*Journal of Perceptual Imaging*® 7: 1–16, 2024.

Erratum: A previous version of this paper erroneously gave the copyright to the Society for Imaging Science and Technology (IS&T).

# How Do Avatars Make a Positive Impression: The Effect of Facial Cues on Avatar Evaluation

**Yury Shevchenko**[*], **Jakob Heisenstein**[†], **and Ulf-Dietrich Reips**[‡]

*Research Methods, Assessment, and iScience, Department of Psychology, University of Konstanz*
*E-mail: yury.shevchenko@uni-konstanz.de*

**Abstract.** *This article examines the influence of facial features on the perception and evaluation of avatars in virtual environments. As people increasingly engage with avatars in virtual spaces, the visual appearance of these digital representations is critical to the design of human-computer interaction. Drawing on research on the evaluation of human faces, this study investigates how facial features influence perceptions of an avatar's attractiveness, trustworthiness, personality traits, and other characteristics. We conducted two factorial experiments that manipulated the avatars' eye size, jaw shape, and hairstyle. It was found that larger eyes conveyed a more positive impression and increased perceptions of attractiveness, sympathy, trustworthiness, extraversion, and openness. Although avatars with prominent jawlines were rated as more attractive, a prominent jawline was associated with a perception of higher dominance and threat. Stylish hairstyles were associated with higher extraversion and openness but also with lower conscientiousness. This study provides important insights into the design of avatars for virtual applications like gaming, e-commerce, and online therapy. It highlights the complex interplay between facial features and perception and contributes to the knowledge of how avatars can be optimally designed to create the desired impressions in virtual environments.*

***Keywords:*** *digital humanities, human-computer interaction, visual cues, avatar, face perception, online methods*

[DOI: 10.2352/J.Percept.Imaging.2024.7.000406]

People are increasingly exposed to virtual environments where they interact with other humans or virtual agents represented by avatars. For example, contacting customer service through a website may open a chat box that displays an animated image of the contact person next to the message. To what extent people evaluate and judge the avatar depends on its visual appearance. This may be relevant in virtual worlds like Second Life or Metaverse, where people may spend more time socially than in real life. Whether research on the evaluation of human faces is relevant to the perception of human-like avatars is debatable. In practice, knowing what visual cues make an avatar more positively perceived can help in designing avatars for specific purposes.

In this study, we capitalize on experiments evaluating human faces to determine how prominent visual cues operate in avatar evaluation. In two studies, we experimentally manipulated features such as eye size, jaw shape, and hairstyle attributes to determine how these cues affect the evaluation of an avatar, such as its attractiveness, trustworthiness, personality traits, and others. We first define an avatar and then review previous research on the effects of visual cues that are used to evaluate humans and avatars. We then describe the experiments and discuss the results.

Nowak and Fox [44] define an *avatar* as a digital representation of a human user for interacting with other users in a computer-mediated environment. An avatar may have a graphical embodiment, such as a two-dimensional drawing of a head or a three-dimensional model of a whole body. It is also possible for an avatar to represent a virtual agent, i.e., a computer algorithm, digital assistant or a game character. In this study, we focus on static avatars that resemble human face.

Avatar applications range from computer gaming [32] to healthcare, where avatars are incorporated into telemedicine and cyber-psychotherapy [46] with virtual reality applications [22]. Avatars are also used in e-learning [41], vocational training [69], health promotion [25, 60], in e-tourism (e.g., as guides in interactive museum tours) [11], workplace communication [27], and e-commerce, including customer support and marketing [24].

Interaction with avatars can evoke emotions and interpersonal attitudes similar to what people feel when interacting with other people [16]. Humans are generally capable of attributing mental states, feelings, and thoughts to others, i.e., theory of mind ability [51]. This ability can also be demonstrated with inanimate objects, such as people who see simple geometric figures moving on a screen (Frith-Happé animations) can attribute complex behavior and relationships to them [1]. A more common example is watching a cartoon with animated characters that can evoke complex and deep emotions and feelings, such as the death of Mufasa in *The Lion King* [3]. The creators of animated films enhance the visual features of characters to make them more appealing, e.g., by using childish features

[*]https://orcid.org/0000-0002-3397-8465
[†]https://orcid.org/0009-0000-1386-6821
[‡]https://orcid.org/0000-0002-1566-4745

Received Jan. 29, 2024; accepted for publication Nov. 13, 2024; published online Dec. 20, 2024. Associate Editor: Kimberly A. Jameson.

2575-8144/2024/7/000406/16/$00.00

[62]. Because anthropomorphic avatars resemble humans, research on human face evaluation can inform research on avatars.

## 1. EFFECT OF VISUAL CUES ON EVALUATION OF HUMAN FACES

McArthur and Baron [36] proposed an *ecological approach* to social perception, according to which the perception and evaluation of a human face is influenced by its adaptive value to individual and species survival. Zebrowitz [75] further specified this theoretical account to explain first impressions of faces and suggested different types of evolutionary important visual cues: Facial features that resemble a baby, features that indicate a high level of fitness or a familiar identity, and features that signal positive emotions. According to Zebrowitz [75], such features generate a positive perception in an individual. Empirical research has also supported the idea that these features are more likely to evoke positive impressions [45, 48, 50, 61]. Dotsch and Todorov [14] determined that the regions of mouth, eye, eyebrows, and hair are most informative in social perception of faces. Of these regions, regions of mouth and eyebrows are related to dynamic expressions of emotions. In this study, we have used static images and focused on the regions of eyes, jaw, and hair.

Eye size appears to have an influence on judging attractiveness, whereby individuals with larger eyes are considered more attractive [13] and more honest [76]. This is presumed to be so, because larger eyes resemble infants and signal a person's attention and arousal [75]. A more prominent jaw gives the impression of higher dominance [74] and may indicate a high level of testosterone during growth [67]. The perception of hair is affected to a lesser extent by its evolutionary fitness, because hair or a hairstyle can be a result of a conscious decision and thus functions as an expression of one's identity or as a way to make impressions on others, such as an unconventional hairstyle [42]. Borkenau and Liebler [10] found positive correlations between stylish hairstyle and ratings of the personality traits of extraversion and openness to experience. Stylish hairstyle may be indicative of higher creativity, which is also positively associated with openness to experience and extraversion and negatively associated with conscientiousness [28]. "Creativity" in this context refers to the perception of one's artistic type, with a good imagination, and a liking for creating works of art (e.g., as measured in the HEXACO-60 personality inventory by Ashton and Lee [6]).

## 2. EFFECT OF VISUAL CUES ON EVALUATION OF AVATARS

Avatars offer more flexibility in manipulating their appearance than human faces, which provides the opportunity to modify visual cues to a degree that is not possible with human faces (e.g., making some features exceptionally large). In addition, the context of the avatar's application should be considered. Previous research has shown that visual features of avatars can not only influence how they are perceived at that instance but also in subsequent engagement. For example, the visual embodiment of an avatar has been shown to affect purchase motivation [68] and student motivation in an e-learning environment [64].

A noteworthy occurrence in the evaluation of avatars' appearance is the *uncanny valley* phenomenon, where avatars that are highly realistic but not quite perfect can elicit feelings of eeriness, discomfort, and revulsion in observers [39]. Several factors contribute to the uncanny valley effect, such as inconsistencies between different aspects of an avatar (e.g., a realistic face with an unrealistic body), unnatural movements or expressions, and a high level of realism that nonetheless deviates from human norms [30]. The uncanny valley effect may decrease trust and affinity towards avatars [72] and reduce their effectiveness in applications like virtual reality or video games [53].

Literature evidence on evaluation of human faces can be applied to the evaluation of anthropomorphic avatars [8, 17]. Schwind and colleagues [62] provide practical advice for avatar creators to use appealing features of a human face such as symmetry and smooth skin. Features that suggest fitness or childishness tend to increase affinity with avatars. With regard to particular visual cues, Ferstl and McDonnell [17] determined that larger avatar eyes elicited perceptions of higher trustworthiness and appeal as well as lower aggressiveness and dominance. Similarly, Wang and colleagues [70] determined that avatars with larger eyes were perceived as more honest. Regarding the association between a prominent jaw and dominance, Schwind and colleagues [63] observed that people who were asked to create an avatar portrayed male villains with more pronounced jawlines. It can be inferred that a more prominent jaw may be associated with a higher perceived threat from avatars. Along the same lines, it was found that toddlers prefer less dominant faces of male avatars [19]. To determine the link between hairstyle and creativity, Bélisle and Bodur [8], used the Brunswik Lens Model, and determined that stylish hairstyles led avatars to be perceived as more extraverted and less emotionally stable.

## 3. PRESENT STUDY

In our study, we systematically varied three types of visual cues; eye size, jawline shape, and hairstyle type, in factorial experiments to determine their effects on avatar evaluation. We measured interpersonal attitudes and judgments about the avatar. General impression has been shown to influence judgments about individual attributes, and described in literature as the "halo effect" [43]. Argyle [5] grouped interpersonal attitudes into two main factors: Affiliation (sympathy versus hostility) and status (dominance versus submission), which align with two main dimensions, valence and dominance, typically used to evaluate human faces [45]. Thus, measures of affiliation (sympathy, attractiveness) and status (dominance) were included in our study. It is possible that people interact with avatars in situations that require building trust (in cyber-psychotherapy), thus, we included measures of trust, calming effect of the avatar, perceived threat, and the readiness to open up to an avatar. Interaction with avatars may depend on "personality". The

Big-Five personality traits proposed by Goldberg [20] are used to evaluate the observer's impression of an avatar's personality. Some of the personality traits, such as high extraversion and agreeableness, have been shown to be rated more positively than their counterparts [61]. Based on literature, we hypothesized that avatars with larger eyes should create a more positive impression than avatars with smaller eyes. Specifically, larger eyes are expected to give the impression of higher trustworthiness, higher agreeableness, higher extraversion, and higher sympathy for the avatar, and it should be easier to open up to the avatar, when it is perceived as less threatening and more attractive.

Two additional hypotheses were formulated regarding the effects of jawline and hairstyle. First, avatars with a square jawline should create the impression of higher dominance and higher perceived threat than avatars with a round jawline. Second, avatars with a stylish hairstyle should evoke the impression of higher creativity than avatars with a plain hairstyle. This hypothesis is based on the idea that a fashionable hairstyle is perceived as a signal of unconventionality or creativity. Previous research has shown that creativity is positively associated with openness to experience and extraversion and negatively associated with conscientiousness [28]. Therefore, a stylish hairstyle is expected to elicit the impression of higher extraversion, higher openness, and lower conscientiousness.

In addition, we aimed to investigate the interactions between different factors that influence the evaluation of avatars. For example, the effect of cues such as eye size, jawline shape and hairstyle could be different for female and male avatars. As this was an exploratory question, we did not formulate any hypotheses for interaction effects.

