Gender- and Age-related Differences in Designing the Characteristics of Stereotypical Virtual Faces

Valentin Schwind[1], **Niels Henze**[1,2]
[1]VIS, University of Stuttgart, Stuttgart, Germany, valentin.schwind@vis.uni-stuttgart.de
[2]University of Regensburg, Regensburg, Germany, niels.henze@ur.de

ABSTRACT
The characteristics of virtual faces can be important factors for avatars and characters in video games. Previous work investigates how users create their own avatars and determines the generally preferred characteristics of virtual faces. However, it is currently unknown how the preferred characteristics of avatar faces depend on the players' age and gender or if these demographics can be predicted based on the data provided by an avatar creation system. In this paper, we investigate the effects of gender and age on the facial characteristics of 4,215 virtual faces created by 1,475 participants (994 male, 481 female) mainly from Central Europe using a web-based avatar creation system and the Caucasian average face. Our results show that with increasing age, men and women design increasingly realistic and less stylized avatars. We also found that young persons design more androgynous avatars, while adults further increase the masculinity and femininity of their avatars. However, women with higher age decrease the femininity and increase the masculinity of stereotypical faces. Based on the data collected during the avatar creation process, we can predict the participants' gender with an accuracy up to 91%, which open up new use cases for video games and avatar creation systems. We discuss potential social, biological, and cognitive explanations for our results and contribute with design implications for games and future avatar customization systems.

CCS Concepts
•**Social and professional topics** → **User characteristics;** •**Human-centered computing** → *HCI design and evaluation methods;*

Author Keywords
Avatars; player demographics; avatar creation; character creation; player classification.

INTRODUCTION
Video games benefit from realistic and visually appealing virtual characters including the player's representation in the virtual world – the *avatar*. Games with human characters such as Lara Croft in the Tomb Raider series [13] or Nathan Drake in the Uncharted series [40] received mainly positive ratings for their outstanding character design and were very successful [24, 2, 4]. It is known that appealing human-like virtual characters are perceived to be more likeable and arousing, and people are more likely to choose to be represented by them [41]. This can partially be explained by research on physical attractiveness, which ascribed more favorable personality traits to humans with physically attractive faces [46, 15, 17]. Such traits may help people to project an idealized self and to have a higher attachment to their avatar [16].

One problem with the acceptance of virtual characters can occur at high levels of realism. The uncanny valley hinders virtual characters from being accepted when they almost look like real humans [39, 47]. Thus, realistic video games such as L.A. Noire [55] or Mass Effect: Andromeda [5] were criticized for the insufficient realism of their main characters and NPCs [54, 22]. Researchers use avatar creation systems to explore which design choices people make to avoid uncanny effects and which facial characteristics they prefer in virtual avatar faces [50, 1]. For example, it has been shown that people use very appealing features such as natural proportions, smooth skin, and youthfulness to prevent uncanny effects of too much realism in virtual faces [51].

For researchers, developers and designers of games, it is not only important to understand the effects of realism but also to learn how individual factors of the users affect the affinity towards game characters they create, like, and play. Particularly, stereotypes can leverage concepts a player already knows and suggest how to interact with them in games [24]. However, attitudes towards stereotypical faces such as heroes and villains are not well explored even when male and female character roles and behaviors are frequently stereotyped in games [18]. Offering the preferred character designs of the own avatar or for different stereotypes as well as customized options for players with the corresponding age and gender could significantly improve the player-avatar connection in games and lead to higher levels of immersion.

Furthermore, it is unknown whether the facial characteristics of an avatar and additional meta-data of its creation can be used to derive a profile of its creator and the stereotype of the avatar. Through determining the players' age, gender, or the kind of avatar, a number of games and further applications could benefit. Role-playing games (RPGs) or life simulations, for

CHI PLAY '18, October 28–31, 2018, Melbourne, VIC, Australia

DOI: https://doi.org/10.1145/3242671.3242692

example, could adapt their storiy lines and adventures based on the demographics of their players and on the appearance of the avatars to provide more immersive and individual experiences.

In this paper, we present the results of a study using the online avatar creation tool *faceMaker* by Schwind et al. [51, 50]. Starting from the Caucasian average face, 1,520 users (994 male, 481 female) distributed across three age groups created 4,215 virtual avatar faces. Analyzing the created faces, we found significant effects of age and gender on the facial characteristics of stereotypical avatar faces. Not only facial preferences depend on age and gender, but also the process how the different user groups create their avatars. Based on the meta-data of the face creation process, we are able to classify the gender of the participants with an accuracy up to 91%.

We discuss social, biological, and cognitive influences explaining these results. We contribute with our findings, the complete data set, design implications for future designs of virtual characters in games, and recommendations for games and further applications where avatar customization is present.

RELATED WORK

Previous work investigates the characteristics of physical attractiveness in faces and the cognitive processes involved in face perception. One cognitive process playing an important role in recognizing and processing faces is *perceptual narrowing* as highlighted in the following section. Furthermore, this work is related to previous research investigating gender- and age-related differences in avatar research and tools developed to understand players and the design process of avatars.

Virtual Face Perception

Nowak and Rauh [41] showed that physical attractiveness of video game characters has a significant impact on gaming. The authors found that people tend to select characters with clear masculinity or femininity, which is also apparent in character designs of current video games. John [26] found in a meta-analysis of game-design magazines that game designers both consciously as well as unconsciously create game characters with physical attractiveness to increase the impact of their game. This was confirmed by Miller [38] who found that male game characters are designed to be more muscular and powerful while female characters were often designed to be physically attractive, innocent, and sexy. Physical attractiveness of virtual faces in games is related to research in evolutionary aesthetics, which found that ratings of facial preferences are consistent between multiple studies. Features such as symmetry [42], sexual dimorphism [43], and averageness [29] are cross-culturally perceived as physically attractive [12]. Interestingly, perfect symmetry is not necessarily desired. Slightly fluctuating symmetries are perceived to be more attractive than faces with perfect symmetry [53].

