



Avatar Embodiment. A Standardized Questionnaire

Tabitha C. Peck^{1*} and Mar Gonzalez-Franco²

¹Department of Mathematics and Computer Science, Davidson College, Davidson, NC, United States, ²Microsoft Research, Redmond, WA, United States

The aim of this paper is to further the understanding of embodiment by 1) analytically determining the components defining embodiment, 2) increasing comparability and standardization of the measurement of embodiment across experiments by providing a universal embodiment questionnaire that is validated and reliable, and 3) motivating researchers to use a standardized questionnaire. In this paper we validate numerically and refine our previously proposed Embodiment Questionnaire. We collected data from nine experiments, with over 400 questionnaires, that used all or part of the original embodiment 25-item questionnaire. Analysis was performed to eliminate non-universal questions, redundant questions, and questions that were not strongly correlated with other questions. We further numerically categorized and weighted sub-scales and determined that embodiment is comprised of interrelated categories of Appearance, Response, Ownership, and Multi-Sensory. The final questionnaire consists of 16 questions and four interrelated sub-scales with high reliability within each sub-scale, Chronbach's α ranged from 0.72 to 0.82. Results of the original and refined questionnaire are compared over all nine experiments and in detail for three of the experiments. The updated questionnaire produced a wider range of embodiment scores compared to the original questionnaire, was able to detect the presence of a self-avatar, and was able to discern that participants over 30 years of age have significantly lower embodiment scores compared to participants under 30 years of age. Removed questions and further research of interest to the community are discussed.

OPEN ACCESS

Edited by:

David Swapp,
University College London,
United Kingdom

Reviewed by:

Takuji Narumi,
The University of Tokyo, Japan
Lorenzo Pia,
University of Turin, Italy

*Correspondence:

Tabitha C. Peck
tapeck@davidson.edu

Specialty section:

This article was submitted to
Virtual Reality and Human Behaviour,
a section of the journal
Frontiers in Virtual Reality

Received: 24 June 2020

Accepted: 21 December 2020

Published: 09 February 2021

Citation:

Peck TC and Gonzalez-Franco M
(2021) Avatar Embodiment. A
Standardized Questionnaire.
Front. Virtual Real. 1:575943.
doi: 10.3389/fvrr.2020.575943

Keywords: embodiment, avatar, virtual reality, rubber hand illusion, body ownership, agency, questionnaire, age

1 INTRODUCTION

Embodiment science is the field inside virtual reality (VR) research that studies and attempts to understand the effects of self-avatars on its users (Spanlang et al., 2014). Embodied avatars are defined to be avatars that are co-located with the user's body and seen from a first person perspective within an immersive virtual environment (VE) (Kiltani et al., 2012).

Research in this field has shown evidence of the importance of being embodied in the self-avatar. Beyond the obvious needs of being virtually represented to interact with others in social VR setups, being embodied has been shown to increase users cognitive abilities (Steed et al., 2016), improve haptic performance (Maselli et al., 2016; Gonzalez-Franco and Berger, 2019) or increase self recognition and identification through enfacement (Gonzalez-Franco et al., 2020b). However, the cognitive load impacts of self-avatars are not well understood and may affect results (Peck et al., 2018; Peck and Tutar, 2020). Lush et al. (2020) has raised concerns that the illusion may be a response to imaginative suggestion and is caused by suggestion. Regardless, being embodied in an

avatar can dramatically change a user's experience in VR, including reducing biases, such as racial bias (Peck et al., 2013), mitigating stereotype threat (Peck et al., 2018; Peck et al., 2020a), responding to a domestic violence scenario (Gonzalez-Liencrez et al., 2020), or affecting how users move and act inside VR (Kilteni et al., 2013; Gonzalez-Franco et al., 2020a).

A standardized measurement of embodiment is needed to be able to compare and replicate experiments across embodiment science. Further, participants are unique and can have significantly different experiences and responses in the same VR setup. For example, avatars have been shown to enhance distance estimation (Ries et al., 2008; Ebrahimi et al., 2018), object size estimation (Jung et al., 2018; Ogawa et al., 2019), and the level of embodiment in an avatar may further affect distance perception (Gonzalez-Franco et al., 2019). A standardized embodiment questionnaire that is sensitive enough to detect individual embodiment differences could aid researchers in better understanding and interpreting the effects of virtual embodiment. Research using self-avatars should measure embodiment on every experiment to rule out and understand intrinsic variables that might be affecting results.

There are challenges to measuring embodiment due to the relation with our own bodies not being something we normally think about. Questions such as "I felt out of my body", can be unrelatable to the many people who have neither experienced nor heard about an autoscopic phenomena (Blanke and Mohr, 2005). There have been attempts to measure embodiment to bridge the gap between the user's physiological experience and reporting of the experience. The most advanced measures include electrophysiological recordings to find quantitative methods (González Franco, 2014; Alchalabi et al., 2019). Previous research has shown that highly embodied participants responded with stronger 400N amplitudes in the parietal cortex when they lost agency over their bodies (Padrao et al., 2016; Pavone et al., 2016), or had stronger P300 responses when their virtual avatar was threatened (González-Franco et al., 2014). In both experiments researchers found correlations between these numerical and physiological responses and a series of embodiment questions. This supports that subjective questionnaires are less cumbersome and yet still a valid form of embodiment evaluation. However, questions need to account for the challenge of asking someone about something they are not able to quantify. A similar challenge was introduced by Slater when addressing why questionnaires cannot fully assess presence in VEs, (Slater, 2004).

With that aim in mind a new questionnaire was introduced by (Gonzalez-Franco and Peck, 2018) "Avatar Embodiment. Toward a Standardized Questionnaire." In this questionnaire the authors proposed 25 questions for an embodiment questionnaire, as collected from the most used embodiment questions in the literature. Starting with the original rubber hand illusion introduced by Botvinick and Cohen (Botvinick and Cohen, 1998) the authors analyzed up to 30 other experiments. The new questionnaire not only included sufficient control questions but also categorized the main

questions into six different recurrent themes in embodiment science (Kilteni et al., 2012, 2015; Maselli and Slater, 2013, Maselli and Slater, 2014) such as: 1) body ownership, 2) agency and motor control, 3) tactile sensations, 4) location of the body, 5) external appearance and 6) response to external stimuli. The paper proposed a calculation for a final embodiment score that could be achieved arithmetically or through an open sourced Principal Component Analysis (PCA).

However, given the nature of the questionnaire based only on a review of previously used questions, the authors carefully titled the paper as an ongoing effort *Toward a standardized questionnaire*, and highlighted that there would be a need to further validate the questions by the community.

In this paper we validate the original questions, improve upon the proposed questionnaire using exploratory factor analysis, and demonstrate the usefulness of the new questionnaire with additional analysis for three studies. Over the course of the last 3 years we collected data from nine experiments that used all or part of the original questionnaire. This accumulated a total of 443 completed questionnaires. We then completed exploratory factor analysis with three main goals:

1. Determine a relevant subset of questions by removing redundant and unrelated questions to create a universally usable questionnaire.
2. Group related questions into sub-scales and provide an accurate and usable embodiment calculation.
3. Validate that the revised questionnaire is reliable and able to better discriminate between individual participant embodiment scores.

Our analysis converged into a reduced questionnaire that proposes 16 questions of the original 25, and four sub-scales instead of the original six. The revised questionnaire is compared to the original on $n = 101$ participants, the results of which are used in the discussion of this paper to further understand the revised embodiment questionnaire. We also discuss potential reasons why some commonly used questions rendered irrelevant, as well as highlight the importance of using a common questionnaire across the field.

