Construction of a Validated Virtual Embodiment Questionnaire

Daniel Roth*  
HCl Group, University of Wrzburg

Marc Erich Latoschik†  
HCl Group, University of Wrzburg

Figure 1: Factors addressed by the VEQ. Virtual body ownership (a), agency (b), and the change in the perceived body schema (c). The user is depicted in grey; the virtual avatar is depicted in orange.

ABSTRACT

User embodiment is important for many virtual reality (VR) applications, for example, in the context of social interaction, therapy, training, or entertainment. However, there is no validated instrument to empirically measure the perception of embodiment, necessary to reliably evaluate this important quality of user experience (UX). To assess components of virtual embodiment in a valid, reliable, and consistent fashion, we developed a Virtual Embodiment Questionnaire (VEQ). We reviewed previous literature to identify applicable constructs and items, and performed a confirmatory factor analysis (CFA) on the data from three experiments ($N = 196$). Each experiment modified a distinct simulation property, namely, the level of immersion, the level of personalization, and the level of behavioral realism. The analysis confirmed three factors: (1) ownership of a virtual body, (2) agency over a virtual body, and (3) change in the perceived body schema. A four study ($N = 22$) further confirmed the reliability and validity of the scale and investigated the impacts of latency jitter of avatar movements presented in the simulation compared to linear latencies and a baseline. We present the final scale and further insights from the studies regarding related constructs.

Index Terms: Human-centered computing—Visualization—Visualization design and evaluation methods

1 INTRODUCTION

User embodiment can be referred to as “the provision of users with appropriate body images so as to represent them to others (and also to themselves)” [7, p.242]. With the availability of consumer VR technology, it became apparent that the relevance of virtual embodiment is not limited to the understanding of cognitive processes. In addition, understanding virtual embodiment has concrete and direct implications for research, design, and development. It is especially relevant for applications that consider therapy [11,16,19,55,57], entertainment [14,15,55], as well as collaboration and social interaction [36,37,59]. Previous research investigated virtual embodiment in various studies. However, a consistent and validated instrument for assessing components of virtual embodiment, to the best of our knowledge, does not exist.

1.1 The Demand for a Standardized Measure

Previous studies adapted measures from originating experiments such as the rubber hand illusion (RHI) [11], for example, individual questionnaires and displacement measures. However, the assessment of concepts such as virtual body ownership (VBO) is not consistent. Effects are often assessed with single items, which was argued to be problematic [19]. Approaches to cross-validating different measures often failed (see [22,58] for discussions). According to Boateng et al. [10], the creation of a rigorous scale undergoes three stages: 1) In the item development stage, the domain is identified, and items are theoretically analyzed regarding their content validity. 2) In the scale development stage, questions are pre-tested, and hypothesized factors are explored with covariance analysis (e.g., exploratory factor analysis), along with a consecutive reduction of items to an item pool that remains internally consistent, followed by the factor extraction. 3) In the scale evaluation stage, the dimensionality is tested with CFA, and consecutively, the reliability (e.g., Cronbach’s alpha) and validity (i.e., relation to other constructs and measurements) are evaluated.

Gonzalez-Franco and Peck emphasized the request for a standardized questionnaire and review assessments [23]. From those, they “identify a set of questions to be standardized for future embodiment experiences” [23, p. 4] which they organized in six experimental interests (body ownership, agency and motor control, tactile sensations, location of the body, external appearance, and response to external stimuli). Progressing beyond a theoretical approach, we propose a scale that is data driven, and proceeded through all stages of rigorous scale development (see also [10,19]), and is generally applicable to many embodiment experiments. A scale should be reliable (produce repeatable results within and across subjects irrespective of the testing conditions), valid (measure the underlying constructs precisely), sensitive (discriminate between multiple outcome levels), and objective (shielded from third-party variable bias) [47,80].

1.2 Contribution

We present the creation of a valid and consistent measurement instrument for assessing three components of virtual embodiment (ownership, agency, and change in perceived body scheme). These components were explored, confirmed, and validated. We investigated the scale performance by exploring it’s relation to related concepts and previous measures applied, and confirm its validity, reliability, and sensitivity. We further provide insights into the impacts of immersion, avatar personalization, behavioral realism, and latency and latency jitter that were investigated with the experiments. The questionnaire can be applied to various VR experiments to assess virtual embodiment.
1.3 Embodiment

Embodiment is a part of self-consciousness and arises through multisensory information processing [21, 22]. Previous literature mainly considered three components of embodiment: A conscious experience of self-identification (body ownership), controlling one’s own body movements (agency) [2], and being located at the position of one’s body in an environment (self-location) [19, 22]. Research also stresses the importance of the perspective with which one perceives the world with (first-person perspective) [3, 2].

Body ownership can be described as the experience and allocation of a bodily self as one’s own body, as “my body,” the particular perception of one’s own body as the source of bodily sensations, unique to oneself so that it is present in one’s mental life [20, 21, 22]. A key instrument in investigations of body ownership is the RHI [11]. The experiment stimulates ownership of an artificial body part in the form of a rubber hand by simultaneous tactile stimulation (visually hidden) of the physical hand combined with a parallel visual stimulation of a rubber hand. Caused by the stimulation, participants start to perceive the rubber hand as part of their body.

Agency, meaning the “experience of oneself as the agent of one’s own actions - and not of others’ actions” ([17] p. 523.) following [20] relates to body ownership [13, 14, 21]. Tsakiris et al. described agency as “the sense of intending and executing actions, including the feeling of controlling one’s own body movements, and, through them, events in the external environment” [22] p. 424).

Previous work showed that bottom-up accounts (multisensory processing and integration) are an important driver [33, 71], and that top-down processes (e.g., form and appearance matching) at least modulate embodiment [13, 24, 22, 23]. Kilteni, et al. [33] reviewed triggers and preventers of body ownership illusions and summarized that cross-modal stimulation, for example, congruent visuomotor and visuoattactile stimulation supported ownership illusions. In contrast, incongruencies (thus counter-acting sensorimotor contingencies) hinder ownership. Further, visuosensorirotor cues, such as perspective shifts or modified distances, as well as semantic modulations can impact ownership illusions [33].

