enhanced body ownership, self-location, and self-identification relative to dissimilar avatars. Conversely, dissimilar avatars produced measurable behavioral changes compared with self-similar ones. Gender cues, however, had little impact in gender-neutral tasks. Overall, the findings show that photorealistic self-similar avatars reinforce embodiment and self-identification. However, VR still falls short of achieving congruence with physical reality, underscoring key challenges for avatar realism and ecological validity.

Index Terms—Virtual reality, avatars, digital twin, cross-reality comparison, presence, sense of embodiment, self-identification, behavior change.

1 INTRODUCTION

Different forms of user representation in virtual reality (VR) have been explored for several decades. Designs range from functional interfaces, such as virtual controllers, to partial renderings, such as hands-only and humanoid full-body avatars [21]. Avatars can appear as robotic or mechanical entities [19], stylized characters inspired by cartoons or comics [8], or highly realistic digital twins replicating a user's physical features with high fidelity [65]. Increasing avatar realism generally enhances user experience, and self-similarity further amplifies this effect [17, 33, 65]. Users interacting with avatars resembling themselves often report stronger self-attribution, projecting beliefs, attributes, and self-concepts onto the virtual body [15, 17]. Conversely, embodying dissimilar avatars can also shape behavior and attitudes [49, 69]. For instance, female participants embodying male avatars were buffered against stereotype threat, showing improved math performance, while male participants embodying female avatars experienced stereotype threat, leading to reduced performance [45, 49].

Only one study has contrasted physical self-representation with a digital twin, which replicated the user's appearance and the physical environment in VR, showing that digitization alone can alter self-perception [13]. This aligns with prior theoretical work [58, 67], which argues that VR still does not constitute a perfect substitution of the physical world because digital twins retain residual sensory gaps. As a result, contemporary VR can be understood as a form of mixed reality, since sensory information is still partly derived from the physical world. However, empirical work did not examine self-identification or behavior change, an important gap given that many VR applications transfer physical-world paradigms into VR (e.g., therapeutic scenarios), where shifts in self-perception may affect outcomes. It is also critical to understand what happens as avatar appearance diverges from the user's own. Self-similar avatars influence self-perception and continuity between physical and virtual selves [17, 55], whereas decreasing similarity reduces shared identity cues, potentially shifting self-perception, behavior, and attitudes toward the avatar's features [17, 49, 69].

---

[*] These authors have contributed equally to this work.

- Marie Luisa Fiedler, Christian Merz and Jonathan Tschanter are with the Psychology of Intelligent Interactive Systems (PIIS) Group and the Human-Computer Interaction (HCI) Group, University of Würzburg, Germany. Carolin Wienrich is with the PIIS Group, Lukas Schach and Marc Erich Latoschik are with the HCI Group.
  E-Mail: firstname.lastname@uni-wuerzburg.de.
- Mario Botsch is with the Computer Graphics Group, TU Dortmund University, Germany. E-Mail: mario.botsch@tu-dortmund.de.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

This motivates our two research questions:

RQ1: Does VR itself alter users' self-perception and behavior compared to performing the same task in physical reality?

RQ2: How do different appearances of realistic avatars affect users' self-perception and behavior?

To address these questions, we designed a four-condition, within-subjects study that progressively reduced self-related identity cues. We aim to lower self-similarity between the user and visual self-representation while keeping the environment largely constant via a digital twin of the physical lab in VR. Participants completed (i) tasks in the physical lab while viewing their physical self-representation, and three VR conditions in the lab's digital twin environment embodying a photorealistic (ii) self-similar avatar, (iii) a dissimilar same-gender avatar, and (iv) a dissimilar opposite-gender avatar. We assessed presence and sense of embodiment and introduced measures not previously evaluated in this context: self-identification with the presented body and subjective and objective measurements of behavior change. Participants reported higher self-perception (presence, sense of embodiment, and self-identification) when observing their physical self-representation than when viewing a self-similar avatar in a digital twin environment in VR. Moreover, they behaved differently in physical reality versus VR. This suggests that current VR does not yet match the visual congruence with one's physical self, contributing to shifts in self-perception and behavior. Within VR, self-similar avatars, which provide stronger visual congruence with the physical self-representation, improved self-perception relative to dissimilar avatars. Self-identification and behavior change did not differ between same-gender and opposite-gender dissimilar avatars, suggesting that gender cues did not play a role in our gender-neutral tasks. Overall, our findings clarify visual incongruences between physical reality and a digital twin environment in VR, as well as among avatar appearances within VR, deepening understanding of how self-related identity cues in visual self-representation shape a user's self-perception and behavior.

2 RELATED WORK

2.1 Comparing Physical and Virtual Reality

Prior work on presence and sense of embodiment highlights both the potential and limits of VR for shaping a user's self-perception and behavior. By substituting visual input, VR can strongly impact self-perception, which often elicits a strong sense of spatial presence, the feeling of "being there" [59, 61]. However, presence is not a reliable indicator of reality: users can feel present while fully aware they are in a simulation [59, 66]. Hence, examining how the virtual body itself influences user experience may be a better indicator of differences between physical and virtual reality. Avatar embodiment research shows that VR can reproduce core body-related experiences such as body ownership [29, 30]. Avatars shape how users experience themselves in VR through the sense of embodiment (SoE), the feeling of owning, controlling, and being located within a virtual body or body parts, typically decomposed into body ownership, agency, and self-location [29]. Visuomotor synchrony between the user's and the avatar's movement reliably elicits SoE for virtual limbs and whole bodies [23, 30, 60]. Virtual mirrors amplify this effect by supporting identification with the virtual body [24, 29]. Moreover, photorealistic, self-similar avatars have been shown to enhance both SoE and spatial presence [16, 65].

Empirical work consistently finds stronger spatial presence and SoE values in the physical reality than in VR [13, 23, 31]. Theory accordingly highlights congruence, defined as the objective match between processed and expected information, as a prerequisite for authentic Virtual, Augmented, and Mixed Reality (XR for short) experiences [34, 58]. Milgram and Kishino's Reality–Virtuality (RV) continuum [41] frames reality and VR as conceptual endpoints. It places the real environment at one end, "consisting solely of real objects" and representing unaltered physical reality (hereafter the physical reality), and the virtual environment at the other, "consisting solely of virtual objects" and excluding all elements of the physical reality, hereafter referred to as virtual reality (VR). However, the field still lacks systematic empirical anchors that directly contrast them. In addition, recent discussions question central aspects of the RV continuum and the spatial presence construct. They argue that state-of-the-art VR does not instantiate a pure endpoint on the RV continuum: residual physical-reality cues (e.g., proprioception, haptics, environmental awareness, rendering) persist, creating incongruencies that make VR systems closer to mixed reality [58, 67]. Accordingly, VR does not fully replace physical reality (yet). This complicates measurement of constructs like spatial presence when physical and virtual cues blend or compete to varying degrees, and underscores the need to quantify the size of these incongruencies rather than assuming a perfect measurement framework. Without matched physical baselines, we cannot tell whether observed differences arise from avatar-specific manipulations or from the shift into a mediated environment that does not (yet) achieve full congruence. Yet most studies contrast only alternative VR conditions and omit equivalent physical baselines. Systematic comparisons with physical reality are therefore essential to isolate the influence of VR on self-perception and behavior. While one recent work compared physical and virtual reality, it omitted important measures of self-identification and behavior [13]. This leaves a research gap in understanding the differences in visual self-representation between physical and virtual reality, and motivates systematic comparisons between artificial stimuli in VR and matched stimuli in physical reality to isolate the unique influence of VR on self-perception and behavior.

2.2 Effect of Self-Similar Avatars

In VR, users can embody avatars that resemble their physical self-representation through congruent visual shared identity cues, increasing self-identification [16, 18]. In this context, self-identification refers to recognizing a representation as oneself [18]. It occurs when a user associates themselves with a virtual representation by recognizing their identity cues, that is, shared internal or external traits, between themselves and an avatar. Internal traits include aspects like personality, values, beliefs, attitudes, and emotions. When users project these traits onto a virtual representation and perceive them as reflected back, the process is called self-attribution. External traits include features such as appearance, facial characteristics, body shape, or clothing. When users perceive these traits as resembling their own, this is called self-similarity. According to Fiedler et al. [16], both self-attribution and self-similarity are core facets of self-identification.

A growing body of research highlights the central role of identity cues in fostering self-identification [12, 15, 17, 18, 55, 63]. For example, Salagean et al. [55] found stronger self-identification for realistic, self-similar avatars. Visuomotor synchrony also acts as a strong cue: Gonzales-Franco et al. [18] showed that even dissimilar avatars elicited stronger self-identification when facial animations mirrored users' expressions. Building on this, Fiedler et al. [15, 17] found that combining self-similarity with visuomotor synchrony between the user's and the avatar's movements produced particularly strong effects on self-identification. Beyond external cues, Döllinger et al. [12] emphasized internal traits, showing that personality similarity expressed through the body language of non-embodied virtual humans reinforced self-identification, especially self-attribution.