2 MATERIALS AND EQUIPMENT

All studies run by either author since the original embodiment questionnaire was proposed that collected data from all or part of the original embodiment questionnaire (Gonzalez-Franco and Peck, 2018) were included in the data analysis. This included nine user studies. There were a total of 443 questionnaires (Table 1) completed by 124 women, 199 men, two non-binary participants, and 118 non-reported. The participant ages ranged from 18 to 76 with an average age of 31.65 ± 11.22 . The diversity of the studies is also of interest to our validation, hence we strived to collect data from experiments in motor control, haptics and perception (Berger and Gonzalez-Franco, 2018; Lee et al., 2019; Peck and Tutar, 2020), distance estimation and locomotion (Abtahi et al., 2019; Gonzalez-

TABLE 1 | Demographic and questionnaire information for each of the nine studies used to revise the embodiment questionnaire. Information includes the range of participant ages, the number of self-identified male, female and non-binary participants, the total number of questionnaires completed, the number of questions used from the original questionnaire, and the number of questions in the revised questionnaire.

Study #	Age	Male	Female	Non-binary	Questionnaires	Original questions	Revised questions
#1	(23, 69)				46	16	10
#2	(25, 55)	76	25		101	9	6
#3	(26, 69)				52	15	9
#4	(18, 24)	15	15		30	11	7
#5	(23, 52)	8	6		14	10	5
#6	(24, 45)	9	2		31	6	4
#7	(22, 53)	26	10		36	8	6
#8	(18, 76)	43	56	2	101	25	16
#9	(25, 61)	22	10		32	16	11

Franco et al., 2019), facial animation (Gonzalez-Franco et al., 2020b) and behavioural applications (Seitz et al., 2020). This included experiments that were within and between participants as well as single condition experiments in which embodiment was measured to find the effects of high and low embodiment on a secondary measure (Gonzalez-Franco et al., 2019; Seitz et al., 2020). Embodiment further varied by including full body-swap illusions, partial body-swap illusions, hand representation illusions and avatar body modifications.

3 METHODS

Exploratory factor analysis using Principal Component Analysis (PCA) was performed following the recommended procedure proposed by Field et al. (Field et al., 2012). A correlation matrix between participant responses of each question was calculated and Bartlett's test of sphericity, χ^2 , was calculated to verify significantly large correlations between items to perform PCA analysis. Kaiser-Meyer-Olkin (KMO) measures were used to verify sampling adequacy and question removal following the recommendation that individual items have to be above the acceptable limit of 0.5. PCA was used to detect the main factors explored by the questionnaire. PCA analysis calculates loadings for more relevant questions that have greater variability among participants, and clusters them based on their algebraic alignment. The factors that emerge from PCA on the questionnaire responses are selected using Kaiser's criterion of one.

4 RESULTS

For the remainder of the paper we will refer to questions from (Gonzalez-Franco and Peck, 2018) as Q# and the revised questions as R#. We first identified questions that could not be universally adapted to embodiment experiments. This included Q4 and Q5 that reference a mirror and Q25 which specifically references harm to the avatar. Other questions should be adaptable to any experiment. For example, VR experiments should at least provide a first-person perspective that controls the view of the scene (agency) even if the avatar does not move.

Additionally some form of passive haptics such as feet touching the floor should exist (tactile). Recommendations about adapting questions to studies can be found in **Section 5**.

After removing these three non-universal questions (Q4, Q5, and Q25) we explored the correlation matrix to determine if questions were correlated with each other and to make sure that no questions had values above 0.9. Three questions were identified that had correlation values above 0.3 with only two other questions, namely Q2, Q7, and Q23. Additionally, Q9 and Q19 had correlation values above 0.3 with only three other questions.

We further removed the three questions that were only correlated with two other questions (Q2, Q7, and Q23). Using the correlation matrix of the remaining 19 questions we verified sampling adequacy for analysis using the Kaiser-Meyer-Olkin (KMO) measure, $KMO = 0.78$ which is considered "good" (Hutcheson and Sofroniou, 1999). However, Q9 had the lowest individual KMO value of 0.56 which is close to the minimal 0.5 cut-off. We additionally remove Q9 due to it having the lowest individual KMO value and not being correlated with many other questions.

Principle component analysis was run on the remaining 18 questions. The KMO measure was 0.79 and all individual item KMO values were above 0.68. Bartlett's test of sphericity indicated that between item correlations were sufficiently large for PCA, $\chi^2(153) = 3159.33, p < 0.0001$.

Eigenvalues were computed for each component of the data. The first five principle components sum-of-squared loadings were greater than one and could be extracted based on Kaiser's criterion of one and this was further supported by a visual examination of the scree plot. These five components explained 64% of the variance in the data. The average communalities was 0.64 which is higher than the necessary 0.6 further supporting the extraction of five factors. Finally, the fit of the PCA model was 0.95 suggesting that the model is a good fit. Fewer than 50% of the model's residuals were greater than 0.05 with a root-mean square of 0.07. The structure matrix was created by rotating using oblique rotation due to the likelihood of sub-scales correlating with each other. Questions were grouped according the factor loadings of the structure matrix that were greater than |0.4|. The five clusterings of questions were categorized as Appearance, Response, Multi-Sensory,

TABLE 2 | Final Analysis with four factors extracted. References in the questionnaire to *body* would need to be updated if only a body part is being explored. Blank spaces ___ depend on the experiment, and the ones referring to touch (marked with *) can be filled with *floor*, if there are no further tactile stimuli.

Rev	Orig	Question	App	Res	Own	MSen
R1	Q15	"I felt out of my body"	0.81	-0.03	0.09	-0.07
R2	Q16	"I felt as if my (real) <i>body</i> were drifting toward the virtual <i>body</i> or as if the virtual <i>body</i> were drifting toward my (real) <i>body</i> "	0.69	0.09	-0.22	0.00
R3	Q8	"I felt as if the movements of the virtual <i>body</i> were influencing my own movements"	0.61	-0.15	-0.08	0.30
R4	Q17	"It felt as if my (real) <i>body</i> were turning into an 'avatar' <i>body</i> "	0.60	0.29	0.09	0.04
R5	Q18	"At some point it felt as if my real <i>body</i> was starting to take on the posture or shape of the virtual <i>body</i> that I saw"	0.57	0.05	0.33	0.07
R6	Q20	"I felt like I was wearing different clothes from when I came to the laboratory"	0.44	0.43	0.12	-0.09
R7	Q24	"I felt as if my <i>body</i> had changed"	0.03	0.84	-0.02	0.11
R8	Q22	"I felt a ___ sensation in my body when I saw ___"	-0.07	0.82	0.13	-0.03
R9	Q21	"I felt that my own <i>body</i> could be affected by ___"	0.15	0.70	-0.10	0.17
R10	Q1	"I felt as if the virtual <i>body</i> was my <i>body</i> "	-0.04	0.06	0.78	0.09
R11	Q19	"At some point it felt that the virtual <i>body</i> resembled my own (real) <i>body</i> , in terms of shape, skin tone or other visual features."	0.11	0.19	0.71	-0.19
R12	Q14	"I felt as if my <i>body</i> was located where I saw the virtual <i>body</i> "	0.13	-0.15	0.57	0.29
R13	Q6	"It felt like I could control the virtual <i>body</i> as if it was my own <i>body</i> "	-0.04	-0.17	0.55	0.32
R14	Q10	"It seemed as if I felt the touch of the ___ in the location where I saw the virtual <i>body</i> touched"	-0.12	0.11	0.21	0.79
R15	Q12	"It seemed as if the touch I felt was caused by the <i>body</i> touching the virtual ___"	0.07	0.29	-0.10	0.72
R16	Q13	"It seemed as if my <i>body</i> was touching the ___"	0.27	-0.10	0.02	0.66
		Eigenvalues	2.78	2.57	2.15	2.18
		Percent of variance	29%	27%	22%	23%

TABLE 3 | The structure matrix of the PCA analysis after performing oblique rotation on 16 questions. Values greater than |0.4| are in bold

Rev	Orig	App	Res	Own	MSen
R1	Q15	0.79	0.24	0.17	0.15
R2	Q16	0.70	0.30	-0.11	0.14
R3	Q8	0.62	0.12	0.06	0.40
R4	Q17	0.72	0.52	0.23	0.28
R5	Q18	0.64	0.31	0.43	0.32
R6	Q20	0.58	0.58	0.23	0.15
R7	Q24	0.35	0.87	0.15	0.31
R8	Q22	0.23	0.82	0.25	0.18
R9	Q21	0.42	0.77	0.08	0.34
R10	Q1	0.10	0.19	0.81	0.30
R11	Q19	0.21	0.30	0.70	0.07
R12	Q14	0.23	0.06	0.64	0.44
R13	Q6	0.05	-0.02	0.61	0.42
R14	Q10	0.14	0.28	0.43	0.84
R15	Q12	0.34	0.47	0.15	0.78
R16	Q13	0.41	0.15	0.22	0.71
	Chronbach's α	0.79	0.82	0.76	0.72

TABLE 4 | Pattern matrix of factor correlations.