1.4 Virtual Embodiment

Virtual embodiment was defined as “the physical process that employs the VR hardware and software to substitute a person’s body with a virtual one” [67] p.1. Embodiment has received ongoing attention in VR research, for example, regarding avatar hand appearance [24, 27, 30, 61] and full body representations [46, 48, 52, 66, 74]. Avatars that represent a user are defined as virtual characters driven by human behavior [5]. Following video-based approaches [40], researchers found that the concept of the RHI also applies to virtual body parts [24], and entire virtual bodies [23, 65, 66].

To investigate and alternate virtual body perception, experimenters have used mirrors in immersive HMD-based simulations [22], and semi-immersive fake (magic) mirror projections [38, 74]. According to Slater, et al. [66], the induction time for body ownership illusions varies between about 10 s and 30 min. Kilteni, et al. [32] summarized findings and measures regarding self-location, agency, and body ownership, and argue for a continuous measurement approach. Similar to Kilteni, et al. [33], Maselli and Slater concluded from their experiments that bottom-up factors, like sensorimotor coherence, and a first-person perspective, are driving factors [46]. They argue that appearance moderates the experience insofar as realistic humanoid textures foster body ownership.

An important effect regarding the embodiment through avatars in VR is the Proteus effect [81, 82]. Yee and Bailenson found a change in behavior, self-perception and participants’ identity when taking the perspective of an avatar with altered appearance. Participants changed their behavior and attitude according to behavior they attributed to their virtual representation.

### Table 1: Proposed items for the generalization and extension.

| Ownership                                                                 | Agency                                                                 | Change (in the perceived body schema) |
|---------------------------------------------------------------------------|------------------------------------------------------------------------|---------------------------------------|
| - It felt like the virtual body was my body.                              | - I felt as if I was controlling the movement of the virtual body.      | - I had the illusion of owning a different body from my own (body). |
| - It felt like the virtual body parts were my body parts.                 | - I felt as if my body really still looked like what I had in mind.     | - I felt the need to check if my body had changed.                   |
| - The virtual body felt like a human body.                                | - I felt the feeling that my height of my body had changed.             | - I had the feeling that the weight of my body had changed.          |
| - It felt like the virtual body was my body.                              | - I had the feeling that the height of my body had changed.             | - I had the feeling that the width of my body had changed.           |

*Note. * Required recoding, † new item.

In summary, the degree and the precision with which sensory stimulation, appearance, and behavior are rendered, as well as the perspective that is presented are important aspects for user-embodiment interfaces [22, 23, 45, 46, 73]. Design choices, for example, the character type, or the realism of the replicated appearance and behavior may strongly influence the perceptual phenomena of embodiment. Measuring embodiment in a valid way, therefore, is crucial to VR applications.

2 Scale Construction

Previous assessments often base on the original RHI experiment [11]. A psychometric approach to assessing levels of embodiment toward an artificial physical body part identified the latent variables of ownership, agency, and location [21]. The assessment of location included items focusing on the coherence between sensation and causation, and locational similarities of artificial physical body parts.

Roth et al. [58] presented a proposal to assess virtual embodiment with 13 questions from previous work [3, 11, 18, 22, 31, 41, 45, 50, 53, 66, 69]. A principal component analysis revealed three factors: acceptance (covering the aspect of ownership perception), control (covering the aspect of agency perception), and change (covering the aspect of a perceived change in one’s own body schema). The last factor may be especially important for studies that make use of altered body appearances and the Proteus effect [81, 82]. The instrument showed good reliability in further assessments [3, 74]. In addition to necessary validation and consistency analysis, two downsides can be identified: 1) The measure is strongly constrained to “virtual mirror” scenarios due to the phrasing, and 2) the components are not balanced regarding the number of items. In particular, the ownership component consists of only three statements. We based our scale construction on their work and performed necessary improvements. Regarding the item development, we first generalized the phrasing of the questions to fit generic scenarios. Second, we added questions to balance the component assessment for ownership (“It felt like the virtual body belonged to another person,” - adapted from [50], “It felt like the virtual body belonged to me”) and agency factors (“The movements of the virtual body were in sync with my own movements”), respectively. The items are shown in Table 1. For the scale development and scale evaluation, CFA was calculated with the data from three studies (N = 196) that assessed the impacts of particular manipulations hypothesized to affect embodiment. The study results are discussed in detail in Section 3.
Table 2: The confirmed embodiment factors (CFA results).

| Ownership | Agency | Change |
|-----------|--------|--------|
| myBody    | .81    |        |
| myBodyParts| .73    |        |
| human     | .53    |        |
| belongsToMe| .71   |        |
| myMovement| .80    |        |
| controlMovements| .70 | |
| causeMovements| .72  | |
| syncMovements| .53  | |
| myBodyChange| .69  | |
| echoHeavyLight| .78 | |
| echoTallSmall| .48  | |
| echoLargeThin| .72  | |

Note. Coefficients represent standardized path coefficients of the CFA. Correlations between factors: Ownership ~ Agency r = .69, Ownership ~ Change r = .17, Change ~ Agency r = -.16.

2.1 Confirmatory Factor Analysis

We performed the CFAs using R with the lavaan package. The reporting of fit indices is based on the recommendation made by Kline [19]. As the assumption of multivariate normality was violated, we conducted a robust maximum likelihood estimation and computed Satorra-Bentler (SB) corrected test statistics (see [14]). The first attempt did not yield an acceptable model fit, and as the modification indices indicated, there were covariations in the error terms of a particular item loading on several factors (“I had the illusion of owning a different body from my own”). Therefore, this item was dropped. A second attempt yielded an acceptable model fit. However, inspection of the modification indices revealed covariations in the error terms of two items loading on several factors (“I had the feeling that the virtual body belonged to another person,” “I enjoyed controlling the virtual body”). Thus, we excluded problematic items. Furthermore, items with a factor loading < .40 were excluded (“I felt the need to check if my body really still looked like what I had in mind”). The third CFA with the remaining 16 items yielded a more parsimonious solution with a good model fit, SB \( \chi^2 = 52.50, df = 51, p = .416 \), root mean square error of approximation \( \text{RMSEA} = .013 \), 90% confidence interval of robust root mean square error of approximation \( [.000; .052] \), standardized root mean square residual \( \text{SRMR} = .047 \), and a robust comparative fit index \( \text{CFI} = .998 \). Thus, the solution was deemed acceptable to characterize these components of virtual embodiment. Table 3 depicts the standardized coefficients. The reliability values for ownership \( (\alpha = .783) \), agency \( (\alpha = .764) \), and change \( (\alpha = .765) \) were acceptable. The resulting scale is depicted in Table 3. Its factors are illustrated in Fig.1. The scale was assessed in German. A professional service was consulted for the translation.