## 4. STUDY 1

In Study 1, we conducted a factorial experiment that manipulated the avatars' eye size, jawline, hairstyle, and gender.

## 5. METHOD

### 5.1 *Participants*

We recruited 55 participants through the online psychology student pool "SONA" at the University of Konstanz. The incentive for participation was course credit, with the additional option to enter a lottery for a 10 Euro gift card. All participants passed the seriousness check [55]. Incomplete datasets due to dropout during the study led to the exclusion of data from three participants. Two measures were implemented to control for careless responses [26]. Based on inconsistency analysis using retest items [31] and intra-individual response variability analysis, four participants were identified as careless responders and their data were excluded from this study. The final sample used for analysis consisted of 48 participants, 36 females and 12 males. The self-reported modal age range was 20–24 years.

### 5.2 *Study Design*

We applied a $2 \times 2 \times 2 \times 2$ full factorial design with within-subject repeated measures. The factors were the avatar's eye size (small and large), jawline (round and square), hairstyle (plain and stylish), and gender of the avatar (female and male). The combination of these four factors resulted in 16 trials. Each participant completed the trials in a randomized order. The study was conducted online using WEXTOR at https://wextor.eu/ [58, 59] and required 30 minutes to complete. The instructions were presented in German language.

### 5.3 *Materials*

#### 5.3.1 *Avatars*

The stimuli consisted of 16 images of avatars generated on the website Avatar Maker (avatarmaker.com). The images depicted the upper body part of an avatar in a frontal view (see Figure 1). Each avatar represented a unique combination of visual cues resulting from four manipulated factors with two levels each. Regarding the manipulated cues, the eye size in the "large eyes" feature was one standard deviation above that of an average European Caucasian face [71]. In "small eyes" feature, eye size was decreased by one standard deviation from an average face. The inter-pupillary distance was kept constant in both conditions. For the jawline factor, a distinct round face shape with little emphasis on the jaw bones was chosen for the "round" feature, while a face shape with an accentuated edged jawline was selected for the "square" feature. Facial breadth measured at ear level and facial height were kept similar in both conditions. The hairstyle in the "plain" feature consisted of a neat, well-groomed hairstyle, while the "stylish" feature represented a stylish, spiky hairstyle. The exposed length of the hair was similar in both conditions. To present avatars with different genders, the default "female avatar" and "male avatar" options on Avatar Maker website were used. The male avatars had a wider and shorter neck and an Adam's apple compared to the female avatars. The male avatars had a slightly larger face, and their eyebrows were denser and lower. The eyelashes were longer and more pronounced on female avatars, and their mouth and lips were slightly smaller than that of male avatars. Apart from the manipulated cues, all other aspects of the avatars were kept constant in all 16 variations. The skin tone for all avatars resembled a typical central European complexion (# ffd0bc), with brown hair color (# 543c32), and the color of the eye irises was light blue (# 7085b3). The background and clothing color were also kept constant with neutral colors for all avatars. Images were centrally presented on the screen at a size of $450 \times 450$ px.

#### 5.3.2 *Measurements*

The measurements consisted of the evaluation of personality traits and interpersonal attitudes attributed to avatars. To measure the personality trait attributions, the observer rating variant of the German version of the Ten-item-personality scale (TIPI-G) was used [21, 40]. Two adjectives were presented for each personality trait and were rated on a 7-point Likert scale from 1 ("Disagree strongly") to 7 ("Agree strongly") regarding how well an adjective described an avatar's personality. The 10 items of the TIPI-G scale were

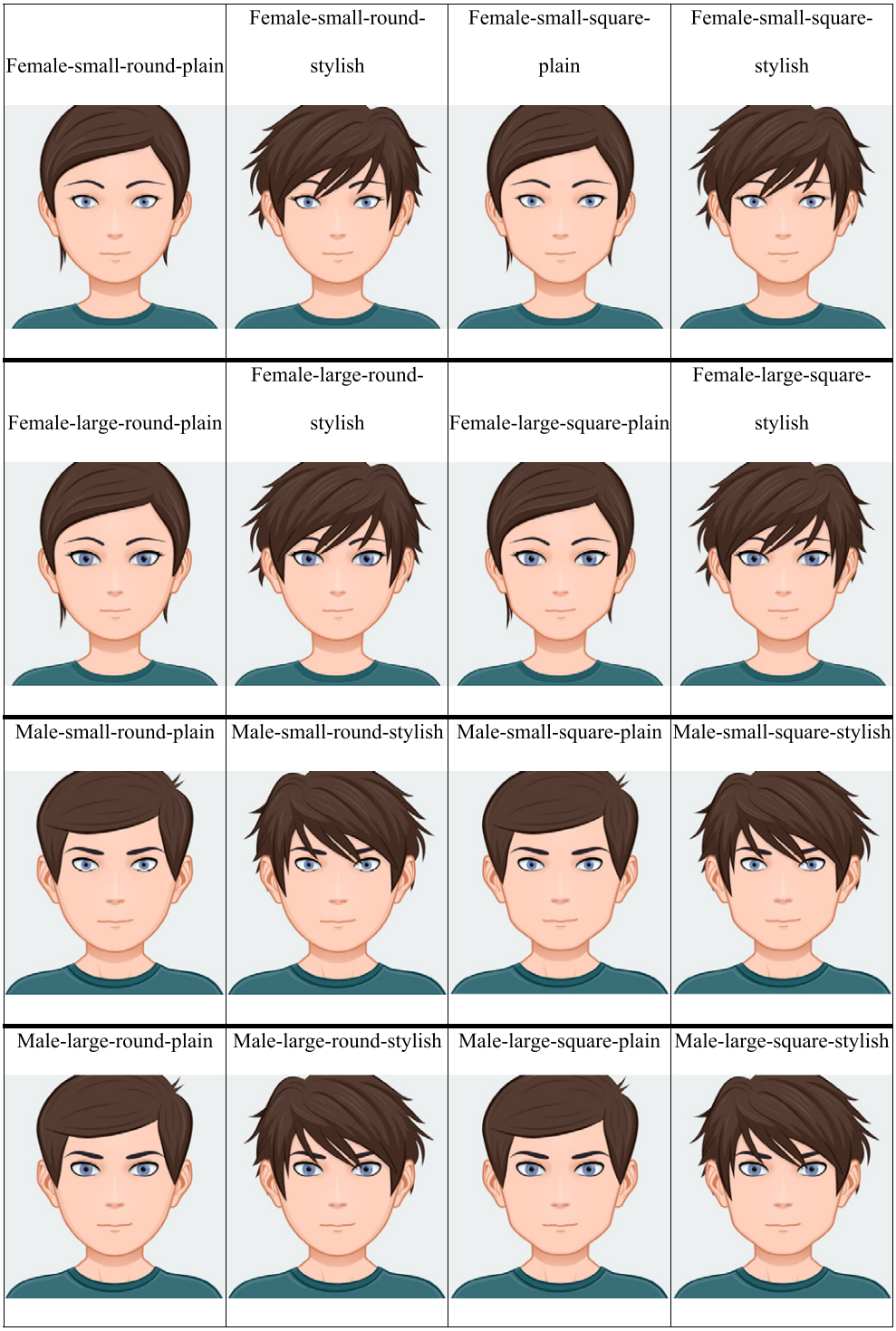


Figure 1. Avatars in Experiment 1. *Note.* The figure displays the stimuli used in Experiment 1 (features described in Section 5.3.1). The adjectives refer to factor combinations ordered by avatars' gender, eye size, jawline, and hairstyle.

presented in blocks of five questions on two subsequent pages, and kept directly below the avatar's image, at all times.

Regarding the interpersonal attitudes that an avatar evokes in a viewer, single-item questions were chosen, similar to previous studies [37, 47] and a visual analog scale (VAS) was used to record responses [18]. The VAS was labeled with "Disagree strongly" at the left end and "Agree strongly" at the right end of the scale. The single-item questions were presented in One-Item-One-Screen (OIOS) format [56]. The statement used for each question was "I feel the avatar is. . ."

followed by an adjective. The adjectives were "trustworthy", "threatening", "sympathetic", "calming", "dominant", and "physically attractive". An additional question asked how easily the participant felt they could open up to the avatar ("I could easily open up to the avatar").

To evaluate the hypotheses, we aggregated measures into three scales: "positive impression", "threatening dominance", and "creativity". All measures were standardized to a scale between 0 and 1. The positive impression scale was calculated as the average of all scores (except dominance), which were considered as positive and socially desirable based on previous studies [12, 61]. The threatening dominance scale was an average of dominance and perceived threat. The creativity-related items scale was derived from an average of extraversion, openness, and reverse-coded conscientiousness scores, which have previously shown to be associated with higher creativity [28]. To reiterate, we are referring to the perception of a person as creative, artistic, and having a good imagination. We did not measure these perceptions directly, but instead used the term "creativity-related scale" to aggregate the personality traits that have been found to be associated with creativity.

### 5.4 *Procedure*

On the first page of the online experiment (The demo version of the experiment is available at https://exp.wextor.eu/demo/avatar_study1/), participants were presented with an informed consent form, general study information, and a seriousness check [57]. The participants then provided their demographic information. Next, participants were shown an instruction page that explained the study procedure and included example images of avatars. The evaluation of the avatars began with a trial in which participants rated one avatar, which was different from the avatars in the main trials. After the trial, participants were asked to rate 16 avatars in random order. Each avatar trial consisted of nine pages. The first two pages presented the avatar's picture and five items of the personality scale. On the following seven pages, the avatar was presented with single-item questions about interpersonal attitudes. After all 16 avatars had been rated, participants again rated three randomly selected avatars from the original sample. Finally, participants were asked questions about their own appearance, including hair color, eye color, hairstyle, and face shape. Participants were then thanked, given the researchers' contact details, and offered the opportunity to register for a raffle.

### 5.5 *Analysis*

Before conducting the main analysis, participants with careless responses were filtered out by analyzing inconsistency of their answers to the same avatar and general response variability. An inconsistency value was calculated by summing the absolute differences between evaluations of three randomly chosen avatars presented twice. Participants with an inconsistency value more than two standard deviations from the mean value were classified as careless responders. To identify participants who gave the same responses to questions, a variant of "Intra-Individual Response Variability (IRV) Index" was used [15]. The IRV index was calculated by determining the standard deviation of each participant's responses across all items. Participants with an IRV index below two standard deviations from the mean IRV were classified as careless responders.

Data pre-processing and main analysis was done in R [52]. Because data was nested within participants, we used *lme*4 package for linear-mixed models [7] and included a random intercept for each participant in the model. Four experimental factors were included as fixed effects. Moreover, to control the effect of participants' gender, we included it as a fixed effect in the model. We evaluated three models, for each dependent variable according to the hypotheses: Positive impression, threatening dominance, and creativity score. To test the interaction effects, we estimated additional models with the interactions between the fixed effects. Additionally, we evaluated the effect of factors on each measurement, calculated the correlation matrix of the measurements and internal consistency of the scales. Data and analysis scripts are accessible at the OSF repository at https://osf.io/nsuzm/.