	Appearance	Response	Ownership	Multi-sensory
Appearance	1.00	0.34	0.13	0.25
Response	0.34	1.00	0.17	0.23
Ownership	0.13	0.17	1.00	0.27
Multi-sensory	0.25	0.23	0.27	1.00

Ownership, and Location. However, the Location factor had low reliability, Chronbach's $\alpha = 0.52$, indicating that the factor and related questions (Q3 and Q11) should be removed from the final questionnaire. Further, it appears that Location is least correlated with the other factors with correlation factors ranging from 0.03 to 0.15 (see **Supplementary Tables S1–S3** of the supplementary material).

PCA was run on the remaining 16 questions, after removing Q3 and Q11 (**Table 2**).

The KMO measure further raised to 0.81, considered "great," Hutcheson and Sofroniou (1999), and all individual item KMO values were above 0.70. Bartlett's test of sphericity continued to indicate that between item correlations were sufficiently large for PCA, $\chi^2(120) = 2662.69$, $p < 0.0001$. Eigenvalues were

computed for each component of the data. The first four principle components sum-of-squared loadings were greater than one and extracted based on Kaiser's criterion of one which was further supported by a visual examination of the scree plot. These four components explained 61% of the variance in the data. The average communalities was 0.61 and the fit of the PCA model was 0.94. 50% of the model's residuals were greater than 0.05 with a root-mean square of 0.07. The residuals were normally distributed with no outliers. The factor loadings after oblique rotation are shown in **Table 2** with the corresponding structure matrix in **Table 3** and pattern matrix of factor correlations in **Table 4** indicating that the four sub-scales are interrelated.

Each of these four factors was inspected and no individual question had an item-rest correlation below 0.30. All questions with factor loadings above |0.40| were included in each sub-scale. Note that numerous questions contributed to two sub-scales further supporting the interrelation of the sub-scales.

Question groups included all questions with a weight above 0.4 as determined from the PCA structure matrix, (see **Table 3**). The score for each sub-group was calculated with equal weight given to each question. These four interrelated question groups each had high reliability, Chronbach's α of 0.79, 0.82, 0.76, and 0.72 respectively. (Gonzalez-Franco and Peck, 2018).

In total 16 of the original 25 questions remained. Nine questions were removed due to low relevance as indicated by the above analysis or not being able to be generally applied to embodiment experiments. All in all, after this thorough analysis and validation the questions removed from original questionnaire were:

- Q2. “It felt as if the virtual *body* I saw was someone else”
- Q3. “It seemed as if I might have more than one *body*”
- Q4. “I felt as if the virtual *body* I saw when looking in the mirror was my own *body*”
- Q5. “I felt as if the virtual *body* I saw when looking at myself in the mirror was another person”
- Q7. “The movements of the virtual *body* were caused by my movements”
- Q9. “I felt as if the virtual *body* was moving by itself”
- Q11. “It seemed as if the touch I felt was located somewhere between my physical *body* and the virtual *body*”
- Q23. “When ____ happened, I felt the instinct to ____”
- Q25. “I had the feeling that I might be harmed by the ____”

5 EMBODIMENT QUESTIONNAIRE

The final questionnaire consists of 16 questions that can be adapted for all embodiment experiments that enable some amount of agency, for example movement of the head or a hand. R7 allows for experiment specifics to customize the question based on the independent variable of the study. i.e. if a specific body swap or if a threat is involved, such as, “I felt as if my body was older” or “I felt as if my hand was attacked.” R8 and R9 can be adapted to non-threat situations such as “I felt a realistic sensation in my body when I saw my hand” or “I felt that my own body could have been affected by the virtual world.” Although we cannot guarantee that these questions are identical, they give freedom to the experimenter to customize the questionnaire for the many varying embodiment studies. The customization supports wider use of a standardized questionnaire and better comparability between experiments. In situations where there are no active touching situations the participant will likely still experience some form of passive haptics such as their feet touching the ground or their hand resting on a table.

We recommend collecting scores using a 7-point Likert scale ranging from never/strongly disagree to always/strongly agree. (This questionnaire is available for download in an editable form from the **Supplementary Material**).

- R1 “I felt out of my body”
- R2 “I felt as if my (real) *body* were drifting toward the virtual *body* or as if the virtual *body* were drifting toward my (real) *body*”
- R3 “I felt as if the movements of the virtual *body* were influencing my own movements”
- R4 “It felt as if my (real) *body* were turning into an “avatar” *body*”
- R5 “At some point it felt as if my real *body* was starting to take on the posture or shape of the virtual *body* that I saw”
- R6 “I felt like I was wearing different clothes from when I came to the laboratory”
- R7 “I felt as if my *body* had changed”
- R8 “I felt a ____ sensation in my *body* when I saw ____”
- R9 “I felt that my own *body* could be affected by ____”
- R10 “I felt as if the virtual *body* was my *body*”
- R11 “At some point it felt that the virtual *body* resembled my own (real) *body*, in terms of shape, skin tone or other visual features.”
- R12 “I felt as if my *body* was located where I saw the virtual *body*”
- R13 “I felt like I could control the virtual *body* as if it was my own *body*”
- R14 “It seemed as if I felt the touch of the ____* in the location where I saw the virtual *body* touched”
- R15 “It seemed as if the touch I felt was caused by the ____* touching the virtual *body*”
- R16 “It seemed as if my *body* was touching the ____*”

References in the questionnaire to *body* would need to be updated only if a body part is being explored. Blank spaces ____ depend on the experiment. The ____ marked with * can refer to the touch of the ground and the feet in touch with the virtual *floor*, this interpretation is preferable to removing the questions if there are no further tactile stimuli.

Ideally the experimental design will include these questions in a randomized order to limit context effects, and using a 7-point Likert-scale directly at the end of the experiment or of each condition if the study is within participants. The Likert-scale should range from:

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
(1)	(2)	(3)	(4)	(5)	(6)	(7)

Alternatively the following scale could be used.

Never	Almost never	Rarely	Half of the time	Often	Most of the time	Always
(1)	(2)	(3)	(4)	(5)	(6)	(7)

At the beginning of the questionnaire, it should be clear that the questions are related to the participants’ experience during the experiment. Starting the questionnaire with a sentence of the style: “During the experiment there were moments in which . . .” could help (see Appendix 2 in **Supplementary Material** for the ready-to-print questionnaire).

5.1 Computing the Score

The final embodiment score will be in a range from 1–7 indicating low to high embodiment. To compute a final embodiment score, calculate each sub-scale by averaging questions within each sub-category. Note that many questions are used in two sub-scales highlighting the correlations between the sub-scales. Average the final sub-scale scores to compute the final embodiment score. This equal weighting of sub-scales contributing to the final embodiment score was based on the percentage of variance of each principle component after applying the oblique rotation being roughly equivalent.

To reach the maximum replicability of future results we recommend the use of the following scores as retrieved from our large scale analysis.