3 Validation

The validation of the VEQ is based on four studies that explored individual aspects of virtual embodiment. We manipulated simulation properties that were hypothesized to affect the perceived embodiment, namely, immersion (Study 1), personalization (Study 2), behavioral realism (Study 3), simulation latency and latency jitter (Study 4). To assess the scale validity, we investigated correlations to related concepts, reliability, and compared the scale performance to related instruments. Participants were individually recruited (i.e., the studies were not performed in block testing fashion) through the recruitment system of the institute for human-computer-media at the University of Würzburg. All studies were approved by the ethics committee of the institute for human-computer-media at the University of Würzburg.

Table 3: The Resulting Virtual Embodiment Questionnaire (VEQ).

| Ownership - Scoring: \( ([\text{OW1}] + [\text{OW2}] + [\text{OW3}] + [\text{OW4}]) / 4 \) |
| OW1. myBody | It felt like the virtual body was my body. |
| OW2. myBodyParts | It felt like the virtual body parts were my body parts. |
| OW3. humanness | The virtual body felt like a human body. |
| OW4. belongsToMe | It felt like the virtual body belonged to me. |
| Agency - Scoring: \( ([\text{AG1}] + [\text{AG2}] + [\text{AG3}] + [\text{AG4}]) / 4 \) |
| AG1. myMovement | The movements of the virtual body felt like they were my movements. |
| AG2. controlMovements | I felt like I was controlling the movements of the virtual body. |
| AG3. causeMovements | I felt like I was causing the movements of the virtual body. |
| AG4. syncMovements | The movements of the virtual body were in sync with my own movements. |

Change - Scoring: \( ([\text{CH1}] + [\text{CH2}] + [\text{CH3}] + [\text{CH4}]) / 4 \)

| CH1. myBodyChange | I felt like the form or appearance of my own body had changed. |
| CH2. echoHeavyLight | I felt like the weight of my own body had changed. |
| CH3. echoTallSmall | I felt like the size (height) of my own body had changed. |
| CH4. echoLargeThin | I felt like the width of my own body had changed. |

Note. Participant instructions: Please read each statement and answer on a 1 to 7 scale indicating how much each statement applied to you during the experiment. There are no right or wrong answers. Please answer spontaneously and intuitively. Scale example: 1–strongly disagree, 4–neither agree nor disagree, 7–strongly agree. A professional service was consulted for the translation.

3.1 Study 1: Impacts of Immersion

Virtual embodiment was previously assessed in immersive head-mounted display (HMD)-based systems in first- and third-person perspective (e.g., [66]), and with less immersive L-Shape displays [76], or fake mirror scenarios [38]. Recent work confirmed that visual immersion (i.e., first-person perspective and HMD) fosters virtual embodiment [74]. Thus, we hypothesized H1.1: A lower immersive display setup is inferior in inducing virtual embodiment. According to the previous literature [32,62,63], we also assumed that H1.2: Higher immersion results in higher presence. The latter was also assessed to investigate the relation between the VEQ factors and presence.

3.1.1 Method

In a one-factor (Medium) between-subjects design modifying the level of immersion, participants were exposed to either a fake-mirror projection, in which case the participant had a visual reference to her or his physical body, or an immersive simulation displayed with an HMD, in which case the participant saw her or his virtual body from a first-person perspective (Fig.2). To provoke the perception of virtual embodiment, participants were asked to perform motions and focus their attention, similar to related experiments [38,74], as described in the following.

Procedure Fig.3 shows the study procedure. Participants were welcomed and informed, before the pre-study questionnaire with demographics and media usage was assessed. Participants were then equipped, calibrated, and given time to acclimate to the simulation. Similar to previous work [38,74], audio instructions were then presented to the users to induce embodiment with a total duration of 150
was the mirror; see Fig. 2). We used avatars created with Autodesk well as a virtual mirror (in the projection condition, the projection was placed in a living room. A reference point was presented, as with your palm facing down"), and further focus on one’s own/the character generator (Fig. 2) which were scaled (uniform) according to the participant’s height.

Character Generator (Fig. 2) which were scaled (uniform) according to the participant’s height.

In the virtual environment, the participant was constrained to the avatar head joint while accounting for the distances accurate mirror image. The virtual camera/tracking point was constrained to the virtual world seem to you?; \( \alpha = .672 \). The responses were given using a 7-point Likert-type scale (see the original source for the anchors). Further assessments are not the subject of the present reporting due to page limitations. No severe sickness effects occurred.

Participants. We excluded participants when problems or severe tracking errors were noted. The final sample consisted of 50 participants (32 female, 18 male, \( M_{\text{age}} = 22.18, SD_{\text{age}} = 2.83 \)). Of those, 49 participants were students, and 46 participants had previous experience with VR technologies. The sample was equally distributed (25 per condition).

Figure 2: Apparatus (Study 1). Left: The projection condition. Center: The HMD condition. Right: The female and male avatars used for Study 1.

Figure 3: Study procedure.