## 6. RESULTS

All three hypotheses were supported by the results (see Table I). Avatars with large eyes made a more positive impression than avatars with small eyes, $b = 0.04$, $SE = 0.01$, $p < 0.001$, $M_{\text{large}} = 0.59$, $SD_{\text{large}} = 0.13$, versus $M_{\text{small}} = 0.55$, $SD_{\text{small}} = 0.14$, $d = 0.25$. Avatars with a square jawline made a stronger impression of threatening dominance than avatars with a round jawline, $b = 0.07$, $SE = 0.01$, $p < 0.001$, $M_{\text{edged}} = 0.37$, $SD_{\text{edged}} = 0.22$, versus $M_{\text{round}} = 0.30$, $SD_{\text{round}} = 0.20$, $d = 0.26$. Avatars with a stylish hairstyle made a stronger impression of creativity than avatars with a plain hairstyle, $b = 0.15$, $SE = 0.01$, $p < 0.001$, $M_{\text{stylish}} = 0.58$, $SD_{\text{stylish}} = 0.15$, versus $M_{\text{plain}} = 0.42$, $SD_{\text{plain}} = 0.17$, $d = 0.68$.

With respect to the interactions between manipulated factors, all interaction effects turned out to be insignificant ($ps > 0.05$), except the interaction effect of large eye size and square jawline on creativity, $b = 0.08$, $SE = 0.02$, $p < 0.001$.

Regarding effects that were not hypothesized, male avatars made a less positive impression than female avatars, $b = -0.07$, $SE = 0.01$, $p < 0.001$ (see descriptive statistics in Table A.1 in Appendix A). Male avatars had a higher threatening dominance and creativity score than female avatars, $b = 0.12$, $SE = 0.01$, $p < 0.001$, and $b = 0.05$, $SE = 0.01$, $p < 0.001$. On the other hand, participant gender was not found to have an effect on avatar ratings nor did it interact significantly with avatar gender ($ps > 0.05$). Other non-hypothesized effects included rating avatars with large eyes or square jawline as more creative ($b = 0.06$, $SE = 0.01$, $p < 0.001$, and $b = 0.05$, $SE = 0.01$, $p < 0.001$), and avatars with a stylish hairstyle as making a more positive impression ($b = 0.05$, $SE = 0.01$, $p < 0.001$).

**Table I.** Regression coefficients for positive impression, threatening dominance, and creativity in Study 1.

| | Positive impression | | | Threatening dominance | | | Creativity | | |
|---|---|---|---|---|---|---|---|---|---|
| Predictors | Estimates | CI | *p* | Estimates | CI | *p* | Estimates | CI | *p* |
| (Intercept) | 0.56 | 0.52–0.60 | <0.001 | 0.23 | 0.16–0.29 | <0.001 | 0.34 | 0.31–0.38 | <0.001 |
| Large eyes (versus small) | 0.04 | 0.03–0.06 | <0.001 | −0.01 | −0.04–0.01 | 0.29 | 0.06 | 0.03–0.08 | <0.001 |
| Square jawline (versus round) | 0.01 | −0.00–0.03 | 0.10 | 0.07 | 0.05–0.10 | <0.001 | 0.05 | 0.03–0.07 | <0.001 |
| Male avatar (versus female) | −0.07 | −0.08–0.05 | <0.001 | 0.12 | 0.10–0.15 | <0.001 | 0.05 | 0.03–0.07 | <0.001 |
| Stylish hairstyle (versus plain) | 0.05 | 0.03–0.06 | <0.001 | −0.01 | −0.04–0.02 | 0.44 | 0.15 | 0.13–0.18 | <0.001 |
| Female participant (versus male) | −0.00 | −0.05–0.04 | 0.83 | 0.03 | −0.04–0.10 | 0.44 | 0.00 | −0.03–0.04 | 0.86 |
| Random effects | | | | | | | | | |
| $\sigma^2$ | | 0.01 | | | 0.03 | | | 0.02 | |
| $\tau^{00}_{participant}$ | | 0.00 | | | 0.01 | | | 0.00 | |
| ICC | | 0.20 | | | 0.23 | | | 0.05 | |
| $N_{participant}$ | | 48 | | | 48 | | | 48 | |
| Observations | | 768 | | | 767 | | | 768 | |
| Marginal $R^2$/Conditional $R^2$ | | 0.11/0.29 | | | 0.11/0.32 | | | 0.26/0.30 | |

**Table II.** Regression coefficients for all measurements in Study 1.

| | Large eyes (versus small) | Square jawline (versus round) | Stylish hairstyle (versus plain) | Male avatar (versus female) | Female participant (versus male) |
|---|---|---|---|---|---|
| Extraversion | 0.08*** | 0.09*** | 0.14*** | 0.09*** | −0.02 |
| Agreeableness | 0.03* | −0.03 | 0.04** | −0.16*** | 0.01 |
| Conscientiousness | −0.03 | −0.02 | −0.15*** | −0.09*** | 0.01 |
| Emotional Stability | 0.03* | 0.06*** | 0.06*** | 0.06*** | 0.02 |
| Openness | 0.07*** | 0.04** | 0.17*** | −0.04* | 0.04 |
| Trust | 0.04** | −0.01 | 0.01 | −0.11*** | −0.00 |
| Threat | −0.02 | 0.05** | −0.02 | 0.13*** | 0.01 |
| Sympathy | 0.07*** | 0.01 | 0.07*** | −0.09*** | −0.01 |
| Calming | 0.03* | −0.03* | 0.04* | −0.12*** | −0.01 |
| Dominance | −0.00 | 0.09*** | 0.00 | 0.12*** | 0.05 |
| Attractiveness | 0.07*** | 0.08*** | 0.08*** | −0.03* | −0.04 |
| Ease of opening up | 0.05*** | −0.01 | 0.05*** | −0.12*** | −0.03 |

*Note*: The regression coefficients are displayed in the table. The significance level is indicated by asterisks: * < 0.05, ** < 0.01, and *** < 0.001.

To investigate further the effect of manipulated factors on each specific scale, we repeated the main analysis for each measurement (see Table II). Avatars with large eyes had higher ratings for extraversion, openness, sympathy, attractiveness, and ease of opening up than avatars with small eyes. Avatars with square jawline had higher ratings for extraversion, emotional stability, attractiveness, and dominance than avatars with round jawline. Avatars with stylish hairstyle had higher ratings for extraversion, emotional stability, openness, sympathy, attractiveness, and ease of opening up, but lower ratings for conscientiousness than avatars with plain hairstyle.

We computed the correlations between measurements to understand the relationships between different measured constructs (see Figure A1 in Appendix A). To address the multiple comparison problem, all p-values were adjusted using Bonferroni correction. Significant and positive correlations ($rs > 0.70$, $ps < 0.001$) were found between measures of sympathy, trust, calming, and ease of opening up. Agreeableness was positively correlated with trust ($r = 0.65$, $p < 0.001$), sympathy ($r = 0.63$, $p < 0.001$), calming ($r = 0.61$, $p < 0.001$), and ease of opening up ($r = 0.62$, $p < 0.001$), and negatively with threat ($r = -0.58$, $p < 0.001$) and dominance ($r = -0.56$,

$p < 0.001$). Extraversion was positively correlated with emotional stability ($r = 0.54$, $p < 0.001$). Attractiveness was positively correlated with sympathy ($r = 0.60$, $p < 0.001$), ease of opening up ($r = 0.53$, $p < 0.001$), and trust ($r = 0.50$, $p < 0.001$). Dominance was positively correlated with threat ($r = 0.61$, $p < 0.001$).

Finally, we computed internal consistency within three scores: Positive impression, threatening dominance, and creativity. It was high for positive impression ($\alpha = 0.84$), moderate for threatening dominance ($\alpha = 0.75$) and acceptable for creativity-related items ($\alpha = 0.67$). An exploratory factor analysis examining item dimensionality identified three factors that generally aligned with the structure of these three scores. Principal component analysis (PCA) was conducted on 12 items with oblique rotation (Promax), and initial analysis of eigenvalues revealed three components exceeding Kaiser's criterion of 1, which was supported by the scree plot's inflections. Pattern matrix revealed that Factor 1 consisted of 5 items (Trust [0.84], Ease to Open [0.84], Sympathy [0.82], Calming [0.79], and Attractiveness [0.73]) explaining 32% of variance; Factor 2 included 3 items (Dominance [0.84], Threat [0.74], and Agreeableness [−0.52]) explaining 15% of variance; and Factor 3 comprised 4 items (Extraversion [0.84], Openness [0.70], Conscienciousness [−0.49], and Emotional stability [0.52]) explaining 15% of variance. The factor correlation matrix showed a moderate negative correlation of -0.48 between Factor 1 and Factor 2, supporting the use of an oblique rotation.

## 7. DISCUSSION

We confirmed that eye size was an important visual cue that influenced the general impression of an avatar. The bigger an avatar's eyes the more positive its impression. This result is consistent with findings on the perception of human faces [47, 76] and avatars [17]. Our results contributed to understanding which judgments and interpersonal attitudes are most influenced by eye size, namely extraversion, openness, sympathy, attractiveness, and ease of opening up. However, Wang et al. [70] detemined that an avatar's large eyes are not always associated with more positive attitudes toward the avatar. A possible explanation for this is that beyond a certain limit in increasing eye size, positive impression is no longer evoked and creates the feeling of eeriness [62]. To investigate this further, we included a condition with unrealistically large eyes in Study 2.

Regarding the effects of jaw shape, we confirmed that a square jawline increased the perception of an avatar's dominance and perceived threat compared to a round jawline. Avatars with a square jawline were also perceived as more extraverted, emotionally stable, and attractive. Square jawline was positively correlated with a higher creativity score, which could be explained by positive associations of square jawline with higher extraversion and openness to experience.

A stylish spiky hairstyle created an impression of a more creative personality than a plain hairstyle. With regard to personality traits, avatars with a stylish hairstyle were perceived as more extraverted, more open to new experience, and less conscientious.

## 8. STUDY 2

The objective of Study 2 was to replicate the main findings using a different avatar tool to create avatars that were more realistic than the avatars in Study 1 (for conceptual replication, see Miccoli & Reips [38]). Additionally, with respect to the eye size feature, we contrasted large eyes with enormous eyes, in order to manipulate the feature to the extent that can remove the positive influence of eye size on the general impression of an avatar.

## 9. METHOD

### 9.1 *Participants*

We recruited 63 participants through the online psychology student pool "SONA" at the University of Konstanz. The incentive for participating was course credit. We applied the same exclusion criteria as in Study 1 to exclude data from participants because of dropouts ($n = 5$), missing data ($n = 1$), or careless responding ($n = 5$). One of the participants who did not complete the experiment did not pass the seriousness check. The final sample used for analysis consisted of 52 participants, 41 females and 11 males. The self-reported modal age range was 20–24 years.