- Appearance = $(R1 + R2 + R3 + R4 + R5 + R6 + R9 + R16)/8$
- Response = $(R4 + R6 + R7 + R8 + R9 + R15)/6$
- Ownership = $(R5 + R10 + R11 + R12 + R13 + R14)/6$
- Multi-Sensory = $(R3 + R12 + R13 + R14 + R15 + R16)/6$
- Embodiment = $(\text{Appearance} + \text{Response} + \text{Ownership} + \text{Multi-Sensory})/4$

Regarding the computation of the final score, we want to note that it is common practice for the creation of questionnaires to produce non-weighted arithmetic for the items (i.e. ignoring the factor loadings) which are summarized with weight one for each question and then only weight the whole sub-scale as a mean (Launois et al., 1996; Boateng et al., 2018). A different arrangement of sub-scales not based on the PCA loadings and components would mean the need of a weighted questionnaire that is much harder to use for the general public. The unweighted approach involves summing standardized item scores or raw item scores, or computing the mean for raw item scores (Armor, 1973).

5.1.1 Multi-Sensory Sub-scores

Our effort aims to create a sensitive and weight-free embodiment questionnaire. However additional sub-scales could be used for the analysis of agency and to enable backward compatibility with the prior questionnaire. These would directly relate to the multi-sensory sub-scale and could be useful for researchers working on this specific topic. The current questionnaire sub-scales re-samples all these questions on to the Multi-sensory score, however, we can put forward the following agency sub-score: Agency = $R3 + R13$.

6 VALIDATION

We test this new questionnaire computation against the original computation on the data collected from nine experiments. In **Figure 1** we can see how the embodiment scores changed in the different experiments between the new refined questionnaire and the original questionnaire. The refined questionnaire responses covers 94% of the scale compared to 87% of the scale with the original questionnaire. Additionally, the refined scale has more dispersion with a significantly wider standard deviation ($M = 1.00$, $SD = 0.24$) for each experiment compared to the

original questionnaire ($M = .76$, $SD = 0.22$), $t(8) = 2.78$, $p = 0.02$, $r = 0.47$. The main aspect to note is that there are no major changes beyond the amplification of the dynamic range of the refined scale. The scores are more spread which indicates more sensitivity of the refined questionnaire providing more granularity to identifying embodiment.

6.1 Category Validation

We ran a correlation study between the different categories that existed in the previous questionnaire (*app*, *own*, *loc*, *tac*, *ag*) and the proposed categories in the revised questionnaire (*appearance*, *response*, *ownership*, *multisensory*). **Figure 2** (left) validates the idea that the previously independent categories of location, tactile and agency are well represented in the *multisensory* aggregated category. From a category perspective the aggregation of sub-categories is supported by previous embodiment research depending on multisensory illusions. These illusions are achieved through multiple simulations including visuo-proprioceptive, visuo-tactile, and visuo-motor. Studies may not include three types of stimulation, and yet will support a multisensory experience.

We further find that the previous *appearance* category is highly correlated with the revised corresponding category with the same name. Note that *appearance* transversely affects the other sub-measures as well. If *appearance* is not supported, embodiment is affected, reducing both *response* and *ownership*. There have been many studies on the importance of visual appearance for the illusions (Ebrahimi et al., 2018; Jung et al., 2018; Ogawa et al., 2019), showing how the illusions did not elicit when participants were presented with incorrect bodies or hands, such as a wooden stick instead of a rubber hand.

A lower, yet still significant Pearson's correlation (0.47) was seen between the previous ownership measure, (*own*) and the new *ownership* sub-measure. Finally our new sub-measure *response*, was mostly affected by the previous *app* sub-measure.

In **Figure 2** Right, we explore the correlations between the new sub-measures, (note this study can be also extracted from our PCA analysis in the previous sections), we find that relatively high correlations (>0.67) were found between appearance and response as well as between the ownership and multisensory sub-measures.

We further select three experiments to compare how the previous and revised questionnaire perform and highlight the relevance of the sub-measures.

6.2 Detail Validation on Study #8

We compared the results of the newly proposed embodiment questionnaire to the previous questionnaire on data collected from study #8. This user study used all the previously proposed questions. The study was a within participants design where each participant was given a full-body avatar that was both gender and race matched to participants (Seitz et al., 2020). Participants were given a mirror in the environment, but they did not walk around. No harm came to the participant during the experiment and nothing in the experiment was designed to modify embodiment. Participants completed all 25 questions from the original questionnaire immediately after the experiment. Data was

collected from $n = 101$ participants (women = 56, men = 43, non-binary = 2), with ages ranging from 18 to 76, ($M = 23.02$, $SD = 9.91$).

The original embodiment questionnaire scores ($M = 4.10$, $SD = 0.59$), were slightly higher than the revised embodiment questionnaire scores, ($M = 3.48$, $SD = 1.02$). However, in line with the changes observed when analyzing the entire set of studies the updated questionnaire provided a wider range of scores covering 74% of the scale compared to only 45% of the scale with the original questionnaire. The wider variation in scores suggests that the updated questionnaire is more sensitive to the full range of subjective embodiment, and that subjective embodiment varies drastically by individuals. Due to the importance of evaluating and testing diverse populations to mitigate underrepresentation within VR (Peck et al., 2020b) we investigated if the revised questionnaire was sensitive enough to detect demographic differences should they exist. Previous findings suggest differences between age (Allen et al., 2000; Moffat et al., 2001). Due to the unequal sample sizes [Age: ≥ 30 ($n = 9$), < 30 ($n = 92$)] analysis was performed using Dunnett’s test. We chose 30 to be the dividing range due to previous work suggesting that personality plasticity changes after 30 years (Terracciano et al., 2006) however further investigations about age and embodiment should be investigated. A significant age (< 30 or ≥ 30) effect was found in the updated questionnaire, $p = .01$, $CI = [0.21, 1.59]$. No significant age effect was found in the original questionnaire, $p = .22$, $CI = [-0.15, 0.67]$, (Figure 3).

Regardless of the interesting results about age and their impact on the scores, the validation is important because it demonstrates that with the previous questionnaire the sensitivity was too low to find differences and with the new questionnaire we could find these differences thanks to the enlarged dynamic range. Even

though a significant difference between age groups was found, the sample size supports that this could be a spurious finding.

This particular research highlights that the revised questionnaire has better discriminatory power compared to the originally proposed questionnaire.

6.3 Detail Validation on Study #4

Study #4 was a between-participant study where participants either saw a collocated self-avatar or did not see a self-avatar. Full details of the experiment design are described in Peck and Tutar (2020). When using the revised questionnaire a significant main effect of avatar presence was found, $F(1, 56) = 15.35$, $p < 0.001$, $\eta^2 = 0.22$. Participants with an avatar had significantly higher embodiment ($M = 3.43$, $SD = 1.30$) compared to participants who did not see a body ($M = 2.08$, $SD = 1.50$) No significant main effect was found when using the original questionnaire, $F(1, 56) = 3.54$, $p = 0.07$, $\eta^2 = 0.06$. The revised scale was able to identify the presence of a self avatar, while the original questionnaire was not. See Figure 4.

6.4 Detail Validation on Study #9

Lee et al. (2019) compared haptic experiences producing embodiment while controlling a virtual hand that grasped objects in VR. Participants (male $n = 22$, female $n = 10$) used either a new haptic device or a regular controller trigger to grasp the objects.

We find an significant interaction between condition and gender for the embodiment score ($F(1,14) = 4.27$, $p = 0.5$), whereas the previous questionnaire wasn’t able to find any traces ($F(1,14) = 0.85$, $p = 0.37$). See Figure 5 Further analysis finds that this effect is mostly driven by the appearance sub-scale. Where a trend is also found for the interaction with gender and condition ($p = 0.074$).