Apparatus. The scenarios and materials are depicted in Fig. 2. Participants were tracked by an OptiTrack Flex 3 tracking system. A Unity3D simulation displayed the stimulus via a fake mirror projection or an Acer H6517ST projector (projection size: 1.31 m high \( \times \) 1.46 m width, active stereo, 480 px \( \times \) 1080 px per eye). The virtual projection was calibrated to match the physical projection preferences, thus allowing a physically accurate mirror image. The virtual camera/tracking point was constrained to the avatar head joint while accounting for the distances of the eyes (head-neck model). The baseline motion-to-photon latency was approximated by video measurement and frame counting (25) to 77 ms for the projection setup, assuming slightly lower values for the HMD setup. In the virtual environment, the participant was placed in a living room. A reference point was presented, as well as a virtual mirror (in the projection condition, the projection was the mirror; see Fig. 2). We used avatars created with Autodesk Character Generator (Fig. 2) which were scaled (uniform) according to the participant’s height.

Measures. We assessed the VEQ (Cronbach’s \( \alpha \geq .751 \)) and the igroup presence questionnaire (IPQ) \( [60] \), which was adapted to fit the presented scenario. The IPQ adaptation assessed general presence (“In the computer-generated world, I had a sense of “being there”), spatial presence (e.g., “I did not feel present in the virtual space” \( \alpha = .786 \)), involvement (e.g., “I was not aware of my real environment” \( \alpha = .851 \)), and realness (e.g., “How real did the virtual world seem to you?” \( \alpha = .672 \)). The responses were given using a 7-point Likert-type scale (see the original source for the anchors). Further assessments are not the subject of the present reporting due to page limitations. No severe sickness effects occurred.

Participants. We excluded participants when problems or severe tracking errors were noted. The final sample consisted of 50 participants (32 female, 18 male, \( M_{\text{age}} = 22.18, SD_{\text{age}} = 2.83 \)). Of those, 49 participants were students, and 46 participants had previous experience with VR technologies. The sample was equally distributed (25 per condition).

3.1.2 Results

Comparisons. T-tests were conducted for each individual measure. The VEQ factors were aggregated according to the scoring depicted in Table 3. In the case of unequal variances (Levene-test), corrected values are reported (Welch-test). The perceived change in body schema was statistically significantly different between the projection condition (M = 2.61, SD = 1.46) and the HMD condition (M = 3.40, SD = 1.31; \( t(48) = -2.013, p = .0498, d = 0.61 \)). As expected, the HMD condition resulted in a stronger perception of perceived change. Neither ownership (projection: M = 4.81, SD = 0.85; HMD: M = 4.71, SD = 1.22) nor agency (projection: M = 6.30, SD = 0.51; HMD: M = 6.18, SD = 0.72) showed a significant difference (\( p_s > .497 \)). Regarding the IPQ, we found a significant difference with general presence between the projection condition (M = 3.88, SD = 1.45) and the HMD condition (M = 5.32, SD = 0.99); t(42.30) = -4.098, \( p < .001 \), d = 1.16. This effect was substantiated by a significant difference in the spatial presence measure between the projection condition (M = 3.88, SD = 1.32) and the HMD condition (M = 5.03, SD = 1.20); t(48) = -3.218, \( p = .002 \), d = 1.19, as well as for the involvement measure between the projection condition (M = 3.45, SD = 1.28) and the HMD condition (M = 5.44, SD = 1.18); t(48) = -5.712, \( p < .001 \), d = 1.62. There was no significant difference in the realness measure t(48) = -0.388, \( p = .70 \).

Correlations. We calculated bivariate Pearson correlations. Table 4 depicts the results. We found significant correlations between ownership and agency, as well as between ownership and general presence, spatial presence, and realness. The change factor was correlated significantly with the general presence assessment. Further, we found correlations within the presence measures.

3.1.3 Discussion

The significantly higher perception in the change of the perceived body schema substantiates previous findings that showed that higher
immersion positively impacts embodiment (H1.1). While Walsmeyer, et al. [24] also found immersion to impact ownership over a virtual body, the proposed scale picked up that the first-person perspective HMD simulation specifically affected the perceived change of one’s own body schema. We interpret this as a result of the presence of a reference to the physical body in the projection condition, and the absence of such in the HMD condition. This indicates a more sensitive pinpointing of this effect by the VEQ. This was confirmed by the expected correlations between ownership and agency as suggested in previous literature, and an absence of correlations of these components with change. Despite relatively high ratings for agency, and above scale mean ratings for ownership, these factors were not significantly impacted between conditions.

Supporting H1.2, several presence dimensions were affected. The HMD condition resulted in greater presence. In addition, we found positive relations between the VEQ and presence measures. Ownership correlated with general presence, spatial presence, and realness, whereas the change factor correlated only with the general presence assessment. Therefore, we assume that presence and embodiment interact positively. However, we cannot provide any insights into the causality based on the analysis. We conclude that immersion is a driving factor for embodiment regarding the fostering of the change in the perceived body schema. As more immersive HMD-driven simulations allow for the perception of a virtual body from a first-person perspective without a physical body reference.

3.2 Study 2: User-Performed Personalization

Previous research investigated the use of photogrammetric avatars and found that personalization can affect the identity with [43], and the ownership over, a virtual body [30,38,74]. Based on these results, we hypothesized H2.1: 

**Perception increases the perception of ownership over a virtual body.** However, whether this assumption holds true using avatar creation tools instead of photogrammetric scanning is an open question and the topic of Study 2.

3.2.1 Method
In a one-factor (Personalization) between-subjects design, we compared the embodiment with an avatar from a creation tool to a personalized avatar created with the same tool. Participants were represented either as a gender-matched generic avatar (Caucasian—as we expected a Caucasian sample) or by a personalized avatar they created as a representation of themselves (see Fig. 4).

**Procedure** The procedure followed the general study procedure (Fig. 1). Participants were welcomed and informed, before the pre-study questionnaire was assessed. Participants were then asked to create an avatar (personalized condition) or to inspect the generic avatar (generic avatar condition). In the personalized condition, participants were taught the avatar generator software and provided help. Participants were permitted to modify the virtual character’s body measures (form, proportions), facial appearance, and hairstyle in 15 min of preparation time. The clothing was kept similar in the two conditions. Following the avatar creation/inspection, a similarity measure was assessed. Participants were then calibrated and had time to acclimatize to the simulation. Similar to Study 1, audio instructions asking participants to perform movements and focus on body parts were used for the embodiment exposition. In addition to the instructions for Study 1, the participants were specifically instructed to step closer to the mirror and pay attention to features of the character’s appearance (e.g., “Look at the eyes of the mirrored self,” “Look at the mouth of the mirrored self,” “Turn 90 degrees and look at the mirrored self from the side”). The instructions lasted for 180 s and are described in detail in the supplementary material.