2.3 Effect of Dissimilar Avatars

Beyond replicas of one's own appearance, avatars in VR can be dissimilar. They can elicit perceptual and behavioral changes due to incongruent visual identity cues between the user and the avatar [17, 20, 43]. For example, Gorisse et al. [20] demonstrated that avatar visual fidelity affects user behavior, with self-similar avatars encouraging protective actions and robotic ones fostering greater risk-taking. Similarly, Oberdörfer et al. [43] found that users adapted their body posture depending on whether their avatar wore sneakers or high heels.

Several psychological theories have been proposed to explain such behavior changes. The most widely adopted is self-perception theory [6], which posits that individuals infer their attitudes and emotions by observing their own behavior. Applied to VR, Yee and Bailenson [69] argued that avatars convey identity cues (e.g., age, gender, skin color, clothing) that shape user behavior, which is called the Proteus effect. Users align their self-concept with behaviors elicited by avatar appearance, adopting traits such as confidence or walking speed [51,69], and conforming to the expectations and stereotypes associated with the avatar's appearance. Identity cues can elicit attitudinal, physiological, behavioral, and cognitive adaptations. Building on this, Ratan et al. [49, 50] integrated priming mechanisms based on situational cues that individuals focus on, suggesting that avatar identity cues activate stereotype-related concepts that guide user behavior. The influence of both identity and situational cues on user behavior depends on how closely the user feels connected to their avatar, often mediated by SoE. Supporting this, several studies show that stronger body ownership over an avatar amplifies avatar-induced behavior change [28, 37, 48, 70], although others report no clear link between body ownership and behavioral effects [49, 51]. The Proteus effect has been demonstrated across numerous studies and contexts [25, 38, 46, 48, 69].

Since behavior changes are often subconscious and therefore difficult to capture with subjective measures alone, researchers have complemented questionnaires measuring a behavior change due to the avatar [44, 53] with more objective approaches, such as association tests [4] or explicit behavioral measures [37]. Furthermore, Kilteni et al. [28] used motion capture to record full upper-body positional data and demonstrated significant behavioral adaptations when avatars' appearances were altered. More recently, Merz et al. [39] introduced an unobtrusive in-situ measurement approach based on deep metric similarity learning, enabling the detection of avatar-induced behavioral changes through biometric user modeling.

2.4 Hypotheses

Overall, we conceptualize a continuum of visual self-representation, from the physical self, through graded levels of self-similarity, to clearly dissimilar avatars. Along this continuum: (1) fewer shared identity cues between user and visual self-representation raise visual incongruence and promote shifts in self-perception and behavior; and (2) more shared identity cues increase visual congruency, thereby dampening such shifts. This continuum of visual self-representation spans across the RV continuum, including physical and virtual reality. Current VR systems do not represent the virtual endpoint of the RV continuum (yet), resulting in fewer shared identity cues than in physical reality. Hence, incongruencies between physical and virtual reality arise from technological limitations of VR systems in visual fidelity (e.g., rendering, resolution, latency, motion/facial capture), and other physical-reality cues (e.g., proprioception, haptics, or awareness of the surrounding environment). In VR, reducing shared identity cues through dissimilar appearance features (e.g., skin tone, hair color, gender, body shape) increases visual incongruencies between physical reality and virtual self-representation (i.e., avatar). What remains unclear is how the position of the visual self-representation on this continuum influences self-perception and behavior. This raises two research goals to investigate whether the proposed continuum can be empirically validated, and linked to perceptual and behavioral outcomes: (1) isolate the effect of VR itself by comparing a user's self-perception and behavior of physical self-representation in physical reality with a self-similar avatar in the physical lab's digital twin environment in VR; and (2) evaluate how realistic avatar appearances with varying levels of shared identity cues affect users' self-perception and behavior.

For the first goal, we formulate the following hypotheses:

H1.1 Spatial Presence is higher when observing the physical self-representation in physical reality than when observing a self-similar avatar in a digital twin environment in VR.

H1.2 Sense of Embodiment (body ownership, agency, self-location) is higher when observing the physical self-representation in physical reality than when observing a self-similar avatar in a digital twin environment in VR.

H1.3 Self-Identification (self-attribution, self-similarity) is higher when observing the physical self-representation in physical reality than when observing a self-similar avatar in a digital twin environment in VR.

Concerning the second goal, our hypotheses are as follows:

H2.1 Spatial Presence is higher when observing a self-similar avatar than when observing dissimilar avatars.

H2.2 Body Ownership is higher when observing a self-similar avatar than when observing dissimilar avatars.

H2.3 Self-Identification (Self-Attribution and Self-Similarity) is higher when observing a self-similar avatar than when observing dissimilar avatars.

H2.4 Perceptual and Behavioral Changes are stronger when observing an opposite-gender dissimilar avatar, followed by a same-gender dissimilar avatar, and weakest when observing a self-similar avatar.

3 METHOD

We conducted a user study to examine how visual self-representation affects a user's self-perception and behavior. To address both research questions and evaluate the proposed continuum, we varied visual self-representation while holding the environment largely constant via a digital twin environment of the physical lab in VR. Participants completed four conditions. One condition occurred in physical reality, where they observed their (i) physical self-representation (PR). The other three conditions were with a photorealistic avatar within a physical lab's digital twin environment in VR: (ii) embodied in a self-similar avatar (SS), (iii) embodied in a dissimilar same-gender avatar (SG), and (iv) embodied in a dissimilar, opposite-gender avatar (OG). To progressively reduce self-related identity cues within VR, we manipulated gender, a common factor in behavior-change research [48] and one with gender-specific gait differences [62], increasing the likelihood of measurable behavioral change.

Overall, the study used a 1×4 within-subjects design with four conditions: PR, SS, SG, and OG. All participants first completed the PR condition (framed as training for the VR tasks), followed by the three VR conditions (SS, SG, and OG) in a counterbalanced order. We measured presence, SoE, self-identification, and subjective and objective measurements of behavior change. The study complied with the Declaration of Helsinki[1] and received approval from the MCM Ethics Committee at the University of Würzburg[2].

3.1 Apparatus

We implemented our system with Unity 6000.0. As a VR HMD, we used the Meta Quest 3 without controllers and integrated it into our Unity application with the Meta SDK version 74 and RealityStack I/O [26]. We connected the Meta Quest 3 via Air Link and Wi-Fi 6 with a Windows 11 PC (Intel Core i9-14900K, Nvidia RTX4090 Ti, 64GB RAM). We changed the settings of the Meta Quest 3, so that it supports 90 fps with a resolution of 3712×1984. For the motion tracking, we used The Captury, which is detailed in Subsubsection 3.1.3. The Unity application ran consistently with 90 fps. We measured the motion-to-photon latency with an iPhone 13 Pro Max recording 240 fps. Using frame-counting [10], the latency in mean was 126.04 ($SD = 9.74$) ms. Additionally, we measured wrist movement by placing two wristbands with AX6 data loggers on the participant. However, the analysis of this was out of scope.

3.1.1 Virtual und Physical Environment

Figure 2 shows the virtual environment, which was a digital twin of the physical environment. For this, we took pictures of the room's equipment, walls, and the ceiling, measured the room and the equipment, and then rebuilt the room using Blender. We imported the digital replica via FBX files into our Unity application. We aligned the virtual environment with the physical one by defining two points in the real world and matching them with corresponding virtual points, and by aligning the virtual ground plane with the physical ground. In this way, both the HMD and motion-tracking coordinate systems were aligned with the physical environment and thus with each other.

---

[1] https://www.wma.net/policies-post/wma-declaration-of-helsinki/

[2] https://www.mcm.uni-wuerzburg.de/forschung/ethikkommission/

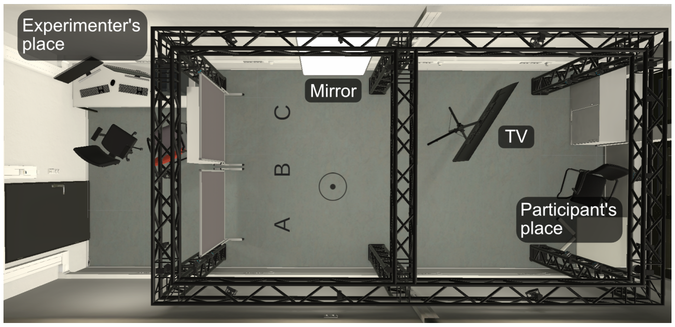


Fig. 2: The figure shows a top-down view of the experimental setup of the digital twin environment in VR during the exposures.