In that experiment the avatar hand was always constant in size. And perhaps too large when compared to the median female

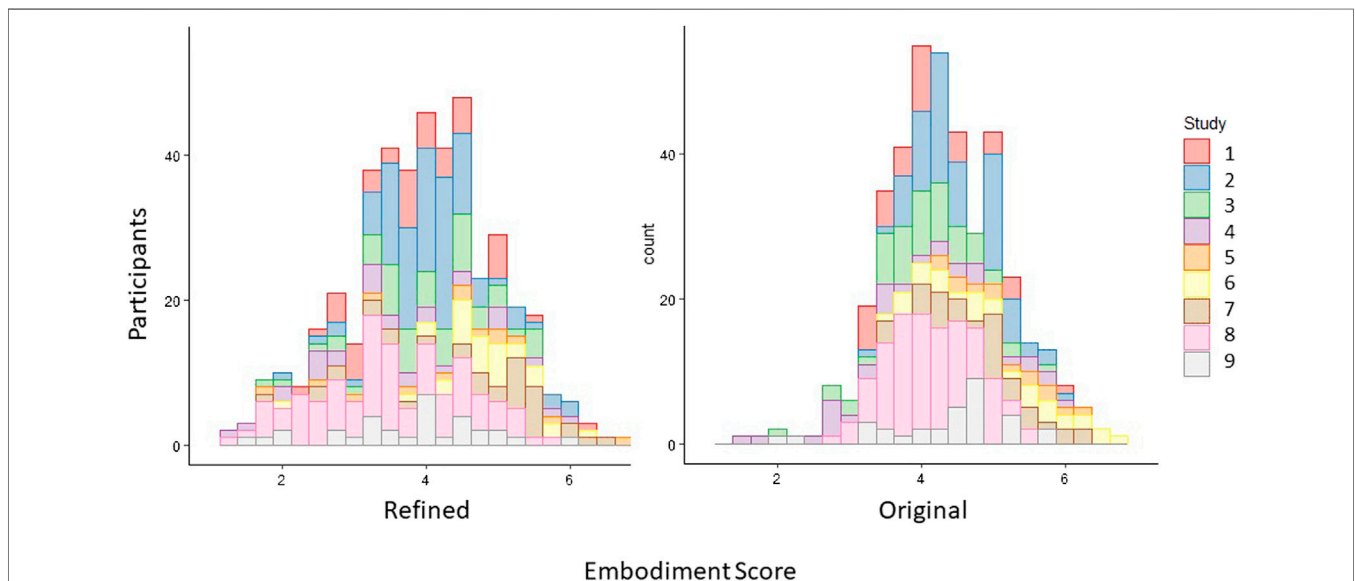
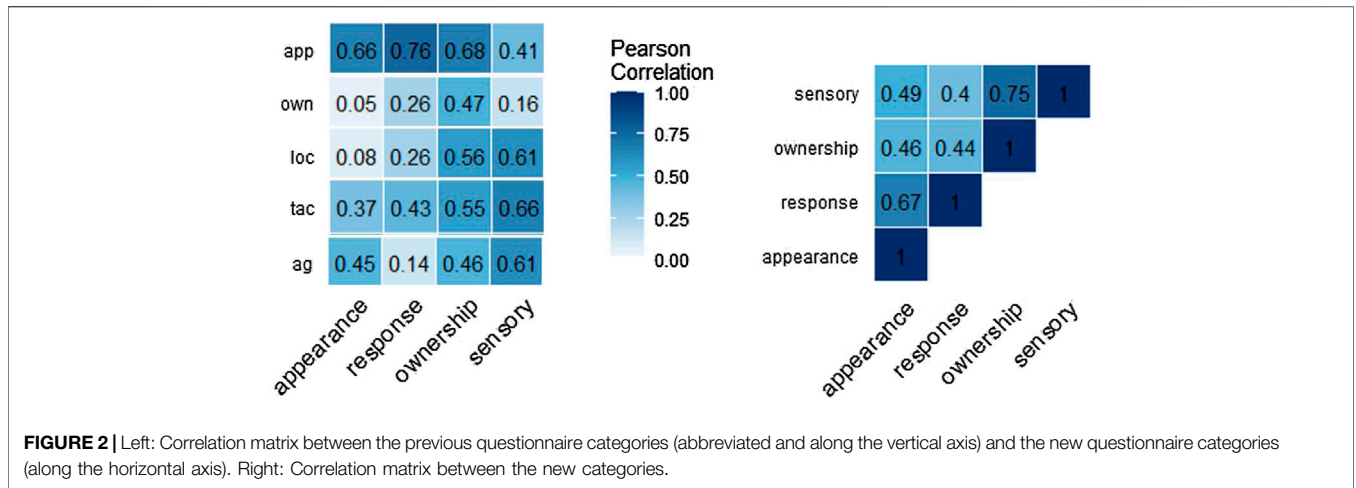


FIGURE 1 | The distribution of scores for the new refined scale and original scale for each of the nine studies.



hand size, therefore affecting the appearance of the hand. This drop in the appearance score might be the underlying reason of the interaction of embodiment and gender for the different controllers.

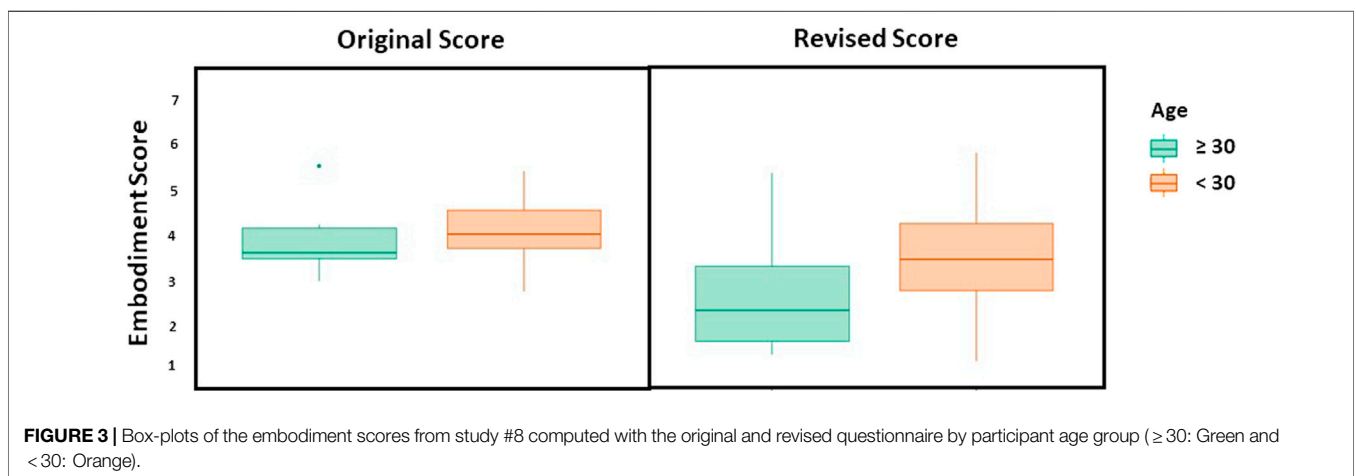
The results highlight not only the higher sensitivity of the new questionnaire, but also the sensitivity at the sub-scale level. The results are in line and further highlight how appearance is a very important trigger for embodiment (Ebrahimi et al., 2018; Jung et al., 2018; Ogawa et al., 2019). This effect was undetected with the previous questionnaire.