Previous work suggests a number of factors influencing the perception of physical attractiveness of faces. One of these factors is the process of perceptual narrowing [12]. The process is a regressive cognitive development and contributes in critical ways to the human ability to recognize faces [31]. It has been shown, that infants lose the ability to discriminate faces of persons with other ethnicities and tune them to their familiar ethnicity. This effect is known as the “own-race bias” [28, 37]. The process tunes infants to recognize mainly adult faces [33]; however, it is known that perceptual narrowing does not end in early years. It has been shown, that the individual notion of a familiar and the average face depends on the population a person has been exposed to over the whole lifetime [37].

Researchers assume that perceptual narrowing could also be responsible for the uncanny valley effect [10, 34, 39]. The effect describes a sudden dip of affinity towards very human-like entities while less or full human-like characters are generally accepted. Perceptual narrowing was brought into connection with the uncanny valley as it enhances the ability to discriminate faces. Thus, abnormalities in virtual renderings seem to be more apparent [10] and inconsistencies in a photo-realistic human face can cause an entity to appear unfamiliar. As some features of the virtual face (*e.g.*, eyes, skin, or hair) are more difficult to create than other facial parts computer-graphics tend to be inherently inconsistent in their realism [10].

Tools to create and customize avatars can help to determine which design choices people make to avoid uncanny effects and which facial characteristics they actually prefer. For example, in an online study using the avatar creation tool *faceMaker*, Schwind et al. [51] found that people do not tend to stylize faces. Instead, they used features such as smooth skin, healthy skin color, ideal proportions, and clear gender cues to create appealing faces. These findings are supported by Nowak and Rauh [41] who found that people are more likely to perceive an avatar as credible when it is anthropomorphic (human-like) and that anthropomorphism as well as lower androgyny of avatars increase physical attractiveness, credibility, and homophily. Chung et al. [11] suggest that the process of avatar creation leads to a stronger sense of self-presence and more identification with the created character. This is supported by findings of Ducheneaut et al. [16], who particularly highlight the importance of changing body shapes, hair styles, and hair colors in avatar creation and identification.

Gender-and Age-related Differences

Prior research showed how avatars foster intrinsic motivation in games [6]; however, a number of age- and gender-related differences seem to exist [41]. In particular, the sexualization of primarily female game characters by game designers is a highly discussed topic in video gaming research [3, 25, 38]. Schwind et al. [49] showed that there are gender-related differences in presence perception of hand avatars in virtual reality. However, not only gaming with avatars but also the creation process of avatars seems to be dependent on gender and age. In a long-term study by Heeter et al. [21], the authors observed teens designing games and describe that girls desire a very high level of avatar customizations while boys preferred to use predefined characters.

John [26] observed a developer team of computer games and found that they had no clear target age in mind for the player. Thus, they tend to design sexualized stereotypes for games assuming that they will be accepted. According to Carrasco [9], avatar studies have typically focused on younger users rather than on older adults. Their findings from an avatar study with older adults show that they prefer aged avatars. The authors also highlight that older adults have specific representational requirements when designing virtual avatars [9]. This is supported by Rice et al. [45] who showed that older adults seek more attractive and expressive features in avatars and have similar interests in customization features of an avatar creation system than younger players.

Summary

Previous work showed that gender and age are important factors in perceiving features of a virtual face [21, 30, 52, 45]. Many insights about avatars are already known through research in attractiveness and from evolutionary aesthetics [43, 42, 29, 12], but virtual avatars do not necessarily look photorealistic or human-like [10, 27, 38]. Research found that avatars are designed to be realistic, attractive and appealing [51], however, players also customize their characters and swap their gender, for example, to gain benefits within the game [23, 52, 35]. As the games compromise the way how people customize their avatar, this research uses a neutral non-gaming environment [50] to investigate how gender and age influence the preferred characteristics of stereotypical avatars. Furthermore, this research aims to investigate if the facial characteristics or additional metadata during the creation process while using the avatar generation system can classify stereotypical faces and the demographics of the player. This knowledge can be crucial for creating dynamically adapting games and providing more immersive experiences for players.

STUDY

The aim of this study is to investigate the effects of age and gender on the preferred characteristics of virtual stereotypical avatar faces. We also investigate if avatar-related factors exist that can predict age or gender of the player. We mainly focus on the realism of the avatar as well as on features that contribute to the affinity towards the avatar faces. Thus, age group and gender are between-subject variables in a multi-factorial mixed-design. Our work is based on web-based avatar generator *faceMaker*, which we already used and validated in prior research [51, 50] and is available at github[1]. The tool is based on the female and male Caucasian average face having the least visual distance to any other faces of this population. The system uses 37 different parameters to change the facial characteristics based on six different tasks [51]. The tasks were introduced to create stereotypical faces of an arbitrary, repulsive, heroine, hero, female villain, and male villain avatar.

Apparatus

We hosted the *faceMaker* system (Figure 1) on our website[2]. The system uses the open Javascript library three.js[3] to render a 3D avatar face based on the male and female Caucasian average face, which was computed by photos of 117 male and 151 female faces. We used the model of *faceMaker* as the provided Caucasian average face belongs to our target group, often appear in recent video games, and is well known from gaming or attractiveness research. The average face has

Figure 1. Screenshot of the WebGL-application and avatar generator *faceMaker* [51].

no individual characteristics. Users can seamlessly blend between different facial characteristics, and can control single parameters without compensating unwanted features. With *faceMaker* the participants can change the average face and manipulate its characteristics using 5 general (avatar gender, style, details, color, hair color) and 32 face specific features (such as mouth size, nose length, eye distance, etc.). A complete list of all parameter states and morphings used in the *faceMaker* application is shown in Table 1.

Each parameter is represented by a slider grouped into a moveable user interface (UI) group box equipped with a help button for further information. UI boxes and slider order within a group are randomized for each participant. Mouse movements above each slider moves the camera to the corresponding facial feature. For example, moving the mouse cursor to the nose length slider rotates the camera to the profile of the face. To the right of each slider is a reset button, where the user could set the slider to the default value. A complete description of the implementation can be found in the original article in our previous work [50]. To avoid biases, no clothes, additional hair styles, scars, tattoos, or equipment were used.