3.1.2 Avatar Generation

We created the self-similar avatars using a method introduced by previous work [1, 5]. The required hardware consists of 97 DSLR cameras (Canon EOS 1300D) mounted on a circular rig and a PC (Ryzen 7u 3800X, NVIDIA RTX2070 Super, 32 GB RAM) running Windows 11. It allows multiple photos to be taken simultaneously from the participants. The photos are then transferred to a high-performance PC (Intel Core i9-14900K, NVIDIA RTX4090, 64 GB RAM) running Ubuntu 24. The experimenter then uses this PC and the photos to generate a self-similar 3D model and high-quality texture that can be quickly imported into Unity using a custom FBX-based runtime importer. No further post-processing was performed. For dissimilar avatars, we used the same approach, using models that reflected the average appearance of our sample.

3.1.3 Motion Tracking and Avatar Animation

To track the participants' full-body motions , without altering the user's appearance with markers or trackers, we used The Captury's markerless tracking system [3], employing eight FLIR Blackfy S BFS-PGE-16S2C-CS RGB cameras attached to the laboratory trusses at 70 Hz. The cameras were connected to a PC (Intel Core i9-11900K, NVIDIA RTX3090, 64 GB RAM, Ubuntu 24) via a Switch (MS510TXPP-100EUS), streaming the cameras' data with a 10Gb/s connection. The PC ran The Captury Live version 270b, and the body pose data was then streamed via a websocket to the Unity application using The Captury's Unity plugin.

The body pose data that was streamed from The Captury was applied to the target avatars using the Character Processor Pipeline [40]. This enabled us to switch between the various target avatars for the VR conditions. Since the dissimilar avatars did not match the users' heights, we applied a vertical offset to the HMD's position to align its eye height with the dissimilar avatars'.

3.2 Tasks

The task design ensured that participants walked extensively during each exposure without becoming bored. The tasks consisted of movement and quiz tasks, presented on a TV screen in the physical or virtual room. In VR, we only used the virtual TV, while we moved the physical TV aside to prevent collisions. Participants received tasks on the TV screen, which was only visible from a specific location, requiring them to walk there to obtain each new task. An audio signal indicated task completion and the availability of the next task.

The movement tasks consisted of simple body-movement exercises performed in front of a physical or virtual mirror while observing their full-body motion from both an egocentric (looking down at their body) and an allocentric (via the mirror) perspective. Exercises included, waving each hand, circling the arms, lifting the legs, and moving the hips from side to side, and were adapted from prior work [11, 65]. For these tasks, participants were instructed to walk to the marker in front of the mirror and to perform the exercise there in a relaxed manner. Each exercise was to be repeated 12 sec until the audio signal indicated the transition to the next task.

Table 1: Summary of the questionnaires used in the study.

Questionnaire	Range	Measure
Presence		
IPQ [56]	[1 – 7]	General Presence
	[1 – 7]	Spatial Presence
Sense of Embodiment (SoE)		
VEQ [53]	[1 – 7]	Virtual Body Ownership
	[1 – 7]	Agency
VEQ+	[1 – 7]	Self-Location
Self-Identification		
VEQ+	[1 – 7]	Self-Attribution
	[1 – 7]	Self-Similarity
Perceptual and Behavioral Change		
VEQ [53]	[1 – 7]	Change
PES	[1 – 7]	Proteus Effect
Controls		
UVI [22]	[1 – 5]	Humanness
	[1 – 5]	Eeriness
FMS [27]	[0 – 20]	Simulator Sickness
RSES [14, 52, 54]	[0 – 30]	Self-Esteem
BFI10 [47]	[1 – 5]	Personality
ITQ [68]	[1 – 7]	Immersive Tendencies

Each quiz task consisted of one question with three possible answers. Participants selected the correct answer by walking to and standing on the field in the room corresponding to the letter of the chosen answer for three seconds, until the audio signal confirmed that the answer had been registered. Participants were instructed that the audio signal indicated when a correct answer had been given. The quiz questions were divided into three categories: identifying a sound (animal or instrument), selecting the word that did not fit the other two, and answering easy general knowledge questions.

3.3 Measures

To test our hypotheses and control for potential confounds, we used a combination of quantitative (subjective and objective) measures and qualitative semi-structured interviews.

3.3.1 Quantitative Subjective Measures

All questionnaires were completed using LimeSurvey 6. To conduct the questionnaires in German, we used either existing validated translated versions or translated the questions using back-and-forth translation. In the PR condition outside of VR, participants were also asked about their experienced SoE, self-identification with their body, and perceived presence. For this purpose, the questionnaire items used in the VR conditions were adapted accordingly. The adapted items are provided in the Supplementary Material. All questionnaires used are summarized in Table 1. The time of measurement is shown in Figure 3.

Revised Version of VEQ+ The Virtual Embodiment Questionnaire+ (VEQ+) [16] extends the VEQ [53] to additionally capture self-location and two facets of self-identification, self-attribution and self-similarity, with an avatar. For this study, an updated (not yet validated) VEQ+ item pool was obtained from the instrument's developers. Administration followed the VEQ's original instructions and response scale. The proposed items for self-location showed good internal consistency, $\alpha = .840$. Self-similarity ($\alpha = .965$) and self-attribution ($\alpha = .947$) showed excellent internal consistency in our experiment. The questionnaire is listed in the Supplementary Material.

Proteus Effect Scale Although behavior changes are often subconscious and difficult to capture with self-report alone, we developed an exploratory Proteus Effect Scale (PES) to assess avatar-induced behavioral change. The items were newly created for this study and rated on a 1–7 Likert scale. The PES demonstrated good internal consistency ($\alpha = .880$). The scale is listed in the Supplementary Material.

---

[3] https://captury.com/real-time-processing/

3.3.2 Quantitative Objective Measures

For the quantitative objective measure, we adopt the methodology introduced by Merz and Schach et al. [39]. This approach measures behavior change across different conditions by analyzing movement data based on head and hand tracking data. We adapt this method by using the same data preprocessing steps and machine learning model. For this, we use The Captury's recorded head and hand movement data from participants and transform it into a body-relative coordinate system, meaning the HMD position and rotation is the origin. This transformation removes environment-specific spatial information. Subsequently, we derive movement velocities from the position data and feed them into a deep metric similarity learning model consisting of a gated recurrent unit (GRU) and a Transformer. For the hyperparameter search, training, and testing, we split our dataset into three subsets: 14 participants for training, 7 for validation, and 14 for testing. Using the test dataset, we compute the identification error rate between all 14 users across different conditions. With the resulting model, we can calculate the identification error rate, which indicates the extent to which the movement patterns of all test users differ across conditions. An identification error rate of 0 indicates no change in movement behavior, whereas a value of 1 signifies complete behavioral change. The identification error rate is derived from user identification performance based on individual movement sequences. For each motion sequence, our model returns an embedding representing this motion. We then find the centroid of the closest user embedding and determine whether it corresponds to the correct user. This allows us to calculate the user identification rate. Hence, this approach enables the detection of behavioral change. However, it cannot determine what, or how, the users' motion changed.

For our first research goal, we examined behavioral change between the PR and the SS conditions. PR served as the reference template. Accordingly, both PR and SS are used as queries to calculate the identification error rate (IER) for SS relative to PR (with PR vs. PR as a quality check). For the second research goal, we analyzed behavioral change among VR conditions by using SS as the reference and calculating IERs for SS, SG and OG conditions relative to SS to quantify changes in user behavior.

3.3.3 Qualitative Interviews

We conducted semi-structured interviews after participants had completed all conditions. The aim was to obtain more detailed insights into participants' experiences during the different exposures and to complement the quantitative measures. The interviews included questions about perceptions of and identification with the visual self-representation, perceptions of the different experimental settings, and perceptions of the different tasks. The complete interview guide is provided in the Supplementary Material.

3.4 Participants

An a priori power analysis was conducted with G*Power 3.1.9.7 for a repeated-measures ANOVA (within-subjects design). Assuming a small-to-medium effect size of $f = 0.2$ for finding a behavior change, as found by previous work [7, 48], an alpha level of .05, and desired power of .80, the analysis indicated a required total sample size of 35 participants. This corresponds to an actual power of .81, ensuring sensitivity to detect the expected within-subject effects.

We recruited 42 participants fulfilling the following inclusion requirements: (1) normal or corrected to normal vision and hearing; (2) at least ten years of experience with German; and (3) no known sensitivity to simulator sickness. Three participants were undergraduate students and received course credit, nine were university employees, and the remaining 30 participants received monetary compensation. We excluded 5 participants due to technical difficulties and two participants with simulator sickness questionnaire values above the cut-off value of 15 [27]. The remaining participants (21 females, 14 males) were aged 18-34 ($M = 25.06$, $SD = 4.19$), all of whom self-reported their ethnicity as white. 31.4% of the participants have used HMDs less than three times, 25.7% have used HMDs between three to ten times, and 42.9% have used HMDs more than ten times.