**Apparatus** The apparatus consisted of a setup identical to the HMD-based setup of Study 1, except that a FOVE 0 HMD (2560 px × 1440 px, 70 Hz, 100 degrees field of view) was used as the display.

**Measures** We measured demographic variables, the VEQ (αs ≥ .744), and affect using the PANAS scale in the short form [70] (PA: α = .840; NA α = .319, dropped from analyses). As we expected that personalization may also have an impact on self-presence, we assessed Ratan and Hasler’s self-presence questionnaire [54], adapted to the context of the study (excluded: “To what extent does your avatars profile info reflect some aspect of your personal identity?” and “To what extent does your avatars name represent some aspect of your personal identity?”). Proto-self-presence (α = .781) was assessed with items such as “How much do you feel like your avatar is an extension of your body within the virtual environment?” Core self-presence (α = .383) was assessed with items such as “When arousing events would happen to your avatar, to what extent do you feel aroused?” Extended self-presence (α = .661) was assessed with items like “To what extent is your avatars gender related to some aspect of your personal identity?” (5-point scale, see [54]). To control for the manipulation, we measured the perceived similarity toward the avatar (generic or personalized) before the exposure (desktop monitor), and after the exposure (reflecting the experience): “Please rate how much the virtual character is similar to you on the following scale, where 1 equals no similarity and 11 equals a digital twin.” Further measures such as sickness and humanness are not part of the present discussion. No severe sickness effects occurred.

**Participants** We excluded participants when problems or severe tracking errors were noted. The final sample consisted of participants (generic avatar: N = 25, personalized avatar: N = 23. 27 female, 21 male, Mage = 21.64, SDage = 2.25). All 48 participants were students, and 45 participants had previous experience with VR.

3.2.2 Results
**Comparisons** T-tests did not reveal significant differences for the VEQ factors. The mean values for ownership were higher in the personalized condition (M = 4.85, SD = 0.84) compared to the
Table 5: Significant Bivariate Pearson Correlations ($r$) – Study 2.

|       | AG     | CH     | PSP    | CSP    | PA     | SPre   | SPo    |
|-------|--------|--------|--------|--------|--------|--------|--------|
| OW    | .51**  | .37**  | .70**  | .33*   |        |        |        |
| AG    | .65**  | .35*   | .43**  | .31*   |        |        |        |
| CH    | .30*   | .37*   | .33*   | .43**  |        |        |        |
| ESP   | .40**  | .36*   | .32*   |        |        |        |        |
| CSP   | .31*   | .35*   | .31*   |        |        |        |        |
| PA    | .34**  |        |        |        |        |        |        |
| SPre  |        |        |        |        |        |        |        |
| SPo   |        |        |        |        |        |        |        |

Note: * $p < .05$; ** $p < .01$. OW ownership, AG agency, CH change, PSP proto-self-presence, CSP core self-presence, PA positive affect, SPre similarity pre-exposure, SPo similarity post-exposure.

3.3 Study 3: Behavioral Realism

Agency specifically relates to controlling one’s own movements [72], we hypothesized H3.1: Increased behavioral realism increases the perceived agency over a virtual body.

3.3.1 Method

In a two-factor (Facial Expressions, Gaze) between-subjects design, we evaluated the impact of behavioral realism. Participants were represented by generic avatars (see Fig. 4 top) and randomly assigned to one of four conditions: body motion only (BO), body and facial motion replication (BF), body and gaze motion replication (BG), and body, gaze, and face motion replication (BFG).

Procedure The procedure followed the general procedure (Fig. 3). Participants were welcomed and informed, before we assessed the pre-study questionnaire. Participants were then calibrated and had time to acclimatize. Audio instructions then asked participants to perform bodily movements and focus on body parts. In Study 3, the participants were specifically instructed to move closer to another marking in front of the mirror, and let their gaze wander to specific focus points, trying to ensure an influence in perception of the manipulation (e.g., “Fixate on the left eye of your mirrored-self,” “Focus on the right ear of your mirrored self,” etc.). For the facial expressions, we asked participants to perform certain expressions (e.g., “Open and close your mouth,” “Try to express happiness by smiling at your mirrored-self,” etc.). The instructions lasted for 219 s and are described in the supplementary material. After the exposure, we assessed the dependent measures.

Apparatus We used the same apparatus and generic avatars as in Study 2, and included the tracking of the participant’s gaze using the FOVE 0’s eye tracking system as well as facial expression tracking performed by a BinaryVR Dev Kit V1. Combining the two individual eye vectors, the FOVE integration calculates the intersection point to estimate the convergence point in the virtual scene, which is the approximate focus point of the user [79]. From this position, we recalculated the eyeball rotation. The BinaryVR Dev Kit was used to gather information about lower facial deformation. The 3D depth-sensing device Pico Flexx (up to 45 Hz) was affixed to the HMDs. Its 2D and depth images were processed by the BinaryVR Dev Kit to generate facial deformation parameters [29, 53]. Tracking expression parameters were mapped to blendshapes, for example, jaw open, and smile. The body tracking, facial expression, and gaze data was fused in the simulation to drive the avatar (see Fig. 4).

Measures We assessed the VEQ ($\alpha_s \geq .732$), a rating of the avatar assessing humanness, eeriness, and attractiveness ($\alpha_s \geq .688$) [26], the self-presence measures previously applied in Study 2 ($\alpha_s \geq .645$) [54], as well as affect ($\alpha_s \geq .674$) [28, 70]. We further asked how real, how natural, and how synchronous the motion behavior of the avatar appeared to the participants: “The movements were realistic,” “The movements were naturalistic,” “The movements were in synchrony to my own movements” (1–strongly disagree, 7–strongly agree). Further measures were excluded due to page limitations. No severe sickness effects occurred.
Participants  We excluded participants when problems or severe tracking errors were noted. The final sample for the analysis consisted of 70 participants (46 female, 24 male, $M_{age} = 21.3$, $SD_{age} = 1.82$, all students), of whom 65 had previous VR experience. 17 participants were assigned to the BO condition, 18 to BF, 18 to BG, and 17 to BFG.

