Achieving Participant Acceptance of their Avatars

Kristopher J. Blom[1], Anna I. Bellido Rivas[1], Xenxo Alvarez[2,3], Ozan Cetinaslan[3], Bruno Oliveira[3], Verónica Orvalho[3], Mel Slater[1,4]

[1] EVENT Lab, Facultat de Psicologia, Universitat de Barcelona, Barcelona, Spain

[2]Face In Motion, Porto, Portugal

[3]Instituto de Telecomunicações – Universidade do Porto, Porto, Portugal

[4] Institució Catalana Recerca i Estudis Avancats (ICREA), Barcelona, Spain

Contact Author: Kristopher J. Blom, blom@ub.edu

Abstract

An experiment was carried out to examine the extent to which an avatar can be perceived by people as similar to themselves, including their face and body. The avatar was judged by the participants themselves rather than by third parties. The experiment was organized in two phases. The initial phase consisted of a forced-choice, paired comparison method used to create a ranking of ten virtual faces in order of preference. This set of faces included a facial mesh, created by a custom software pipeline to rapidly generate avatars that resemble the experimental participants. Six more faces, derived from the participants own face, were also shown in order to gain insight into the acceptance of a variety of facial similarities. In the second phase, full body avatars with the most and least preferred faces were presented along with the direct pipeline output. Participants rated their level of satisfaction with those avatars as virtual self-representations and provided the level of perceived resemblance to themselves. Results show that our avatars are perceived to be similar

to the self, rated at 7.5/10. Those avatars with faces derived from the participants face mixed with an ethnically similar face were also rated with high scores. These results differ significantly from how arbitrary avatars are perceived. Therefore, reasonably physically similar avatars can also be expected to be perceived as similar by participants.

1 Introduction

Imagine that you were to look in a mirror and staring back at you, you see not yourself but the image of another person. You move and the image moves like you, reinforcing the idea that the body you see is, in actuality, yours. In virtual reality such occurrences have become the foundation of recent research investigating body ownership (Llobera et al., 2013; Maselli & Slater, 2013; Normand, Giannopoulos, Spanlang, & Slater, 2011; Slater, Spanlang, Sanchez-Vives, & Blanke, 2010). A strong illusion that the virtual body is one’s own body can be induced, even though it is visually dissimilar (Banakou, Groten, & Slater, 2013; Peck, Seinfeld, Aglioti, & Slater, 2013). During our own experiments, we have observed that people who, through chance circumstances perceived themselves as having a strong visual likeness to the avatar, tended to report a high level of body ownership. This is only an anecdotal finding, but it seems reasonable to expect that an avatar that one perceives to look like, or at least resemble, oneself may deepen the sensation of body ownership. In this work we are interested in the initial step, verifying that a look-alike avatar is perceived by that person as visually similar to the self. At the same time we investigate how close that appearance has to be their true appearance to be perceived by the *participants themselves* as visually similar.

Viewing an avatar with an appearance that is intended to look like oneself is presumed to create a connection with that avatar. Such avatars have been referred to as virtual doppelgängers (Bailenson & Segovia, 2010). In a series of studies, Bailenson and colleagues have investigated the effect of a 'look-alike' avatar seen in a third person perspective on perceptions of that avatar and whether those avatars can have a greater impact on the intentions and behaviors of the participant than an arbitrary avatar. These avatars have been created to have the participants' facial appearance by directly using photos of the person. They validated their look-alike heads (without bodies) by third party evaluation of the similarity of the head to the target person (Bailenson, Beall, & Blascovich, 2003; Bailenson, Beall, Blascovich, & Rex, 2004). In later studies they attached the heads to articulated bodies. In one such study that investigated vicarious reinforcement, they asked participants to evaluate how much they thought the full body avatars looked like themselves as a control item (Fox & Bailenson, 2009). Although significantly higher levels of resemblance for the similar avatars were reported than for an arbitrary avatar of appearance that was not reported in the paper, the overall resemblance of the full body avatar that was supposed to be visually similar was rated very closely to the midpoint between dissimilar and similar (mean 2.7/5). Therefore, it is not clear whether participants truly perceived the avatars as look-alike for themselves. The various experiments, summarized in (Bailenson & Segovia, 2010), show a positive effect of receiving advice and advertising from an avatar designed to look like the subject vs. an arbitrary dissimilar avatar. (van Vugt, Bailenson, Hoorn, & Konijn, 2008) also reported finding comparable effects for avatars with facial appearance modified from the subject's avatar in order to remove conscious awareness of facial similarity.

We formally investigate here whether a full-body look-alike avatar is identified by participants themselves as being *visually* similar, i.e. that participants *themselves* view the avatar as looking like themselves, rather than reliance on third party opinions. This is important because our motivation was that when participants perceive their avatars as looking like themselves this could have an impact on their level of body ownership. That question itself is not addressed in this paper, but rather only the issue of perceived similarity to the self. This differs critically from the works cited above, where formal validated of their avatars was performed via third person judgment of similarity. An external person's evaluation is perhaps more objective, while in the self-evaluation case resemblance corresponds to the level of match between the avatar and the person's own body image, our primary concern. This is particularly true for one's own face, of which a person has only mediated visual experiences, e.g. mirrors or photos.