3.5 Procedure

Our study followed a standardized procedure, visualized in Figure 3. The total duration averaged 111 min, with each exposure lasting approximately 12 min. The experiment was always conducted by two experimenters: one scanned the participants to create the self-similar avatars, and the other conducted the main study. We had three experimenters who rotated doing these two parts. Participants first provided informed consent for study participation and data storage. The experimenter measured each participant's body height, as required for the body scan. Before the scan, participants were instructed to wear tight-fitting, non-monochromatic clothing and to remove glasses or other accessories that could lead to reconstruction errors. While the avatar generation pipeline was running, participants completed the pre-questionnaires on a dedicated questionnaire computer, as listed in Figure 3. Before the first condition began, participants were equipped with two wrist-worn sensors to record their hand movements on both wrists, which they wore throughout all conditions.

During all four exposures (outside and inside VR), instructions were automatically delivered as pre-recorded voice prompts. Participants then performed the PR condition first, framed as a training trial for the following VR conditions, and completed the corresponding post-questionnaires, as shown in Figure 3. Subsequently, participants performed the three VR conditions (SS, SG, and OG) in counterbalanced order. For the VR exposures, participants received instructions from the experimenter on how to wear the VR HMD, and the avatar calibration was conducted. Afterwards, a pre-programmed experimental procedure was initiated, and participants entered the virtual environment. After leaving VR, participants completed the post-experience questionnaires on the questionnaire computer. The second and third VR conditions followed. Finally, participants were interviewed about their experiences. Participants completed 40 tasks per condition (10 movement tasks, 30 quiz tasks with 10 tasks per category). The tasks were performed as described in Subsection 3.2 and were displayed on the TV screen.

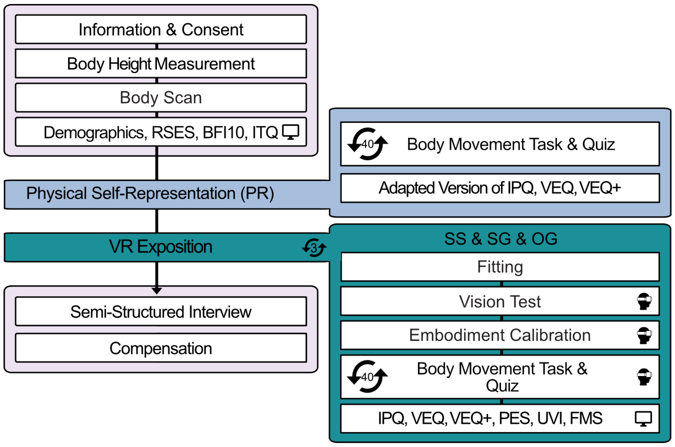


Fig. 3: The figure shows the complete experimental procedure on the left. A detailed overview of the physical self-representation condition and the VR conditions is shown in two boxes on the right.

4 RESULTS

For data aggregation, the machine learning model, and plots, we used Python 3.9. We used R version 2024.04.1[4] for statistical analysis. All tests were conducted with $\alpha = .05$. Because the data violated normality and sphericity assumptions, we used nonparametric repeated-measures tests. For presence, SoE, and self-identification, we calculated 1×4 Friedman tests, including the PR condition. For perceptual/behavioral change, uncanniness, and simulator sickness, we calculated 1×3 Friedman tests on the VR-conditions only. Post-hoc analyses used Wilcoxon signed-rank tests with Bonferroni adjustments.

---

[4] https://www.R-project.org/

Because the a priori power analysis was based on a repeated-measures ANOVA, analyzing the same sample size with the Friedman test is more conservative [71]. Therefore, to demonstrate the sensitivity of our conclusions, we report test-specific effect sizes (Kendall's W for Friedman tests and rank-biserial correlation r for Wilcoxon signed-rank tests). Since we had no expectations for the PR condition relative to the dissimilar avatar conditions, we conducted pairwise comparisons of PR vs. SS, SS vs. SG, SS vs. OG, and SG vs. OG, but not PR vs. SG/OG. All descriptive data and post hoc comparisons are in Table 2, Friedman tests are reported in the following subsections, and plots are shown in Figure 4. To control for demographics in our analysis, we added gender of the participants and the experimenter of the main study as a between-factor, and age as a covariate, which resulted in no significant effects of those factors.

4.1 Presence

We found a significant main effect of our conditions on spatial presence, $\chi^2(3) = 26.11$, $p < .001$, $W = 0.25$. For **H1.1**, pairwise post hoc tests indicated greater spatial presence when observing the self-similar avatar than when observing the physical self-representation, contrary to the hypothesis (large effect). However, a reliability analysis for spatial presence in the PR condition revealed a very poor fit ($\alpha = .249$), indicating that the items do not work as intended in this condition. General presence also differed by condition, $\chi^2(3) = 23.90$, $p < .001$, $W = 0.23$. Pairwise post hoc tests showed a higher general presence for the PR condition than the SS condition (large effect), supporting **H1.1**. For **H2.1**, VR-condition-only pairwise post hoc tests showed no differences, so the hypothesis **H2.1** was not supported.

4.2 Sense of Embodiment

Friedman tests revealed significant effects for body ownership, $\chi^2(3) = 89.46$, $p < .001$, $W = 0.85$, agency, $\chi^2(3) = 50.29$, $p < .001$, $W = 0.48$, and self-location, $\chi^2(3) = 20.80$, $p < .001$, $W = 0.20$. For **H1.2**, pairwise post hoc tests revealed higher body ownership and agency when observing the physical self-representation than when observing the self-similar avatar (very large effects), whereas self-location did not differ between the two conditions, thus partially confirming **H1.2**. For the VR-condition-only comparisons relevant to **H2.2**, pairwise post hoc tests showed higher body ownership for the self-similar avatar than for both dissimilar avatars, while the same-gender exceeded the opposite-gender avatar, confirming **H2.2**. Agency did not differ among avatar types, and self-location was higher for the self-similar avatar than for both dissimilar avatars (not hypothesized).

4.3 Self-Identification

Friedman tests indicated significant effects for both scales, self-attribution, $\chi^2(3) = 85.52$, $p < .001$, $W = 0.81$, and self-similarity, $\chi^2(3) = 92.58$, $p < .001$, $W = 0.88$. For **H1.3**, pairwise post hoc tests showed higher self-attribution and self-similarity ratings for the physical self-representation than the self-similar avatar (very large effects), confirming **H1.3**. For the VR-only comparison, **H2.3** was supported: pairwise post hoc tests indicated higher self-attribution and self-similarity for the self-similar avatar than for both dissimilar avatars, while they did not differ significantly.

4.4 Perceptual and Behavioral Change

For the Proteus effect scale, the Friedman test showed no significant effect of condition, $\chi^2(2) = 2.27$, $p = .321$, $W = 0.03$.

Perceived change did differ significantly between conditions, $\chi^2(3) = 42.71$, $p < .001$, $W = 0.41$, with post hoc tests indicating greater perceived change for the self-similar avatar than for the physical self-representation. However, pairwise post hoc tests did not differ between all VR-conditions, so **H2.4** was not confirmed.

The behavioral change measured with the identification error rate is shown in Table 3. In Subsubsection 3.3.2, we describe how we determined the behavior change.

4.5 Control Variables

Humanness differed significantly by condition, $\chi^2(2) = 13.42$, $p = .001$, $W = 0.19$, as did eeriness, $\chi^2(2) = 11.31$, $p = .004$, $W = 0.16$. Pairwise post hoc tests showed higher humanness and eeriness for the self-similar avatar than for either dissimilar avatar, while the dissimilar avatars did not differ.

Simulator sickness also showed a significant effect, $\chi^2(2) = 10.09$, $p = .006$, $W = 0.14$. Pairwise post hoc tests indicated higher reports for the opposite-gender avatar than for both self-similar and same-gender avatars, whereas they did not differ. However, these outcomes had a small effect size with low descriptive values, suggesting that the main psychological results are unlikely to be artifacts of simulator sickness.

4.6 Exploratory Analysis

We calculated correlations and adjusted α with Bonferroni correction. Since previous work showed that immersive tendencies can impact various VR qualia [35], we calculated correlations with all our dependent variables. We found a significant positive correlation between immersive tendencies and perceived change, $r = .450, p = .036$. Self-esteem can influence preferences for avatar appearance [19]. Hence, we calculated correlations with all our dependent variables. There is a significant negative correlation between self-esteem and the Proteus effect scale, $r = -.480, p = .022$. Prior work showed that users' personality traits can influence their VR experience [3], but we found no significant correlations between participants' personality traits and the dependent variables measured.

4.7 Qualitative Analysis

We analyzed the interviews using a thematic analysis with clustered responses on sticky notes [9]. The following summarizes participant feedback on (1) comparing the digital twin environment in VR to the physical reality environment, (2) perceiving avatars with varying levels of self-related identity cues, and (3) perceiving behavioral changes when embodying avatars.