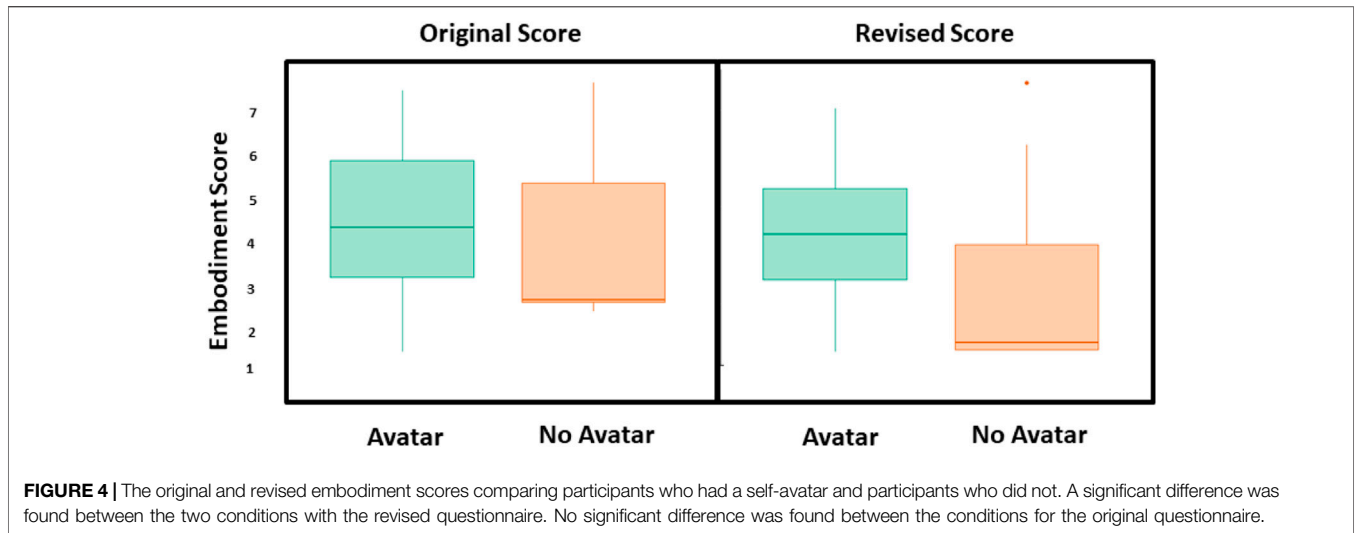
7 DISCUSSION

The current paper has validated many of the original questions proposed to measure embodiment from (Gonzalez-Franco and Peck, 2018). The exploratory factor analysis of almost 450 questionnaires has fine tuned the original questionnaire and removed nine of the original questions. This included questions that could not be universally applied to embodiment scenarios (Q4, Q5, Q25), questions that were correlated with only a couple of other questions or had low

KMO scores (Q2, Q7, Q23, Q9) and questions that comprised an unreliable factor (Q3, Q11). Many of the removed questions were originally added to be dual balanced statements that control for statement bias (Malhotra, 2006) (i.e. Q2,Q3,Q5, Q9,Q11), such as asking if participants felt they had more than one body, that their body was someone else, or that they did not control the body. The prevalence of these dual balanced control statements in embodiment questionnaires follows from standard questionnaire design. However, the ambiguity of these statements likely confused participants and led to additional noise within the data. Interestingly, Q2 and Q3 were previously found unreliable (Peck and Tutar, 2020) and our analysis further supports that these are questions of concern. This highlights the importance of this work since these questions are commonly used in embodiment questionnaires (Gonzalez-Franco and Peck, 2018) and use of these questions may be disadvantaging researchers from fully understanding and interpreting their embodiment results. All in all, our analysis supports that 16 of the 25 questions are relevant for measuring embodiment.

Our findings on the number of components also reduced the sub-scales from 6 to 4. The categories originally labeled: agency





and localization; are no longer presented as independent, but rather integrated into the other four categories: Appearance, Response, Ownership, and Multi-sensory. This does not mean that agency, touch or localization are not important for embodiment, (Kilteni et al., 2012), but rather that they are related to other senses and instead contribute to one of the four prominent embodiment categories. The questions on motor control and agency were mostly assigned to the Response category. This makes sense as motor actions can be considered as yet another type of efferent response. The location and tactile questions fell into the multi-sensory experience. This makes sense too, as touch and proprioception are very related in those questions, such as R14: "It seemed as if I felt the touch of the ___ in the location where I saw the virtual *body* touched."

The revision of the questionnaire highlighted the inter-relatedness of the previous categories and supports the use of sharing questions to measure each category. This inter-relation

between embodiment sub-measures is supported in previous work. Techniques shown to increase embodiment include using mirrors, self-location of avatars, and synchronous movement. A reframing of a questionnaire does not nullify previous work. Instead, the reframing of the categories proposes new challenges and interesting insights for future embodiment research. For example, previous work supports the importance of agency for inducing embodiment illusions. The previously proposed questionnaire highlighted questions that were believed directly related to this topic. However, participant responses for two of the four questions were unreliable (low KMO value and low correlations with other questions). Instead, the two remaining "agency" questions are each used in two of the three sub-measures.

The inter-dependence of these categories and complicated nature of embodiment is further highlighted by the percentage of variance accounted for in each sub-scale, the weights of the pattern matrix, see

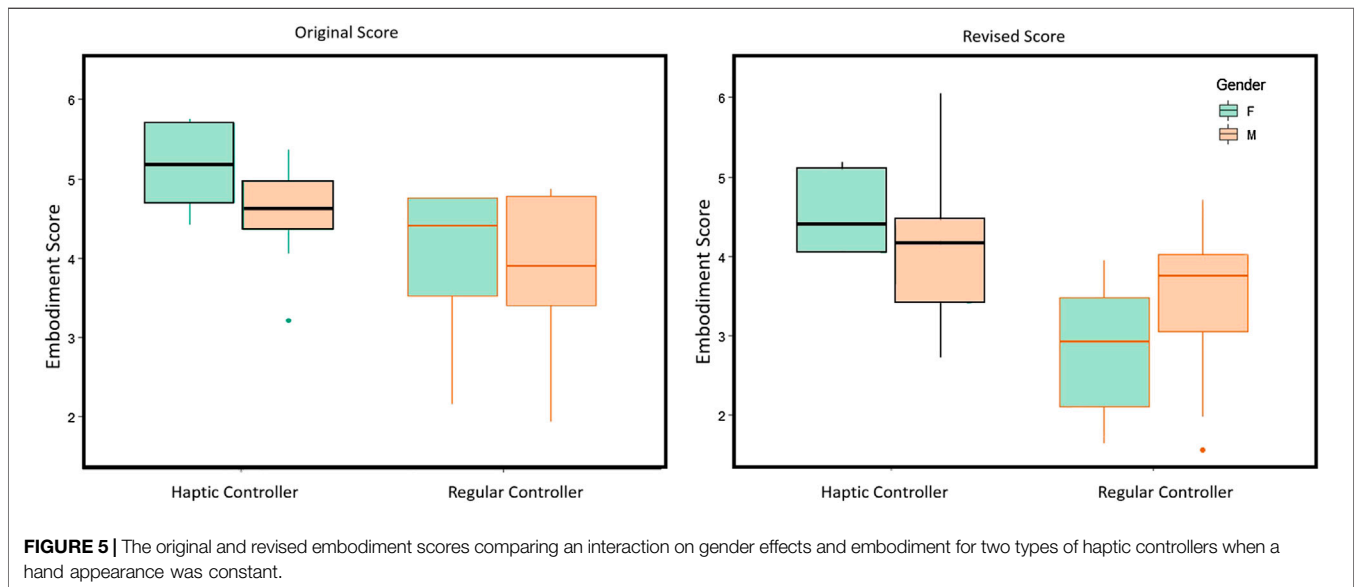


Table 4, and the weights of the structure matrix used to cluster questions, see **Table 3**. For example, Appearance and Response are the highest correlated sub-scales ($r = 0.34$) and share three questions, R4, R6, and R9. In fact, only Ownership and Response do not directly have overlapping questions. The interdependence and roughly equal contribution of each sub-scale to overall embodiment highlights the complicated nature of the sensation and the importance of using the complete questionnaire to measure and further the understanding of embodiment.