Research from a number of distinct areas provide further evidence that stresses the importance of investigating first person evaluation of the avatars, by highlighting the differences between third person and self-evaluation of visual similarity. Moreover, we concentrate not just on the face but on the whole body. Recent neuroscience literature shows that processing of self related images occurs through different pathways than processing of others (Keyes, Brady, Reilly, & Foxe, 2010; Sui, Zhu, & Han, 2006). Various forms of body image disturbances are well known in psychology and only affect self-image (Cash & Pruzinsky, 2002). Outside of clinical cases, a recent large scale study of university students in Spain found that 46.5% falsely believed their body size to be different than their actual body size (Pimenta, Sánchez-Villegas, Bes-Rastrollo, López, & Martínez-González, 2009). This provides evidence that the inclusion of the body may be significant in similarity assessment, as well as how prevalent differences are in objective and self-evaluation. Even self-esteem has been shown to influence visual self-identification (Richetin, Xaiz, Maravita, & Perugini, 2012). Based on these factors, the *evaluations of visual similarity by third parties cannot be relied on as an indicator of perceived visual similarity to the self.*

Beyond differences in self and third party evaluation, there are also indications that avatar appearances that differ from objective representations may be accepted as the self quite readily. The research of van Vugt et al. (2008) indicates that participants may evaluate avatars with appearances that deviate from their own image as themselves, which we have also seen in our experiments. Other research supports the idea that the algorithmically best avatar, satisfying objective criteria of similarity to the real face or body, may not necessarily be the appearance that the person will perceive as the most like themselves. An example of this is the self-enhancement effect, where participants evaluate themselves as more attractive than they are (Epley & Whitchurch, 2008; Rhodes, 2006). Much of the research in this area relies on the evaluation of images of the self and others. The third person judgment is generally considered stable and used as the comparative standard. Studies have shown that morphed images of the self with others cannot be detected and can even be selected as being more like oneself than a more faithful image. Importantly, this research shows that people tend to identify themselves as more attractive than they are, when identifying morphed photographs. For instance, in (Epley & Whitchurch, 2008) participants were shown faces that included their own image as well as morphs towards an ugly face and the average face, i.e. a universally attractive face. The study showed a skew in the selection towards the more attractive (average) face.

To enable a broader framework of investigations into the impact of having an avatar that is perceived to be similar to the self, we have created a pipeline for the generation of ‘look-alike’ avatars. We noticed during prototyping of our system that, while evaluations by others of the resemblance of an avatar to the target person of the avatars were positive, the target person sometimes felt uncomfortable viewing the same avatar. Many felt the self-likeness was poor, even though third person evaluations indicated that the likeness was strong. These experiences, though anecdotal, reinforce the previous point that self-evaluation of the look-alike avatar may be significantly different from third person evaluation, and the most faithful replication possible might not be perceived as the most visually similar to the self.

We examine a range of possible likenesses to determine the preferred virtual appearance and the range of appearances that may be acceptable for producing a feeling of visual similarity to the self. Based on the self-enhancement effect literature, we chose to focus on the face for this part of the study. The face is not only the most salient part of the body, but is also a critical means for identifying people. As in the self-enhancement effect literature, we tested a range of faces based on a morphing technique. We morphed the look-alike face with each of three external faces, the average face and two control faces. The self-enhancement effect predicts that morphs towards the average face will be preferred over the most faithful reproduction of the target face. Two control faces were selected, an average Northern European face and an average Southern European face, to guarantee that at least one would have a significantly different appearance than the participant. We extend the existing facial self-enhancement literature base, in that the morphs and viewing of the faces were done in 3D space. This was done both because our intended use for these avatars is in

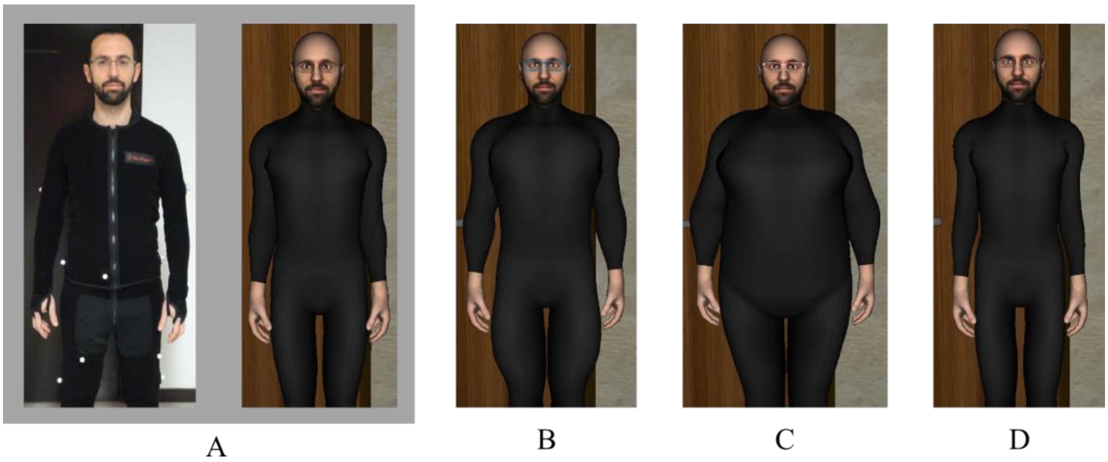


Figure 1 Avatars made with the pipeline for the creation of avatars similar to the self: (a) the participant, wearing the tracking suit used in our experiments and the look-alike avatar (b-d) demostrations of the capabilities of alternative body shapes and adjustments used for roughly matching the avatar's body shape to the participant.

immersive VR and also because there is evidence that 3D cues influence facial recognition (Chelnokova & Laeng, 2011; Liu & Ward, 2006).

Although the face is important, we were interested in the perceived similarity to the self of a full-body avatar. There are several reasons for this. Not only will full bodied avatars be used in our future research, but the similarity effect we have seen in our experiments has sometimes been based on body and clothing similarities. The body may contribute to the level of perceived similarity to the self of the full avatar. Fox and Bailenson (2009) noted that the body may have been a factor in the low resemblance scores they found, though it is not clear what component of the body might have been important; the quality of the body in their work was quite low, the clothing was not controlled, and the avatar was animated with motion that was not that of the subjects. Therefore, in a second part to the experiment presented, participants evaluated a small set of static avatars with bodies that matched their own body shape, were clothed the same, and with faces selected from those evaluated in the first part of the experiment. Figure 1 demonstrates a look-alike avatar and some of the adaptations available to match face and body to the participants.