Measures

We measured the 37 parameter values of each avatar face and additional metadata during the creation. Metadata include task completion time, mouse clicks, mouse movements, and resets of the sliders. Mouse clicks indicate if the parameter of a slider was changed. Mouse movements indicate how often the mouse was above a slider. Resets were counted to measure how often a value was set to the default value. Before submitting a face, we asked participants to assess their confidence in succeeding in the task (very successful/very unsuccessful) as well as the likeability (very likable/very unlikeable), attractiveness (very attractive/very unattractive), and gender affiliation (very masculine/very feminine) of the created avatar using 7-point semantic differentials. The success question was used to ensure that users completed the tasks correctly. We asked for likeability and attractiveness to operationalize the preferred avatar design. We used gender affiliation as the control variable for gender manipulation.

[1] https://github.com/valentin-schwind/facemaker
[2] http://facemaker.uvrg.org
[3] https://threejs.org

parameter	−	values	+	type
face gender	female	androgynous*	male	t m
face style	realistic*		cartoon	m
face details	none	half*	full	t
skin color	black	average*	white	t
hair color	black	brun. med.blonde* red	bright blonde	t
eyes color	black	brown* amber blue lt. blue green	grey	t
eyes shape	droopy	down round oval* almond up	asian	m
eyes opening	narrow	average*	wide	m
eyes size	small	average*	big	m
eyes height	up	average*	down	m
eyes distance	narrow	average*	wide	m
eyes orbit	bulgy	average*	covernous	m
eyes rotation	in	average*	out	m
eyebrows color	black	brun. med.blonde* red	bright blonde	t
eyebrows shape	pointedstraight	average*	roundhooked	m
eyebrows strength	thin	average*	thick	t
nose shape	snub	average*	hooked	m
nose length	short	average*	long	m
nose width	thin	average*	thick	m
nose bridge	thin	average*	thick	m
nose cartilage	round	average*	flat	m
forehead size	down	average*	up	m
ear size	small	average*	big	m
throat size	thin	average*	thick	m
jaw shape	triangle	average*	squared	m
jaw length	long	average*	short	m
chin shape	pointed	average*	cleft	m
cheeks shape	full	average*	scraggy	m
lips volume	thin	average*	full	m
lips size ratio	upper lip	average*	lower lip	m
mouth shape	down	average*	up	m
mouth width	wide	average*	narrow	m
mouth height	up	average*	down	m
mouth depth	backwards	average*	forwards	m
eyes shadow	none*		full	t
lipstick	none*		full	t
rouge	none*		full	t

* default, blending types: t = texture blending, m = mesh morphing

Table 1. Overview about the 37 parameter scales. The parameters of the interfaces were grouped into the groups: common, eyes, eyebrows, nose, outer face, cheeks/jaw, mouth, and make-up.

Tasks
Our study was designed to understand how men and women of three different age groups design different stereotypical avatars. In line with our previous work related to this study [51, 50], we asked participants to create six different faces: An arbitrary avatar face, an uncanny, repulsive face, a heroine, a hero, a female villain and a male villain.

Procedure
An online session started with demographic questions about gender, age, game and movie consumption. Minors were advised to receive the informed consent of their parents or legal guardians before using the application. After confirming the terms of use, the application showed the instructions and tasks. All geometries and texture blendings start from the androgynous middle between the male and female average face with all default values as shown in Table 1. Before submitting an avatar, participants had to answer the four questions about the created face (see Measures). After submitting a face, the application started with a new task until all of them were completed. The tasks were presented in a balanced 6×6 Latin square order, starting with the first task containing the fewest data entries in our database. Each task could only be processed once. Session timer was paused when a participant closed the browser tab. A participant could continue the study within seven days. Participants were able to download their avatars as images or 3D models after completing all tasks.

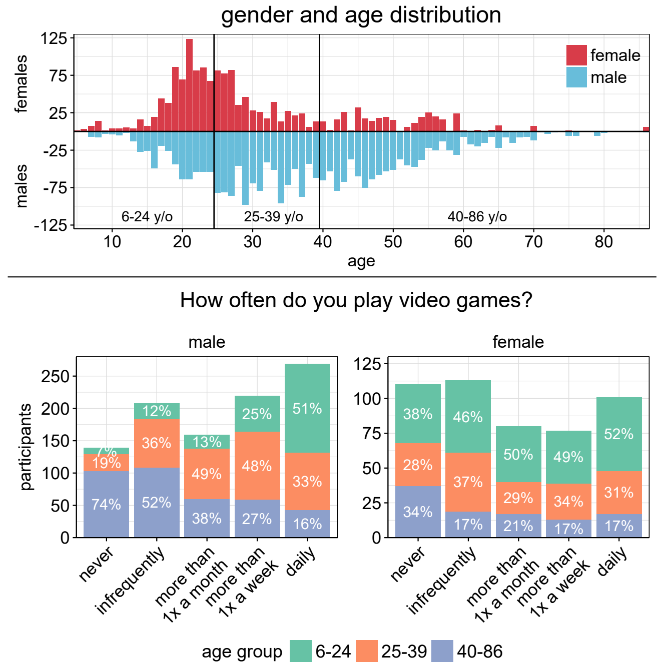

Figure 2. Age histogram (top) and statements about video game consumption (bottom) of male and female participants.

Participants
In total, 1,519 participants (994 male, 481 female, 44 other/not specified) took part in our study. They were recruited via social media, mailing lists, advertisements, and word of mouth. To perform further statistical analyses we removed the 44 other/not specified participants. The remaining 1,475 participants were between 6 and 86 years old ($M = 33.8$, $SD = 15.3$). Three age groups were assigned based on the tertile of the participants' age (see Figure 2): 6-24 year old (y/o) (472 participants), 25-39 y/o (526 participants), and 40-86 y/o (477 participants). An overview about the statements how often the participants play video games can be found in Figure 2. The majority, 940 participants, stated that their origin country is Germany, 59 participants came from the U.S., 57 from Poland, 31 from Austria, 29 from the U.K., 23 from Italy, 22 from Spain, 20 from Canada, 17 from Russia, 12 from France, and further 277 participants from Australia and further 84 countries in Europe, America, and Asia. The study received ethics clearance according to the ethics and privacy regulations of our institution.

