The Influence of Avatar Personalization on Emotions in VR

Rivu Radiah [1] , **Daniel Roth** [2], **Florian Alt** [1] **and Yomna Abdelrahman** [1,3]

1   University of the Bundeswehr, Munich, Germany; {firstname.lastname}@unibw.de
2   Friedrich-Alexander University Erlangen-Nürnberg, Germany; d.roth@fau.de
3   European Universities in Egypt, Cairo, Egypt; {firstname.lastname}@unibw.de

Abstract: In this paper, we investigate the impact of avatar personalization on perceived emotions. Avatar embodiment is a crucial aspect of collaborative and social virtual reality (VR) systems. Previous research found that avatar appearance impacts the acceptability of the virtual body and changes users' behavior. While virtual embodiment has been extensively investigated, we know very little about how embodiment affects users' experienced emotions. In a user study (N=40), we applied an autobiographical recall method to evoke happiness and investigated the influence of different types of avatar embodiment (personalized same-gender, personalized opposite-gender, non-personalized same-gender, and non-personalized opposite-gender) on participants' perceived emotion. We recorded both self-reported assessments and physiological data to observe participants' emotional responses resulting from the emotions elicited by the use of different avatars. We found significant differences in happiness with the personalized same-gender avatar and the personalized opposite-gender avatar. We provide empirical evidence, demonstrating the influence of avatar personalization on emotions in VR. We conclude with recommendations for users and designers of virtual reality experiences.

Keywords: Avatars, Virtual Reality, Emotions, Embodiment

Figure 1. We investigate how different avatar embodiment influences user emotions in VR. We used the participant's photograph by Ready Player Me[1] (a), to generate four types of embodiment: (b) personalized, same-gender avatar based on (a), (c) personalized opposite-gender avatar based on (a), (d) non-personalized, same-gender avatar, and (e) non-personalized opposite-gender avatar.

1. Introduction

Emotion is a human state that has been subject to extensive research in different fields, including but not limited to psychology [1–3], neuroscience [4], and human-computer interaction [5]. As a result of the growing interest in virtual reality, emotions shifted into the focus of virtual reality research [6,7]. To this end, prior research looked at the role of embodiment in VR [8–10], and, more specifically, how different approaches to embodiment influence emotions [11]. At the same time, there is no comprehensive investigation into how avatar personalization and gender influences our emotions in VR.

This knowledge is beneficial for several reasons. Social VR platform designers could provide more suitable avatar options to users, for example, in terms of experience, enjoyment, and efficiency. Therapy sessions conducted in VR might lead to better outcomes with knowledge of the emotional effects avatars have on users when resembling themselves. And finally, researchers working on avatar embodiment could be provided with insights into how avatar designs may affect participants' affective states in a VR user study.

VR applications often support the embodiment of a user through virtual avatars, making avatars an integral part of VR environments. Applications making extensive use of avatars include gaming [12,13], training [14], and socializing [15]. Avatars are defined as virtual characters driven by human behavior [16] and acting as a user's surrogate in virtual environments (VEs) [17]. Avatar appearances and gender were also shown to have an impact on the actual perception of oneself and may result in a change in behavior [18].

Research showed ways in which the behavior of digital representations can be altered to induce positive effects [19]. Even though there is an understanding of the effects of avatar appearance (e.g, full-body avatars create a stronger presence in VR [20], physically attractive male and female avatars are rated more socially competent [21]) we lack knowledge on how different avatar appearances affect our emotions in VR. This paper addresses this research gap by exploring how different avatar appearances affect emotion and emotion elicitation in VR. Our work is guided by the following research question: *How does the embodied avatar appearance based on personalization and gender affect user emotions in VR?*

To answer this research question we explore four different avatar appearances (personalized same-gender, personalized opposite-gender, non-personalized own gender, and non-personalized opposite-gender) in a user study with 40 participants. Using a within-subject design, we investigate the effect of avatar appearance on emotion and emotion elicitation. To evoke emotions among the participants we used the autobiographical recall method [22]. Emotions were measured using the self-assessment manikin [23] and physiological measures. We found significant differences in happiness with personalized same-gender avatars and personalized opposite-gender avatars.

Contribution Statement. We contribute towards understanding the effect of avatar appearance and embodiment on user emotions in VR through conducting a within-subject user study with 40 participants. Findings from qualitative and quantitative data analysis show that all autobiographical recall methods can successfully elicit happiness among the participants. From four different kinds of avatars (personalized own- and opposite-gender, non-personalized own- and opposite-gender), significant differences in the emotional state were found with personalized same-gender avatars. Participants felt the strongest embodiment with a personalized same-gender avatar. We present qualitative feedback collected from participants to improve the design of personalized avatars.

The article is structured as follows: In section 2, we present prior research relevant to our work. Section 3 presents our methodology and the conducted user study, describing the data collection alongside a description of the tasks used for emotion elicitation. Furthermore, we present our analysis and results in section 4, followed by a discussion of the findings in section 5. We conclude with directions for future work in section 7.

2. Related Work

We review work on avatars and embodiment, avatar realism, and on emotions in VR.

2.1. Avatars and Embodiment

Prior work extensively studied avatar embodiment [24–26]. Latoschik et al. [27] implemented a fake mirror system to serve as a tool for embodiment studies. The authors presented a markerless face and body tracking to animate avatars and evaluation of the system shows high acceptance of the setup. Roth et al. studied virtual body ownership based on different avatar appearances [28]. The authors created four different avatars (wooden mannequin, robot, male/female non-personalized human, and individualized human) to understand how the participants perceive themselves in a virtual mirrored image. A study to evaluate whether participants' belief in interacting with an avatar or an agent leads to different social effects is presented by Pütten et al. [29]. Lugrin et al.[14] investigated how different avatar appearances (realistic, non-realistic, no avatar) affects the illusion virtual body ownership when using a virtual mirror for fitness training. Their findings show that same-gender realistic human avatar appearance elicited a higher level of illusion.

2.2. *Avatar Appearance & Emotions*

Lugrin et al. [30] studied the impact of avatar anthropomorphism on the fear of heights and their preliminary results show that there is no difference in stress levels between human and non-human avatars. Using enhanced avatars to affect emotion was studied by Oh et al. [31] where the authors used avatars with an enhanced smile. The evaluation showed that such enhancements lead to more positive affect and stronger social presence compared to normal smile conditions. Waltemate et al. [10] studied the effect of avatar penalization on immersion, body ownership, and emotional response. The authors implemented various avatars (non-personalized hand-modeled, generic scanned, individualized scanned versions) and evaluated these over CAVE and HMD immersion. Findings show that personalized avatars increase body ownership, presence, and dominance. The proteus effect as investigated by Yee & Bailenson [18] validates how factors such as attractiveness and height can alter our behavior within the virtual environment and how that can have subsequent changes in face-to-face interactions. Prior research also investigated how culture, physical appearance, and nonverbal behavior of avatars affect the positive and negative perceptions among users [32]. The affective embodiment was also studied by Pandita [33]. The authors explore how avatar appearance and posture may affect the users emotionally.

2.3. *Emotion & Emotion Elicitation*

In 1992 Ekman described the six characteristics for an emotion to be listed as a basic emotion [34]. He stated that anger, fear, disgust, and sadness can be listed as basic emotions while emotions such as embarrassment, awe, and excitement can also be considered to be basic emotions. In 1995 Lang characterized emotions in terms of judged valence and arousal, where valence measures whether the current state is positive or negative while arousal indicates whether the user is calm or triggered [35].