2 Methods and Materials

16 males were recruited to participate in the experiment. Participants were limited to males due to limitations of the implementation of the look-alike avatar pipeline available at the time. See the Annex for details of the avatar pipeline as it was used. The average age was 28 (SD 5.6). All reported their ethnicity as Caucasian, except for one Asian and one 'other.' Participants gave informed written consent and were paid 10€. The experiment was approved by the Bioethics Commission of University of Barcelona.

The experimental design fulfilled two purposes. The primary investigation was to determine the extent to which participants identified the look-alike avatars as matching their own physical appearance and whether the avatars could readily be accepted as self-representations. A secondary purpose of the experiment was to determine whether avatar appearances that deviated from their physical appearance would be preferred. A two-part experiment was designed.

In the first part, participants compared ten faces derived from the participant's face in a forced-choice Paired Comparison (*PC*) task (Brown & Peterson, 2009). The *PC* method builds a full ranking of faces ordered by preference by asking the participant to choose between two virtual faces at a time, instead of providing an absolute valuation of each face. We implemented a forced-choice paired-comparison method following (Peterson & Brown, 1998). The use of a relative valuation method is important in a repeated measures design in order to achieve statistical stability and with such an abstract concept. Participants may use very different criteria in their determination and, as we have noted in the introduction, various factors may exist that impact the participant's internal image of their own face. Two situations we wished to avoid were where participants might choose either: a desired image ('how I want to look') or something that emphasizes some personal quality, instead of exactly how they think they look, as they might do in online worlds (Martey & Consalvo, 2011). Therefore, we instructed participants to "Imagine a close friend is going to meet you in a crowded virtual café. Please select the avatar which your friend is most likely to pick out as you." This question required the participant to choose the representation that most resembled how they believed they look to an external person. This approach is reminiscent of the "the looking-glass self," which hypothesizes that we define our self by how we imagine our appearance to be in the eyes of others (Cooley, 1964).

The *PC* method creates a relative ordered ranking, but without further information it was not known how much the participant felt that the likeness was similar to their own likeness, as all the information was only comparative. An absolute valuation of specific likenesses from the ranking, now with the face attached to a full body adjusted to match their body shape, was provided for a small subset of the faces. Questions to elicit how well the participants felt the avatar captured their likeness were asked while immersed and viewing the avatar. To get an idea of the spectrum of the acceptability of the different avatars, two avatars were accessed: the first (1^{st}) and last avatars in the ranking, i.e. the avatars perceived by the participant as being the most and least visually similar to themselves. Additionally, we were interested in how the participant would evaluate the avatar that was objectively the most faithful reproduction, i.e. the avatar that was the direct output of the look-alike avatar pipeline without morphing. In cases where this avatar was not the first ranked avatar (11/16), participants additionally evaluated it.

2.1 Experimental Avatars and Stimuli

To determine the best avatars for perceived visual similarity to the self, we created a range of avatars using the pipeline described in detail in the Annex. A set of

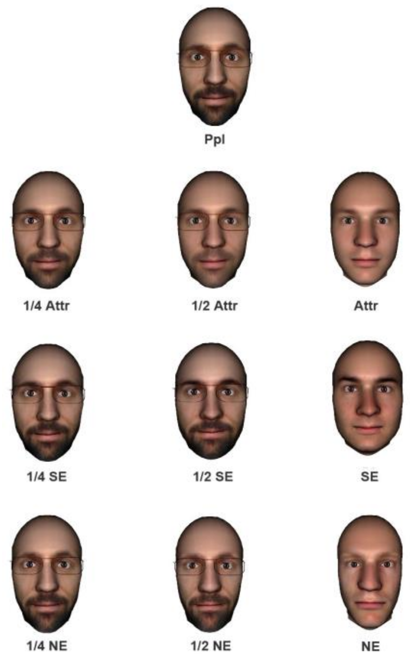


Figure 2 The faces used in the Paired Comparison tests. Ppl = pipeline results ; Attr = Attractive/Average face; SE = Southern European face; NE = Northern European face.

basis faces were selected for the avatars. One avatar is the direct result of our pipeline without modification (Ppl). We used the average European face as the universally attractive base face (Attr), as most participants were expected to be European. We additionally wanted a face that was distinctly not visually similar to the participants face. By having two faces, we could guarantee that for each participant at least one was not visually similar. Since the majority of subjects were Spanish, we chose two Caucasian faces, one with a more Southern European appearance (SE) and one having more of a Northern European appearance (NE). The selected faces can be seen in the rightmost column of Figure 2.

For each basis face, we created morphs of Ppl towards that face. These morphs use the morphing capabilities of Singular Inversion's FaceGen, which morphs the geometry within 'face space' as well as the base color and detailed textures of the target faces. The use of FaceGen in our pipeline is discussed in the Annex. Morphing

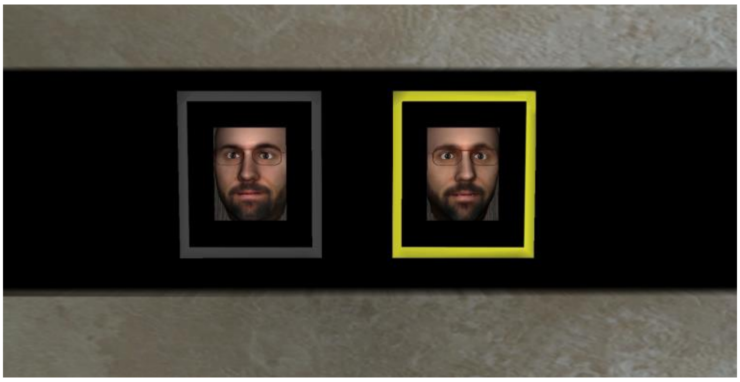


Figure 3: View of the scenario while comparing the faces in the PC part of the experiment. The framing lit up to indicate active selection of that face, still requiring confirmation by the participant.

in FaceGen was controlled by an automated script to guarantee matching morph percentages and to accelerate the process.