We further validated the new questionnaire with a study with 101 participants that used the previous 25 questions. That user study results further showed that the current questionnaire is a continuation and improvement of the previously proposed questionnaire (Gonzalez-Franco and Peck, 2018). The general measure of embodiment was similar, however the sensitivity of the scale was drastically improved. The updated computation, the smaller and more relevant question set, and the interrelationship between questions and embodiment components provide a more robust alternative and a stronger framework for understanding embodiment. Additionally, the validation highlights the importance of using sub-scales to more deeply explore the importance of the different components that may affect embodiment, such as age. This was seen when evaluating the difference in embodiment scores for participants over and under 30 years of age. A significant difference was found in the sub-scales of Appearance, Ownership, and Multi-Sensory, however no significant difference was found in Response. When considering the low score on Appearance this could be due to the avatars not being age matched to the older participants highlighting the importance of having age-matched avatars for participants over 30.

Regarding the computation of the final embodiment scores, as opposed to the recommendation in the original paper (Gonzalez-Franco and Peck, 2018) to use PCA, we now recommend the use of our arithmetic summation. This is because of the validation of the questions and grouping with our large data set. Using the computation of embodiment as presented will benefit the community as it will enable more comparable results with other studies and research. The questions have been arithmetically grouped and the percentage of variance of each sub-scale is roughly even. The inter-relationship of sub-scales was accounted for by using oblique rotation and further validated by the correlations between sub-scales and the shared questions when computing sub-scale scores.

Some questions are still open, such as the optimal Likert use of the questions. Questionnaires can be presented as time related never/always response versus an agreement/disagreement scale. We also think the community would benefit from additional research demonstrating if more quantitative metrics such as electrophysiology correlate with this new refined embodiment questionnaire.

8 CONCLUSION

This paper has presented a thorough verification of the previous embodiment questionnaire that was collected through review of

common questions used in previous embodiment studies (Gonzalez-Franco and Peck, 2018). From our analysis of a collection of nine experiments using the originally proposed questionnaire in full or part (totalling almost 450 questionnaires), we have now been able to validate and streamline it from 25 to 16 critical questions, and from 6 sub-scales to 4. The four sub-scales, Appearance, Response, Ownership, and Multi-Sensory, further the understanding of embodiment by clearly defining the most important aspects of the phenomenon as well as their interrelations.

We now ask the VR and Embodiment Science community to widely use this questionnaire whenever a participant is given a self-avatar. This is of special importance as we know that there are large inter-individual differences between participants of studies even when presented with the exact same conditions (Gonzalez-Franco et al., 2019). The differences in embodiment scores by age further highlights the necessity of evaluating diverse participant samples to limit bias being added into research results (Peck et al., 2020b). Further, our understanding of the cognitive effects of embodiment is not well understood, (Peck et al., 2018; Peck and Tutar, 2020). This suggests that embodiment can be a factor for explaining many of the results or differences in studies including avatars or collocated body parts (such as hands).

Furthermore, the new scale of embodiment calibrated in this paper supports the use of the results between 1–7 as an absolute and comparable value. This means that users of this questionnaire can safely claim that a seven score meant their participants were highly embodied or on the contrary, that a one score meant their participants did not experience embodiment of their self-avatars regardless of a control disembodied condition in experiments. Additionally, researchers do need to run their own PCA analysis, as the arithmetic computation proposed here would be sufficient, and in fact desirable for future comparability.

The work here intends to simplify the use of embodiment questionnaires through this standard validated questionnaire. The use of the questionnaire at large will not only help further the understanding of the effects that derive from embodiment of avatars, but also can aid in the replication and comparison of future studies. Both of these aspects will become more relevant as the democratisation of avatar use for self-representation in VR becomes more mainstream.

DATA AVAILABILITY STATEMENT

The original data presented in this work are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by institutional review boards at either Microsoft Research or Davidson College. The patients/participants

provided their written informed consent to participate in the above studies.

AUTHOR CONTRIBUTIONS

TP conceived the paper and ran the main analysis. TP and MG-F contributed the data and wrote the paper. This topic of research has been of interest to both authors for over a decade and they have collaborated on previous proposals of embodiment questionnaires.

FUNDING

The research reported in this paper was supported in part by a grant from the National Science Foundation (#1942146).

REFERENCES

- Abtahi, P., Gonzalez-Franco, M., Ofek, E., and Steed, A. (2019). "I'm a giant: walking in large virtual environments at high speed gains," in Proceedings of the 2019 CHI conference on human factors in computing systems, Glasgow Scotland, United Kingdom, October 17, 2019 (New York, NY: Association for Computing Machinery), 1–13.
- Alchalabi, B., Faubert, J., and Labbe, D. R. (2019). "Eeg can be used to measure embodiment when controlling a walking self-avatar," in 2019 IEEE conference on virtual reality and 3D user interfaces (VR), Osaka, Japan, March 23–27, 2019 (Piscataway, NJ: IEEE), 776–783.
- Allen, R. C., Singer, M. J., McDonald, D. P., and Cotton, J. E. (2000). Age differences in a virtual reality entertainment environment: a field study. *Proceedings of the human factors and ergonomics society annual meeting* 44, 542–545.
- Armor, D. J. (1973). Theta reliability and factor scaling. *Socio. Methodol.* 5, 17–50. doi:10.1177/154193120004400515
- Berger, C. C., and Gonzalez-Franco, M. (2018). "Expanding the sense of touch outside the body," in Proceedings of the 15th ACM symposium on applied perception, New York, NY, United States, August 10, 2018 (New York, NY: Association for Computing Machinery), 1–9.
- Blanke, O., and Mohr, C. (2005). Out-of-body experience, heautoscopy, and autoscopic hallucination of neurological origin Implications for neurocognitive mechanisms of corporeal awareness and self-consciousness. *Brain Res Brain Res Rev* 50, 184–199. doi:10.1016/j.brainresrev.2005.05.008
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., and Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front. Public Health* 6, 149. doi:10.3389/fpubh.2018.00149
- Botvinick, M., and Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature* 391, 756. doi:10.1038/35784
- Ebrahimi, E., Hartman, L. S., Robb, A., Pagano, C. C., and Babu, S. V. (2018). "Investigating the effects of anthropomorphic fidelity of self-avatars on near field depth perception in immersive virtual environments," in 2018 IEEE Conference on virtual Reality and 3D user interfaces (VR), Reutlingen, Germany, March 18, 2018 (Piscataway, NJ: IEEE), 1–8.
- Field, A., Miles, J., and Field, Z. (2012). *Discovering Statistics Using R*. Thousand Oaks, CA: Sage publications.
- González Franco, M. (2014). *Neurophysiological Signatures of the Body Representation in the Brain Using Immersive Virtual Reality*. Barcelona, Spain: Universitat de Barcelona.
- Gonzalez-Franco, M., and Berger, C. C. (2019). Avatar embodiment enhances haptic confidence on the out-of-body touch illusion. *IEEE Trans Haptics* 12, 319–326. doi:10.1109/TOH.2019.2925038
- Gonzalez-Franco, M., Cohn, B., Ofek, E., Burin, D., and Maselli, A. (2020a). "The self-avatar follower effect in virtual reality," in 2020 IEEE conference on virtual reality and 3D user interfaces (VR), Atlanta, GA, March 18, 2018 (Piscataway, NJ: IEEE), 18–25.
- Gonzalez-Franco, M., Steed, A., Hoogendyk, S., and Ofek, E. (2020b). Using facial animation to increase the enfacement illusion and avatar self-identification. *IEEE Trans. Visual. Comput. Graph.* 26, 2023–2029. doi:10.1109/TVCG.2020.2973075
- Gonzalez-Franco, M., Abtahi, P., and Steed, A. (2019). "Individual differences in embodied distance estimation in virtual reality," in 2019 IEEE conference on virtual reality and 3D user interfaces (VR), Osaka, Japan, March 23–27, 2019 (Piscataway, NJ: IEEE), 941–943.
- Gonzalez-Franco, M., and Peck, T. C. (2018). Avatar embodiment towards a standardized questionnaire. *Front. Robot. AI* 5, 74. doi:10.3389/frobt.2018.00074
- González-Franco, M., Peck, T. C., Rodríguez-Fornells, A., and Slater, M. (2014). A threat to a virtual hand elicits motor cortex activation. *Exp. Brain Res.* 232, 875–887. doi:10.1007/s00221-013-3800-1
- Gonzalez-Lienres, C., Zapata, L. E., Iruetagoena, G., Seinfeld, S., Perez-Mendez, L., Arroyo-Palacios, J., et al. (2020). Being the victim of intimate partner violence in virtual reality: first- versus third-person perspective. *Front. Psychol.* 11, 820. doi:10.3389/fpsyg.2020.00820
- Hutcheson, G. D., and Sofroniou, N. (1999). *The Multivariate Social Scientist: Introductory Statistics Using Generalized Linear Models*. Thousand Oaks, CA: Sage publications.
- Jung, S., Bruder, G., Wisniewski, P. J., Sandor, C., and Hughes, C. E. (2018). "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, Berlin, Germany, October 13–14, 2018 (New York, NY: Association for Computing Machinery), 60–68.
- Kiltner, K., Bergstrom, I., and Slater, M. (2013). Drumming in immersive virtual reality: the body shapes the way we play. *IEEE Trans. Visual. Comput. Graph.* 19, 597–605. doi:10.1109/TVCG.2013.29
- Kiltner, K., Maselli, A., Kording, K. P., and Slater, M. (2015). Over my fake body: body ownership illusions for studying the multisensory basis of own-body perception. *Front. Hum. Neurosci.* 9, 141. doi:10.3389/fnhum.2015.00141
- Kiltner, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. *Presence Teleoperators Virtual Environ.* 21, 373–387. doi:10.1162/PRES_a_00124
- Launois, R., Reboul-Marty, J., and Henry, B. (1996). Construction and validation of a quality of life questionnaire in chronic lower limb venous insufficiency (civiq). *Qual. Life Res.* 5, 539–554. doi:10.1007/BF00439228
- Lee, J., Sinclair, M., Gonzalez-Franco, M., Ofek, E., and Holz, C. (2019). "Torc: a virtual reality controller for in-hand high-dexterity finger interaction," in Proceedings of the 2019 CHI conference on human factors in computing systems, Glasgow Scotland, United Kingdom, October 17, 2019 (New York, NY: Association for Computing Machinery), 1–13.
- Lush, P., Vazire, S., and Holcombe, A. (2020). Demand characteristics confound the rubber hand illusion. *Collabra: Psychology* 6, 83. doi:10.1525/collabra.325