### 9.2 *Study Design*

We used a $2 \times 2 \times 2$ full factorial design with within-subject repeated measures, similar to Study 1. The factors included the avatar's eye size (large and enormous), jawline (round and square), and hairstyle (plain and stylish) (In contrast to Study 1, the factor of avatar's gender was not included in Study 2, so all avatars were presented as female avatars. The reason for not considering this factor was that gender cues have a strong impact on the rating of avatars, but cannot always be freely chosen when an avatar is supposed to represent a user. In addition, following feedback on Experiment 1 regarding long duration and exhaustion, which could affect data quality, options for shortening the experiment were considered.). The combination of these three factors resulted in 8 trials, which were presented in a random order and lasted 15 minutes.

### 9.3 *Materials*

#### 9.3.1 *Avatars*

Study 2 used eight avatar images created with Character Creator 3 [54]. These images showed the upper body of a female avatar from the front in a central position (see Figure 2). Each avatar represented a unique combination of eye size, jawline, and hairstyle. The eye size in "large eyes" feature was one standard deviation above that of an average European Caucasian face [71]. For "enormous eyes" feature, eye size was increased by three standard deviations over an average face. Manipulation of the jawline was similar to that

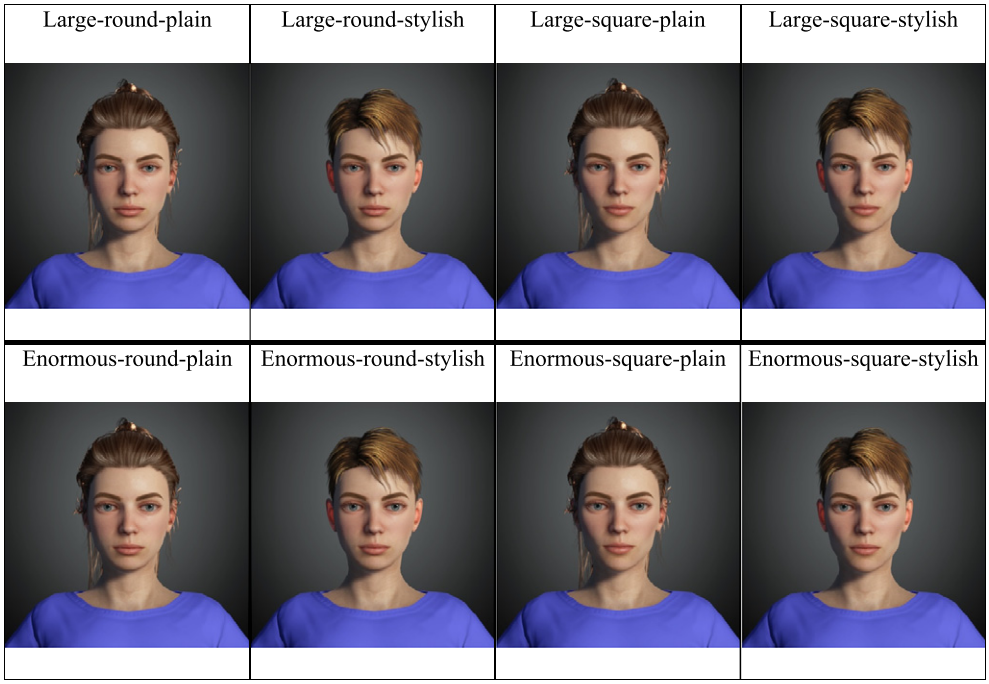


Figure 2. Avatars in Experiment 2. *Note*: The figure displays the stimuli used in Experiment 2 with the conditions described in section 9.3.1. The adjectives refer to factor combinations ordered by eye size, jawline, and hairstyle.

in Study 1. Regarding hairstyle, avatars with "plain" feature had a bun hairstyle, whereas avatars with "stylish" feature had a spiky hairstyle with few strands over the forehead. The hair appeared shorter in "stylish" feature than in "plain" feature. Other visual characteristics, such as skin tone, hair color, and eye color, remained the same for all avatars. Background and clothing colors were neutral. Images were presented centrally on the screen at a size of 450 × 450 pixels.

#### 9.3.2 *Measurements*

The measurements in Study 2 were identical to those in Study 1, except that the question of whether it is easy to open up to an avatar was omitted as redundant because it had high correlations with trust, sympathy, and calming in Study 1.

### 9.4 *Procedure*

The procedure in Study 2 was similar to the one in Study 1, except that we used only two randomly selected avatars for retesting after the main trials in order to reduce the length of the experiment (The demo version of the experiment is available at https://exp.wextor.eu/demo/avatar_study2/).

## 10. RESULTS

The hypotheses were supported by the results (see Table III). The positive impressions by avatars with enormous eyes were not different from avatars with large eyes, $b = 0.00$, $SE = 0.01$, $p = 0.95$, $M_{\text{enormous}} = 0.55$, $SD_{\text{enormous}} = 0.12$, versus $M_{\text{large}} = 0.55$, $SD_{\text{large}} = 0.12$, $d < 0.01$. Avatars with a square jawline made a stronger impression of threatening dominance than avatars with a round jawline, $b = 0.06$, $SE = 0.02$, $p < 0.001$, $M_{\text{edged}} = 0.45$, $SD_{\text{edged}} = 0.21$, versus $M_{\text{round}} = 0.39$, $SD_{\text{round}} = 0.21$, $d = 0.26$. Avatars with a stylish hairstyle made a stronger impression of creativity than avatars with a plain hairstyle, $b = 0.05$, $SE = 0.01$, $p < 0.001$, $M_{\text{stylish}} = 0.52$, $SD_{\text{stylish}} = 0.14$, versus $M_{\text{plain}} = 0.47$, $SD_{\text{plain}} = 0.12$, $d = 0.30$.

Regarding the interactions between manipulated factors, all interaction effects turned out to be insignificant ($ps > 0.05$), except the interaction effect of enormous eye size and stylish hair on positive impression, $b = 0.04$, $SE = 0.02$, $p = 0.042$, and creativity, $b = -0.06$, $SE = 0.03$, $p = 0.018$. The significant interaction effect of eye size and jawline on creativity found in Study 1 was not confirmed in Study 2, $b = 0.00$, $SE = 0.02$, $p = 0.93$.

Regarding non-hypothesized effects, avatars with enormous eyes had a lower threatening dominance score than avatars with large eyes (see descriptive statistics in Table B.1 in Appendix B). Similar to Study 1, avatars with a square jawline had higher creativity ratings, and avatars with a stylish hairstyle made a more threatening dominance impression. In contrast to Study 1, in which the stylish hairstyle left a more positive impression, in Study 2 the stylish hairstyle made a less positive impression than plain hairstyle. As in Study 1, the gender of the participants had no influence on the evaluation of the avatars.

We repeated the main analysis for each measurement to evaluate the effect of manipulated factors on a specific scale (see Table IV). Avatars with enormous eyes had lower ratings for dominance than avatars with large eyes. Avatars with square jawline had higher ratings for extraversion, con-

**Table III.** Regression coefficients for positive impression, threatening dominance, and creativity in Study 2.

| | Positive impression | | | Threatening dominance | | | Creativity | | |
|---|---|---|---|---|---|---|---|---|---|
| Predictors | Estimates | CI | *p* | Estimates | CI | *p* | Estimates | CI | *p* |
| (Intercept) | 0.55 | 0.50–0.59 | <0.001 | 0.37 | 0.30–0.44 | <0.001 | 0.46 | 0.42–0.51 | <0.001 |
| Enormous eyes (versus large) | 0.00 | −0.02–0.02 | 0.95 | −0.07 | −0.10–0.04 | <0.001 | 0.01 | −0.02–0.03 | 0.61 |
| Square jaw (versus round) | 0.05 | 0.03–0.06 | <0.001 | 0.06 | 0.03–0.09 | <0.001 | 0.05 | 0.03–0.08 | <0.001 |
| Stylish hairstyle (versus plain) | −0.05 | −0.07–0.03 | <0.001 | 0.09 | 0.06–0.12 | <0.001 | 0.05 | 0.03–0.08 | <0.001 |
| Female participant (versus male) | 0.00 | −0.04–0.05 | 0.94 | 0.01 | −0.07–0.08 | 0.86 | −0.03 | −0.07–0.02 | 0.21 |
| Random effects | | | | | | | | | |
| $\sigma^2$ | 0.01 | | | 0.03 | | | 0.01 | | |
| $\tau^{00}_{participant}$ | 0.00 | | | 0.01 | | | 0.00 | | |
| ICC | 0.26 | | | 0.33 | | | 0.16 | | |
| $N_{participant}$ | 52 | | | 60 | | | 52 | | |
| Observations | 416 | | | 473 | | | 416 | | |
| Marginal $R^2$/Conditional $R^2$ | 0.08/0.33 | | | 0.09/0.39 | | | 0.09/0.23 | | |

**Table IV.** Regression coefficients for all measurements in Study 2.

| | Enormous eyes (versus large) | Square jawline (versus round) | Stylish hairstyle (versus plain) | Female participant (versus male) |
|---|---|---|---|---|
| Extraversion | −0.03 | 0.13*** | 0.06** | −0.06 |
| Agreeableness | 0.05** | −0.00 | −0.07*** | −0.01 |
| Conscientiousness | −0.04* | 0.05** | −0.05** | 0.03 |
| Emotional Stability | −0.05** | 0.11*** | 0.02 | 0.05 |
| Openness | 0.00 | 0.08*** | 0.05** | 0.01 |
| Trust | 0.02 | 0.02 | −0.09*** | −0.01 |
| Threat | −0.05** | 0.05* | 0.08*** | 0.02 |
| Sympathy | 0.02 | 0.03* | −0.09*** | −0.00 |
| Calming | 0.01 | 0.01 | −0.10*** | −0.02 |
| Dominance | −0.09*** | 0.09*** | 0.10*** | 0.02 |
| Attractiveness | −0.03 | 0.05** | −0.17*** | 0.04 |

*Note.* The regression coefficients are displayed in the table. The significance level is indicated by asterisks: * <0.05, ** <0.01, and *** <0.001.

scientiousness, emotional stability, openness, and dominance than avatars with round jawline. Avatars with stylish hairstyle had lower ratings for agreeableness, trust, sympathy, calming, and attractiveness and higher ratings for perceived threat and dominance than avatars with plain hairstyle.