We broadly explored the morphing effect by creating Ppl morphs at 1/4 and 1/2 towards each basis faces. As an example, the 1/4 morphed face towards Attr was 3/4*Ppl + 1/4*Attr. In total, ten faces were seen by each participant. A complete set of faces can be seen in Figure 2.

The stimuli for the Paired Comparison part of the experiment were created based on the methodology widely used in existing literature investigating self-identification with photo-based methods (Rhodes, 2006). As in those studies, we “cut” off the faces, such that only the main features of the face were displayed, as seen in Figure 3. However, our method using a ‘cutting plane’ maintained the spatial facial geometry. The stimuli for the valuation part of the experiment were the full-body avatars. An approximation of the shape of the participant was applied to the avatar, e.g. slimmer participants had slimmer avatars. The base shapes shown in Figure 1, and any mixture of those shapes could be generated. Clothing was matched by having both avatar and participants wear a tracking suit that is usually used in experiments in our lab, as seen in Figure 1.

2.2 Measures

The PC method implemented permits derivation of a ranking of the stimuli per participant and an internal reliability per participant (Peterson & Brown, 1998). Based on a complete set of two comparisons per pair, a matrix of preferences can be created, where each cell codifies the preference choices comparing the two faces. Inconsistencies, in the form of cycles, called triads, are identified in the initial matrix. These cycles result from inconsistent rating of avatars. As per Peterson and Brown's (1998) method the inconsistencies can be broken by presenting once more stimuli from the triad to break the cycles . The final matrix can then be used to determine a relative ranking of the faces. Using the triads that were required to be solved, a measure of the internal reliability of the *PC* results for each participant can be calculated. The reliability statistic ranges between 0 and 1, 1 being the most reliable.

In the second phase, four questions were administered orally to provide a subjective absolute valuation of the full avatars. Responses were given on a scale of 1 (not at all) to 10 (completely). The first question directly assessed how well the participants thought the avatar resembled themselves. An additional question related to this (the fourth question) was similar to the decision criterion used in the *PC* phase of the experiment. It provided a second measure of the perceived avatar similarity. Two further questions addressed the participant's willingness to use that avatar, in general and in a formal social setting, like a meeting. Exact wordings can be seen in the table in .

The post questionnaire included a section about the experience while evaluating the avatars. Participants were asked to rate their experience of the comparison phase of the experiment on the following statement on a 5 point Likert scale (disagree completely to completely agree) "I had difficulty telling the difference between some faces." The remaining questions, listed in the table in Figure 6, addressed the experience of viewing what they perceived as the best of the full body avatars. Three of the questions elicited additional information about the extent to which they perceived the full-body avatar to be visually similar. The remaining questions addressed any uneasy feelings when viewing the avatar.

2.3 Experimental Procedure

Prior to the experimental session, three pictures of the participant were acquired for creating the avatars. The process to create the seven unique avatars for each participant took between 45 minutes and one hour - the time reduction over a single avatar creation time of approximately ten minutes was due to parallelization of the process. Written consent was given for use of the photos to generate the avatar.

An information sheet was provided and informed consent for the experiment was solicited at the time of the experimental session. The procedure of the session was explained to the participants and they were asked if they had any questions and reminded that they could quit at any time, without giving any reasons. They put on the tracking suit, assuring that all participants were wearing the same clothing as the avatars. The experimental session, which lasted approximately 10 minutes, started with the mounting of the HMD. The HMD calibration was performed, using a method adapted from (Grechkin, Nguyen, Plumert, Cremer, & Kearney, 2010). During this period, the height of the viewing slot was adjusted to the height of the participant using head tracking information. The virtual environment was then shown.

The Paired Comparison task was performed first. The participants could move freely in the environment with a first person view of the 3D faces. After selection and confirmation, the blinders over the face slots closed, the next comparison was cued up, and the blinders reopened automatically. An initial round of comparisons between each possible combination of all of the stimuli (10 faces) was performed. Each comparison was performed twice, once in each order to balance any positional biases, and the order of presentation was randomized.

After the complete set of 90 comparisons was evaluated, an on-line algorithm calculated triads in the matrix and produced pairings to retest. The participant was not aware of this process, as the inconsistency comparisons followed without visible interruption. On average, the participants made 6.1 (SD 2.9) additional comparisons. After the final comparison, the blinders remained closed. The complete *PC* task took approximately 8 minutes to complete.

The first full body avatar then loaded automatically, positioned to the right of the user out of their field of view. The participant was requested to turn to the right, where the avatar was located. For each avatar, the participant was encouraged to approach to it and inspect it closely. It was explained to the participants that all avatars would be wearing the same clothing, the black suit, and would also be bald. After a moment to investigate the avatar, the four oral questions were asked. The 1^{st}, *last* and *Ppl* avatars were how visually similar to themselves in randomized order. The HMD was removed. The participants answered the post-test questionnaire and were paid.

2.4 Equipment & Scenario

Our future plans for the look-alike avatars include experiments that will be performed in an immersive 3D setting. Displaying avatars in an immersive 3D setting, rather than on a desktop display, also provides spatial cues to assist in the assessment of the likeness. To that end, we displayed the virtual environment via a stereo NVIS nVisor SX111 head-mounted display (HMD). It has dual SXGA displays with 76°

Hx64°V field of view (FOV) per eye, totaling to a 111° horizontal FOV, with 50° of binocular overlap. The displays were driven at 60Hz. Head tracking was performed by a 6-DOF Intersense IS-900 device. The participants wore a black tracking suit, used as part of a full-body tracking in other experiments; however, full-body tracking of the participants was not performed in this experiment.

The virtual environment was created in Maya and consisted of a single room with a “line-up” slot on one wall, plants in the corners, and doors at either end of the elongated dimension of the room. The virtual environment was displayed using the XVR software system (Tecchia et al., 2010) and the avatars were loaded using the HALCA software system (Gillies & Spanlang, 2010). Input for the selection of the faces in the first phase was provided via the Intersense wand, using the thumb joystick for selection and the trigger for accepting the selection. The currently selected face was indicated by a frame around the face which lit up, as seen in Figure 3.