Prior work suggests that VR can act as an effective elicitation medium [6]. A particular focus has been on understanding different stimuli that can generate emotions in VR [36,37]. Emotion has been extensively researched in various fields such as psychology. This is a fundamental characteristic of human behavior and is crucial for communication. Thus we explore how emotions are affected when we embody different avatars.

Ratan [38] highlights the importance of emotion in connecting to our virtual self. We extend this work by investigating the emotions of the users when they perceive their embodied avatars. The relationship between us and our self-representations is a complex one. Yet, there is little research on the relationship between avatar appearance and affective states. We close this gap by investigating how avatar appearance, specifically the level of personalization in avatars, may affect our affective states.

3. Methodology

We designed a within-subject user study (N=40) to investigate the effect of avatar appearance and gender on emotions in VR. Participants experienced four different avatars of two different genders – their own gender (own) and the opposite gender (opposite). Each participant embodied all four avatars. To study the effect of different avatar appearances, we designed a total of four different avatars as follows:

- **P**ersonalized Avatar with the **S**ame **G**ender (**PSG**)
- **P**ersonalized Avatar with the **O**pposite **G**ender (**POG**)
- **N**on-**P**ersonalized Avatar with the **S**ame **G**ender (**NPSG**)
- **N**on-**P**ersonalized Avatar with the **O**pposite **G**ender (**NPOG**)

(a) Initial Room setup

(b) Avatar in game mode

(c) Participant in front of mirror

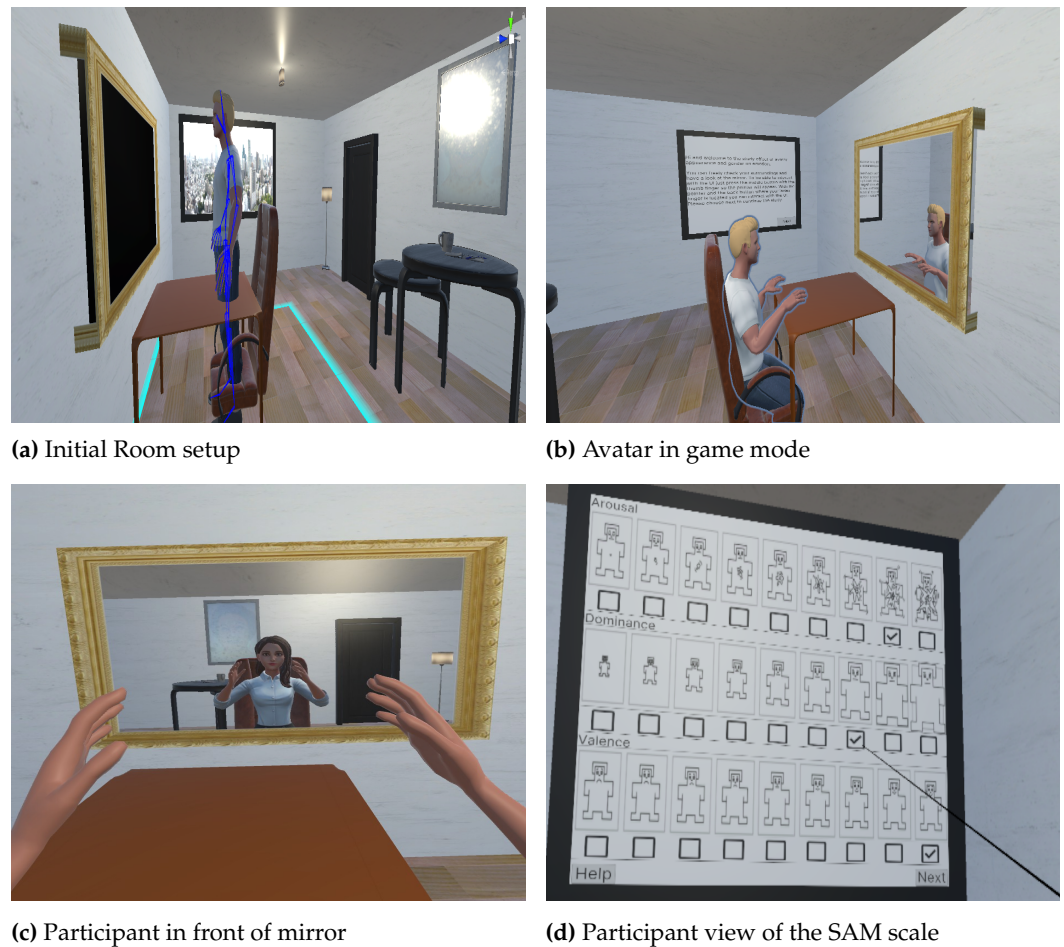**(d)** Participant view of the SAM scale

Figure 2. Virtual room setup from different perspectives

3.1. Apparatus

3.1.1. VR Environment

We created the VR environment using Unity3D and presented it to participants via an HTC VIVE Pro headset (2880 × 1600 pixels, 90 Hz, 110° fov), using the SteamVR plugin[2]. The VR environment was a virtual room where the participants were seated in front of a table and mirror. The room had additional props, e.g., a window and a framed picture to make the room more lively. To avoid distraction these props were made very simple. Using the mirror, participants could view their embodied avatars. In addition, the mirror enhances avatar embodiment within the participants. On the right side of the room, instructions were displayed to participants. The room setup is shown in Figure 2.

3.1.2. Autobiographical Recall Method

Autobiographical Recall involves recalling intense emotional experiences and is a widely used method to experimentally induce emotions [39,40]. This method was selected as it does not require additional material participants need to interact with. This allowed participants to focus on their avatar appearance reflected in the mirror. The happy emotion was chosen as autobiographical recall and has been found to be effective to elicit happiness [41], both in VR and in the real world [7]. It was also found that happy memories occur proportionally more often than unhappy memories [22].

---

2   https://assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647, last accessed March 24, 2023

3.1.3. Avatar Design

The avatars were generated by *Ready Player Me*[3], a tool to generate avatars based on photographs. It integrates with Unity and eases generating multiple avatars of the own and opposite-gender. Mouth movement from *Ready Player Me* was included to enhance embodiment. We created two different kinds of avatars.

Non-personalized Avatars To create non-personalized avatars, we used photographs of a male and female face from the study by Braun et al. [42]. For the hair and eye color, the most common colors [43] in Europe were chosen for the non-personalized avatars, that is blonde hair and blue eyes. The human models and the non-personalized avatars are shown in Figures 1–d and 1–e.

Personalized Avatars To generate a personalized avatar, a photograph of each participant was used. Minor adjustments were made to ensure that the generated avatar looks as similar as possible to the provided picture of the participant. The same picture was also used to generate the avatar of the opposite-gender. The facial features were chosen mainly by *Ready Player Me*. Examples are shown in Figures 1–b and 1–1c.

3.2. Measures

In order to answer our research question we collected the following data:

Demographic Information We asked for participants' ages, gender, profession, education, and prior experience with VR (scale from 'no experience' to 'owner of a headset').

Emotions have been measured traditionally in two ways: (1) by subjective self-reports and (2) by observing physiological responses while being exposed to an emotion elicitation task. To answer our research question, we collected both self-reported assessments and physiological data. All the scales used are standardized well-established questionnaires. Additionally, the physiological responses considered are inspired by the literature [44,45].