The correlations between measurements were computed with p-values adjusted using Bonferroni correction (see Figure B1 in Appendix B). We found significant and positive correlations ($rs > 0.60$, $ps < 0.001$) between measures of sympathy, trust, and calming. Agreeableness was positively correlated with trust ($r = 0.57$, $p < 0.001$), calming ($r = 0.57$, $p < 0.001$), sympathy ($r = 0.56$, $p < 0.001$), and negatively with threat ($r = -0.55$, $p < 0.001$) and dominance ($r = -0.48$, $p < 0.001$). Extraversion was positively correlated with emotional stability ($r = 0.47$, $p < 0.001$) and openness ($itr = 0.47$, $p < 0.001$). Attractiveness was positively correlated with sympathy ($r = 0.61$, $p < 0.001$), trust ($r = 0.56$, $p < 0.001$), and calming ($r = 0.50$, $p < 0.001$). Finally, dominance was positively correlated with threat ($r = 0.62$, $p < 0.001$).

Regarding internal consistency of the scales, it was moderate for positive impression ($\alpha = 0.78$) and threatening dominance ($\alpha = 0.75$) and low for creativity items ($\alpha = 0.38$). Three factors that were generally consistent with the structure of these three scores were identified in an exploratory factor analysis that examined the dimensionality of the items. Principal component analysis (PCA) was conducted on 11 items with oblique rotation (Promax), and based on the scree plot and Kaiser's criterion, three factors were retained in the final analysis. Pattern matrix revealed that Factor 1 consisted of 5 items (Attractiveness [0.79], Sympathy [0.74], Trust [0.72], Calming effect [0.64], and Conscientiousness [0.36]) explaining 24% of variance; Factor 2 included 3 items (Dominance [0.89], Threat [0.74], and Agreeableness [−0.54]) explaining 18% of variance; and Factor 3 comprised 3 items (Extraversion [0.96], Openness [0.59], and Emotional stability [0.47]) explaining 15% of variance. The factor correlation matrix showed a moderate negative correlation of −0.42 between Factor 1 and Factor 2 and a weak positive correlation of 0.38 between Factor 1 and Factor 3, supporting the use of an oblique rotation.

## 11. DISCUSSION

Using a different avatar creation tool, we replicated the results of Study 1 with respect to the effects of jawline on the perception of threatening dominance and stylish hairstyle

**Table V.** Summary of study results and previous research.

| Factor | Our findings | Previous research on avatars | Previous research on human faces |
|---|---|---|---|
| Eye size: large versus small | Larger eyes were associated with a more positive impression and greater attractiveness. | Ferstl & McDonnell [17]; Wang et al. [70] | Cunningham [13]; Zebrowitz et al. [76] |
| Jawline: square versus round | A square jawline had higher ratings of threatening dominance. | Schwind et al. [63] | Swaddle & Reierson [67]; Windhager et al. [74] |
| Hairstyle: stylish versus plain | A stylish hairstyle was linked to higher ratings of openness to new experiences and lower conscientiousness. | Bélisle & Bodur [8] | Borkenau & Liebler [10] |

on the perception of creativity. With respect to eye size, we found that increasing the eye size to an extreme degree did not uniformly improve the overall positive impression of the avatar, as measured by the average of all scores excluding dominance. Specifically, avatars with enormous eyes were perceived as more agreeable, but, on the other hand, less conscientious and emotionally stable. These results help reconcile the seemingly contradictory findings of a positive effect of eye size in some studies (e.g., Ferstl & McDonnell [17]) but not in others (e.g., Wang et al. [70]). Increasing the size of the eyes in an avatar can help create a more positive impression, but only to a certain limit, which is one standard deviation from the average eye size, according to this study.

With respect to jaw shape, avatars with square jawline had higher ratings of threatening dominance than avatars with round jawline. The effect size in Study 2 ($b = 0.06$) was similar to the that of Study 1 ($b = 0.06$), even though Study 2 presented only female avatars while Study 1 had both female and male avatars.

Stylish hairstyle of an avatar made a more creative impression than plain hairstyle. However, the effect size in Study 2 ($b = 0.05$) was smaller than in Study 1 ($b = 0.15$). Stylish hairstyle also had a stronger positive effect on the threatening dominance score in Study 2 ($b = 0.09$) than in Study 1 ($b = 0.02$). Additional analysis revealed that avatars with stylish hairstyle in Study 2 were perceived as less attractive, agreeable, calming, and more dominant than avatars with plain hairstyle, which was not the case in Study 1. This can be related to the perception of the stylish hair in Study 2 as short hair.

## 12. GENERAL DISCUSSION

Through two studies we experimentally investigated, which visual cues elicited certain impressions about avatars. To summarize the results of both studies, eye size, jawline, and hairstyle may affect an avatar's evaluation (see Table V). Large eyes may evoke a more general positive impression and make the avatar being perceived as more attractive. These results may inform virtual environments where the goal is to create a trusting relationship with a virtual agent, such as online therapy, customer support, digital health services, etc. It is worth noting that increasing the eye size is effective to a certain extent only, above which the avatar may give an eerie impression.

The effect of a square jawline on higher ratings of threatening dominance was confirmed in both studies and is generally consistent with findings from the research on the perception of human faces [67, 74]. Although a square jawline may look more attractive, it may elicit a higher perception of potential threat. This means that visually emphasizing the prominent jaw is not recommended for virtual environments that are intended to create a trusting relationship. On the other hand, in other virtual environments that focus more on competition (e.g., online games), a prominent jaw may be more effective as it conveys dominance and a potential threat – qualities that may be important in a competitive environment.

Manipulation of an avatar's hairstyle is another visual cue that can influence evaluation. Our finding was that a more stylish hairstyle gave the impression of a more creative personality, as measured by openness to new experiences and lower conscientiousness. This is consistent with impression management theory in the perception of human faces, where hairstyle is an easier feature to manipulate (compared to eye size and jaw shape) that people can use to represent their social identity. However, certain aspects of perception that are affected by hairstyle can be influenced by the avatar creator's tools and the avatar's attractiveness. Applying a stylish hairstyle to an avatar in the virtual environment that is supposed to create trusting relationships can backfire, as that hairstyle can be viewed as a potential threat and reduce trust in a virtual agent. On the other hand, virtual environments where creativity is emphasized as a value (such as a dating website or a social network) may provide more uses for stylish hairstyles.

The results of our study can be used either for designing avatars representing humans or creating avatars as virtual agents. For example, with avatar creation tools such as *Bitmoji* [65], users can create an avatar by uploading their photo. Further customization of the avatar allows one to adjust its features, e.g., the eyes could be slightly enlarged or the jaw shape changed towards the desired effect without losing resemblance to the user. If we were to put together a suitable avatar image for a virtual environment that would

promote trust in a virtual agent, it would be a female avatar with large eyes, a round jawline, and a plain haircut. Our study elucidated the aspects of avatar evaluation that are influenced by the visual cues.

## 13. STUDY LIMITATIONS

Our study's limitations primarily stem from the restricted scope of manipulated factors and their respective conditions, particularly concerning eye size and avatar gender. In Study 1, we intentionally created a distinct contrast between two conditions (small versus large eyes) to enhance statistical power. However, this approach may have oversimplified the nuanced effects of eye size variation. Future research should consider incorporating a condition representing average eye size to provide a more comprehensive understanding of the eye size effect continuum. Study 2 focused exclusively on female avatars, a decision made with specific research goals in mind. Future studies should systematically examine both female and male avatars to investigate potential interactions between avatar gender, observer characteristics, and other theoretically relevant variables that may influence avatar perception.

In this context, Peterson and colleagues [49] provide an example of a different research methodology that can integrate a wide range of factors and inferred attributes in studies of human faces. The authors used deep neural networks to train models that associate facial representations with various attributes people infer. An interesting outcome of this research is a model that captures patterns in human judgments and can generate new facial photographs designed to evoke specific impressions.

Our study's scope was also limited to anthropomorphic static avatars representing only the facial area, which, while informative, presents opportunities for more comprehensive future research. Previous studies have shown that other cues can also influence social evaluations. Dynamic cues, such as emotional expressions, can provide richer information for evaluating people [23] and avatars [66]. At the same time, tools have been developed that allow users to create dynamic avatars that can mimic their facial expressions, such as Memoji [4]. Other aspects of the avatars' body, such as clothing or posture, can provide information for social evaluations [33, 34]. Finally, other communication channels, such as audio, can contribute to avatar evaluation [9], which, combined with recent advances in speech generation [2], offers interesting opportunities for future research.

## 14. METHODOLOGICAL NOTES

Both of our studies were conducted as online experiments. In an online study, it is important to have response quality control, which we did by performing a seriousness check and identifying careless responders. Participants in the experiments were presented with all conditions in a within-subjects repeated-measures design. This was possible because the number of avatars to be assessed was relatively small. If the number of avatars is large, other experimental designs should be considered. For example, one might consider random assignment of a subset of avatars to each participant, which would then be modeled as a random effect of the stimulus in a mixed model [29].

## 15. OUTLOOK

Future research in this area should explore the impact of dynamic cues, such as emotional expressions and body language, on avatar evaluations. Investigating how avatars with varying clothing styles and postures are perceived can also provide valuable insights. Additionally, exploring the role of audio communication channels in avatar evaluation, especially in combination with recent advances in speech generation, presents exciting opportunities. Moreover, studies could delve into the interplay of multiple visual and dynamic cues to better understand their combined effects on avatar perception. As technology continues to advance, further research can help refine avatar design principles to cater to specific virtual environments and objectives, ultimately enhancing user experiences and interactions in virtual spaces.

Future studies should consider the significant role that individual differences among observers play in avatar evaluation, particularly in relation to face preferences. Research in human face perception has already established important links between individual characteristics and facial preferences. For instance, personality traits, such as extraversion, have been associated with women's preferences for facial masculinity [73], suggesting that an observer's personality may influence their perception and evaluation of avatar features. Additionally, desired partner personality traits have been shown to correlate with judgments of attractiveness [35], implying that an individual's ideal partner characteristics might affect how they perceive and rate avatar attractiveness. Extending these insights to avatar studies could reveal complex interactions between observer characteristics and avatar design.

## 16. CONCLUSION

In summary, our two factorial experimental studies highlight the significance of visual cues like eye size, jaw shape, and hairstyle in shaping impressions of avatars. Large eyes have a positive effect, but the size should be moderated to a certain extent to avoid giving an eerie impression. A strong jaw emphasizes dominance and can also convey potential threat. Stylish hairstyle signals creativity, but should be used cautiously in virtual environments that require building trust with users. These insights have implications for various virtual settings, from therapy to gaming, where the design of avatars has an impact on users.

## ACKNOWLEDGMENT

This article is based on the Master's thesis completed by Jakob Heisenstein in 2022.