3 Results

3.1 Avatar Ranking

The ranking produced by the Paired Comparison method created an ordering of the faces for each participant, i.e. the first ranked to the tenth ranked for each face. The *PC* ranking results across all participants are shown in Figure 4. The average internal reliability score per participant was 0.97 (SD .05).

The inter-rater agreement on the ranking data, as measured by Krippendorff’s alpha, was (α=.78). Manual inspection of the data found one subject’s responses were slightly different in respect to the morphed SE and NE faces. Further investigation revealed that the participant was a Northern European Caucasian, where the others

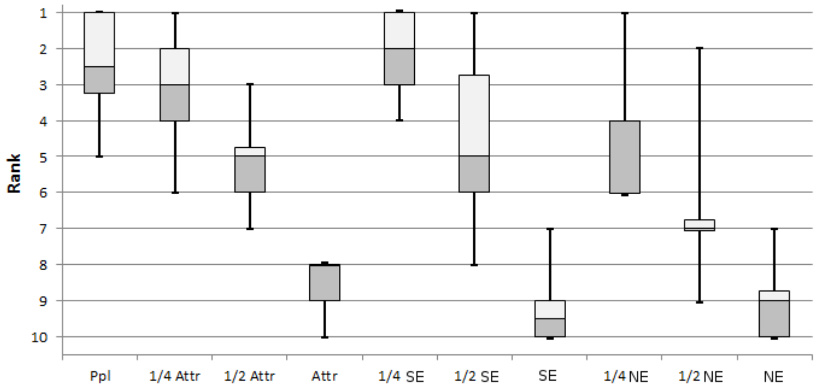
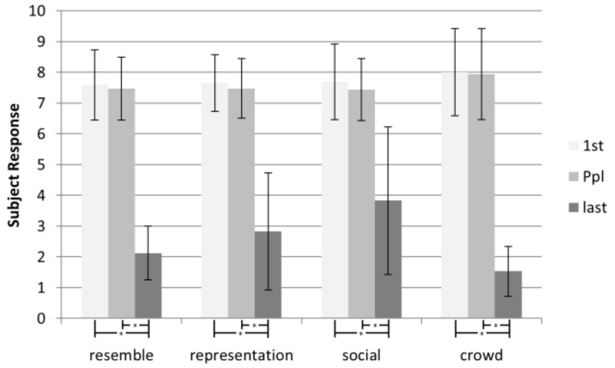


Figure 4: Box plot of the rankings of the avatars. 1st rank is the most similar to the self; 10th rank is the least similar to the self. Ppl=direct pipeline results; Attr=attractive/average; SE=Southern European; NE=Northern European

were Southern European. NE was a closer match to the participant's own facial structure. Not including this participant, the inter-rater agreement was (α=.82). The majority of participants had a facial structure similar to the avatar SE, which explains why the avatar 1/4 SE rated higher than 1/4 NE.

3.2 Avatar Rating

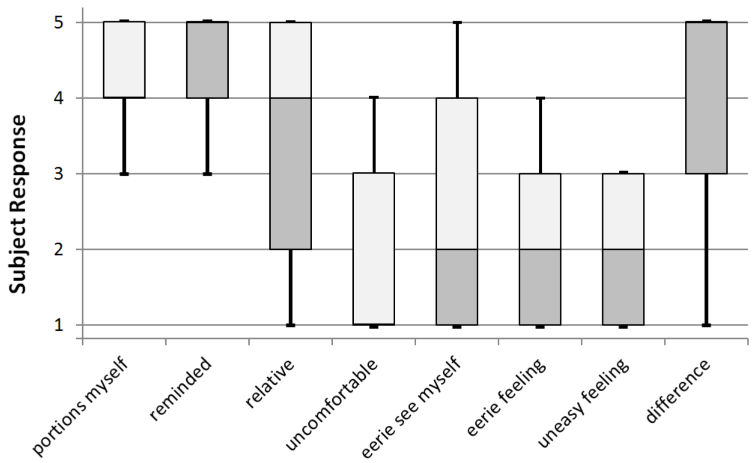
The ratings of the avatars in the second phase of the experiment can be seen in Figure 5. In the five cases where Ppl was ranked 1^{st}, the avatar was only evaluated once and the scores were replicated in both categories. For all questions, seen in the table of Figure 5, a Friedman Rank Sum repeated measures ANOVA was performed with a null-hypothesis that the distributions were the same across avatars. The null-hypothesis was rejected for all questions at the p<.001 level. PostHoc Fisher's LSD tests showed that the avatar with the face *last* was rated significantly different than the 1^{st} and Ppl for all questions at the p<.001 level. Ppl and 1^{st} responses did not differ significantly.



resemble	How much do you feel the avatar resembles you?
repre-sentation	How satisfied are you with the avatar as your representation?
social	How willing would you be to use the avatar in a formal social setting? (for example, a teleconference meeting for work)
crowd	How likely do you think it is that someone who knows you could pick this avatar out of a crowd when searching for you?

Figure 5: Study responses to the questions about the full avatars. Error Bars are the Standard Deviation. * is significant at the p<.001 level. The questions were administered orally when viewing the full body avatars on a response scale of 1-10.

The post questionnaire responses are shown in Figure 6. Participants generally responded that the avatar reminded them of themselves, as during the session. They reported having difficulty telling the faces from each other during the *PC* portion of the study. Only a single participant indicated no problems. This is also reflected in the variation in the rankings obtained from the *PC* method for the first rated faces.



portions myself	I saw portions of myself in the avatar.
reminded	The avatar reminded me of myself.
relative	The avatar looked like he could be a close relative.
uncomfortable	I felt uncomfortable looking at the avatar.

eerie see myself	It was eerie to see myself standing there. Eerie is defined as "causing fear because of strangeness"
eerie feeling	Looking at the avatar gave me an eerie feeling.
uneasy feeling	The avatar gave me an uneasy feeling.
difference	I had difficulty telling the difference between some faces.