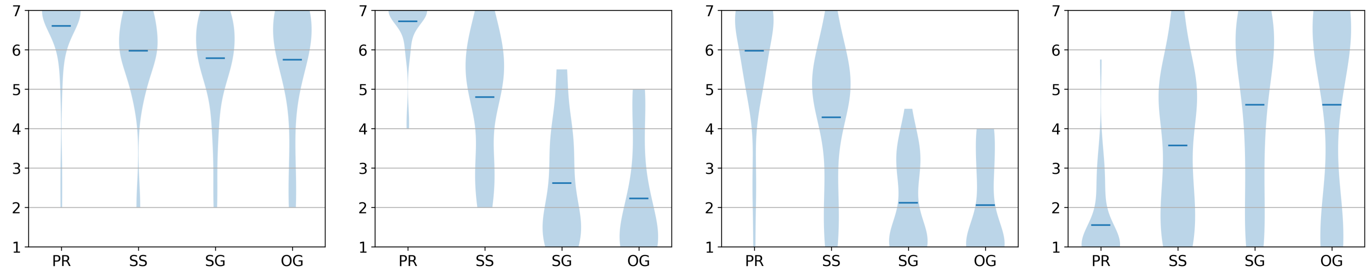
Most participants perceived the VR environment as broadly comparable to the physical reality, with six noting strong similarity and nine describing it simply as the same setting, only in VR. Some differences were highlighted: delayed movements (4 participants), lower resolution, missing level of detail ("the experimenter was missing"; 7), and different appearance of the avatar compared to the physical self-representation (7). Furthermore, most participants (26) reported observing their visual self-representation in the mirror more frequently in VR than in physical reality. Of these, 10 participants stated that they looked in the mirror more often while embodying their self-similar avatar than the dissimilar avatars.

Self-identification increased with the level of self-related identity cues. The self-similar avatar elicited the strongest self-identification, either through visual resemblance (12 participants) or knowledge of being body-scanned. However, reduced texture quality, lack of facial expressions, and posture mismatches hindered self-identification (16), while six participants explicitly noted that knowing "this is not real" limited their sense of self-identification. Dissimilar avatars were associated with weaker identification, though some participants described role-taking (3) or relating through shared movements (4). Identification was stronger for the same-gender avatar compared to the non-matching one due to similarities in gender and body shape (6).

All participants correctly recognized the gender of the dissimilar avatars based on external appearance. Most reported no change in gender identity or behavior due to embodying an avatar of the other gender (28 participants). Still, a few described disorientation when embodying an opposite-gender avatar, particularly when body proportions did not match typical gender expectations (e.g., "this was a male avatar with female-like height"; 3). Two participants noted that the avatar's appearance helped them "get into character". Subtle shifts were also mentioned: one female participant felt taller and more muscular when embodying the male avatar, while one male participant reported moving slightly differently as a female avatar. Overall, gender-related effects were rare and difficult for most participants to articulate.

Table 2: This table shows the descriptive statistics and Wilcoxon paired post-hoc tests for all questionnaires. As reported in Section 4, there is a significant main effect for each variable except for the Proteus effect scale, which is therefore marked as not significant (*ns*) for the post hoc tests.

	Descriptive values: Mean (SD)				**Wilcoxon post hoc tests: p (rank-biserial *r*)**			
Variable	PR	SS	SG	OG	PR–SS	SS-SG	SS-OG	SG-OG
Presence								
General Presence	6.60 (1.12)	5.97 (1.20)	5.79 (1.25)	5.74 (1.58)	0.014 (0.75)	1.000 (0.33)	1.000 (0.33)	1.000 (0.03)
Spatial Presence	4.49 (0.96)	5.51 (1.84)	5.41 (1.59)	5.49 (1.62)	0.019 (0.58)	1.000 (0.24)	1.000 (0.09)	1.000 (0.30)
Sense of Embodiment								
Body Ownership	6.72 (0.72)	4.80 (1.59)	2.61 (1.49)	2.23 (1.48)	< .001 (1.00)	< .001 (0.93)	< .001 (0.99)	0.020 (0.70)
Agency	6.74 (0.65)	5.46 (1.54)	5.06 (1.00)	5.14 (1.29)	< .001 (0.99)	0.588 (0.37)	1.000 (0.26)	1.000 (0.27)
Self Location	5.75 (0.77)	5.26 (1.79)	4.46 (1.91)	4.49 (1.99)	0.600 (0.38)	0.006 (0.78)	0.013 (0.79)	1.000 (0.13)
Self-Identification								
Self-Attribution	5.97 (1.45)	4.29 (1.78)	2.11 (1.16)	2.06 (1.24)	< .001 (0.95)	< .001 (1.00)	< .001 (1.00)	1.000 (0.13)
Self-Similarity	6.60 (0.89)	5.69 (0.87)	2.07 (0.90)	1.71 (0.96)	< .001 (0.79)	< .001 (1.00)	< .001 (1.00)	0.064 (0.59)
Perceptual and Behavioral Change								
Change	1.55 (1.04)	3.57 (1.93)	4.60 (2.18)	4.60 (2.57)	< .001 (0.93)	0.077 (0.54)	0.163 (0.48)	1.000 (0.05)
Proteus Effect	–	2.97 (1.92)	2.84 (1.59)	3.11 (1.84)	–	*ns*	*ns*	*ns*
Uncanniness								
Humanness	–	3.31 (1.18)	2.47 (0.87)	2.49 (0.95)	–	0.003 (0.67)	0.003 (0.70)	1.000 (0.09)
Eeriness	–	2.73 (0.69)	2.28 (0.83)	2.29 (0.96)	–	0.001 (0.79)	0.009 (0.70)	1.000 (0.08)
Control								
Simulator Sickness	–	3.43 (3.61)	3.74 (3.46)	4.31 (3.96)	–	0.789 (0.31)	0.024 (0.78)	0.007 (0.87)



(a) General Presence

(b) Body Ownership

(c) Self-Attribution

(d) Perceived Change

Fig. 4: These are violin plots for some of our dependent variables. Significant differences are not shown in the plots but reported in Table 2.

Table 3: Behavior change measured as mean identification error rate in % with (SD).

Reference	PR	SS	SG	OG
PR	21.00(8.00)	52.30(21.45)	–	–
SS	–	12.80(4.10)	27.10(13.60)	26.70(11.60)

5 DISCUSSION

Our work provides a first systematic investigation into the influence of the visual self-representation as a continuum from a physical self-representation through graded levels of visual self-similarity to clearly having a dissimilar self-representation in VR. In a within-subjects design, participants first solved quiz tasks in physical reality with a physical self-representation, and then, in a counterbalanced order, in three VR conditions with a self-similar avatar, a dissimilar same-gender avatar, and a dissimilar opposite-gender avatar within a digital twin environment of the physical lab. Some of our control variables showed a significant influence on our measurements. Motion sickness was significantly different between our conditions. However, further analysis showed no correlation of motion sickness with spatial presence or sense of embodiment, indicating that it was not a confounding factor with overall low descriptive scores. Humanness and eeriness were significantly higher for the self-similar avatar than for the dissimilar avatars. Previous work documented such an effect as well, which points toward an uncanny valley effect [42].

5.1 Incongruencies of Digital Twins

We first contrasted physical reality against meticulously matched digital twin avatar and environment in VR to isolate how VR affects a user's self-perception and behavior when the environment and tasks are held constant. Hence, we compared the results of observing the physical self-representation in physical reality to observing the self-similar avatar in a digital twin environment in VR. In line with **H1.1**, general presence was higher in physical reality than in VR. However, the IPQ Spatial Presence subscale favored VR over the physical self-representation condition in contrast to recent empirical work [13]. One possible reason for this finding in spatial presence lies in the limited reliability of the IPQ [56] Spatial Presence scale within the physical reality. While we initially applied this scale to test **H1.1**, reliability analysis indicated a poor fit in the physical self-representation condition ($\alpha = .249$), likely due to contextual adaptation of the items, as already noted in prior work [59, 64, 67]. As these works noted, applying the same questionnaire to VR and physical reality contexts is problematic. Our findings support this view, as the IPQ Spatial Presence scale failed to capture presence reliably in the physical self-representation condition. We therefore interpret only the general presence pattern as broadly consistent with a residual gap favoring the physical reality, while noting the measurement caveat for spatial presence outside VR.

Consistent with our hypotheses **H1.2** and **H1.3**, participants reported a higher sense of embodiment (ownership, agency) and self-identification (self-attribution, self-similarity) for their physical self-representation than for a self-similar avatar in VR. Moreover, the results of self-similarity fulfill our manipulation check. Together, this confirms that even under tightly controlled, visually faithful conditions, VR introduces incongruencies that reduce key facets of physical self-representation in comparison to the self-similar photorealistic avatar. Our objective measure of user behavior supports this. The identification error rate was higher in VR when embodying a self-similar avatar than in the physical self-representation condition. This indicates that the user's movement pattern changed, as did their self-perception, stressing that a residual gap remains between "being in the physical reality" and "being in its virtual replica".