RESULTS
We received a total of 4840 avatar faces. We removed 625 faces due to the following issues: cookies were deleted during the browser session, no parameters were changed during the session, session time exceeded 6 hours, when the browser did not support the English user interface of *faceMaker*, too low screen resolutions (<1280p). For statistical evaluation, we used 4,215 avatar faces. Based on our research questions, the analyses of our results highlight the following points: face realism, gender affiliation, facial preferences, parameter usage, and classification.

Realism

We used the parameters *face style* and *skin details* to determine the realism of the avatar faces. Increased *face style* morphed the model to a stylized, cartoon face, *skin details* increases the visibility of a photo-realistic skin texture. We would like to point out that other over-exaggerated atypical features deviating from the human norm may also cause to appear a face less realistic, which is not considered in this part of the analysis.

For each measure, we conducted a three-way analysis of variance (ANOVA). For *skin details* there were significant main effects for GENDER, $F(1,4179) = 31.510$, $p < .001$, AGE, $F(2,4179) = 17.471$, $p < .001$, FACE $\times$ GENDER $F(5,4143) = 4.959$, $p < .001$, FACE $\times$ AGE $F(10,4143) = 1.898$, $p = .040$, GENDER $\times$ AGE $F(2,4143) = 3.005$, $p = .049$. There was no three-way interaction effect ($p = 0.715$). For testing the hypothesis that more skin details were added with increased age, we conducted Bonferroni-corrected pairwise t-tests between the average values of each age group (gender and stereotype were aggregated). Post-hoc tests revealed significant differences of participants between 6-24 y/o ($M = .475$, $SD = .304$) and 25-39 y/o ($M = .526$, $SD = .315$), $t(2625.947) = 4.283$, $p < 0.001$, between 6-24 y/o and 40-86 y/o ($M = .557$, $SD = .307$), $t(2559.241) = 6.831$, $p < 0.001$, and between 25-39 y/o and 40-86 y/o, $t(2999.272) = 2.772$. $p = 0.02$. Thus, significantly more skin details were used when participants were older and significantly more skin details were used by male ($M = .541$, $SD = .306$) than by female participants ($M = .489$, $SD = .316$).

For *face style* there were no significant main effects for GENDER, $F(1,4179) = 2.263$, $p = .133$ but for AGE, $F(2,4179) = 21.967$, $p < .001$, and FACE $\times$ GENDER, $F(5,4143) = 2.753$, $p = .017$. There were no further interactions effects (all with $p > .130$). Post-hoc tests revealed no significant differences of participants between 6-24 y/o ($M = .100$, $SD = .247$) and 25-39 y/o ($M = .082$, $SD = .222$), $t(2433.582) = -2.061$, $p = .118$; however, between 6-24 y/o and 40-86 y/o ($M = .052$, $SD = .174$), $t(2104.844) = -5.702$, $p < .001$, and between 25-39 y/o and 40-86 y/o, $t(2948.262) = -4.095$. $p < .001$. Thus, the face style parameter significantly decreased with age, especially between 6-24 y/o and 25-39 y/o participants.

We summarize that both *skin details* and *face style* operationalized to determine facial realism were significantly affected by gender and age. While participants increased the photo-realistic textural details on their avatars with increased age, the stylized geometrical cartoon shape was decreased.

Gender Affiliation

We used the parameter *face gender* to determine if an avatar face was designed rather feminine or masculine. We conducted a three-way ANOVA to determine the effects of AGE, GENDER, and STEREOTYPE. We found significant effects for AGE, $F(1,4179) = 2.678$, $p < .001$, GENDER, $F(1,4179) = 35.921$, $p < .001$, STEREOTYPE, $F(5,4179) = 906.276$, $p < .001$, AGE $\times$ STEREOTYPE, $F(10,4179) = 3.848$, $p < .001$, GENDER $\times$ STEREOTYPE, $F(5,4179) = 5.124$, $p < .001$, AGE $\times$ GENDER, $F(2,4179) = 9.799$, $p < .001$. There was no significant three-way interaction effect for AGE $\times$ GENDER $\times$ STEREOTYPE, $F(10,4179) = 0.950$, $p < .485$.

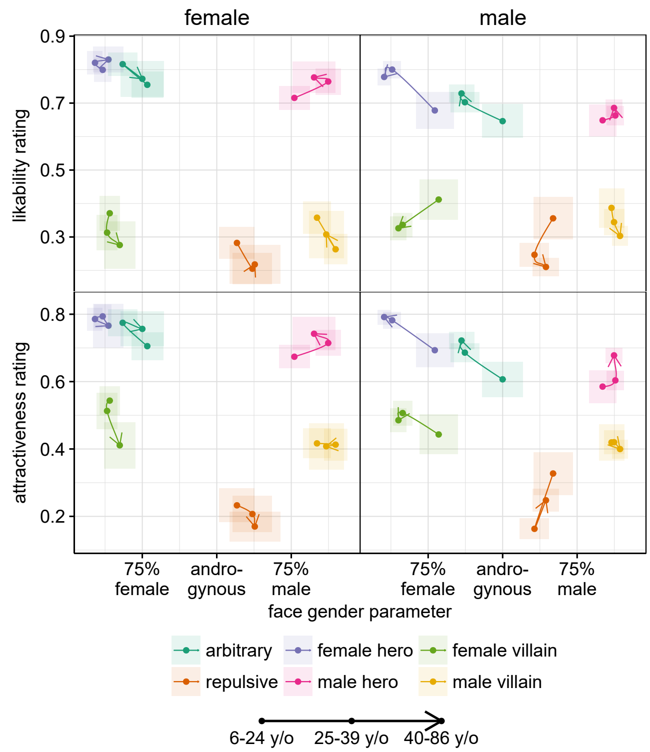

Figure 3. The *faceGender* parameter related to the perceived likability and attractiveness of the stereotypical faces. Markers along the arrow segments show the mean rating among the three age groups. Rectangles show CI95 in each direction.