Figure 6: Subject responses to the post session questions about the full body avatar perceived to be the most visually similar to the self. Responses were on an anchored 5pt Likert scale (disagree completely 1 …completely agree 5).

4 Discussion

The results indicated that participants reported that the look-alike avatars did appear to be visually similar to themselves. This extends existing findings (Bailenson et al., 2003, 2004) to show that look-alike avatars are perceived as visually similar in 1st person evaluation. This validates their use in research investigating the impact of having a look-alike avatar. Moreover, our results indicate that perceived similarity with the self might be achieved with avatars with lesser degrees of objective likeness to the person, as avatars morphed 25% towards arbitrary, but racially similar avatar were often chosen as the most visually similar to the self (11/16). These results also speak in favour of being very cautious with the avatars chosen in any experimental context, as slight morphs towards a racially matched face produced similar results. Even an incidental physical similarity between participant and avatar may be a confounding factor.

The control question in Fox and Bailenson (2009) that addressed participant self resemblance scored at the midpoint of the 5 point Likert scale they used. In contrast, the look-alike avatars in our study scored well above the midpoint (7.5/10) of the scale, indicating higher agreement with avatar's visual appearance being like oneself. It is difficult to pinpoint why such a difference might exist. While the visual quality of the faces seem comparable in quality, a variety of other factors such as hair, the exact questions, clothing match/mismatch, visual quality of the body, matching of the avatar body size to the participant, and the static vs. moving bodies make it hard to compare. A comparison with earlier validation studies of their look-alike avatars (Bailenson et al., 2003, 2004) is, however, problematic since they used 3rd person evaluation in those studies whereas we have performed 1st person evaluation in this study. These issues remain open for future studies.

An important result of this experiment is the recognition of the breadth of acceptance of avatar likenesses. The rankings indicate that the faces made from morphs 1/4 to others produced a perception of visual similarity to the self much the same as the direct pipeline results. Most participants indicated that they had difficulties telling some of the faces apart (average score of 4 out of 5). Eleven participants experienced both the Ppl avatar and a morphed avatar that was ranked 1*st* in the full-body phase of the experiment. The scoring on all questions indicated that there were no differences in the perception of the self likeness between the two. We believe these results indicate that people do not require avatars that are identical to themselves, but rather ones that are "close enough." Not only are the direct pipeline results approximations, but the rankings and ratings of avatars that are morphed mixes of other faces indicate that there is a lot of leeway.

We believe that this self-appearance flexibility is not specifically tied to our pipeline but they are generally applicable. Our results are consistent with the self-enhancement effect found with morphed photos. For instance, in (Epley & Whitchurch, 2008) the highest likelihood morph for being rated as one's own was the 20% morph, and the 30% morph still maintained a likelihood of 40%. However, the fact that the 1/4 morphed avatars ended up as high in the ranking as the direct pipeline result may alternatively be an indication of the shortcomings of the current pipeline.

Fine details of the person are somewhat washed out in the avatars. This may make the differences less pronounced. Participants indicated difficulties seeing the differences in some face pairs, consistent with the facial recognition literature. Also, the similarity scores were quite high for the 1*st* ranked avatars, which included 1/4 morphed avatars, and informally we found accurate third person evaluations similar to (Bailenson et al., 2003, 2004), indicating that this is not a general issue with our pipeline.

A possible confounding factor in our study was that participants viewed the avatar from a third person viewpoint, which resulted in feelings of discomfort for some of the participants. Although we are used to seeing ourselves from a third person perspective in mirrors, seeing a static, life-sized 3D representation visually similar to the self is extremely rare. In the case of a body ownership experiment, which was the motivation for this work, the avatar would be seen only from 'within' the avatar, for instance in a virtual mirror. We are hopeful that those participants who experienced uncomfortable sensations viewing the static avatar will be more at ease when the avatar, moving as they do, is viewed from a first person perspective in a mirror. We are currently investigating our technology in the context of virtual embodiment.

We believe this work also shows the potential for using virtual avatars to investigate self-image issues. The issue of choice of representation for virtual environments has already received some attention (Ducheneaut, Wen, Yee, & Wadley, 2009; Inkpen & Sedlins, 2011), without yet investigating avatars that are similar to the self. Recent work has started to explore the impact with third person avatars, for instance (Kim & Sundar, 2012). When combined with embodiment, a number of different research avenues can be identified. Research to gain a better understanding of the mechanisms of self-representation could be performed, for example on neurological mechanisms (Blanke & Metzinger, 2009; Uddin, 2011). Other research could deal with body-image issues, where direct and controlled manipulation of the visually perceived self appearance could be performed.

Acknowledgements

This research was supported by the FP7 EU collaborative project VERE (257695), FP7 Marie Curie IAPP project GOLEM (251415), and the ERC project TRAVERSE (227985).

References

Bailenson, J. N., Beall, A. C., & Blascovich, J. (2003). Using Virtual Heads for Person Identification: An Empirical Study Comparing Photographs to Photogrammetrically-Generated Models. *Journal of Forensic Identification*, *53*(6), 722–728.

Bailenson, J. N., Beall, A. C., Blascovich, J., & Rex, C. (2004). Examining Virtual Busts: Are Photogrammetrically Generated Head Models Effective for Person Identification? *Presence: Teleoperators and Virtual Environments*, *13*(4), 416–427. MIT Press.

Bailenson, J. N., & Segovia, K. Y. (2010). Virtual Doppelgangers: Psychological Effects of Avatars Who Ignore Their Owners. In W. S. Bainbridge (Ed.), *Online Worlds: Convergence of the Real and the Virtual* (pp. 175–186). Springer London.

Banakou, D., Groten, R., & Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences*, *110*(31), 12846–12851.

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

Brown, T. C., & Peterson, G. L. (2009). *An Enquiry Into the Method of Paired Comparison*. RMRS-GTR-216WWW, U.S. Department of Agriculture.