3.3.2 Results

Comparisons  We calculated two-factor (gaze, facial expression) analyses of variance (ANOVAs). Although the BFG condition was rated highest in ownership ($M_s$ between 4.28 and 4.68) and highest in control ($M_s$ between 5.65 and 5.99), the differences were not significant. We found a significant effect of gaze on the perceived humanness; $F(1, 66) = 7.826$, $p = .007$, $\eta^2_p = .106$, indicating that the perceived humanness was higher in the conditions with enabled gaze tracking and replication ($M = 3.07$, $SD = 0.78$) compared to the conditions with disabled gaze tracking and replication ($M = 2.60$, $SD = 0.60$). No further significant effects were observed.

3.3.3 Discussion

We did not find supporting evidence for H1.3. Increased behavioral realism did not significantly impact the perceived embodiment between conditions. The additional gaze replication resulted in a greater perception of humanness. Regarding the perceived agency, we interpret that the facial expression and gaze replication did not have a strong influence in comparison to body movement, which was always present and represents greater motion dynamics, and arguably evokes a stronger visual stimulation. Future research should investigate body movement as an additional manipulation factor. Some limitations arise. The detection of facial and gaze behavior was limited by the visual resolution of the display. However, the participants stood especially close (about 50 cm) to the virtual mirror during parts of the induction phase. Third, the induction phase might have been too short for some participants, and as a result, the time spent on the gaze and facial expression interaction might not have been sufficient.

3.4 Study 4: Latency and Latency Jitter

In an independent fourth study, we further validated the scale, specifically targeting the agency factor. Previous work suggested that latency negatively impacts virtual embodiment [77]. Thus, we hypothesized H4.1: Latency negatively impacts components of virtual embodiment. Although linear latencies were subject to previous investigations [77], the impact of latency jitter has not been assessed.

3.4.1 Method

In a one-factor (Latency Level) repeated-measures design, we evaluated the impact of latency and latency jitter, meaning the non-periodic spontaneous peaks of latency, on the perception of the factors of the VEQ. Participants were exposed to four conditions of delayed and jitter-delayed simulation display (baseline, LB; small latency, LS; large latency, LL; latency jitter, LJ).

Procedure  The procedure followed the procedure depicted in Fig. 3 in repeated fashion. We welcomed participants and informed them, before assessing the pre-study questionnaire. The exposure conditions were then presented to the participants in randomized order. In each trial, the participants were calibrated, exposed to the simulation and induction, followed by an assessment of the dependent measures. In the audio instructions, the participants were specifically instructed to perform more rapid and fluid movements (e.g., “Rubber hand moved just like I wanted it to, as if it was obeying my will”). The complete audio instructions are described in the supplementary material. The instructions lasted for 282 s.

Apparatus  We used a similar apparatus Study 2. By buffering the tracking input data from the motion tracking system, artificial delays were introduced into the simulation by buffering and delaying the replication of the body motion tracking data. We prevented biasing sickness effects and influenced only the delay of the body movements, but did not influence the delay of the virtual camera (head movements, respectively), which, therefore, transformed according to the raw system delay without further modifications. Thus, the body motion was delayed, whereas the camera/head pose was not. We adapted the procedure described by Stauffert, et al. [65] to introduce latency jitter, which uses a stochastical model for latency distribution, introducing high-latency spikes into the simulation. Motion-to-photon latency of the resulting simulations was approximated by video frame counting (Canon, 1000 Hz). The measures resulted in $M = 90.12$ ms ($SD = 16.14$ ms for the simulation baseline $L_B$, $M = 206.93$ ms ($SD = 16.47$ ms) for the small delay $L_S$, $M = 353.07$ ms ($SD = 15.38$ ms) for the larger delay $L_L$, and $M = 102.58$ ms ($SD = 49.71$ ms) for $L_J$, $L_B$, $L_S$, and $L_J$ were measured with 60 samples (see Fig. 6). Despite measuring $N = 165$ repetitions in the jitter condition $L_J$, the resulting mean and SD may not accurately reflect the induced jitter, due to some spikes that could not be captured using the motion-apex measurement applied. Users were embodied with the generic avatars (see Fig. 4).

Measures  We assessed the VEQ ($\alpha \geq .771$) and performed a comparison to the questionnaire developed for RHI experiments by Kalckert and Ehrsson (KE) [31] (partly adapted from [41]), that measures ownership ($\alpha \geq .801$), ownership control ($\alpha \geq .655$), agency ($\alpha \geq .636$), and agency control ($\alpha$ between .239 and .563) with a 7-point Likert-type scale (1—strongly disagree, 7—strongly agree). Questioning was adapted to fit the scenario (e.g., “The rubber hand moved just like I wanted it to, as if it was obeying my will”). As manipulation control, we asked how realistic, natural, and synchronous the movements of the avatar appeared to the participants: “The movements were realistic,” “The movements were naturalistic,” “The movements were in synchrony to my own movements” (1—strongly disagree, 7—strongly agree). Further measures are not part of the present discussion. No severe sickness effects occurred.

Participants  We excluded participants when problems or severe tracking errors occurred. The final sample consisted of 22 participants (17 female, 5 male, $M_{age} = 21.77$, $SD_{age} = 3.62$, 22 students), of whom 21 had previous experience with VR.

3.4.2 Results

Comparisons  We calculated repeated-measures ANOVAs for each dependent variable. Where the assumption of sphericity was violated, we report Greenhouse–Geisser corrected values. Fig. 6 depicts the descriptive results. We found a significant main effect for ownership; $F(3, 63) = 3.57$, $p = .024$, $\eta^2_p = .138$. Similarly, agency measure was affected; $F(1.741, 63.553) = 7.50$, $p = .003$, $\eta^2_p = .263$. No significant impacts were observed for change. We did not observe any significant main effects in the agency and ownership measures of the scale adapted from KE [31]. The synchronicity assessment showed a significant main effect $F(1.905, 40.005) = 7.077$, $p = .003$, $\eta^2_p = .252$. In contrast, neither the realism, nor the naturalness of behavior showed a significant main effect. The strongest linear latency injection (LLJ) yielded to the lowest scoring of ownership and agency (see Fig. 6). The jitter condition was similarly affected, but resulted in significantly better ratings than the condition with the largest latency injection. The lower level linear latency (LS) had comparable results to the jitter-injected condition. Similarly, the synchronicity ratings were affected (see Fig. 6).