## APPENDIX A.

**Table A.1.** Means and standard deviations of all dependent variables for factors; eye size, jawline, hairstyle, and avatar's gender in Experiment 1.

| | Small eyes ($N = 384$) | Large eyes ($N = 384$) | Round jawline ($N = 384$) | Square jawline ($N = 384$) | Plain hair ($N = 384$) | Stylish hair ($N = 384$) | Female avatar ($N = 384$) | Male avatar ($N = 384$) | Overall ($N = 768$) |
|---|---|---|---|---|---|---|---|---|---|
| **Positive impression** | | | | | | | | | |
| Mean (SD) | 0.55 (0.14) | 0.59 (0.13) | 0.57 (0.14) | 0.58 (0.14) | 0.55 (0.14) | 0.60 (0.13) | 0.61 (0.13) | 0.54 (0.13) | 0.57 (0.14) |
| Range | 0.08–0.90 | 0.15–0.90 | 0.08–0.90 | 0.16–0.90 | 0.08–0.87 | 0.25–0.90 | 0.20–0.90 | 0.08–0.90 | 0.08–0.90 |
| **Threatening dominance** | | | | | | | | | |
| N-Miss | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| Mean (SD) | 0.34 (0.22) | 0.32 (0.21) | 0.30 (0.20) | 0.37 (0.22) | 0.34 (0.22) | 0.33 (0.20) | 0.27 (0.19) | 0.39 (0.22) | 0.33 (0.21) |
| Range | 0.01–0.86 | 0.01–0.88 | 0.01–0.83 | 0.01–0.88 | 0.01–0.86 | 0.01–0.88 | 0.01–0.84 | 0.01–0.88 | 0.01–0.88 |
| **Creativity** | | | | | | | | | |
| Mean (SD) | 0.47 (0.18) | 0.53 (0.17) | 0.48 (0.18) | 0.53 (0.17) | 0.42 (0.17) | 0.58 (0.15) | 0.48 (0.17) | 0.53 (0.18) | 0.50 (0.18) |
| Range | 0–0.94 | 0.03–0.92 | 0.03–0.92 | 0–0.94 | 0–0.92 | 0.14–0.94 | 0.08–0.94 | 0–0.92 | 0–0.94 |
| **Extraversion** | | | | | | | | | |
| N-Miss | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| Mean (SD) | 3.92 (1.51) | 4.38 (1.39) | 3.87 (1.46) | 4.42 (1.42) | 3.72 (1.48) | 4.58 (1.32) | 3.89 (1.45) | 4.41 (1.44) | 4.15 (1.47) |
| Range | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 |
| **Agreeableness** | | | | | | | | | |
| Mean (SD) | 4.40 (1.34) | 4.60 (1.31) | 4.57 (1.36) | 4.42 (1.28) | 4.38 (1.37) | 4.61 (1.27) | 4.96 (1.21) | 4.03 (1.27) | 4.50 (1.33) |
| Range | 1.50–7 | 1–7 | 1–7 | 1.50–7 | 1–7 | 1.50–7 | 1.50–7 | 1–7 | 1–7 |
| **Conscientiousness** | | | | | | | | | |
| Mean (SD) | 4.69 (1.30) | 4.54 (1.29) | 4.66 (1.28) | 4.57 (1.31) | 5.07 (1.21) | 4.16 (1.21) | 4.90 (1.17) | 4.33 (1.35) | 4.61 (1.29) |
| Range | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 | 1.50–7 | 1.50–7 | 1–7 | 1–7 |
| **Emotional stability** | | | | | | | | | |
| N-Miss | 1 | 1 | 1 | 1 | 2 | 0 | 2 | 0 | 2 |
| Mean (SD) | 4.53 (1.14) | 4.69 (1.17) | 4.44 (1.17) | 4.78 (1.12) | 4.43 (1.19) | 4.79 (1.10) | 4.42 (1.16) | 4.79 (1.13) | 4.61 (1.16) |
| Range | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 |
| **Openness to experience** | | | | | | | | | |
| N-Miss | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| Mean (SD) | 4.29 (1.36) | 4.69 (1.26) | 4.36 (1.31) | 4.62 (1.33) | 3.98 (1.30) | 5.00 (1.15) | 4.60 (1.34) | 4.38 (1.31) | 4.49 (1.33) |
| Range | 1–7 | 1–7 | 1–7 | 1–7 | 1–7 | 1.50–7 | 1.50–7 | 1–7 | 1–7 |
| **Trust** | | | | | | | | | |
| N-Miss | 0 | 2 | 0 | 2 | 0 | 2 | 1 | 1 | 2 |
| Mean (SD) | 137.20 (54.35) | 147.47 (50.21) | 143.28 (52.35) | 141.36 (52.78) | 140.89 (53.19) | 143.77 (51.91) | 156.20 (49.62) | 128.44 (51.77) | 142.32 (52.54) |
| Range | 6–249 | 5–250 | 5–250 | 6–248 | 5–250 | 10–249 | 18–250 | 5–248 | 5–250 |
| **Calming** | | | | | | | | | |
| N-Miss | 3 | 2 | 3 | 2 | 4 | 1 | 1 | 4 | 5 |
| Mean (SD) | 121.39 (58.06) | 128.49 (56.00) | 128.80 (56.79) | 121.11 (57.24) | 120.56 (57.12) | 129.30 (56.84) | 139.81 (56.54) | 109.97 (53.74) | 124.95 (57.11) |
| Range | 2–249 | 4–249 | 3–249 | 2–246 | 4–249 | 2–248 | 5–249 | 2–248 | 2–249 |
| **Sympathy** | | | | | | | | | |
| N-Miss | 2 | 1 | 0 | 3 | 3 | 0 | 2 | 1 | 3 |
| Mean (SD) | 137.98 (57.17) | 154.20 (48.60) | 144.92 (55.35) | 147.30 (51.90) | 137.24 (54.56) | 154.90 (51.28) | 157.84 (51.24) | 134.39 (53.48) | 146.10 (53.63) |
| Range | 2–247 | 4–247 | 2–247 | 2–247 | 2–247 | 5–246 | 2–247 | 2–243 | 2–247 |
| **Ease of opening up** | | | | | | | | | |
| N-Miss | 3 | 1 | 0 | 4 | 1 | 3 | 2 | 2 | 4 |
| Mean (SD) | 118.48 (59.98) | 131.45 (55.56) | 125.95 (58.43) | 124.00 (57.89) | 118.64 (57.35) | 131.35 (58.29) | 139.90 (57.51) | 110.05 (54.89) | 124.98 (58.13) |
| Range | 0–249 | 2–243 | 0–249 | 1–240 | 0–249 | 1–247 | 2–249 | 0–243 | 0–249 |

**Table A.1.** Continued.

| | Small eyes ($N = 384$) | Large eyes ($N = 384$) | Round jawline ($N = 384$) | Square jawline ($N = 384$) | Plain hair ($N = 384$) | Stylish hair ($N = 384$) | Female avatar ($N = 384$) | Male avatar ($N = 384$) | Overall ($N = 768$) |
|---|---|---|---|---|---|---|---|---|---|
| **Attractiveness** | | | | | | | | | |
| N-Miss | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| Mean (SD) | 124.50 (61.37) | 141.51 (59.62) | 122.57 (59.13) | 143.45 (61.24) | 123.36 (60.93) | 142.71 (59.71) | 136.66 (60.29) | 129.39 (61.68) | 133.02 (61.06) |
| Range | 0–249 | 1–248 | 1–237 | 0–249 | 1–249 | 0–243 | 1–248 | 0–249 | 0–249 |
| **Dominance** | | | | | | | | | |
| N-Miss | 4 | 1 | 3 | 2 | 3 | 2 | 2 | 3 | 5 |
| Mean (SD) | 96.80 (63.77) | 95.85 (63.93) | 84.48 (59.56) | 108.13 (65.76) | 96.29 (65.12) | 96.35 (62.56) | 81.49 (59.31) | 111.19 (64.77) | 96.32 (63.81) |
| Range | 0–249 | 1–249 | 0–249 | 1–249 | 1–249 | 0–249 | 0–233 | 1–249 | 0–249 |
| **Threat** | | | | | | | | | |
| N-Miss | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| Mean (SD) | 72.16 (56.82) | 65.92 (54.50) | 63.31 (53.03) | 74.80 (57.80) | 71.65 (58.43) | 66.44 (52.82) | 53.11 (48.68) | 85.02 (57.79) | 69.05 (55.72) |
| Range | 1–210 | 0–236 | 0–236 | 0–221 | 0–236 | 0–221 | 0–221 | 0–236 | 0–236 |

*Note*: The marginal aggregated and raw scores are presented. The personality attribution ratings were done on a 7-point Likert scale. The interpersonal attitudes were measured on a visual analog scale that for measurement purposes is divided into 250 measurement points.

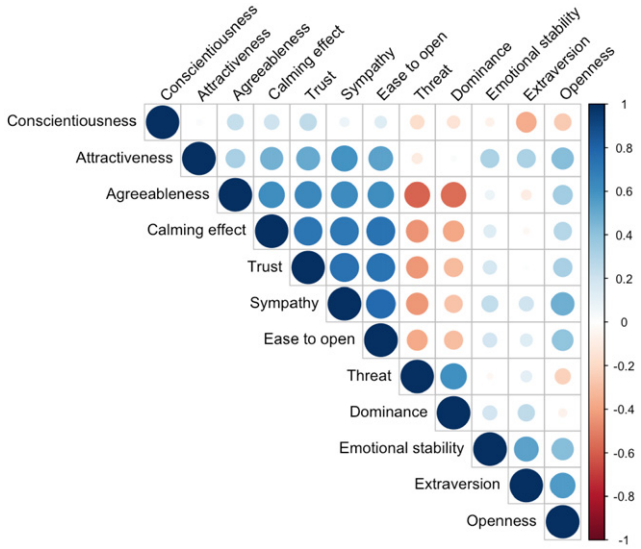


Figure A1. Correlations between measurements in Experiment 1.