Cash, T. F., & Pruzinsky, T. (2002). *Body image: A handbook of theory, research, and clinical practice*. Guilford press New York.

Chelnokova, O., & Laeng, B. (2011). Three-dimensional information in face recognition: An eye-tracking study. *Journal of vision*, *11*(13). Association for Research in Vision and Ophthalmology.

Cooley, C. H. (1964). *Human nature and the social order*. Transaction Publishers.

Ducheneaut, N., Wen, M.-H., Yee, N., & Wadley, G. (2009). Body and mind: a study of avatar personalization in three virtual worlds. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’09)* (pp. 1151–1160). ACM.

Epley, N., & Whitchurch, E. (2008). Mirror, mirror on the wall: Enhancement in self-recognition. *Personality and Social Psychology Bulletin, 34*, 1159–1170.

Fox, J., & Bailenson, J. N. (2009). Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors. *Media Psychology, 12*(1), 1–25. Taylor & Francis.

Gillies, M., & Spanlang, B. (2010). Comparing and evaluating real time character engines for virtual environments. *PRESENCE: Teleoperators and Virtual Environments*, *19*(2), 95–117.

Grechkin, T. Y., Nguyen, T. D., Plumert, J. M., Cremer, J. F., & Kearney, J. K. (2010). How does presentation method and measurement protocol affect distance estimation in real and virtual environments? *ACM Transactions on Applied Perception (TAP)*, *7*(4), 26.

Inkpen, K. M., & Sedlins, M. (2011). Me and my avatar: exploring users' comfort with avatars for workplace communication. *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 383–386). New York, NY, USA: ACM.

Keyes, H., Brady, N., Reilly, R. B., & Foxe, J. J. (2010). My face or yours? Event-related potential correlates of self-face processing. *Brain and cognition*, *72*(2), 244–254. Elsevier.

Kim, Y., & Sundar, S. S. (2012). Visualizing ideal self vs. actual self through avatars: Impact on preventive health outcomes. *Computers in Human Behavior*, *28*(4), 1356–1364.

Liu, C. H., & Ward, J. (2006). The use of 3D information in face recognition. *Vision research*, *46*(6), 768–773. Elsevier.

Llobera, J., González-Franco, M., Perez-Marcos, D., Valls-Solé, J., Slater, M., & Sanchez-Vives, M. V. (2013). Virtual reality for assessment of patients suffering chronic pain: a case study. *Experimental Brain Research*, *225*(1), 105–117.

Luginbühl, T., Delattre, L., & Gagalowicz, A. (2011). Towards the automatic generation of 3D photo-realistic avatars using 3D scanned data. *Computer Vision/Computer Graphics Collaboration Techniques*, 192–203. Springer.

Martey, R. M., & Consalvo, M. (2011). Performing the looking-glass self: Avatar appearance and group identity in Second Life. *Popular Communication*, *9*(3), 165–180. Taylor & Francis.

Maselli, A., & Slater, M. (2013). The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience, 7*, 83.

Normand, J.-M., Giannopoulos, E., Spanlang, B., & Slater, M. (2011). Multisensory Stimulation Can Induce an Illusion of Larger Belly Size in Immersive Virtual Reality. *PLoS ONE*, *6*(1), e16128.

Orvalho, V., Bastos, P., Parke, F., Oliveira, B., & Alvarez, X. (2012). A Facial Rigging Survey. *in Proc. of the 33rd Annual Conference of the European Association for Computer Graphics - Eurographics* (pp. 10–32). ACM.

Orvalho, V., Zacur, E., & Susin, A. (2008). Transferring the Rig and Animations from a Character to Different Face Models. *Computer Graphics Forum*, *27*(8), 1997–2012.

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787.

Peterson, G. L., & Brown, T. C. (1998). Economic valuation by the method of paired comparison, with emphasis on evaluation of the transitivity axiom. *Land Economics*, *74*(2), 240–261.

Pimenta, A. M., Sánchez-Villegas, A., Bes-Rastrollo, M., López, C. N., & Martínez-González, M. Á. (2009). Relationship between body image disturbance and incidence of depression: the SUN prospective cohort. *BMC public health*, *9*(1), 1. BioMed Central Ltd.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review Psychology*, *57*, 199–226.

Richetin, J., Xaiz, A., Maravita, A., & Perugini, M. (2012). Self-body recognition depends on implicit and explicit self-esteem. *Body image*, *9*(2), 253–260. Elsevier.

Slater, M., Spanlang, B., Sanchez-Vives, M. V, & Blanke, O. (2010). First Person Experience of Body Transfer in Virtual Reality. *PLoS ONE*, *5*(5), e10564.

Sui, J., Zhu, Y., & Han, S. (2006). Self-face recognition in attended and unattended conditions: An event-related brain potential study. *NeuroReport*, *17*(4), 423–427. LWW.

Tecchia, F., Carrozzino, M., Bacinelli, S., Rossi, F., Vercelli, D., Marino, G., Gasparello, P., et al. (2010). A flexible framework for wide-spectrum VR development. *PRESENCE: Teleoperators and Virtual Environments*, *19*(4), 302–312.

Uddin, L. Q. (2011). Brain connectivity and the self: The case of cerebral disconnection. *Consciousness and Cognition*, *20*(1), 94–98.

Van Vugt, H. C., Bailenson, J. N., Hoorn, J. F., & Konijn, E. A. (2008). Effects of facial similarity on user responses to embodied agents. *ACM Trans. Comput.-Hum. Interact.*, *17*(2), 7:1–7:27. New York, NY, USA: ACM.

Wang, H., Sun, S., Shu, T., Shi, F., & Wu, J. (2009). Survey: Parameterized 3d human body modeling and geometric deformation technology. *Computer-Aided Industrial Design & Conceptual Design* (pp. 1486–1493). IEEE.