Consequently, the question remains: Why does the gap persist? Our design held auditory, olfactory, and spatial layout cues constant and manipulated the visual channel via HMD-mediated rendering and avatar embodiment. Qualitative reports point to residual visual incongruence between participants' physical self-representation and their self-similar avatars: Despite praising the close match between the physical lab and its digital twin environment, they noted discrepancies in avatar appearance, movement mirroring, and absent facial mimicry. This indicates that environmental replication had little impact on the gap. Instead, it likely stems primarily from visual incongruencies between the participant's physical self-representation and the self-similar avatar, which shrink SoE and self-identification. This is in line with theories that treat congruence as a prerequisite for authentic XR experiences [34, 58] and that state-of-the-art VR does not represent an endpoint on the RV continuum since technical limitations persist [58, 67].

Overall, this offers a practical contribution by informing how reliably physical reality interventions can be transferred into virtual contexts. Our findings raise awareness of "what is getting lost" between physical and virtual reality and inform how reliably interventions can be transferred. For example, for body-related therapy and training (e.g., physiotherapeutic movement exercises), practitioners should expect reduced self-identification, sense of embodiment, and adapted behavior (e.g., different movement patterns), and thus potentially a different treatment effect than in physical reality. More broadly, our findings underscore the need to quantify incongruencies between physical and virtual reality with objective movement metrics and self-reports when evaluating transfer from physical to virtual contexts. These incongruencies are caused by technological limitations of VR systems in visual fidelity (e.g., rendering, resolution, latency, motion/facial capture), and other physical-reality cues (e.g., proprioception, haptics, or awareness of the surrounding environment). Given these current VR limitations, perfect congruence is not yet achievable. Until VR systems improve, careful benchmarking against physical baselines remains essential.

5.2 Avatar Appearance Shapes Perception & Behavior

Within VR, spatial presence did not differ between avatar appearances, providing no support for **H2.1** and suggesting a partial dissociation once a sufficient immersive baseline is reached. Moreover, avatar appearance systematically shaped both self-perception and behavior. Self-similar avatars strengthened body ownership and self-location, and increased self-identification relative to dissimilar avatars, while agency remained unchanged. Thus, **H2.2** and **H2.3** were supported, and agency functioned as a useful technical control. These outcomes align with prior work indicating that shared identity cues, combining visual similarity with visuomotor congruence, strengthen SoE and self-identification without necessarily elevating presence [17, 29, 65].

Contrary to **H2.4** and our manipulation check of perceived self-similarity, we did not observe the predicted falling gradient in subjective measures from self-similar to same-gender to the opposite-gender avatar appearance. The two dissimilar avatars were statistically comparable and both significantly lower than the self-similar avatar on subjective measures of SoE and self-identification (also applying to the manipulation test of perceived self-similarity). Interview data are consistent with this pattern: most participants reported no change in gender identity or gendered behavior, and when changes were mentioned, they were subtle and hard to articulate. This is in line with self-perception accounts and priming mechanisms of the Proteus effect-related theories, which emphasize that avatar identity cues bias attitudes and behavior primarily when they fit the context [6, 49, 50]. Our gender-neutral task context likely accounts for the lack of a gender-based gradient between the dissimilar avatars, despite their clear difference from the self-similar avatar. Previous work likewise highlights the importance of cue–context fit to elicit a Proteus effect (e.g., protective actions with self-similar or humanlike avatars, greater risk-taking with robotic ones, and induction or mitigation of stereotype threat with gendered avatars) [20, 45]. Objective movement analysis provided a complementary view. Relative to the self-similar avatar, both dissimilar avatars produced higher identification error of similar magnitude, indicating condition-dependent behavioral deviations even when subjective reports did not register attitudinal or behavioral change. This reconciles recurring gaps between questionnaires and action-level evidence and echoes prior work showing that subtle, often non-conscious adaptations are better captured by objective measures than by self-report alone [28, 37, 39, 44, 53]. Qualitative reports, high variance in our subjective and objective change measures, and exploratory correlations (i.e., immersive tendency with perceived change; self-esteem with Proteus effect ratings) point toward individual sensitivity.

To sum up, self-similar avatars strengthened body ownership, self-location, and self-identification without altering agency, supporting **H2.2** and **H2.3**. Dissimilar avatars, regardless of gender match, produced an objectively measurable behavior change that self-reports did not uniformly show. Reducing shared identity cues via appearance manipulations (e.g., skin tone, hair color, gender, body shape) did induce visual incongruencies. Hence, we partially confirm our proposed continuum of visual self-representation, from the physical self, through graded self-similarity, to clearly dissimilar avatars. Perceived self-similarity followed the intended order (physical self-representation $>$ self-similar avatar $>$ dissimilar avatars). However, we did not observe the predicted decline in the gradient from same-gender to opposite-gender avatar appearance. Even so, our results support the continuum's directional claim: The better the visual self-referential cues match physical self-representation, the lower the visual incongruencies, thereby mitigating such shifts in self-perception and behavior (and vice versa). In practice, we recommend using photorealistic self-similar avatars when identity continuity is the goal. To intentionally shape behavior, pair specific identity cues with task contexts that make those cues consequential, and verify effects additionally with objective movement measures rather than depending on questionnaires alone.

5.3 Limitations and Future Work

Our manipulations focused on binary gender (self-similar vs. dissimilar same-gender vs. dissimilar opposite-gender avatars). This choice was driven by sample size and task neutrality, but limits generalizability to other identity cues (e.g., age, ethnicity, body shape, disability) and to gender-salient contexts. While our sample had a gender imbalance, adding gender as an additional factor to our analysis did not reveal significant gender effects. Since our experiment only included one cultural and ethnic background, larger, more diverse samples and multifactorial cue manipulations are needed to test interactions and intersectional effects. The avatar choice was motivated by previous work that utilized these avatars as generic representations of a similar white sample [15, 16]. Future work should pretest whether used avatars are perceived as dissimilar ones. Additionally, we could not perfectly replicate the physical world and the physical self-representation condition in the VR conditions. Future work should evaluate whether our avatar-generation method led participants to dislike their avatars, as shown in previous work [36].

A further limitation is our use of the IPQ [56] to measure presence in the PR condition. Since the IPQ was developed for mediated virtual environments, several items may be a poor fit for physical reality and likely require contextual reinterpretation. Accordingly, the IPQ, particularly the Spatial Presence subscale, which showed poor reliability in PR, may not have captured presence as intended, limiting cross-condition comparability. However, there is currently no widely established measure that validly assesses presence in both VR and physical reality, making direct comparisons methodologically challenging. Future work should therefore develop and validate cross-context presence measures to establish construct equivalence across realities.

Moreover, our design was not fully randomized because the PR condition was always performed first. We chose this order to provide a baseline for familiarization and to help participants become comfortable with the task before entering VR. However, this fixed ordering may have introduced an anchor effect. While such an anchor would not contradict our expectation of PR as an upper bound, it could still influence subjective ratings and potentially amplify perceived differences between PR and VR. Future work should therefore replicate the study with fully randomized condition orders.

Wearing an HMD introduced experiential and measurement confounds (facial occlusion, added mass), potentially affecting head-movement metrics and animation realism. The animation realism of avatars can impact users' perception [2]. Additionally, the avatars lacked real-time facial expressions (only random blinks), which participants cited as reducing self-identification, consistent with evidence that facial animation increases embodiment [32]. Together, these factors underscore that our comparison spans physical reality and a high-fidelity digital twin avatar and environment in VR, which, in practice, is closer to mixed reality [57, 67]. Future work should add real-time facial expression capture to improve visual congruence to physical reality and sample additional levels along the RV continuum (e.g., augmented reality and augmented virtuality variants). They would provide a more complete map of how visual congruence varies across the RV continuum and how it shapes self-perception and behavior.

Our manipulation was not sensitive enough to detect differences between same- and opposite-gender dissimilar avatars. Future work should probe the continuum with finer-grained manipulations of visual self-representation and test how individual characteristics and context affect these dynamics.

Finally, we note that the mirror in the physical setting was smaller than the mirror in the digital twin environment in VR. Because body perception (e.g., SoE, self-identification) was nevertheless higher in physical reality, there is little indication that this size difference materially shrank the physical reality advantage; however, subtle effects cannot be ruled out. Future work should equalize mirror size and exposure across conditions.