## APPENDIX B.

**Table B.1.** Means and standard deviations of all dependent variables for factors; eye size, jawline, hairstyle, and avatar's gender in Experiment 2.

| | Large eyes ($N = 208$) | Enormous eyes ($N = 208$) | Round jawline ($N = 208$) | Square jawline ($N = 208$) | Plain hairstyle ($N = 208$) | Stylish hairstyle ($N = 208$) | Overall ($N = 416$) |
|---|---|---|---|---|---|---|---|
| **Positive impression** | | | | | | | |
| Mean (SD) | 0.55 (0.12) | 0.55 (0.12) | 0.52 (0.12) | 0.57 (0.12) | 0.57 (0.12) | 0.52 (0.12) | 0.55 (0.12) |
| Range | 0.24–0.87 | 0.22–0.86 | 0.22–0.86 | 0.26–0.87 | 0.22–0.86 | 0.24–0.87 | 0.22–0.87 |
| **Threatening dominance** | | | | | | | |
| Mean (SD) | 0.45 (0.23) | 0.38 (0.20) | 0.38 (0.21) | 0.45 (0.21) | 0.37 (0.20) | 0.46 (0.22) | 0.42 (0.21) |
| Range | 0.01–0.96 | 0–0.83 | 0–0.92 | 0.02–0.96 | 0–0.92 | 0–0.96 | 0–0.96 |

**Table B.1.** Continued.

| | Large eyes ($N = 208$) | Enormous eyes ($N = 208$) | Round jawline ($N = 208$) | Square jawline ($N = 208$) | Plain hairstyle ($N = 208$) | Stylish hairstyle ($N = 208$) | Overall ($N = 416$) |
|---|---|---|---|---|---|---|---|
| **Creativity** | | | | | | | |
| Mean (SD) | 0.49 (0.13) | 0.50 (0.14) | 0.47 (0.13) | 0.52 (0.12) | 0.47 (0.12) | 0.52 (0.14) | 0.50 (0.13) |
| Range | 0.14–0.89 | 0.17–0.83 | 0.14–0.89 | 0.22–0.83 | 0.14–0.81 | 0.17–0.89 | 0.14–0.89 |
| **Extraversion** | | | | | | | |
| Mean (SD) | 4.11 (1.28) | 3.95 (1.31) | 3.63 (1.27) | 4.43 (1.20) | 3.85 (1.22) | 4.22 (1.34) | 4.03 (1.30) |
| Range | 1–7 | 1–6.50 | 1–7 | 1.50–7 | 1–7 | 1–7 | 1–7 |
| **Agreeableness** | | | | | | | |
| Mean (SD) | 4.02 (1.21) | 4.31 (1.08) | 4.17 (1.15) | 4.16 (1.16) | 4.39 (1.07) | 3.94 (1.19) | 4.16 (1.16) |
| Range | 1.50–6.50 | 1.50–7 | 1.50–7 | 1.50–6.50 | 2–6.50 | 1.50–7 | 1.50–7 |
| **Conscientiousness** | | | | | | | |
| Mean (SD) | 4.74 (1.10) | 4.49 (1.11) | 4.45 (1.13) | 4.78 (1.08) | 4.76 (1.07) | 4.47 (1.14) | 4.62 (1.11) |
| Range | 2–7 | 2–7 | 2–7 | 2–7 | 2.50–6.50 | 2–7 | 2–7 |
| **Stability** | | | | | | | |
| Mean (SD) | 4.50 (1.25) | 4.19 (1.18) | 4.01 (1.30) | 4.68 (1.05) | 4.29 (1.19) | 4.40 (1.26) | 4.34 (1.23) |
| Range | 1–7 | 1.50–7 | 1–7 | 1.50–6.50 | 1–7 | 1–7 | 1–7 |
| **Openness** | | | | | | | |
| Mean (SD) | 4.53 (1.11) | 4.55 (1.12) | 4.30 (1.14) | 4.78 (1.03) | 4.39 (1.08) | 4.69 (1.13) | 4.54 (1.11) |
| Range | 1.50–7 | 1.50–7 | 1.50–7 | 2–7 | 1.50–7 | 2–7 | 1.50–7 |
| **Trust** | | | | | | | |
| N-Miss | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| Mean (SD) | 131.49 (48.82) | 136.41 (47.76) | 131.35 (48.20) | 136.57 (48.38) | 145.88 (46.95) | 122.09 (46.78) | 133.95 (48.30) |
| Range | 1–225 | 35–240 | 1–240 | 18–237 | 39–240 | 1–223 | 1–240 |
| **Calming** | | | | | | | |
| Mean (SD) | 110.71 (52.84) | 113.93 (50.20) | 111.00 (52.32) | 113.64 (50.76) | 124.51 (50.30) | 100.13 (49.89) | 112.32 (51.50) |
| Range | 9–238 | 11–245 | 10–245 | 9–238 | 11–245 | 9–224 | 9–245 |
| **Sympathy** | | | | | | | |
| Mean (SD) | 132.26 (50.73) | 136.54 (48.25) | 130.12 (49.32) | 138.68 (49.41) | 146.18 (47.20) | 122.62 (49.03) | 134.40 (49.49) |
| Range | 4–249 | 24–240 | 4–249 | 11–246 | 39–249 | 4–237 | 4–249 |
| **Attractiveness** | | | | | | | |
| Mean (SD) | 129.77 (65.37) | 123.50 (63.93) | 119.77 (64.55) | 133.50 (64.18) | 147.52 (59.29) | 105.75 (63.15) | 126.64 (64.65) |
| Range | 1–249 | 5–242 | 1–242 | 3–249 | 5–249 | 1–246 | 1–249 |
| **Dominance** | | | | | | | |
| Mean (SD) | 130.36 (63.31) | 107.50 (54.83) | 108.26 (59.44) | 129.61 (59.29) | 106.60 (56.41) | 131.26 (61.56) | 118.93 (60.25) |
| Range | 0–245 | 0–222 | 0–234 | 5–245 | 2–226 | 0–245 | 0–245 |
| **Threat** | | | | | | | |
| Mean (SD) | 96.23 (61.50) | 83.31 (56.26) | 84.11 (58.37) | 95.43 (59.66) | 80.02 (55.74) | 99.51 (61.10) | 89.77 (59.22) |
| Range | 1–240 | 0–237 | 0–237 | 0–240 | 0–237 | 2–240 | 0–240 |

*Note.* The marginal aggregated and raw scores are presented. The personality ratings were done on a 7-point Likert scale. The interpersonal attitudes were measured on a visual analog scale that for measurement purposes is divided into 250 measurement points.

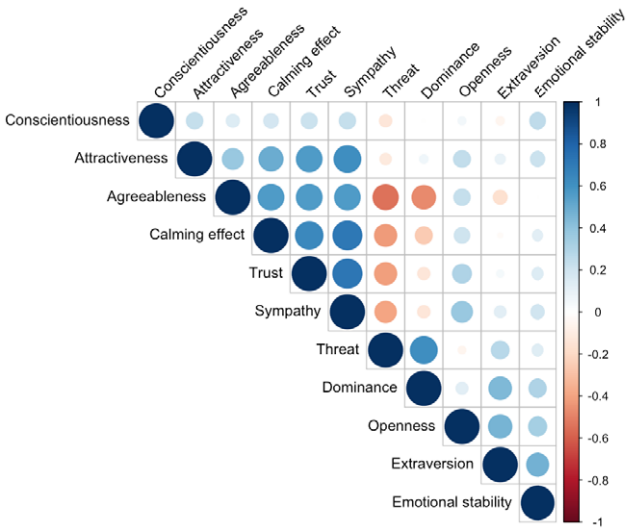


Figure B1. Correlations between measurements in Experiment 2.

## REFERENCES


[1] F. Abell, F. Happe, and U. Frith, "Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development," Cogn. Dev. **15**, 1–16 (2000).

[2] S. Alexanderson, É. Székely, G. E. Henter, T. Kucherenko, and J. Beskow, "Generating coherent spontaneous speech and gesture from text," *Proc. 20th ACM Int'l. Conf. on Intelligent Virtual Agents* (ACM, New York, NY, 2020), pp. 1–3.

[3] R. Allers and R. Minkoff (Directors), *The Lion King* [Film]. Buena Vista Pictures (1994).

[4] Apple (2024) *Use Memoji on your iPhone or iPad Pro* https://support.apple.com/en-us/111115.

[5] M. Argyle, "Non-verbal communication and language," R. Inst. Phil. Suppl. **10**, 63–78 (1976).

[6] M. C. Ashton and K. Lee, "The HEXACO–60: A short measure of the major dimensions of personality," J. Pers. Assess. **91**, 340–345 (2009).

[7] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," J. Statistical Software **67**, 1–48 (2015).

[8] J.-F. Bélisle and H. O. Bodur, "Avatars as information: Perception of consumers based on their avatars in virtual worlds," Psychol. Mark. **27**, 741–765 (2010).

[9] G. Bente, S. Rüggenberg, N. C. Krämer, and F. Eschenburg, "Avatar-mediated networking: Increasing social presence and interpersonal trust in net-based collaborations," Hum. Commun. Res. **34**, 287–318 (2008).

[10] P. Borkenau and A. Liebler, "Observable attributes as manifestations and cues of personality and intelligence," J. Pers. **63**, 1–25 (1995).

[11] A. Cafaro, H. H. Vilhjálmsson, and T. Bickmore, "First impressions in human-agent virtual encounters," ACM Trans. Comput. Hum. Interact. **23**, 1–40 (2016).

[12] A. Cafaro, H. H. Vilhjálmsson, T. Bickmore, D. Heylen, K. R. Jóhannsdóttir, and G. S. Valgarðsson, "First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters," *Proc. 12th Intelligent Virtual Agents: 12th Int'l. Conf. IVA 2012* (Springer, Cham, 2012), pp. 67–80.

[13] M. R. Cunningham, "Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty," J. Pers. Soc. Psychol. **50**, 925 (1986).

[14] R. Dotsch and A. Todorov, "Reverse correlating social face perception," Soc. Psychol. Pers. Sci. **3**, 562–571 (2012).

[15] A. M. Dunn, E. D. Heggestad, L. R. Shanock, and N. Theilgard, "Intra-individual response variability as an indicator of insufficient effort responding: Comparison to other indicators and relationships with individual differences," J. Bus. Psychol. **33**, 105–121 (2018).

[16] N. Epley, A., Waytz, and J. T. Cacioppo, "On seeing human: A three-factor theory of anthropomorphism," Psychol. Rev. **114**, 864 (2007).

[17] Y. Ferstl and R. McDonnell, "A perceptual study on the manipulation of facial features for trait portrayal in virtual agents," *Proc. 18th Int'l. Conf. on Intelligent Virtual Agents* (ACM, New York, NY, 2018), pp. 281–288.

[18] F. Funke and U.-D. Reips, "Why semantic differentials in web-based research should be made from visual analogue scales and not from 5-point scales," Field Methods **24**, 310–327 (2012).

[19] C.-I. Galusca, M. Mermillod, J.-C. Dreher, J.-B. van der Henst, and O. Pascalis, "Toddlers' sensitivity to dominance traits from faces," Sci. Rep. **13**, 2229 (2023).

[20] L. R. Goldberg, "An alternative 'description of personality': The Big-Five factor structure," J. Pers. Soc. Psychol. **59**, 1216–1229 (1990).

[21] S. D. Gosling, P. J. Rentfrow, and W. B. Swann Jr, "A very brief measure of the Big-Five personality domains," J. Res. Pers. **37**, 504–528 (2003).

[22] A. Gruber and R. Kaplan-Rakowski, "User experience of public speaking practice in virtual reality," in *Cognitive and Affective Perspectives on Immersive Technology in Education*, edited by Robert Z. Zheng (IGI Global, Hershey, PA, 2020), pp. 235–249.

[23] E. Hehman, J. K. Flake, and J. B. Freeman, "Static and dynamic facial cues differentially affect the consistency of social evaluations," Pers. Soc. Psychol. Bull. **41**, 1123–1134 (2015).

[24] P. Hemp, "Avatar-based marketing," Harv. Bus. Rev. **84**, 48–57 (2006).