Annex: Look-alike Avatar Creation

The custom pipeline used to create the look-alike avatars is detailed here. The presented materials are intended to illustrate a method that enables the production of visually similar avatars to a target person in a short time span, as was used in the experiment presented. A design decision to treat the head and body as two different entities enabled us to tackle the issue of having a high quality face coupled with a lower polygon count body, while still using only unregistered images. For a review of full body modeling see (Wang, Sun, Shu, Shi, & Wu, 2009). Solutions such as high-end scanners can produce the quality required, but require extensive manual work to reduce the body mesh (Luginbühl, Delattre, & Gagalowicz, 2011). The division of body and face enabled the usage of Singular Inversion's FaceGen software for facial capture. However, this introduced a new problem; the generated head mesh had to be incorporated with the body mesh, in a process known as stitching. This is a difficult process and typically requires an expert artist, since the vertices in the head have to be matched and welded to the vertices in the neck. This operation also impacts the process of rigging of the avatar (see (Orvalho, Bastos, Parke, Oliveira, & Alvarez, 2012) for a review of the rigging process and methods), since not only must the meshes be joined, but also the two control structures.

We overcome this issue by controlling the template mesh to which the unregistered images are matched, during the model acquisition step. By providing a mesh with an invariant topology, we constrain both the stitching and the rigging, creating the necessary conditions for an automated process. An expert artist created a single full body rigged avatar, which acts as a template and only had to be created once. At the time this experiment was performed, we only had a male mesh available.

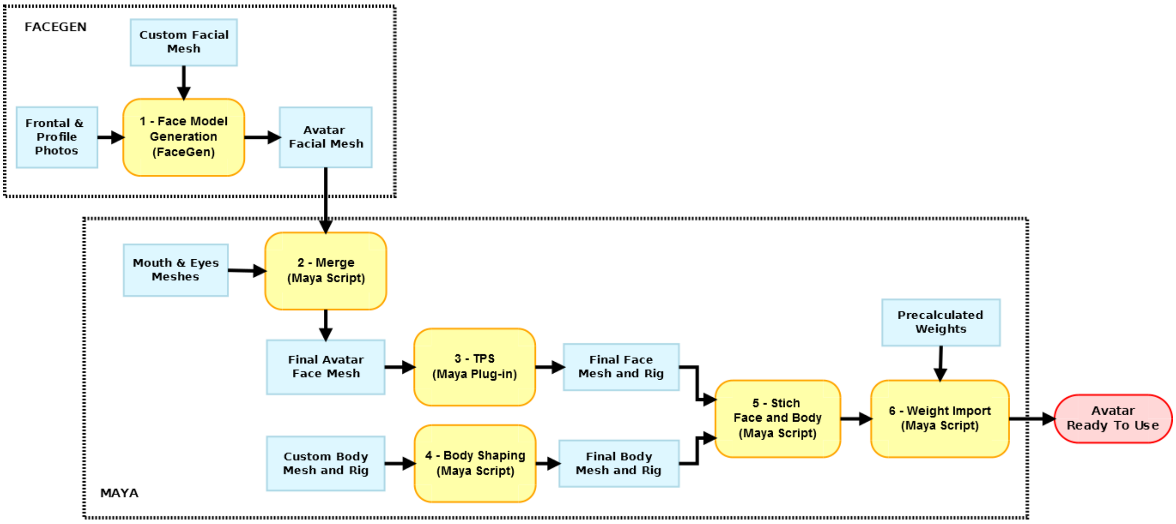


Figure 8 Diagram of the process of the proposed look-alike avatar pipeline. The hatched portion of the process realized in Autodesk Maya is performed in a single off-line process, performed once per base avatar body.

The head was then split from the body by a carefully chosen cutting plane, providing a sound stitching during the avatar creation process.

We now present an outline of the pipeline, visualized in Figure 7, and a description of the operation of key points. The process begins by taking three unregistered images of the target person, one frontal and two profile, which serve as input to FaceGen. The output mesh from FaceGen is then loaded into the modelling software package Autodesk Maya, where a custom plugin helps the user perform the remaining steps. Maya was chosen because it offers easy customization in the form of plugins and is one of the most commonly used modelling packages. The plugin consists of a number of scripts which perform mesh and skeleton manipulations necessary to combine the head and body and to customize the final body. The initial code appends eye balls and the interior of the mouth, complete with tongue and teeth, to the head mesh; this provided better control the autorigging task. Although automatically positioned, the eye balls occasionally require minor manual positioning adjustments for best visual appearance, which can be optionally performed with simple translate operations. The rigging follows, using the Thin Plate Spline (TPS) method (Orvalho, Zacur, & Susin, 2008). TPS transfers the rig from the template mesh created earlier. We have implemented the TPS algorithm in Maya. Since the topology is strictly controlled, the TPS method no longer requires manual landmarking; thus, we were able to achieve completely automated autorigging with no further intervention required. At this point, two steps down the pipeline, the head is ready.

Next, the operator sets up the body of the target person. With a set of sliding controls, the operator sets the body mass and the fitness level of the individual. Figure 1 depicts a set of different body shapes for the same target person. Additionally, it is possible to create several versions of the target person's body, which can be later morphed between in real-time in the running application, making experimental setups such as the one in (Normand et al., 2011) possible with avatars visually similar to the self. For our experimental purposes, we dressed the avatar with the same outfit that the target person is wearing during the experiment, a black suit for motion capture.

Once the avatar creator is satisfied with the virtual body, a single command in the plugin triggers a code segment that stitches the two pieces, head and body, together. Since the head and body mesh separated from an original full-body mesh, this consists of reconnecting the edges, but also requires handling the difference in size caused by the body shape morphing. Once this automated task is complete, the avatar is ready to be used. In order to give the avatar an extra level of likeness, it is possible to additionally enhance the model with accessories, such as eye glasses. The avatar creator selects from a set of different prefabricated eye glasses with extra textures applicable to each glasses style, and the plugin automatically positions them in place. As with any automated process, some manual adjustments may be required. This concludes the avatar creation.