Mean comparisons reveal that the average *face gender* values tend to be more androgynous when the face has the opposite gender of the participants (see Figure 4). We used Bonferroni-corrected pairwise comparisons to determine between which age groups and face types significant differences of the parameter *face gender* exist. We found no significant differences between the male age groups and the mean values of the male hero or the male villain. This was also the case for the female age groups and the mean values of the female hero and the female villain. However, we found significant differences of the parameter *face gender* between the male 6-24 y/o age group and the mean values of 25-39 y/o, $t(165.622) = 1.319$, $p = .002$, and 40-86 y/o, $t(160.378) = 1.193$, $p < .001$ for the female hero and between 6-24 y/o age group and 25-39 y/o, $t(129.054) = 1.319$, $p = .017$, or 40-86 y/o, $t(341.120) = 1.193$, , $p = .004$, for the female villain. Furthermore, there were significant differences between the female 6-24 y/o age group and mean value of the 25-39 y/o group for the male hero, $t(200.513) = 2.801$, $p = .017$. All mean values are shown in the face gender row in Figure 4.

We summarize that avatar faces created by both younger male as well as female age groups show significantly lower manifestations of the participants' opposite gender than older age groups. Furthermore, previous work showed that women tend to swap the gender with their avatars [23]. Therefore, we were interested in the difference between the means of the *face gender* parameters of the arbitrary face. The overall differ-

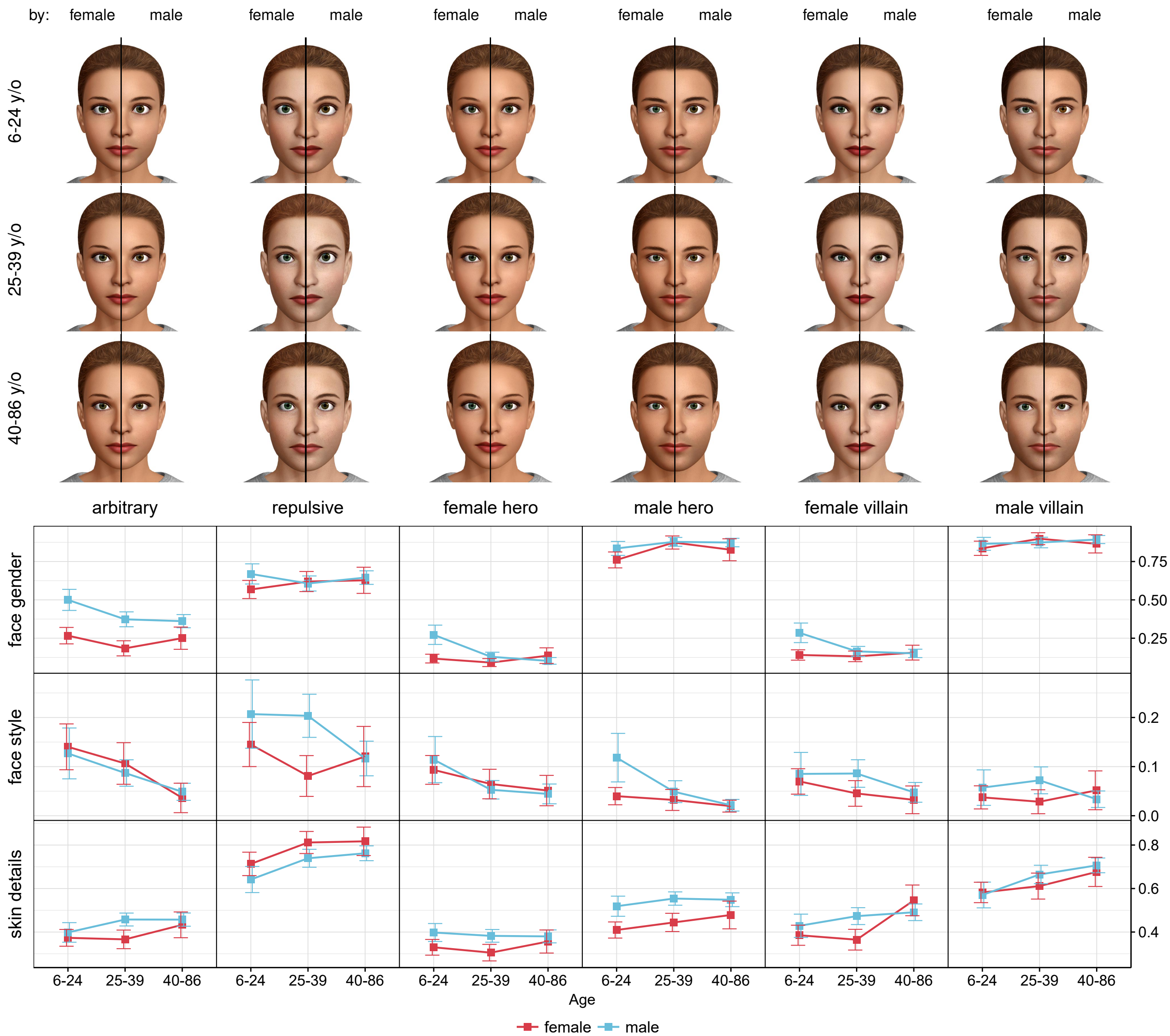

Figure 4. Average faces of the six stereotypes created by women (left image sides) and men (right image sides) and the different age groups. Below are the parameter plots of face style, and skin details. Error bars of the plot show 95% confidence interval (CI95)

ence between both gender was significant, $t(593.924) = 5.983$, $p < .001$. For male participants there were significant differences of the *face gender* parameter between 6-24 y/o and 25-39 y/o, $t(176.246) = -2.496$, $p = .040$, between 6-24 y/o and 40-86 y/o, $t(159.179) = -2.827$, $p = .016$, but not between 25-39 y/o and 40-86 y/o, $t(350.952) = -.287$, $p = 1.00$. There were no significant differences of the means of the *face gender* parameter between the female age groups (all $p > .180$). Thus, female participants prefere more feminine faces than men.