Correlations  Pearson correlations are depicted in Table 6. The synchronicity assessment showed large correlations with the agency
Although we found expected correlations between the ownership body part) scenarios. Although the results confirmed the expected factor of the VEQ and the agency factor of the scale from KE. Furthermore, synchronicity showed a medium to large correlation with the ownership factor, and the KE ownership factor. In turn, ownership showed medium to large correlations with agency, and a large correlation with ownership (KE) and ownership control (KE) across all measures. Interestingly, we did not find stable correlations between the perceived naturalness, synchronicity, and realism of the movement, providing room for further discussion.

### 3.4.3 Discussion

The results support H4.1, that embodiment is negatively affected by simulation delays, which is in line with previous findings. The descriptive results reveal that a jitter (in the applied spectrum) can result in a decreased perception of ownership and agency compared to an average latency of approximately 207 ms, which emphasizes the negative impact of latency jitter. However, the total delay of about 353 ms performed worse than the jitter condition, which quantifies the impact of jitter to some extent. The correlations confirmed that the VEQ picks up modifications in movement synchronicity. Although we found expected correlations between the agency factor and the additionally assessed agency (KE) and ownership (KE) measures, variance analysis revealed that the VEQ showed greater sensitivity in the present scenario. This may result from the fact that the (KE) questions were developed for the RHI (i.e., physical world body part) scenarios. Although the results confirmed the expected correlations between ownership and agency, we could show that the change factor was not affected by modifications in the motion dimension, and thus, the VEQ validly discriminates the factors to this regard. One limitation is that the agency control measure adapted from the (KE) questionnaire had overall low reliabilities, which is, in turn, a further argument for the proposed VEQ. In conclusion, we found significant impacts of latency and jitter on ownership and agency, and could further quantify the impact of jitter.

### 4 General Discussion and Conclusion

Overarching the individual findings of the studies, our high-level goal was the construction and validation of the Virtual Embodiment Questionnaire (VEQ). The proposed scale is not merely theoretical, but instead, its factors are statistically confirmed through rigorous scale development. The CFA confirmed the previously explored factor of the change in the perceived body schema, which could especially be relevant for VR applications in the context of therapy and disorder treatment, as we assume it is a potential predecessor of the Proteus effect. Future research may investigate potential moderations, e.g., whether higher ownership fosters a change perception. Regarding the performance of the VEQ, acceptable to good reliabilities were confirmed with analyses of Cronbach’s α throughout all studies. In contrast to previous component identifications, the VEQ focuses on virtual experiences. Further, the VEQ showed higher sensitivity to manipulations compared to a previous scale constructions that aimed at assessing physical world RHI-like experiences. Compared to related measures for physical world experiments (KE), the VEQ reacted more sensitive to performed latency manipulations. The VEQ also reacted validly and sensitively in the regard that we could confirm and further specify the impact of immersion: immersion specifically addresses the perception of a change in the perceived body schema.

We further showed that the proposed measure to part correlates with related constructs of presence or self-presence but discriminates those from virtual embodiment components. By investigating internal correlations of the VEQ, we could also confirm a correlation between agency and ownership as mentioned in previous works. The change component assessing the perceived change of one’s own body schema seems to be separable, in most regards. We interpret the confirmation of this component to be specifically relevant for VR. We confirmed that immersion (first-person perspective and HMD display) is a driver of this component, and thus desktop games or simulations or third-person VR simulations may not evoke this perception.

Some limitations arise. In contrast to the suggestions by Gonzalez-Franco and Peck, the VEQ does not include “external appearance,” or “response to external stimuli,” which may be the subject of future work. However, as also noted by Gonzalez-Franco and Peck and reviewed in Section 1, the previous literature topics body ownership, agency, and self-location in the context of embodiment. The VEQ lacks the assessment of self-location. The initial and updated questionnaire did not strictly concentrate on this factor. We argue that disturbed self-location (self-localization) is a factor typically present with disorders (paroxysmal illusions) evoking disrupted body perception, such as heautoscopy, out-of-body experiences, or autoscopic hallucinations. Although self-location discrepancies can be assessed in RHI exper-
ments [11] or out-of-body experiences [4], it is typically not the goal of VR applications to evoke such effects but rather to accurately reassemble the user’s location, first-person perspective (see [32] for a discussion), and behavior. Therefore, self-location may be subject to future extensions or alternatively assessed with implicit measures, such as displacement measures [11]. The present findings are also limited by the applied scenarios and the collected sample form and size. Future work should consequently assess the reliability and consistency of the VEQ, along with its application to different (non-mirror) scenarios, such as user-embodying games, and more abstract avatar types with varying form and appearances. In this regard, we can only assume that the measure is objective (i.e., shielded from third-party bias) on the basis of our results, as shown in a multi-study setting. The questionnaire is available at [url] in [languages]. The simulation codes are available upon request.

In conclusion, we presented the creation of a validated virtual embodiment questionnaire (VEQ) that can be applied to various VR experiments to assess latent factors of virtual embodiment.

**References**

[1] P. L. Anderson, M. Price, S. M. Edwards, M. A. Obasaju, S. K. Schmertz, E. Zimand, and M. R. Calamaras. Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. 81(5):751, 2013. doi: 10.1037/a0033559

[2] F. Argelaguet, L. Hoyet, M. Trico, and A. Lécuyer. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In 2016 IEEE Virtual Reality (VR), pp. 3–10. IEEE, 2016. doi: 10.1109/VR.2016.7504682

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

[4] L. Aymerich-Franch, D. Petit, G. Ganesh, and A. Kheddar. Embodiment of a humanoid robot is preserved during partial and delayed control. In 2015 IEEE International Workshop on Advanced Robotics and Its Social impacts, 2015. doi: 10.1109/ARSO.2015.7428218

[5] J. N. Bailenson and J. Blascovich. Avatars, 2004.