Big Five Inventory-10 (BFI-10) The BFI-10 questionnaire was used to assess participants' personalities. The questionnaire evaluates Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness [46] on a 5-point Likert scale.

Presence Questionnaire To assess how participants felt in the virtual world, we used a presence questionnaire [47]. Specifically, we used the version revised by the UQO Cyberpsychology Lab 2004[4], consisting of 19 questions.

Embodiment Questionnaire To measure the degree of the experienced embodiment with each avatar, a virtual embodiment questionnaire (VEQ) [48] was used. The VEQ measures three main factors: the perceived body ownership over a virtual avatar (e.g, "It felt like the virtual body was my body"), the agency over a virtual avatar (e.g., "I felt like I was controlling the movements of the virtual body."), and the perceived change in the body schema (e.g. "I felt like the form or appearance of my own body had changed."). Each factor consists of 4 statements on a 7-point Likert scale.

Self Assessment Manikin (SAM) We assessed participants' emotional responses, using the SAM scale, which is a pictorial self-assessment technique. It consists of three dimensions: valence, arousal, and dominance [23], rated on a 9-point Likert scale.

Color Representation Previous work showed a correlation between emotions and colors [49]. In our work, we investigate if there is any correlation between color and the elicited emotion in VR. Hence, participants were asked to choose a color that best represents how they feel. It was shown to them in the interface and they had to select it from the available palette, using their controller. The choices available were yellow, red, green, blue, purple, and orange.

---

3 https://readyplayer.me, last accessed March 24, 2023
4 https://marketinginvolvement.files.wordpress.com/2013/12/pq-presence-questionnaire.pdf, last accessed March 24, 2023

Avatar Similarity The participants were asked to rate how well they could identify with the shown avatar. The scale was from 1 to 7 (1=not similar at all; 7=perfect virtual representation). They were also asked what they liked and disliked about the avatar and how the avatar could be improved for a better rating.

Physiological Data In addition to self-report assessments, physiological data were recorded to observe physiological responses resulting from the emotions elicited by the use of different avatars. We measured the heart rate (HR), Electrodermal Activity (EDA), and temperature with the E4 Empatica wristwatch and processed the data using the Empatica analysis software service[5].

Post-Study Interview We conducted a short interview after the study, where participants reflected and talked about their experience with the embodied avatars to see if they felt any differences. The interview was audio recorded and later transcribed.

3.3. Study Procedure

Prior to the study, each participant was asked to prepare four different happy stories so they can elicit happiness with four different avatar appearances. In addition, they were asked to send a photo of themselves to create a personalized avatar.

As participants arrived at the lab, they gave consent to data collection, (demographics, self-assessments, the physiological data, and the audio of the interviews). Then, they provided demographic data and filled in the personality questionnaire (BFI-10). Afterwards, they were shown the avatars they would embody during the study. For the personalized avatar, we let participants rate how well they could identify themselves with the shown avatar. The experimenter instructed them to recall their happy memory, taking as much time as they needed. Furthermore, they were instructed to look at the mirror at least once before emotion elicitation, so as to be aware of which avatar they currently represent.

We used multiple neutralization videos to help participants reach a neutral state after embodying each avatar [50]. After entering the VR scene, the first neutralization video was played. Afterwards, they had to fill in the SAM questionnaire and color picker. They were then asked to recall their first happy memory. Then the participants needed to answer the SAM scale, the color representing their feeling, and the embodiment questionnaire. After this, a new neutralization video was played. During the neutralization video, the avatar was switched. This process was repeated for the other 3 avatars. The order was counterbalanced using a Latin square. After the last elicitation, participants filled in the presence questionnaire. When participants left the VR scene, a post-interview was conducted. Participants recalled their experiences and we asked them how they felt.The study took approximately 30 minutes.

4. Analysis and Results

We present the analysis of the data collected. Embodiment, SAM scale, and the physiological data were first tested for normality with the Shapiro-Wilk test before further analysis. If not normally distributed ($p < .05$), a non-parametric test was chosen.

4.1. Participants

A total of 40 participants (21 females, 19 males) with an average age of $M = 25.5$ years ($SD = 3.8$, $Min = 17$, $Max = 39$) were recruited for the study through mailing lists and social media platforms. About half of the sample ($N = 22$) consisted of students. 17 participants had a bachelor's degree (42,5%), and seven had a master's degree (17,5%).

Regarding prior VR use, 16 participants reported that they had used VR only once (40%), 10 reported an occasional use (25%), nine had no experience (22,5%), four used it often (10%) and one was a VR device owner (2.5%). All participants were healthy individuals.

---

[5] https://www.empatica.com/research/, last accessed March 24, 2023

	Mean ± SD	Reference Mean ± SD
Extraversion	3.24 ± .84	3.93± .83
Agreeableness	3.68 ± .80	3.44 ±.71
Conscientiousness	3.30 ± .89	3.91 ± .83
Neuroticism	3.18 ± .92	2.25 ± .69
Openness	3.49 ± .94	3.73 ± .90

Table 1. Result of BFI-10 and a reference value of the respective group.

Avatar	Ownership		Agency		Change		Total
	Mean	Median	Mean	Median	Mean	Median	Mean
PSG	4.67 ± 1.74	5.25	5.76 ± 0.90	6.00	3.25 ± 1.51	3.75	13.68
POG	3.93 ± 1.68	4.50	5.42 ± 0.99	5.50	3.64 ± 2.03	4.25	12.99
NPSG	3.26 ± 1.33	3.50	5.11 ± 1.23	5.25	3.45 ±1.88	3.75	11.82
NPOG	3.35 ± 1.38	3.50	5.39 ± 1.03	5.50	3.50 ± 2.05	4.00	12.24

Table 2. Virtual Embodiment Questionnaire results (both genders).

4.2. Big Five Inventory

The BFI-10 [51] results and reference group are shown in Table 1. The reference group is 18–35 years old, with high education, and mixed gender which fits the demographics of our participants. Compared to the reference group, our participants had on average a lower extraversion and conscientiousness score but a higher neuroticism score.

4.3. Presence

To assess how participants perceived the environment, the total *Presence* score was calculated. The results show a *Presence* score of $M = 4.92$ ($SD = 0.67$) out of 7, which indicates that participants felt quite present in the virtual environment during the study and, thus, no particular biases is expected.

4.4. Embodiment

The descriptive embodiment scores are depicted in Table 2. The sample for the embodiment data analysis had a similar distribution to the main sample, with an average age of $M = 25.37$ years ($SD = 4, 43$, $Min = 17$, $Max = 39$), 12 males and 15 females[6]. We conducted the analysis with $N = 27$ for the embodiment measures.

Shapiro-Wilk tests showed that data was not normally distributed for some of the embodiment factors assessed. Therefore, we performed individual pairwise comparisons and report Wilcoxon-Signed Rank tests for the comparisons where the normality assumption was violated, and t-tests for those where normality could be assumed.

4.4.1. Effects of Avatar Personalization on Embodiment

For this, the personalized same-gender (PSG) and non-personalized same-gender (NPSG) avatars are compared. Wilcoxon Signed-Rank test showed a significant difference ($Z = -3.748$, $p < .001$) for *Ownership*, with the personalized condition PSG ($Mdn = 5.25$) being higher rated than the non-personalized condition NPSG ($Mdn = 3.5$). A significant difference was also found for agency ($Z = -3.733$, $p < .001$). Again, PSG ($Mdn = 6$) is higher rated than ($Mdn = 5.25$). We did not find a significant difference for the change factor ($Z = 0.504$, $p = .615$), with the median for PSG and for GSG being equal; $Mdn = 3.75$.