6 CONCLUSION

This work provides a systematic examination of visual self-representation as a continuum spanning physical reality, self-similar avatars, and dissimilar avatars in virtual reality. By directly comparing a physical self-representation with a highly controlled digital twin of this self-representation in VR, we demonstrate that even photorealistic, self-similar avatars do not fully reproduce the levels of presence, sense of embodiment, and self-identification (summarized as self-perception) experienced in physical reality. This persistent gap highlights residual visual incongruencies and supports the view that current VR does not yet constitute an endpoint on the reality–virtuality continuum. In VR, decreasing visual self-similarity reliably reduced embodiment and self-identification, resulting in measurable behavioral changes. While self-similar avatars reinforced identity continuity, dissimilar avatars induced subtle but detectable behavioral adaptations, even when subjective reports showed limited sensitivity. Together, these findings underscore the importance of visual congruence in maintaining self-perception and highlight the significance of objective measures in evaluating behavioral effects. Overall, our results have practical implications for the transfer of interventions from physical reality to VR. Designers and practitioners should anticipate reduced self-identification and altered behavior as visual incongruence increases, and favor self-similar avatars when continuity of the self is desired.

ACKNOWLEDGMENTS

This research has been funded by the Bavarian State Ministry for Digital Affairs in the project XR Hub (A5-3822-2-16).

REFERENCES

[1] J. Achenbach, T. Waltemate, M. E. Latoschik, and M. Botsch. Fast generation of realistic virtual humans. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10. ACM, 2017. doi: 10.1145/3139131.3139154 4

[2] N. Amadou, K. I. Haque, and Z. Yumak. Effect of appearance and animation realism on the perception of emotionally expressive virtual humans. In *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents*, IVA '23. ACM, New York, NY, USA, 2023. doi: 10.1145/3570945.3607307 9

[3] R. V. Aranha, R. Nakamura, R. Tori, and F. L. Nunes. Personality traits impacts in virtual reality's user experience. In *2018 20th Symposium on Virtual and Augmented Reality (SVR)*, pp. 47–56, 2018. doi: 10.1109/SVR.2018.00019 6

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

[5] A. Bartl, S. Wenninger, E. Wolf, M. Botsch, and M. E. Latoschik. Affordable but not cheap: A case study of the effects of two 3D-reconstruction methods of virtual humans. *Frontiers in Virtual Reality*, 2, 2021. doi: 10.3389/frvir.2021.694617 4

[6] D. J. Bem. Self-perception theory. vol. 6 of *Advances in Experimental Social Psychology*, pp. 1–62. Academic Press, 1972. doi: 10.1016/S0065-2601(08)60024-6 2, 8

[7] D. Beyea, R. R. Ratan, Y. S. Lei, H. Liu, G. E. Hales, and C. Lim. A new meta-analysis of the proteus effect: Studies in VR find stronger effect sizes. *PRESENCE: Virtual and Augmented Reality*, 31:189–202, 2022. doi: 10.1162/pres_a_00392 5

[8] P. Bimberg, M. Feldmann, B. Weyers, and D. Zielasko. The influence of environmental context on the creation of cartoon-like avatars in virtual reality. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 298–308. IEEE, 2024. doi: 10.1109/VR58804.2024.00052 1

[9] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006. doi: 10.1191/1478088706qp063oa 6

[10] S. T. Bryson and S. S. Fisher. Defining, modeling, and measuring system lag in virtual environments. In *Stereoscopic displays and applications*, vol. 1256, pp. 98–109. SPIE, 1990. doi: 10.1117/12.19894 3

[11] T. D. Do, C. I. Protko, and R. P. McMahan. 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):2434–2443, 2024. doi: 10.1109/TVCG.2024.3372067 4

[12] N. Döllinger, J. Topel, M. Botsch, C. Wienrich, M. E. Latoschik, and J.-L. Lugrin. Exploring agent-user personality similarity and dissimilarity for virtual reality psychotherapy. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 424–427. IEEE, 2024. doi: 10.1109/VRW62533.2024.00082 2

[13] N. Döllinger, E. Wolf, M. Botsch, M. E. Latoschik, and C. Wienrich. Are embodied avatars harmful to our self-experience? The impact of virtual embodiment on body awareness. In *Proceedings of the 2023 ACM Conference on Human Factors in Computing Systems*. ACM, 2023. doi: 10.1145/3544548.3580918 1, 2, 7

[14] D. Ferring and S.-H. Filipp. Messung des Selbstwertgefühls: Befunde zu Reliabilität, Validität und Stabilität der Rosenberg-Skala. *Diagnostica-Goettingen*, 42:284–292, 1996. 4

[15] M. L. Fiedler, M. Botsch, C. Wienrich, and M. E. Latoschik. Self-similarity beats motor control in augmented reality body weight perception. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–11, 2025. doi: 10.1109/TVCG.2025.3549851 1, 2, 8

[16] M. L. Fiedler, E. Wolf, N. Döllinger, M. Botsch, M. E. Latoschik, and C. Wienrich. Embodiment and personalization for self-identification with virtual humans. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 799–800. IEEE, 2023. doi: 10.1109/VRW58643.2023.00242 2, 4, 8

[17] M. L. Fiedler, E. Wolf, N. Döllinger, D. Mal, M. Botsch, M. E. Latoschik, and C. Wienrich. From avatars to agents: Self-related cues through embodiment and personalization affect body perception in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–11, 2024. doi: 10.1109/TVCG.2024.3456211 1, 2, 8

[18] M. Gonzalez-Franco, A. Steed, S. Hoogendyk, and E. Ofek. Using facial

animation to increase the enfacement illusion and avatar self-identification. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):2023–2029, 2020. doi: 10.1109/TVCG.2020.2973075 2

[19] G. Gorisse, O. Christmann, S. Houzangbe, and S. Richir. 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*, pp. 1–5. ACM, 2018. doi: 10.1145/3206505.3206525 1, 6

[20] G. Gorisse, O. Christmann, S. Houzangbe, and S. Richir. 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, 2019. doi: 10.3389/frobt.2019.00008 2, 8

[21] D. Hepperle, C. F. Purps, J. Deuchler, and M. Wölfel. Aspects of visual avatar appearance: Self-representation, display type, and uncanny valley. *The Visual Computer*, 38(4):1227–1244, 2022. doi: 10.1007/s00371-021-02151-0 1

[22] C.-C. Ho and K. F. MacDorman. Measuring the uncanny valley effect. *International Journal of Social Robotics*, 9(1):129–139, 2017. doi: 10.1007/s12369-016-0380-9 4

[23] W. A. IJsselsteijn, Y. A. W. de Kort, and A. Haans. Is this my hand I see before me? The rubber hand illusion in reality, virtual reality, and mixed reality. *Presence: Teleoperators & Virtual Environments*, 15(4):455–464, 2006. doi: 10.1162/pres.15.4.455 2

[24] Y. Inoue and M. Kitazaki. Virtual mirror and beyond: The psychological basis for avatar embodiment via a mirror. *Journal of Robotics and Mechatronics*, 33(5):1004–1012, 2021. doi: 10.20965/jrm.2021.p1004 2

[25] P. Jacquemin, M. Gräf, and M. Shayan. The proteus effect: Exploring the influence of avatar characteristics on user behavior in virtual environments. *AMCIS 2025 Proceedings*, 2025. 3

[26] F. Kern and M. E. Latoschik. Reality stack i/o: A versatile and modular framework for simplifying and unifying XR applications and research. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 74–76. IEEE, 2023. doi: 10.1109/ISMAR-Adjunct60411.2023.00023 3

[27] B. Keshavarz and H. Hecht. Validating an efficient method to quantify motion sickness. *Human Factors*, 53(4):415–426, 2011. doi: 10.1177/0018720811403736 4, 5

[28] K. Kilteni, I. Bergstrom, and M. Slater. Drumming in immersive virtual reality: The body shapes the way we play. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):597–605, 2013. doi: 10.1109/TVCG.2013.29 3, 8

[29] K. Kilteni, R. Groten, and M. Slater. The sense of embodiment in virtual reality. *Presence: Teleoperators & Virtual Environments*, 21(4):373–387, 2012. doi: 10.1162/PRES_a_00124 2, 8

[30] K. Kilteni, A. Maselli, K. P. Kording, and M. Slater. Over my fake body: Body ownership illusions for studying the multisensory basis of own-body perception. *Frontiers in Human Neuroscience*, 9, 2015. doi: 10.3389/fnhum.2015.00141 2

[31] M. Kocur, A. Kalus, J. Bogon, N. Henze, C. Wolff, and V. Schwind. The rubber hand illusion in virtual reality and the real world - comparable but different. In *Proceedings of the 2022 ACM Conference on Human Factors in Computing Systems*, pp. 1–12. ACM, New York, NY, USA, 2022. doi: 10.1145/3562939.3565614 2

[32] P. Kullmann, T. Schell, T. Menzel, M. Botsch, and M. E. Latoschik. Coverage of facial expressions and its effects on avatar embodiment, self-identification, and uncanniness. *IEEE Transactions on Visualization and Computer Graphics*, 31(5):3613–3622, 2025. doi: 10.1109/TVCG.2025.3549887 9