[6] S. Beck, A. Kunert, A. Kulik, and B. Froehlich. Immersive group-to-group telepresence. 19(4):616–625, 2013. doi: 10.1109/TVCG.2013.33

[7] S. Benford, J. Bowers, L. E. Fahlen, C. Greenhalgh, and D. Snowdon. User embodiment in collaborative virtual environments. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 242–249. ACM Press/Addison-Wesley Publishing Co., 1995. doi: 10.1145/232904.232935

[8] O. Blanke. Multisensory brain mechanisms of bodily self-consciousness. 13(8):556, 2012. doi: 10.1006/dgdh.3

[9] O. Blanke and T. Metzinger. Full-body illusions and minimal phenomenal selfhood. 13(1):7–13, 2009. doi: 10.1016/j.tics.2008.10.003

[10] G. O. Boateng, T. B. Neillands, A. E. Frongillo, H. R. Melgar-Quionez, and S. L. Young. Best practices for developing and validating scales for health, social, and behavioral research: A primer. Frontiers in Public Health, 6:149, 2018. doi: 10.3389/fpubh.2018.00149

[11] M. Botvinick and J. Cohen. Rubber hands’ feel’touch that eyes see. 17(2):523–534, 2008. doi: 10.1016/j.concog.2008.03.004

[12] H. H. Ehrsson. The experimental induction of out-of-body experiences. 317(5841):1048–1048, 2007. doi: 10/cb9w6h

[13] M. Furr. Scale construction and psychometrics for social and personality psychology. SAGE Publications Ltd, 2011.

[14] S. Gallagher. Philosophical conceptions of the self: Implications for cognitive science. 4(1):14–21, 2000. doi: 10.1016/S1364-6613(99)01417-5

[15] S. Gallagher. How the Body Shapes the Mind. Clarendon Press, 2006.

[16] M. González-Franco, D. Pérez-Marcos, B. Spanlang, and M. Slater. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In Virtual Reality Conference (VR), 2010 IEEE, pp. 111–114. IEEE, 2010. doi: 10.1109/VR.2010.5444805

[17] M. G.-F. Gonzalez-Franco and T. C. Peck. Avatar Embodiment. Towards a Standardized Questionnaire. 5:74, 2018. doi: 10.3389/frobt.2018.00074

[18] A. Haans, W. A. Jisselesteijn, and Y. A. de Kort. The effect of similarities in skin texture and hand shape on perceived ownership of a fake limb. 5(4):389–394, 2008. doi: 10.1016/j.bodyim.2008.04.003

[19] D. He, F. Liu, D. Pape, G. Dawe, and D. Sandin. Video-based measurement of system latency. In International Immersive Projection Technology Workshop, 2000.

[20] C.-C. Ho and K. F. MacDormain. Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. 26(6):1508–1518, 2010-11. doi: 10.1016/j.chb.2010.05.015

[21] L. Hoyet, F. Argelaguet, C. Nicole, and A. Lécuyer. “Wow! I have six Fingers!”: Would You accept structural changes of Your hand in VR? 3:27, 2016. doi: 10.3389/frobt.2016.00027

[22] S. Janke and A. Göckler-Rist. Deutsche Version der Positive and Negative Affect Schedule (PANAS). In Zusammenstellung Sozialwissenschaftlicher Items Und Skalen. Doi, vol. 10, p. 6102, 2014.

[23] Y. U. Jihun and P. Jungwoon. Head-mounted display with facial expression detecting capability, 2017-03.

[24] S. Jung, G. Bruder, P. J. Wisniewski, C. Sandor, and C. E. Hughes. Over my hand: Using a personalized hand in vr to improve object size estimation, body ownership, and presence. In Proceedings of the Symposium on Spatial User Interaction, pp. 60–68. ACM, 2018.

[25] A. Kalckert and H. H. Ehrsson. Moving a rubber hand that feels like your own: A dissociation of ownership and agency. 6, 2012. doi: 10.3389/fnhum.2012.00040

[26] K. Kilteni, R. Groten, and M. Slater. The sense of embodiment in virtual reality. 21(4):373–387, 2012. doi: 10.1162/PRES_a.00012

[27] 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. 9, 2015. doi: 10.3389/fnhum.2015.00141

[28] R. Kleijer. Enter the avatar: The phenomenology of prosthetic telepresence in computer games. In The Philosophy of Computer Games, pp. 17–38. Springer, 2012.

[29] R. B. Kline. Principles and Practice of Structural Equation Modeling 3 Rd Ed. New York, NY, The Guilford Press, 2010.

[30] A. Kilik, A. Kunert, S. Beck, C.-F. Matthes, A. Schollmeyer, A. Kreskowski, B. Fröhlich, S. Cobb, and M. D’Cruz. Virtual Valcarmónica: Collaborative exploration of prehistoric petroglyphs and their surrounding environment in multi-user virtual reality. 26(03):297–321, 2018. doi: 10.1162/prea_a_00297

[31] M. 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 ACM Symposium on Virtual Reality Software and Technology, pp. 39:1–39:10, 2017. doi: 10.1145/3139131.3139156

[32] M. E. Latoschik, J.-L. Lugrin, and D. Roth. FakeMi: A fake mirror system for avatar embodiment studies. In Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology, pp. 73–76. ACM, 2016. doi: 10.1145/2903369.2903399

[33] B. Lenggenhager, M. Mouthon, and O. Blanke. Spatial aspects of
[82] N. Yee and J. N. Bailenson. Walk a mile in digital shoes: The impact of embodied perspective-taking on the reduction of negative stereotyping in immersive virtual environments. pp. 24–26, 2006.

[83] J. Yu and J. Park. Real-time facial tracking in virtual reality. In SIGGRAPH ASIA 2016 VR Showcase, p. 4. ACM, 2016. doi: 10.1145/2996376.2996390