We also compared the personalized opposite-gender (POG) and non-personalized opposite-gender (NPOG) avatars. A Wilcoxon Signed-Rank test shows no significant difference ($Z = -1.766$, $p = .077$) in ownership between POG ($Mdn = 4.50$) and NPOG ($Mdn = 3.50$). For agency, there is no significant difference ($Z = -.263$, $p = .793$), with median for POG and NPSG being $Mdn = 5.50$. For change, there is no significant difference ($Z = -.238$, $p = .812$) with the median for POG ($Mdn = 4.25$) and NPOG ($Mdn = 4.00$).

---

6   Due to an issue with data recording, the embodiment data of the first 13 participants were lost.
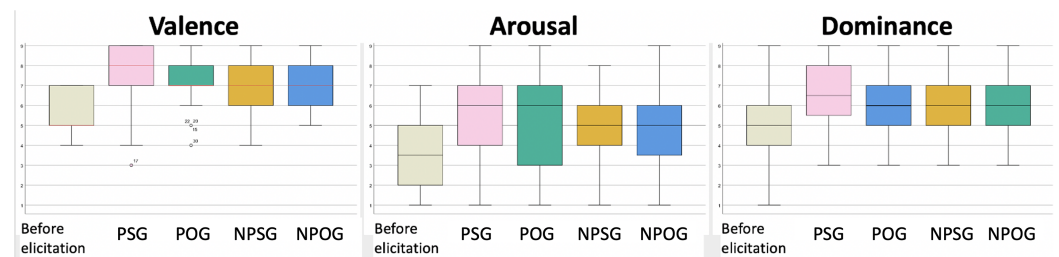
Figure 3. Box plot SAM Value (Valence, Arousal and Dominance). The median is shown in red.

In summary, there is a stronger embodiment perception for the personalized same-gender avatar compared to the non-personalized same-gender avatar in ownership and agency. No significant difference was found between non-personalized or personalized avatars of the opposite-gender.

4.4.2. Effects of Avatar Gender on Embodiment

We first compared the personalized same-gender (PSG) and personalized opposite-gender (POG) avatars. Wilcoxon Signed-Rank test showed a significant difference ($Z = -2.473$, $p = .013$) in ownership, with the median for PSG ($Mdn = 5.25$) being higher than for POG ($Mdn = 4.50$). For agency, there is a significant difference ($Z = -2.811$, $p = .005$), with PSG ($Mdn = 6$) being higher rated than POG ($Mdn = 5.25$). For the change, we found no significant difference ($Z = .786$, $p = .432$; $Mdn = 3.75$ for PSG, $Mdn = 4.25$ for POG).

When comparing the avatars for non-personalized same-gender (NPSG) and non-personalized opposite-gender avatar (NPOG), we found no significant difference ($t(26) = -.368$, $p = .716$) for ownership. Similarly, no significant difference was found for agency ($Z = 1.150$, $p = .250$; $Mdn = 5.25$ for NPSG, $Mdn = 5.50$ for NPOG) or change ($Z = .310$, $p = .757$); $Mdn = 3.75$ for NPSG, $Mdn = 4.00$ for NPOG).

In summary, significant differences were found between avatar gender for the personalized avatars. A personalized avatar of the same-gender is favored regarding ownership and agency. However, there is no significant difference in embodiment between avatar gender when comparing the non-personalized avatars.

An exploratory 2 (avatar gender)$\times$ 2 (avatar personalization) analysis of variance (ANOVA) corroborates the above results for gender and personalization and showed respective main effects ($p < .05$). Pairwise comparisons revealed an influence of avatar gender only in the personalized avatar conditions ($p < .05$), whereas an influence of personalization was only visible in conditions where the avatar had the same-gender as the participant ($p < .05$). Controlling for the participant gender using analysis of covariance (ANCOVA) did not show any significant influence of the participant gender on the results.

4.5. Emotion Elicitation

During the study, we collected a total of five SAM assessments: one before the emotion elicitation and the other four for each avatar after the emotion elicitation. Box-plot diagrams for each dimension of the SAM assessments are shown in Figures 3. Shapiro-Wilk tests reveal that data was not normally distributed for some of the SAM assessments. We performed individual pairwise comparisons and report Wilcoxon-Signed Rank tests for the comparisons where the normality assumption was violated, and t-tests for those where normality could be assumed on the basis of the results of the Shapiro-Wilk tests.

4.5.1. Baseline

We first compared the emotional rating for each condition to the baseline. Wilcoxon Signed-Rank tests showed that all emotion dimensions were significantly different from the baseline in all conditions ($p < .05$).

4.5.2. Effects of Avatar Personalization on Emotion

First, the personalized same-gender (PSG) and non-personalized same-gender avatars (NPSG) are tested for significant differences. A Wilcoxon Signed-Rank test ($N = 40$) shows significant differences ($Z = -2.912$, $p = .004$) in valence, with PSG ($Mdn = 8$) being higher rated than the NPSG ($Mdn = 7$). For arousal, there is a significant difference ($Z = -2.433$, $p = .015$), with the median of PSG ($Mdn = 6$) being higher than the median of NPSG ($Mdn = 5$). For dominance, there was no significant difference ($Z = -1.885$, $p = .059$) with the median of PSG ($Mdn = 6.5$) being slightly higher than the median of NPSG ($Mdn = 6$).

A Wilcoxon signed rank test shows no significant difference ($Z = -.583$, $p = .560$) in valence, with POG ($Mdn = 7$) being higher than NPOG ($Mdn = 6$). For arousal, there is a significant difference ($Z = -2.304$, $p = .021$) with the median of POG ($Mdn = 6$) being higher than NPOG ($Mdn = 5$). For dominance, there is no significant difference with ($Z = -.494$, $p = .621$) with POG and NPOG resulting an a median of $Mdn = 6$.

In summary, there is a significant difference between the personalized avatar and non-personalized avatar of the same avatar gender as the participant, which is shown by a stronger emotion elicitation for the personalized avatar in valence and arousal with a mean increase of 1 in valence and 1 in arousal. For the opposite gender, a significant difference was only found for arousal, where an increase of 1 in arousal can be observed for the personalized opposite-gender avatar compared to its non-personalized counterpart.

4.5.3. Effects of Avatar Gender on Emotion

We first compared the two personalized avatars of the same and opposite gender. A Wilcoxon signed rank test shows no significant difference ($Z = -1.406$, $p = .160$) in valence, with the median for PSG of $Mdn = 7$ being higher than for POG ($Mdn = 6$). For arousal, there is no significant difference ($Z = .136$, $p = .892$), with the median of PSG ($Mdn = 6$) being higher than POG ($Mdn = 5$). For dominance, there is no significant difference ($Z = -1.556$, $p = .120$), the median being the same for PSG and POG ($Mdn = 6$).

Furthermore, we compared the two non-personalized avatars types. A Wilcoxon signed rank test shows no significant difference ($Z = .788$, $p = .431$) in valence, with the median of NPSG ($Mdn = 7$) being higher than NPOG ($Mdn = 6$). For arousal, there is no significant difference ($t(39) = -.361$, $p = .720$), with the median of NPSG ($Mdn = 6$) being higher than NPOG ($Mdn = 5$). For dominance, there seems to be no significant difference with ($t(39) = .405$, $p = .687$), with the median for NPSG and NPOG being $Mdn = 6$.