We asked for the subjective gender affiliation of an avatar face to determine which facial characteristic contributes to the femininity or masculinity of a face. Thus, a linear regression model was performed, $R^2 = .727$, $R^2_{Adj.} = .596$, $SE = .210$, $F(1367, 2847) = 5.554$, $p < .001$, $d = 1.980$, which showed significant effects. No auto correlations were found. Visual cues provided by the scatterplots of the standardized residuals (not illustrated) indicate that the data met the assumptions of homogeneity of variance, linearity, and homoscedasticity for the regression. Facial manipulation *face gender* showed the highest parameter coefficients among all parameters ($\beta = .579$, $p < .001$). Further significant coefficients were *throat size* ($\beta = .095$, $p < .001$), *skin details* ($\beta = .095$, $p < .001$), *eye rotation* ($\beta = .041$, $p < .001$), and *jaw shape* ($\beta = .041$, $p < .001$). We conclude that the *face gender* parameter was the most significant parameter contributing to the masculinity or femininity of a face.

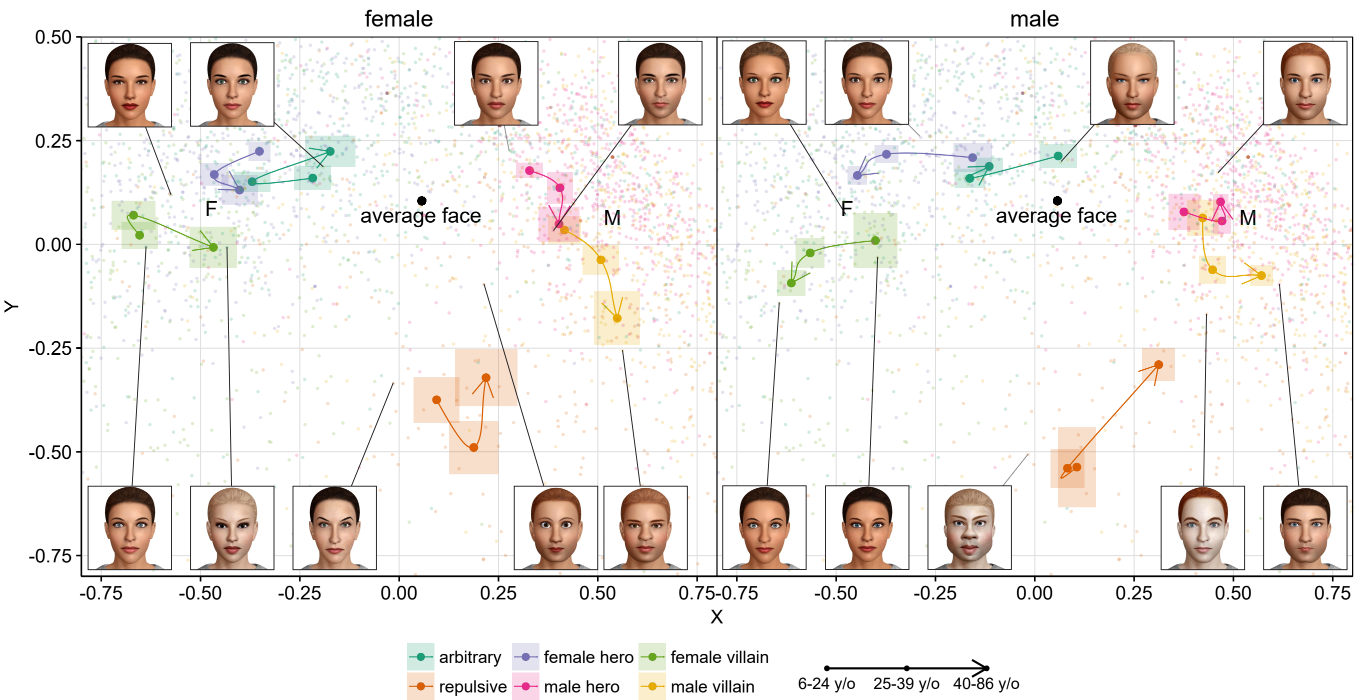

Figure 5. MDS mapping of the avatar faces among all stereotypical faces. The center positions of the face clusters change between gender and ages. The average face is located close to the zero point of the map. The horizontal axis correspond to the gender affiliation of the stereotypes. Cluster centers of male faces (M) can be found right to average, female faces (F) are left to the average. Points on the arrow segments show the cluster centers among the three age groups. Rectangles show CI95. Face renderings show face samples from the corresponding areas.

Stereotypes

Our research aimed to determine how facial characteristics of the stereotypes differ between gender and age. Due to the high number of different facial characteristics, we reduced the number of dimensions and visualized their similarities using visual mapping. First, we conducted a three-way multivariate analysis of variance (MANOVA) to determine if the 37 latent parameters were significantly affected by STEREOTYPE, AGE, and GENDER. Main effects in all separate univariate ANOVAs showed significant effects on all 37 facial characteristics (all with $p < .05$). Thus, all facial parameters were included in dimension reduction. Second, we calculated the euclidean distances between the parameters using the R cmdscale function[4]. Again, we conducted a three-way MANOVA and subsequent univariate ANOVAs on the two dimensions from MDS, which were significant for all main and interaction effects between the independent variables (all with $p < .05$), except for y-values with GENDER, STEREOTYPE×GENDER, AGE×GENDER, and STEREOTYPE×GENDER×AGE. Third, we used the euclidean distances calculated by the MDS algorithm to visualize the faces using a spatial 2D mapping (stress value = .154). We determined two clusters containing male and female faces using hierarchical clustering (HC) and visualized their centers ("M" and "F" in Figure 5). By averaging the positions of the faces designed by male, female, and participants of three different age groups we determined the centers of the corresponding groups on the map. For the sake of clarity, we split the map into a male and female panel for the gender of the participants and connected the average stereotypical faces the age groups with arrow paths as shown in Figure 5.

Visual inspection revealed a filament structure in the upper part of the map where male and female stereotypical faces shape two corresponding main clusters. Interestingly, the MDS algorithm constructed a similar arrangement of averaged faces as in the plot in Figure 3. Male faces are right of the average faces and located at positive x-values, female faces are left of the average and have rather negative x-values on the map. Values on the y-axis correspond with likeability and attractiveness measures of the virtual faces. Stereotypes with lower y-values received more skin details, had an increased skin brightness (*cf.* Figure 4), and more or exaggerated characteristics deviating from the human average. Furthermore, there is an offset between heroes and villains and a similar development of their characteristics between the age groups.