[33] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, VRST '17. ACM, New York, NY, USA, 2017. doi: 10.1145/3139131.3139156 1

[34] M. E. Latoschik and C. Wienrich. Congruence and plausibility, not presence: Pivotal conditions for XR experiences and effects, a novel approach. *Frontiers in Virtual Reality*, 3, 2022. doi: 10.3389/frvir.2022.694433 2, 8

[35] Y. Ling, H. T. Nefs, W.-P. Brinkman, C. Qu, and I. Heynderickx. The relationship between individual characteristics and experienced presence. *Computers in Human Behavior*, 29(4):1519–1530, 2013. doi: 10.1016/j.chb.2012.12.010 6

[36] S. Liu, K. I. Haque, and Z. Yumak. "i don't like my avatar": Investigating human digital doubles. In *Proceedings of the 2025 18th ACM SIGGRAPH Conference on Motion, Interaction, and Games*, MIG '25. ACM, New York, NY, USA, 2025. doi: 10.1145/3769047.3769061 8

[37] D. Mal, E. Wolf, N. Döllinger, C. Wienrich, and M. E. Latoschik. The impact of avatar and environment congruence on plausibility, embodiment, presence, and the proteus effect in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2358–2368, 2023. doi: 10.1109/TVCG.2023.3247089 3, 8

[38] A. Martin Coesel, B. Biancardi, and S. Buisine. A theoretical review of the proteus effect: Understanding the underlying processes. *Frontiers in Psychology*, 15, 2024. doi: 10.3389/fpsyg.2024.1379599 3

[39] C. Merz, L. Schach, M. L. Fiedler, J.-L. Lugrin, C. Wienrich, and M. E. Latoschik. Unobtrusive in-situ measurement of behavior change by deep metric similarity learning of motion patterns, 2025. doi: 10.48550/arXiv.2509.04174 3, 5, 8

[40] C. Merz, J. Tschanter, F. Kern, J.-L. Lugrin, C. Wienrich, and M. E. Latoschik. Pipelining processors for decomposing character animation. In *30th ACM Symposium on Virtual Reality Software and Technology*, VRST '24, pp. 1–2. ACM, 2024. doi: 10.1145/3641825.3689533 4

[41] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and Telepresence Technologies*, vol. 2351, pp. 282–293. International Society for Optics and Photonics, 1995. doi: 10.1117/12.197321 2

[42] M. Mori, K. F. MacDorman, and N. Kageki. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2):98–100, 2012. doi: 10.1109/MRA.2012.2192811 7

[43] S. Oberdörfer, S. Birnstiel, and M. E. Latoschik. Proteus effect or bodily affordance? The influence of virtual high-heels on gait behavior. *Virtual Reality*, 28(2):81, 2024. doi: 10.1007/s10055-024-00966-5 2

[44] T. C. Peck and M. Gonzalez-Franco. Avatar embodiment. A standardized questionnaire. *Frontiers in Virtual Reality*, 1, 2021. doi: 10.3389/frvir.2020.575943 3, 8

[45] T. C. Peck, J. J. Good, and K. A. Bourne. Inducing and mitigating stereotype threat through gendered virtual body-swap illusions. In *Proceedings of the 2020 ACM Conference on Human Factors in Computing Systems*, pp. 1–13. ACM, 2020. doi: 10.1145/3313831.3376419 1, 8

[46] A. S. Praetorius and D. Görlich. How avatars influence user behavior: A review on the Proteus effect in virtual environments and video games. In *International Conference on the Foundations of Digital Games*, pp. 1–9. ACM, New York, NY, USA, 2020. doi: 10.1145/3402942.3403019 3

[47] B. Rammstedt and O. P. John. Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german. *Journal of research in Personality*, 41(1):203–212, 2007. doi: 10.1016/j.jrp.2006.02.001 4

[48] R. Ratan, D. Beyea, B. J. Li, and L. Graciano. Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: A meta-analysis of the proteus effect. *Media Psychology*, 23(5):651–675, 2020. doi: 10.1080/15213269.2019.1623698 3, 5

[49] R. Ratan and Y. J. Sah. Leveling up on stereotype threat: The role of avatar customization and avatar embodiment. *Computers in Human Behavior*, 50:367–374, 2015. doi: 10.1016/j.chb.2015.04.010 1, 3, 8

[50] R. A. Ratan and M. Dawson. When mii is me: A psychophysiological examination of avatar self-relevance. *Communication Research*, 43(8):1065–1093, 2016. doi: 10.1177/0093650215570652 3, 8

[51] R. Reinhard, K. G. Shah, C. A. Faust-Christmann, and T. Lachmann. Acting your avatar's age: Effects of virtual reality avatar embodiment on real life walking speed. *Media Psychology*, 23(2):293–315, 2020. doi: 10.1080/15213269.2019.1598435 3

[52] M. Rosenberg. *Society and the adolescent self-image*. Princeton University Press, Princeton, New Jersey, USA, 2015. 4

[53] D. Roth and M. E. Latoschik. Construction of the virtual embodiment questionnaire (VEQ). *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3546–3556, 2020. doi: 10.1109/TVCG.2020.3023603 3, 4, 8

[54] M. Roth, O. Decker, P. Y. Herzberg, and E. Brähler. Dimensionality and norms of the Rosenberg self-esteem scale in a german general population sample. *European Journal of Psychological Assessment*, 24(3):190–197, 2008. doi: 10.1027/1015-5759.24.3.190 4

[55] A. Salagean, E. Crellin, M. Parsons, D. Cosker, and D. Stanton Fraser. 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 ACM Conference on Human Factors in Computing Systems*. ACM, 2023. doi: 10.1145/3544548.3581182 1, 2

[56] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of pres-

ence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001. doi: 10.1162/105474601300343603 4, 7, 8

[57] R. Skarbez, J. Frederick P. Brooks, and M. C. Whitton. A survey of presence and related concepts. *ACM Computing Surveys*, 50(6):96, 2017. doi: 10.1145/3134301 9

[58] R. Skarbez, M. Smith, and M. C. Whitton. Revisiting milgram and kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, 2:647997, 2021. doi: 10.3389/frvir.2021.647997 1, 2, 8

[59] M. Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3549–3557, 2009. doi: 10.1098/rstb.2009.0138 2, 7

[60] M. Slater, D. Pérez Marcos, H. Ehrsson, and M. V. Sanchez-Vives. Inducing illusory ownership of a virtual body. *Frontiers in Neuroscience*, 3, 2009. doi: 10.3389/neuro.01.029.2009 2

[61] M. Slater and S. Wilbur. A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6(6):603–616, 1997. doi: 10.1162/pres.1997.6.6.603 2

[62] L. K. Smith, J. L. Lelas, and D. C. Kerrigan. Gender differences in pelvic motions and center of mass displacement during walking: Stereotypes quantified. *Journal of women's health & gender-based medicine*, 11(5):453–458, 2002. doi: 10.1089/15246090260137626 3

[63] A. Thaler, M. N. Geuss, S. C. Mölbert, K. E. Giel, S. Streuber, J. Romero, M. J. Black, and B. J. Mohler. Body size estimation of self and others in females varying in BMI. *PLOS ONE*, 13(2), 2018. doi: 10.1371/journal.pone.0192152 2

[64] M. Usoh, E. Catena, S. Arman, and M. Slater. Using presence questionnaires in reality. *Presence: Teleoper. Virtual Environ.*, 9(5):497–503, 2000. doi: 10.1162/105474600566989 7

[65] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1643–1652, 2018. doi: 10.1109/TVCG.2018.2794629 1, 2, 4, 8

[66] S. Weber, D. Weibel, and F. W. Mast. How to get there when you are there already? Defining presence in virtual reality and the importance of perceived realism. *Frontiers in Psychology*, Volume 12 - 2021, 2021. doi: 10.3389/fpsyg.2021.628298 2

[67] C. Wienrich, P. Komma, S. Vogt, and M. E. Latoschik. Spatial presence in mixed realities – Considerations about the concept, measures, design, and experiments. *Frontiers in Virtual Reality*, 2:141, 2021. doi: 10.3389/frvir.2021.694315 1, 2, 7, 8, 9

[68] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998. doi: 10.1162/105474698565686 4

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

[70] N. Yee, J. N. Bailenson, and N. Ducheneaut. The Proteus effect: Implications of transformed digital self-representation on online and offline behavior. *Communication Research*, 36(2):285–312, 2009. doi: 10.1177/0093650208330254 3

[71] D. W. Zimmerman and B. D. Zumbo. Relative power of the wilcoxon test, the friedman test, and repeated-measures anova on ranks. *The Journal of Experimental Education*, 62(1):75–86, 1993. doi: 10.1080/00220973.1993.9943832 6