To conclude, no significant differences were found for the avatar gender, neither when comparing the personalized avatars nor when comparing the non-personalized avatars.

The above-stated effects were exploratorily tested using a 2 (avatar gender) × 2 (avatar personalization) ANOVA. The results corroborate the above-stated findings with the main effects of Avatar personalization. Pairwise comparisons confirmed that for arousal, personalization led to a significant effect in both gender types ($p < .05$). However, the effects of valence and dominance were only present in the same-gender avatars ($p < .05$). Exploratory ANCOVAs using gender, age, and the dimensions of the BIG5 inventory as covariates revealed a significant impact of Agreeableness for the arousal ratings, a significant impact of extraversion and neuroticism for the dominance ratings, and a significant impact of conscientiousness for the valence rating (all $p < .05$).

4.6. Color Choice

Since there were six choices for the colors, an equal distribution would mean color was picked in 16.7 % of the cases by the participants by chance. Before emotion elicitation, the colors appear to be distributed rather evenly, with the top 3 being orange (25%), followed by blue (22.5%) and yellow (20%). After elicitation, generally blue seems to be chosen less often compared to green. This is especially noticeable for the personalized same-gender avatar, where the most frequently picked color was green (30%), followed by yellow (25%). In contrast, the other three avatars seem to share a similar pattern, with yellow being the most picked color. For the personalized opposite-gender avatar, yellow (25%) is followed by

Avatar	Before VR		Baseline		After VR	
	Mean (%)	Med.(%)	Mean (%)	Med.(%)	Mean (%)	Med.(%)
PSG	3.41 ± 4.1	2.26	0.24 ± 0.5	0.12	-1.23 ± 2.33	-0.56
POG	3.74 ± 4.05	2.96	0.53 ± 0.88	0.20	-0.9 ± 2.7	-0.90
NPSG	3.43 ± 3.63	3.04	0.78 ± 1.6	0.23	-1.2 ± 2.16	-0.90
NPOG	3.35 ± 3.84	2.55	0.25 ± 0.69	0.15	-1.27 ± 2.67	-1.30

Table 3. Difference of mean temperature of each avatar and the different segment of the study. The average temperature of the segment was used as the base value.

Avatar	Before VR		Baseline		After VR	
	Mean (%)	Med.(%)	Mean (%)	Med.(%)	Mean (%)	Med.(%)
PSG	-2.85 ± 8.16	-3.98	-1.10 ± 6.73	-0.64	-1.99 ± 8.8	-1.92
POG	-2.8 ± 11.12	-3.98	0.02 ± 9.04	-0.45	-2.07 ± 9.91	-2.88
NPSG	-2.86 ± 10.47	-4.73	-0.03 ± 10.84	-0.49	-2.00 ± 10.88	-2.42
NPOG	-1.10 ± 12.18	-2.52	1.35 ± 10.44	0.60	-0.03 ± 11.73	-1.21

Table 4. Difference of the mean heart rate of each avatar and the different segments of the study. The average heart rate of the segment was used as the base value.

orange (22.5%). For the non-personalized same-gender avatar, yellow (27.5%) is followed by orange (22.5%) and green (22.5%). For the non-personalized opposite-gender avatar, yellow (35%) is followed by green (22.5%) and orange (17.5%). In comparison, the second least picked color is red. The least picked color was purple, ranked last for every avatar in this study. In summary, after the participants elicit their happy memory, they favor the colors green, orange and yellow compared to the other colors.

4.7. Physiological Data

The time intervals for each avatar were calculated with the help of time stamps. Physiological data were averaged for the different time sections of the study, which are: (1) The time when participants elicit their emotion, (2) the baseline before the emotion elicitation, which is the time of the previous neutralization videos before each elicitation, (3) the time before they entered the VR and (4) the time after they left the VR.

Since each participant has different absolute physiological values, changes were calculated in percentages. For calculation, the given time section (before VR, baseline, and after VR) was used as the base value. The results are shown in Tables 4, 3 and 5. Shapiro Wilk tests showed, that the majority of measures for temperature, heart rate, and EDA deviate from a normal distribution (p's $< .05$). We performed Friedman tests to analyze the data.

We found that for all participants temperature increased during emotional elicitation using autobiographical memory, compared to the temperature before entering the virtual environment. A Friedman test was used and showed no significant differences before VR, baseline, and after VR between the avatars ($p > .05$).

An analysis of the heart rate showed that there is a decrease in heart rate when emotion elicitation happens compared to the heart rate before entering VR. The heart rate during elicitation is also lower compared to the after-VR heart rate. Friedman tests showed no significant differences in any of the avatar conditions between the time sections ($p > .05$).

The strongest change can be found for EDA. A decrease can be observed in EDA when recalling a memory compared to before. For the baseline, the difference is less than one percentage for POG and NPOG but a decrease of around 4 to 5 percent can be seen for PSG and NPSG. The EDA during elicitation is also lower compared to the after-VR EDA. The largest difference here can be found for NPSG with a 26.54 percent difference. Friedman tests showed no significant differences between the avatars for all time sections ($p > 05$).

4.8. Avatar Similarity

For the personalized avatar of the same and opposite gender, participants were asked how well they can identify with the avatar. The mean value is $M = 4.95$ ($SD = 1.28$) for the avatar with the same-gender and $M = 4.72$ ($SD = 1.62$) for the avatar with the opposite

Avatar	Before VR		Baseline		After VR	
	Mean (%)	Med.(%)	Mean (%)	Med.(%)	Mean (%)	Med.(%)
PSG	-11.97 ± 59.50	-16.06	-8.84 ± 17.82	-4.07	-15.10 ± 40.90	-10.90
POG	-5.02 ± 72.02	-12.20	2.21 ± 32.36	-0.91	5.25 ± 105.03	-21.12
NPSG	-11.31 ± 74.42	-27.01	-8.24 ± 20.32	-5.35	-9.57 ± 90.48	-26.54
NPOG	-0.72 ± 70.82	-16.48	9.97 ± 37.54	0.79	5,31 ± 82.33	-14.63

Table 5. Difference of the mean electrodermal activity of each avatar and the different segments of the study. The average electrodermal activity of the segment was used as the base value.

Rating Score	Both		Male		Female	
	PSG	POG	PSG	POG	PSG	POG
1	1	0	1	0	0	0
2	1	3	0	0	1	3
3	2	7	0	5	2	2
4	7	11	3	3	4	8
5	17	4	8	3	9	1
6	8	7	4	4	4	3
7	4	8	4	5	0	3
Mean	4.95 ± 1.28	4.72 ± 1.62	5.00 ± 1.29	4.89 ± 1.56	4.90 ± 1.30	4.57 ± 1.69
Median	5	4	5	5	5	4

Table 6. Mean, Median and occurrence of avatar rating grouped by both male and female gender

gender. A breakdown of the rating is shown in Table 6. For both genders, the personalized same-gender avatar has a higher score than the personalized opposite-gender avatar. The data suggests that male participants give slightly higher scores for the avatar with the opposite gender compared to female participants.