[25] M. Horne, A. Hill, T. Murells, H. Ugail, R. Chinnadorai, and M. Hardy, "Using avatars in weight management settings: A systematic review," Internet Interventions **19**, 100295 (2020).

[26] J. L. Huang, P. G. Curran, J. Keeney, E. M. Poposki, and R. P. Deshon, "Detecting and deterring insufficient effort responding to surveys," J. Bus. Psychol. **27**, 99–114 (2012).

[27] K. M. Inkpen and M. Sedlins, "Me and my avatar: Exploring users' comfort with avatars for workplace communication," *Proc ACM 2011 Conf. on Computer Supported Cooperative Work* (ACM, New York, NY, 2011), pp. 383–386.

[28] M. Jirásek and F. Sudzina, "Big five personality traits and creativity," Qual. Innov. Prosperity **24**, 90–105 (2020).

[29] C. M. Judd, J. Westfall, and D. A. Kenny, "Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem," J. Pers. Soc. Psychol. **103**, 54 (2012).

[30] J. Kätsyri, K. Förger, M. Mäkäräinen, and T. Takala, "A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness," Frontiers Psychol. **6** (2015).

[31] J. W. Keeley, C. Webb, D. Peterson, L. Roussin, and E. H. Flanagan, "Development of a response inconsistency scale for the personality inventory for DSM-5," J. Pers. Assess. **98**, 351–359 (2016).

[32] R. Klevjer, "Enter the avatar: The phenomenology of prosthetic telepresence in computer games," in *The Philosophy of Computer Games*,

edited by J. R. Sageng, H. Fossheim, and T. M. Larsen (Springer, Cham, 2012), pp. 17–38.
[33] D. Küster, E. G. Krumhuber, and U. Hess, "You are what you wear: Unless you moved—Effects of attire and posture on person perception," J. Nonverbal Behav. **43**, 23–38 (2019).
[34] K. Legde and D. W. Cunningham, "Evaluating the effect of clothing and environment on the perceived personality of virtual avatars," *Proc. 19th ACM Int'l. Conf. on Intelligent Virtual Agents* (ACM, New York, NY, 2019), pp. 206–208.
[35] A. C. Little, D. M. Burt, and D. I. Perrett, "What is good is beautiful: Face preference reflects desired personality," Pers. Individual Differences **41**, 1107–1118 (2006).
[36] L. Z. McArthur and R. M. Baron, "Toward an ecological theory of social perception," Psychol. Rev. **90**, 215 (1983).
[37] M. Mehu, A. C. Little, and R. I. M. Dunbar, "Sex differences in the effect of smiling on social judgments: An evolutionary approach," J. Soc. Evol. Cult. Psychol. **2**, 103 (2008).
[38] M. Miccoli and U.-D. Reips, "Exact versus conceptual replication: Internet-based research investigating the replicability of cognitive effects." [Manuscript submitted for publication] J. Perceptual Imaging (Special Issue) (2024).
[39] M. Mori, K. F. Macdorman, and N. Kageki, "The uncanny valley [From the Field]," IEEE Robot. Autom. Mag. **19**, 98–100 (2012).
[40] P. M. Muck, B., Hell, and S. D. Gosling, "Construct validation of a short five-factor model instrument," Eur. J. Psychol. Assess. **23**, 166–175 (2007).
[41] S. Mystakidis, E., Berki, and J.-P. Valtanen, "Deep and meaningful e-learning with social virtual reality environments in higher education: A systematic literature review," Appl. Sci. **11**, 2412 (2021).
[42] L. P. Naumann, S. Vazire, P. J. Rentfrow, and S. D. Gosling, "Personality judgments based on physical appearance," Pers. Soc. Psychol. Bull. **35**, 1661–1671 (2009).
[43] R. E. Nisbett and T. D. Wilson, "The halo effect: Evidence for unconscious alteration of judgments," J. Pers. Soc. Psychol. **35**, 250 (1977).
[44] K. L. Nowak and J. Fox, "Avatars and computer-mediated communication: A review of the definitions, uses, and effects of digital representations," Rev. Commun. Res. **6**, 30–53 (2018).
[45] N. N. Oosterhof and A. Todorov, "The functional basis of face evaluation," *Proc. National Academy of Sciences* (NAS, Washington, DC, 2008), 105, pp. 11087–11092.
[46] C. Pagliari, C. Burton, B. Mckinstry, A. Szentatotai, D. David, L. Ferrini, S. Albertini, J. C. Castro, S. Estevez, and M. Wolters, "Psychosocial implications of avatar use in supporting therapy for depression," Annu. Rev. Cybertherapy Telemedicine **2012**, 329–333 (2012).
[47] S. V. Paunonen, K. Ewan, J. Earthy, S. Lefave, and H. Goldberg, "Facial features as personality cues," J. Pers. **67**, 555–583 (1999).
[48] I. S. Penton-Voak, N. Pound, A. C. Little, and D. I. Perrett, "Personality judgments from natural and composite facial images: More evidence for a "kernel of truth" in social perception," Soc. Cognit. **24**, 607–640 (2006).
[49] J. C. Peterson, S. Uddenberg, T. L. Griffiths, A. Todorov, and J. W. Suchow, "Deep models of superficial face judgments," *Proc. National Academy of Sciences* (NAS, Washington, DC, 2022), 119, p. e2115228119.
[50] R. Petrican, A., Todorov, and C. Grady, "Personality at face value: Facial appearance predicts self and other personality judgments among strangers and spouses," J. Nonverbal Behav. **38**, 259–277 (2014).
[51] D. Premack and G. Woodruff, "Does the chimpanzee have a theory of mind?," Behav. Brain Sci. **1**, 515–526 (1978).
[52] R Core Team (2023) *R: A Language and Environment for Statistical Computing* (Version 4.3.2) https://www.r-project.org/.
[53] D. Ratajczyk, "Uncanny valley in video games: An overview," Homo Ludens **1**, 135–148 (2019).
[54] Reallusion (2022) *Character Creator 3 [Computer Software]* [Computer software] https://www.reallusion.com/character-creator/.
[55] U.-D. Reips, "Internet experiments: Methods, guidelines, metadata," in *Human Vision and Electronic Imaging XIV*, edited by B. E. Rogowitz and T. N. Pappas (SPIE, Bellingham, WA, 2009), Vol. 7240, pp. 53–61.
[56] U.-D. Reips, "Web-based research in psychology: A review," Z. für Psychol. **229**, 198–213 (2021).
[57] U.-D. Reips and J. H. Krantz, "Conducting true experiments on the Web," in *Advanced Methods for Conducting Online Behavioral Research*, edited by S. D. Gosling and J. A. Johnson (APA, Washington, DC, 2010), pp. 193–216.
[58] U.-D. Reips and C. Neuhaus, "WEXTOR: A web-based tool for generating and visualizing experimental designs and procedures," Behav. Res. Methods Instrum. Comput. **34**, 234–240 (2002).
[59] U.-D. Reips and Y. Shevchenko, "How to conduct online experiments [How-to Guide]," SAGE Res. Methods: Doing Research Online (2023).
[60] M. Rheu, Y. Jang, and W. Peng, "Enhancing healthy behaviors through virtual self: A systematic review of health interventions using avatars," Games Health J. **9**, 85–94 (2020).
[61] D. F. Sacco and M. Brown, "Preferences for facially communicated big five personality traits and their relation to self-reported big five personality," Pers. Individual Diff. **134**, 195–200 (2018).
[62] V. Schwind, K. Wolf, and N. Henze, "Avoiding the uncanny valley in virtual character design," Interactions **25**, 45–49 (2018).
[63] V. Schwind, K. Wolf, N. Henze, and O. Korn, "Determining the characteristics of preferred virtual faces using an avatar generator," *Proc. 2015 Annual Symposium on Computer-Human Interaction in Play* (ACM, New York, NY, 2015), pp. 221–230.
[64] Y. Shiban, I. Schelhorn, V. Jobst, A. Hörnlein, F. Puppe, P. Pauli, and A. Mühlberger, "The appearance effect: Influences of virtual agent features on performance and motivation," Comput. Hum. Behav. **49**, 5–11 (2015).
[65] Snap Inc. (2021) *Bitmoji (Version 10.31) [Mobile app]* [Computer software] https://www.bitmoji.com.
[66] M. Suda and M. Oka, "Evaluation of the effect of mimicry on facial expression in avatar-mediated communication," Artif. Life Conf. Proc. 33 **2021**, 48 (2021).
[67] J. P. Swaddle and G. W. Reierson, "Testosterone increases perceived dominance but not attractiveness in human males," *Proc. R. Soc. London Series B* (The Royal Society, London, 2002), Vol. 269, pp. 2285–2289.
[68] K. Terada, L. Jing, and S. Yamada, "Effects of agent appearance on customer buying motivations on online shopping sites," *Proc. 33rd Annual ACM Conf. Extended Abstracts on Human Factors in Computing Systems* (ACM, New York, NY, 2015), pp. 929–934.
[69] J. Thompson, S. White, and S. Chapman, "Interactive clinical avatar use in pharmacist preregistration training: Design and review," J. Med. Internet Res. **22**, e17146 (2020).
[70] Y. Wang, J., Geigel, and A. Herbert, "Reading personality: Avatar versus Human faces," *2013 Humaine Association Conf. on Affective Computing and Intelligent Interaction* (IEEE, Piscataway, NJ, 2013), pp. 479–484.
[71] S. M. Weinberg, Z. D. Raffensperger, M. J. Kesterke, C. L. Heike, M. L. Cunningham, J. T. Hecht, C. H. Kau, J. C. Murray, G. L. Wehby, L. M. Moreno, and M. L. Marazita, "The 3D facial norms database: Part 1. A web-based craniofacial anthropometric and image repository for the clinical and research community," The Cleft Palate-Craniofacial J. **53**, 185–197 (2016).
[72] W. D. Weisman and J. F. Peña, "Face the Uncanny: The effects of doppelganger talking head avatars on affect-based trust toward artificial intelligence technology are mediated by uncanny valley perceptions," Cyberpsychology Behav. Soc. Netw. **24**, 182–187 (2021).
[73] L. L. M. Welling, L. M. Debruine, A. C. Little, and B. C. Jones, "Extraversion predicts individual differences in women's face preferences," Pers. Individual Differ. **47**, 996–998 (2009).
[74] S. Windhager, K. Schaefer, and B. Fink, "Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity," Am. J. Hum. Biol. **23**, 805–814 (2011).
[75] L. A. Zebrowitz, "The origin of first impressions," J. Cultural Evolutionary Psychol. **2**, 93–108 (2004).
[76] L. A. Zebrowitz, L., Voinescu, and M. A. Collins, "'Wide-eyed' and 'crooked-faced': Determinants of perceived and real honesty across the life span," Pers. Soc. Psychol. Bull. **22**, 1258–1269 (1996).