ACKNOWLEDGMENTS

We mainly want to thank the co-authors that helped run many of the studies that we have now used to establish the large database of questionnaires used on our validation. This includes: Christopher Berger, Anthony Steed, Eyal Ofek, Antonella Maselli, Parastoo Abtahi, Brian Cohn, Jessica J. Good, and the students of Davidson College's DRIVE Lab.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frvir.2020.575943/full#supplementary-material>.

- Malhotra, N. K. (2006). "Questionnaire design and scale development," in *The handbook of marketing research: Uses, misuses, and future advances*. Editors R. Grover and M. Vriens (Thousand Oaks, CA: Sage publications), 83–94.
- Maselli, A., Kilteni, K., López-Moliner, J., and Slater, M. (2016). The sense of body ownership relaxes temporal constraints for multisensory integration. *Sci. Rep.* 6, 30628. doi:10.1038/srep30628
- Maselli, A., and Slater, M. (2014). Sliding perspectives: dissociating ownership from self-location during full body illusions in virtual reality. *Front. Hum. Neurosci.* 8, 693. doi:10.3389/fnhum.2014.00693
- Maselli, A., and Slater, M. (2013). The building blocks of the full body ownership illusion. *Front. Hum. Neurosci.* 7, 83. doi:10.3389/fnhum.2013.00083
- Moffat, S. D., Zonderman, A. B., and Resnick, S. M. (2001). Age differences in spatial memory in a virtual environment navigation task. *Neurobiol. Aging* 22, 787–796. doi:10.1016/s0197-4580(01)00251-2
- Ogawa, N., Narumi, T., and Hirose, M. (2019). "Virtual hand realism affects object size perception in body-based scaling," in 2019 IEEE Conference on virtual Reality and 3D user interfaces (VR), Osaka, Japan, March 23–27, 2019 (Piscataway, NJ: IEEE), 519–528.
- Padrao, G., Gonzalez-Franco, M., Sanchez-Vives, M. V., Slater, M., and Rodriguez-Fornells, A. (2016). Violating body movement semantics: neural signatures of self-generated and external-generated errors. *Neuroimage* 124, 147–156. doi:10.1016/j.neuroimage.2015.08.022
- Pavone, E. F., Tieri, G., Rizza, G., Tidoni, E., Grisoni, L., and Aglioti, S. M. (2016). Embodying others in immersive virtual reality: electro-cortical signatures of monitoring the errors in the actions of an avatar seen from a first-person perspective. *J. Neurosci.* 36, 268–279. doi:10.1523/JNEUROSCI.0494-15.2016
- Peck, T. C., Doan, M., Bourne, K. A., and Good, J. J. (2018). The effect of gender body-swap illusions on working memory and stereotype threat. *IEEE Trans. Visual. Comput. Graph* 24, 1604–1612. doi:10.1109/TVCG.2018.2793598
- Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Conscious. Cognit.* 22, 779–787. doi:10.1016/j.concog.2013.04.016
- Peck, T. C., Good, J. J., and Bourne, K. A. (2020a). "Inducing and mitigating stereotype threat through gendered virtual body-swap illusions," in Proceedings of the 2020 CHI conference on human factors in computing systems, Honolulu, HI, United States, April 25, 2020 (New York, NY: Association for Computing Machinery), 1–13.
- Peck, T. C., Sockol, L. E., and Hancock, S. M. (2020b). Mind the gap: the underrepresentation of female participants and authors in virtual reality research. *IEEE Trans. Visual. Comput. Graph* 26, 1945–1954. doi:10.1109/TVCG.2020.2973498
- Peck, T. C., and Tutar, A. (2020). The impact of a self-avatar, hand collocation, and hand proximity on embodiment and stroop interference. *IEEE Trans. Visual. Comput. Graph* 26, 1964–1971. doi:10.1109/TVCG.2020.2973061
- Ries, B., Interrante, V., Kaeding, M., and Anderson, L. (2008). "The effect of self-embodiment on distance perception in immersive virtual environments," in Proceedings of the 2008 ACM symposium on Virtual reality software and technology, New York, NY, United States, October 27, 2008 (New York, NY: Association for Computing Machinery), 167–170.
- Seitz, K. R., Good, J. J., and Peck, T. C. (2020). "Shooter bias in virtual reality: the effect of avatar race and socioeconomic status on shooting decisions," in 2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW), Atlanta, GA, March 22–26, 2020 (Piscataway, NJ: IEEE), 607–608.
- Slater, M. (2004). How colorful was your day? why questionnaires cannot assess presence in virtual environments. *Presence Teleoperators Virtual Environ.* 13, 484–493.
- Spanlang, B., Normand, J.-M., Borland, D., Kilteni, K., Giannopoulos, E., Pomés, A., et al. (2014). How to build an embodiment lab: achieving body representation illusions in virtual reality. *Front. Robot. AI* 1, 9. doi:10.3389/frobt.2014.00009
- Steed, A., Pan, Y., Zisch, F., and Steptoe, W. (2016). "The impact of a self-avatar on cognitive load in immersive virtual reality," in 2016 IEEE virtual reality (VR), Greenville, SC, March 19–23, 2016 (Piscataway, NJ: IEEE), 67–76.
- Terracciano, A., Costa, P. T., Jr, and McCrae, R. R. (2006). Personality plasticity after age 30. *Pers. Soc. Psychol. Bull.* 32, 999–1009. doi:10.1177/0146167206288599
- Conflict of Interest:** Author MGF was employed by company Microsoft.
- The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- The handling editor declared a past co-authorship with one of the authors MGF.
- Copyright © 2021 Peck and Gonzalez-Franco. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.