4.9. Participant Feedback

Most participants felt more comfortable with their personalized same-gender avatars. This was followed by the personalized opposite-gender avatar. Opinions regarding the generic avatars were mixed. Some participants mentioned that when they saw the non-personalized avatar in the mirror, they felt like there was another person in the room. P02 stated: "At first I thought that I can generally identify more with males but when I saw the non-personalized man I had the feeling that it was a second person in the room and not me. I had this feeling [of identity] more towards the female avatars.". P05 said: "Sometimes I did not feel like myself when looking in the mirror.".

Participants were also asked to rank the avatars based on how well participants can identify with the avatars after recalling a happy moment with each of them. The data suggest that people identify with personalized avatars more. The personalized same-gender avatar was placed first the most followed by the personalized avatar of the opposite-gender.

The opinions on the avatar's clothes were generally positive or indifferent. Some female participants noted that the clothing for the female was too serious and, thus, they lowered the rating since the avatar felt very stiff to them. The body size was considered positive but some participants mentioned that they think the avatars look too muscular or too thin, depending on the participant's own body. The feedback on the proposed hairstyle for the avatar was overall very positive. Face features that needed improvements were eyebrows, a slimmer face, a better jawline, and a more similar-looking nose.

As seen in Table 6, the avatar rating for the opposite-gender was on average lower compared to the avatar of the same-gender. A common remark was that participants felt the avatar of the opposite-gender should automatically receive a lower identity score compared to the avatar of their own gender. Another reason for a lower score was that the participants had a harder time identifying with the shown avatar because they felt no connection to it since they have no comparison of how their real-life opposite-gender appearance would look like.

5. Discussion

We conducted a VR experiment to investigate the effect of different avatar appearances in terms of personalization and the avatar's gender on embodiment and emotion.

To confirm previous findings and to investigate interactions with gender, we compared the two levels of personalization as well as the two levels of avatar gender with regard to embodiment. From the literature, we expected that personalization leads to a stronger embodiment perception, specifically with regard to the perception of virtual body ownership [10,48]. We found such an impact for body ownership, as expected, but also on the level of agency. Since we did not change any control paradigms, the latter may be explained by the strong association and link between the concepts of body ownership and agency [48,52]. One of the most interesting findings of our study is that these effects seem to be limited to the avatars of the same gender but to be rather strong here, whereas this effect was not visible when the avatars had the opposite gender of the participant. In a very similar fashion, the avatar gender was influencing the level of embodiment significantly, but only in the personalized avatar condition.

This is particularly interesting because suggests an "ideal" avatar to be of the same-gender and to be personalized. In this way a high score with regard to embodiment can be achieved, which might lead to positive aspects, for example, in therapy, social VR applications, and the like.

Furthermore, we compared the effects of different avatar appearances on emotion using the autobiographical recall method, eliciting a happy emotion. This study answers an interesting question about emotion elicitation. Prior research investigated the autobiographical recall method in VR using the think-aloud method in the absence of any avatars. Findings suggest that users find it difficult and also silly at some point to speak aloud without anyone present virtually. In our study, participants recalled past life events in front of a mirror reflection that embodied either their own self or a non-personalized avatar.

Two findings are interesting. Firstly, the comparison to the baseline indicates that emotion can be successfully elicited using the recall method when an avatar is present in the virtual scene, which extends previous findings. Results show that this method successfully elicited the desired emotions among the avatars types.

We found that the level of personalization affects the elicited emotion. Based on our findings, we believe using self-reflection can be a successful method of eliciting emotions in VR environments. Qualitative feedback from participants also suggests that participants feel comfortable using this method.

We consider three measures to understand the effectiveness of this method. The SAM scale, the physiological data, and the gathered feedback from the interview. Every avatar led to an increase in valence, arousal, and dominance compared to values obtained before the elicitation. The physiological data, consisting of heart rate, EDA, and temperature, can be found in Table 4, 3 and 5. To see if emotion elicitation caused a physiological response, the average of the physiological data gathered during emotion elicitation is compared with the data obtained (1) before the participants entered VR and (2) the baseline, which is the time frame of the previous neutralization video before emotion elicitation. Kreibig's review of the autonomic nervous system shows that eliciting happiness should cause an increase in temperature, heart rate, and EDA [53].

The interview revealed that nearly all participants felt happy when recalling their memories with the different avatars. While some noted that with some avatars it was easier or more intense, the general opinion was that emotion elicitation works with every avatar. Using the autobiographical recall method, eliciting happiness is possible with all avatars.

Generally, the personalized same-gender avatar seems to perform best in various aspects. It had the highest embodiment score, its avatar rating compared to the personalized opposite-gender avatar was higher, its SAM score regarding valence was the highest and the interview showed that the avatar was often placed first when asked to rank all embodied avatars. These results show that, if an avatar is to be used, it should closely resemble the user for eliciting the happy emotion.

The average age of participants was rather young (25.55) and many participants mentioned how they are used to embodying different avatars from video games. Therefore, different-gender avatars might not trigger any discomfort. This absence of discomfort when embodying different avatars might also explain the average presence score of 4.92 out of 7 which means that participants, despite embodying various avatars in a short period of time, felt predominantly present during the study.

Significant differences were found between the avatars for embodiment and SAM score. Regarding embodiment, the personalized same-gender avatar has a significantly higher score for acceptance and control compared to the non-personalized same-gender. There is also a significant difference between both personalized avatars. The avatar of the same gender was found to have a significantly higher acceptance and control score. For SAM values, there are also significant differences found between the personalized and non-personalized avatars of the same gender where valence and arousal are significantly higher for the personalized same-gender avatar. A significant difference was also found between the personalized and non-personalized avatars of the opposite-gender where arousal levels are higher for the personalized avatar of the opposite gender.

In summary, to answer our research question, the data in this study show that personalized avatars perform better than non-personalized avatars regarding embodiment score and SAM score, and they were rated higher regarding identity. In addition, this study found significant differences between avatars for embodiment and SAM score. Furthermore, the data hint at a correlation between color and embodiment for some avatars. Also, there are additional correlations found between personality traits, SAM score, and embodiment.

6. Limitations

We acknowledge that our work has limitations. Firstly, while physiological data from the participants were recorded through the E4 watch, this data might contain inaccuracies as participants were constantly moving their arms during the study which might cause the sensor to record inaccurate data. The psycho-physiological literature discusses a range of challenges related to the (real-time) assessment of a person's physiological information. Most important are the 'Baseline Problem', the 'Timing of Data Assessment Problem', and the 'Intensity of Emotion Problem [54].

Secondly, Autobiography recall for the same emotion might be limited to a maximum number of uses. Some participants felt that they were already happy and recalling another happy memory did not cause a further change in mood.

Thirdly, *Ready Player Me* was used for the avatar creation. This stylized avatar has its advantages as it tries to avoid the Uncanny Valley effect, an effect where an increase in robot or avatar realism might actually lead to an acceptance decrease [55]. However, *Ready Player Me* also struggles to create avatars with more unusual faces, for example, smaller eyes or a bigger nose. As a result, the avatars from *Ready Player Me* looked somewhat similar and differences between the avatars might be subtle.

Fourthly, we conducted our study with 40 people only. While future work could explore how the findings generalize to larger populations, for example by means of out-of-the-lab studies [56], our exploration still demonstrates the influence of avatar personalization on emotions in VR. Moreover, we did not consider the impact of the game itself nor the different groups of game players on the attitude towards the specific avatar.