As shown in Figure 5, the arbitrary faces are close to the locations of the female heroes, which indicates that the participants tend to design female faces with physical attractiveness when they had the free design choice. As the arbitrary faces also contains male ones, it is closer to the average face than the other face groups. Design choices of arbitrary and female hero faces follow a similar pattern and development with increased age of the participants. Similar to Figure 4, abnormal or deviating faces from the repulsive face category can be found at the bottom of the MDS map. Participants use multiple and very atypical features, which causes a larger distance from

[4]MDS cmdscale for R: https://stat.ethz.ch/R-manual/R-devel/library/stats/html/cmdscale.html

the human average. Older participants tend to create more realistic and less deviating repulsive faces. Thus, their faces were more masculine and villain-like, which gives them a more reasonable and realistic appearance.

The aggregated dimensions of the stereotypes designed by females showed highest feminine traits when they were created by 25-39 y/o participants. Thus, female participants of the 6-24 and 40-86 y/o age groups design heroes and villains of their own gender less feminine than 25-39 y/o participants. This was not the case for male participants either for their designs of female or for male stereotypes. Furthermore, masculinity of male heroes and villains generally increases with the age of the female participants. This is in contrast to stereotypes designed by the male participants, who designed female heroes and villains more and more feminine when they were older.

We summarize that our participants differentiate facial preferences between stereotypical faces. The effect of stereotypes exist after a reduction from 37 to 2 dimensions, which could be visualized using MDS. The MDS map showed a similar arrangement as the relative arrangement of gender and preference as shown in Figure 3. By comparing the average positions between the faces, we found that female participants designed all stereotypical faces more masculine. Male participants as they get older design their faces less average, except the repulsive faces, which were designed to be more realistic.

Parameter Usage

In terms of avatar and face customization research, it is also crucial to learn how much effort male, female, and participants within the different age groups put on creating their faces. We conducted a three-way ANOVA on the log-transformed time participants needed to create a face. There was a significant effect of GENDER, $F(1,3987) = 37.519$, $p < .001$, AGE, $F(2,3987) = 18.998$, $p < .001$, and a significant interaction effect between GENDER $\times$ AGE, $F(10,3987) = 5.765$, $p = .003$. There were no main or further interaction effects with STEREOTYPE (all with $p > .607$).

Consequently, with an increased TCT the average number of clicks is also higher for female than for male participants. We found significant effects of GENDER, $F(1,4179) = 99.888$, $p < .001$, AGE, $F(2,4179) = 7.681$, $p < .001$, without further interaction effects. Figure 6 shows significant differences of the average number of mouse clicks and the task completion time (TCT). We summarize that female and older participants spent more time and effort into creating their avatars. The stereotype of the face had no effect on the TCT, which means that there were no significant differences between the average times participants needed to create the six stereotypical faces.

We also examined the differences of the parameter usage between male and female participants. An overview about the parameter changes as indicator of their facial attention is shown in Figure 7. The average number of parameter manipulations showed a high usage of the face gender parameter, which can be explained by our tasks. For all parameters, female performed more mouse clicks than male participants participants (see vertical lines in Figure 7). Additionally, the *hair color* parameter was often manipulated by male and females

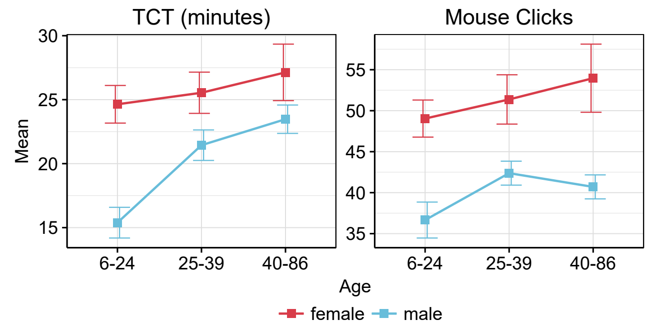

Figure 6. Task completion time (TCT) and number of all mouse clicks of male and female participants between the different age groups. Error bars show CI95.

participants. Female participants were further interested in the brightness of the skin, eye colors, eyebrows, cheeks, and lips volume. Men were interested in manipulating cheeks, nose shape, eye depth, eyebrows line, skin details, and the brightness of the skin.

Classification

For games providing avatar customization, useful mechanisms or scenarios exist in which determining the players' age, gender, and the facial stereotype of the avatar can be an important and helpful feature. However, it is currently unknown if such classifications can be performed only based on the facial characteristics of an avatar as well as on the metadata of the parameter usages (duration, clicks, resets, mouse moves) during the creation process. Classification can also provide further insights into how similar facial stereotypes are and how they were created by the participants. Therefore, we conducted a classification analysis with threefold benefits: It (1) supports the overall validation of the data set, (2) aims to classify the membership of faces and participants to their corresponding groups, and (3) visualizes the similarities of the classified faces according to gender and age.

We used support vector machine (SVM) classification, which is a reliable and robust probabilistic non-linear classification method for high dimensional data using supervised learning. We used multiple C-SVM with radial kernels. The best hyperparameters were determined using grid search and a 5-fold cross-validation. To avoid biases and overfitting of the SVMs, the data set was split into balanced training sets with an equal number of samples. Remaining samples were used for validation. The accuracy of the classifier was defined as the probability of correctly assigned faces in percent.

Age Classification

For age classification the data was split into balanced training samples to determine the accuracy of age classification (1,000 for each group). Remaining samples were used for testing (6-24 y/o: 201, 25-39 y/o: 577, 40-86 y/o: 437). Using the facial characteristics only, the age groups of the participants could be classified with an accuracy of 40.3%, ($\kappa = .112$, McNemar's $p < .001$). By including additional metadata we could increase the accuracy up to 44.2%, ($\kappa = .171$, McNemar's $p < .001$) for age classification.