Lastly, we considered healthy individuals only, leaving an investigation of users with potentially influencing factors for future work.

7. Conclusion and Future Work

In this work, we observed that an avatar's appearance can influence the user's emotion elicitation in virtual reality. Our work is an initiative to explore avatar appearance and its effect on emotions. The personalized avatar in this study elicited a stronger emotional response in valence and arousal compared to the non-personalized avatar only for the happy emotion. In the future, we plan to research other emotions such as sadness and fear.

Further research could explore additional features of avatar appearance, for example, race and age, or could even look at non-human avatars, for example, an animal. Furthermore, ethical aspect [57] could be subject to future research.

We hope our findings on the interplay between embodiment and emotion will enable other researchers to design their studies in a way such that they lead to more predictable and explainable results. Furthermore, we hope our findings to be applied more broadly in research in Mixed Reality [58] as well as in other areas in which human behavior plays an important role.

Informed Consent Statement: Informed consent was obtained from all study participants.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Anderson, P.; Rothbaum, B.O.; Hodges, L.F. Virtual reality exposure in the treatment of social anxiety. *Cognitive and Behavioral Practice* **2003**, *10*, 240–247.
2. Aymerich-Franch, L.; Bailenson, J. The use of doppelgangers in virtual reality to treat public speaking anxiety: a gender comparison. In Proceedings of the Proceedings of the International Society for Presence Research Annual Conference. Citeseer, 2014, pp. 173–186.
3. Repetto, C.; Gaggioli, A.; Pallavicini, F.; Cipresso, P.; Raspelli, S.; Riva, G. Virtual reality and mobile phones in the treatment of generalized anxiety disorders: a phase-2 clinical trial. *Personal and Ubiquitous Computing* **2013**, *17*, 253–260.
4. Geiger, A.; Bente, G.; Lammers, S.; Tepest, R.; Roth, D.; Bzdok, D.; Vogeley, K. Distinct functional roles of the mirror neuron system and the mentalizing system. *NeuroImage* **2019**, *202*, 116102.
5. Kosch, T.; Hassib, M.; Reutter, R.; Alt, F. Emotions on the Go: Mobile Emotion Assessment in Real-Time Using Facial Expressions. In Proceedings of the Proceedings of the International Conference on Advanced Visual Interfaces; Association for Computing Machinery: New York, NY, USA, 2020; AVI '20. https://doi.org/10.1145/3399715.3399928.
6. Goetz, L.; Rivu, R.; Alt, F.; Schmidt, A.; Maekelae, V. Methods for Autobiographical Recall in Virtual Reality. In Proceedings of the Proceedings of the 12th Nordic Conference on Human-Computer Interaction:; Association for Computing Machinery: New York, NY, USA, 2022; NordiCHI '22.
7. Rivu, R.; Zhou, Y.; Welsch, R.; Mäkelä, V.; Alt, F. When Friends become Strangers: Understanding the Influence of Avatar Gender On Interpersonal Distance Between Friends in Virtual Reality **2021**.
8. Kilteni, K.; Groten, R.; Slater, M. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments* **2012**, *21*, 373–387.
9. Bovet, S.; Debarba, H.G.; Herbelin, B.; Molla, E.; Boulic, R. The critical role of self-contact for embodiment in virtual reality. *IEEE Transactions on Visualization and Computer Graphics* **2018**, *24*, 1428–1436.
10. Waltemate, T.; Gall, D.; Roth, D.; Botsch, M.; Latoschik, M.E. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE transactions on visualization and computer graphics* **2018**, *24*, 1643–1652.
11. Gall, D.; Roth, D.; Stauffert, J.P.; Zarges, J.; Latoschik, M.E. Embodiment in virtual reality intensifies emotional responses to virtual stimuli. *Frontiers in Psychology* **2021**, *12*.
12. Zyda, M. From visual simulation to virtual reality to games. *Computer* **2005**, *38*, 25–32. https://doi.org/10.1109/MC.2005.297.
13. Koulouris, J.; Jeffery, Z.; Best, J.; O'neill, E.; Lutteroth, C. Me vs. Super (wo) man: Effects of Customization and Identification in a VR Exergame. In Proceedings of the Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–17.
14. Lugrin, J.L.; Landeck, M.; Latoschik, M.E. Avatar embodiment realism and virtual fitness training. In Proceedings of the 2015 IEEE Virtual Reality (VR). IEEE, 2015, pp. 225–226.
15. Maloney, D.; Freeman, G. Falling asleep together: What makes activities in social virtual reality meaningful to users. In Proceedings of the Proceedings of the Annual Symposium on Computer-Human Interaction in Play, 2020, pp. 510–521.
16. Bailenson, J.N.; Blascovich, J. Avatars. Encyclopedia of human-computer interaction. *Berkshire Publishing Group* **2004**, *64*, 68.
17. Szolin, K.; Kuss, D.J.; Nuyens, F.M.; Griffiths, M.D. Exploring the user-avatar relationship in videogames: A systematic review of the Proteus effect. *Human–Computer Interaction* **2022**, pp. 1–26.
18. Yee, N.; Bailenson, J. The Proteus effect: The effect of transformed self-representation on behavior. *Human communication research* **2007**, *33*, 271–290.
19. Bailenson, J.N.; Blascovich, J.; Guadagno, R.E. Self-Representations in Immersive Virtual Environments 1. *Journal of Applied Social Psychology* **2008**, *38*, 2673–2690.
20. Heidicker, P.; Langbehn, E.; Steinicke, F. Influence of avatar appearance on presence in social VR. In Proceedings of the 2017 IEEE Symposium on 3D User Interfaces (3DUI), 2017, pp. 233–234. https://doi.org/10.1109/3DUI.2017.7893357.
21. Merola, N.; Pe, J. The effects of avatar appearance in virtual worlds. *Journal For Virtual Worlds Research* **2009**, *2*.
22. Berntsen, D.; Rubin, D. Emotionally Charged Autobiographical Memories Across the Life Span: The Recall of Happy, Sad, Traumatic, and Involuntary Memories. *Psychology and aging* **2003**, *17*, 636–52. https://doi.org/10.1037/0882-7974.17.4.636.

23. Bradley, M.M.; Lang, P.J. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* **1994**, *25*, 49–59. https://doi.org/https://doi.org/10.1016/0005-7916(94)90063-9.

24. Ratan, R.; Sah, Y.J. Leveling up on stereotype threat: The role of avatar customization and avatar embodiment. *Computers in Human Behavior* **2015**, *50*, 367–374.

25. Falconer, C.J.; Rovira, A.; King, J.A.; Gilbert, P.; Antley, A.; Fearon, P.; Ralph, N.; Slater, M.; Brewin, C.R. Embodying self-compassion within virtual reality and its effects on patients with depression. *BJPsych open* **2016**, *2*, 74–80.

26. Michalak, J.; Troje, N.F.; Fischer, J.; Vollmar, P.; Heidenreich, T.; Schulte, D. Embodiment of sadness and depression—gait patterns associated with dysphoric mood. *Psychosomatic medicine* **2009**, *71*, 580–587.

27. Latoschik, M.E.; Lugrin, J.L.; Roth, D. FakeMi: A Fake Mirror System for Avatar Embodiment Studies. In Proceedings of the Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology; Association for Computing Machinery: New York, NY, USA, 2016; VRST '16, p. 73–76. https://doi.org/10.1145/2993369.2993399.

28. Roth, D.; Lugrin, J.L.; Latoschik, M.E.; Huber, S. Alpha IVBO - Construction of a Scale to Measure the Illusion of Virtual Body Ownership. In Proceedings of the Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2017; CHI EA '17, p. 2875–2883. https://doi.org/10.1145/3027063.3053272.

29. Von der Puetten, A.M.; Krämer, N.C.; Gratch, J.; Kang, S.H. "It doesn't matter what you are!" explaining social effects of agents and avatars. *Computers in Human Behavior* **2010**.

30. Lugrin, J.L.; Polyschev, I.; Roth, D.; Latoschik, M.E. Avatar Anthropomorphism and Acrophobia. In Proceedings of the Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology; Association for Computing Machinery: New York, NY, USA, 2016; VRST '16, p. 315–316. https://doi.org/10.1145/2993369.2996313.

31. Oh, S.Y.; Bailenson, J.; Krämer, N.; Li, B. Let the avatar brighten your smile: Effects of enhancing facial expressions in virtual environments. *PloS one* **2016**, *11*, e0161794.

32. Bente, G.; Dratsch, T.; Rieger, D.; Al-Issa, A. Emotional contagion with artificial others. Effects of culture, physical appearance, and nonverbal behavior on the perception of positive/negative affect in avatars. In Proceedings of the International Conference on Social Computing and Social Media. Springer, 2014, pp. 411–420.

33. Pandita, S. Affective Embodiment: The effect of avatar appearance and posture representation on emotions in VR. In Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 2020, pp. 539–540. https://doi.org/10.1109/VRW50115.2020.00121.

34. Ekman, P. An argument for basic emotions. *Cognition & emotion* **1992**, *6*, 169–200.

35. Lang, P.J. The emotion probe: studies of motivation and attention. *American psychologist* **1995**, *50*, 372.

36. i Badia, S.B.; Quintero, L.V.; Cameirao, M.S.; Chirico, A.; Triberti, S.; Cipresso, P.; Gaggioli, A. Toward emotionally adaptive virtual reality for mental health applications. *IEEE journal of biomedical and health informatics* **2018**, *23*, 1877–1887.

37. Felnhofer, A.; Kothgassner, O.D.; Schmidt, M.; Heinzle, A.K.; Beutl, L.; Hlavacs, H.; Kryspin-Exner, I. Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. *International journal of human-computer studies* **2015**, *82*, 48–56.

38. Ratan, R. Self-presence, explicated: Body, emotion, and identity extension into the virtual self. In *Handbook of research on technoself: Identity in a technological society*; IGI Global, 2013; pp. 322–336.

39. Bradley, M.M.; Lang, P.J. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry* **1994**, *25*, 49–59.

40. Mills, C.; D'Mello, S. On the validity of the autobiographical emotional memory task for emotion induction. *PloS one* **2014**, *9*, e95837.

41. Siedlecka, E.; Denson, T. Experimental Methods for Inducing Basic Emotions: A Qualitative Review. *Emotion Review* **2018**, *11*, 175407391774901. https://doi.org/10.1177/1754073917749016.

42. Braun, C.; Gruendl, M.; Marberger, C.; Scherber, C. Beautycheck, Ursachen und Folgen von Attraktivität, 2003.

43. Katsara, M.A.; Nothnagel, M. True colors: A literature review on the spatial distribution of eye and hair pigmentation. *Forensic Science International: Genetics* **2019**, *39*, 109–118. https://doi.org/https://doi.org/10.1016/j.fsigen.2019.01.001.

44. Hassib, M.; Buschek, D.; Wozniak, P.W.; Alt, F. HeartChat: Heart Rate Augmented Mobile Chat to Support Empathy and Awareness. In Proceedings of the Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2017; CHI '17, p. 2239–2251. https://doi.org/10.1145/3025453.3025758.

45. Ioannou, S.; Gallese, V.; Merla, A. Thermal infrared imaging in psychophysiology: potentialities and limits. *Psychophysiology* **2014**, *51*, 951–963.

46. Rammstedt, B.; Kemper, C.; Klein, M.C. 'e.l.; Beierlein, C.; Kovaleva, A. A short scale for measuring the f "u nf dimensions of personality ö nuity: big-five-inventory-10 (BFI-10). *methods, data, analyzes (mda)* **2013**, *7*, 233–249.

47. Witmer, B.G.; Singer, M.J. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* **1998**, *7*, 225–240, [https://direct.mit.edu/pvar/article-pdf/7/3/225/1836425/105474698565686.pdf]. https://doi.org/10.1162/105474698565686.

48. Roth, D.; Latoschik, M.E. Construction of the Virtual Embodiment Questionnaire (VEQ). *IEEE Transactions on Visualization and Computer Graphics* **2020**, *26*, 3546–3556. https://doi.org/10.1109/TVCG.2020.3023603.

49. Kołakowska, A.; Landowska, A.; Szwoch, M.; Szwoch, W.; Wróbel, M., Modeling emotions for affect-aware applications; 2015; pp. 55–67.

50. Samson, A.; Kreibig, S.; Soderstrom, B.; Wade, A.; Gross, J. Eliciting positive, negative, and mixed emotional states: A film library for affective scientists. *Cognition and Emotion* **2015**, *30*. https://doi.org/10.1080/02699931.2015.1031089.

51. Rammstedt, B.; Kemper, C.J.; Klein, M.C.; Beierlein, C.; Kovaleva, A. Big Five Inventory (BFI-10). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS), 2014. https://doi.org/10.6102/zis76. Accessed: 01.11.2021.

52. Kalckert, A.; Ehrsson, H.H. Moving a rubber hand that feels like your own: a dissociation of ownership and agency. *Frontiers in human neuroscience* **2012**, *6*, 40.

53. Kreibig, S.D. Autonomic nervous system activity in emotion: A review. *Biological Psychology* **2010**, *84*, 394–421.

54. Levenson, R.W. The autonomic nervous system and emotion. *Emotion review* **2014**, *6*, 100–112.

55. Mori, M.; MacDorman, K.F.; Kageki, N. The Uncanny Valley [From the Field]. *IEEE Robotics Automation Magazine* **2012**, *19*, 98–100. https://doi.org/10.1109/MRA.2012.2192811.

56. Rivu, R.; Mäkelä, V.; Prange, S.; Rodriguez, S.D.; Piening, R.; Zhou, Y.; Köhle, K.; Pfeuffer, K.; Abdelrahman, Y.; Hoppe, M.; et al. Remote VR Studies – A Framework for Running Virtual Reality Studies Remotely Via Participant-Owned HMDs. *ACM Transactions on Computer-Human Interaction (ToCHI)* **2021**, *28*. https://doi.org/10.1145/3472617.

57. Delgado Rodriguez, S.; Rivu, R.; Mäkelä, V.; Alt, F. Challenges in Virtual Reality Studies: Ethics and Internal and External Validity. In Proceedings of the Proceedings of the Augmented Humans International Conference 2023; Association for Computing Machinery: New York, NY, USA, 2023; AHs '23, p. 105–111. https://doi.org/10.1145/3582700.3582716.

58. Rauschnabel, P.A.; Felix, R.; Hinsch, C.; Shahab, H.; Alt, F. What is XR? Towards a Framework for Augmented and Virtual Reality. *Computers in Human Behavior* **2022**, *133*, 107289. https://doi.org/https://doi.org/10.1016/j.chb.2022.107289.