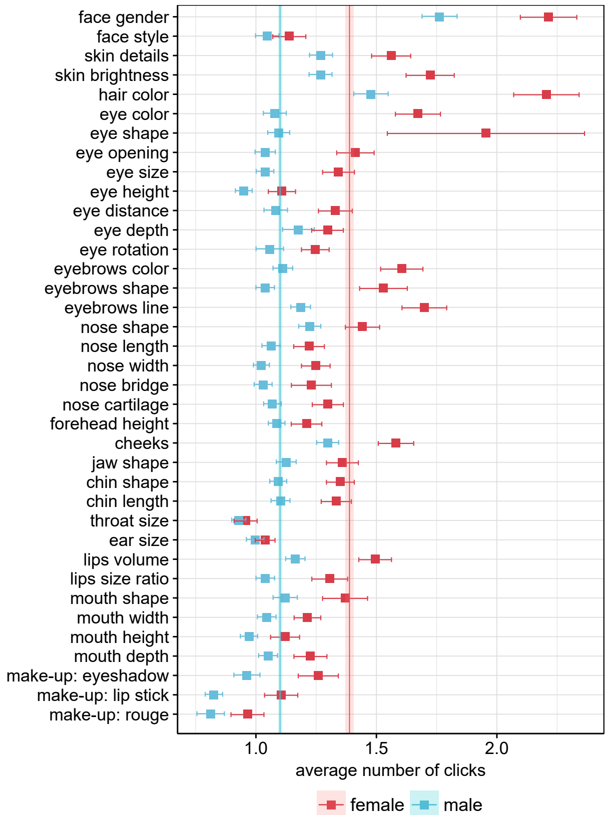

Figure 7. Average number of parameter manipulations by male and female participants. Error bars show CI95.

Gender Classification

For gender classification the data was split into a balanced training set including an equal number of male and female participants ($N = 1,300$ for each group). Remaining samples were used to determine the accuracy of gender classification (male: 1,377, female 238). Using the facial characteristics only, the gender of the participants could be classified with an accuracy of 63.0%, ($\kappa = .122$, McNemar's $p < .001$). By including the metadata of the avatar creation the accuracy of gender classification increased to 91.1%, ($\kappa = .692$, McNemar's $p < .001$).

Face Classification

To determine the accuracy of classification for an avatars' stereotype, data was split into 90% for training (ranging from 548 to 572 samples per category) and 10% for validation (ranging from 130 to 154 samples per category). Using the facial characteristics only, the created faces could be classified with an accuracy of 62.0%, ($\kappa = .544$, McNemar's $p < .001$). By including the metadata of the avatar creation the classification accuracy decreases to 54.9%, ($\kappa = .182$, McNemar's $p < .001$).

We summarize that the highest classification accuracy could be achieved for participants' gender. Using facial characteristics and additional metadata, the prediction was correct in 91.1% of cases. Classification accuracy for facial stereotypes showed the highest accuracy when only facial characteristics were used.

DISCUSSION

In this work, we investigate the effects of age and gender on the preferred characteristics of virtual avatar faces. The browser-based avatar creation system *faceMaker* provides a wide range of facial parameters which were used to determine the preferred characteristics of six stereotypical faces. First, we found that with increased age participants increased the realism of their avatar. Second, we found that male and female participants between 6-25 y/o designed more androgynous faces with less feminine or masculine features than the 25-39 y/o age group. Male participants in the 40-86 y/o age group further increased masculinity and femininity of their avatars, while 40-86 y/o females designed more masculine faces. Third, we can confirm previous studies [21, 51] showing that female participants put more effort into the creation of their avatars. Finally, we used classification to investigate if the age or gender of a participant could be predicted using the facial characteristics and the parameter usage during the creation process. We show that using SVM classification, participants' gender can be predicted with an accuracy of 91%, age with 44.2%, and avatar stereotypes with 62.0%.

Face realism increased with the age of the participants. Assuming the participants do not perceive realism differently, the results could potentially be explained by the process of perceptual narrowing. People are generally exposed to faces within their own age groups including themselves. Thus, reduced skin details and increased stylization were used by young participants, as children and adolescents have bigger eyes and smoother skin related to their head size. Older participants used the skin detail parameters to add more skin structures building their avatars more mature and similar to the faces to which they are typically exposed to. They used less stylization as they are potentially less in contact with puppets, toys, cartoons, or comics. As perceptual narrowing does not stop in early years [37], it is conceivable that more face details and less stylization are caused by the notion of human faces attuned to age. This could also explain findings by Carrasco et al. [9] showing that older adults rather identify more with aged avatars. As sensitivity in face discrimination due to perceptual narrowing [12, 37] could also explain the effect of the uncanny valley [39, 34] or why it is more difficult to create realistic looking virtual characters than objects or scenes [10], we assume that perceptual narrowing potentially plays a fundamental role in the perception and identification of faces of virtual avatars.

Investigations of the virtual avatars' gender revealed that younger participants created more androgynous faces of the opposite gender. Older participants enhanced cues of the opposite gender and increased the physical attractiveness of their faces. This could be explained by sexual development and increased expectations in femininity or masculinity of potential mates [48, 8, 14] and also support findings by Rice et al [45]. However, we found that female avatars created by 40-86 y/o female participants are less feminine and resemble the average face more. This "U-turn" in terms of femininity was not the case with avatars created by men who generally increased gender cues of their avatars. One possible explanation is that femininity becomes less relevant in their biological context [7, 32] and is, thus, not further exaggerated. It is also conceivable that this is caused by the process of perceptual narrowing and the exposure of females to their own age group while perceiving a decrease in exaggerated femininity. The results could also be connected to findings from psychology showing that women are not necessarily attracted to people who create physiological arousal in them [20]. It is conceivable that our subjects did not create avatars they would actually be attracted to and this should be the subject of future work.

Furthermore, we analysed how the male and female participants design faces when they had the free design choice without further social or competitive influences of game environments (*e.g.*, through gender swapping "to gain benefits from the other gender" [52]). We found that female participants prefer to design avatars of their own gender than male participants, which support findings by Nowak and Rauh [41]. The reason for this is not entirely clear. As women possess an advantage over men in emotional functioning to promote attachment [44], we assume that in order to promote attachment and identification with their own avatar, they tend to design the face according to their own gender representing a more idealized version of themselves (*cf.* Nowak and Rauh [41] und Veltri et al. [56]).

Characteristics of avatars not only depend on age and gender, but also on the desired stereotype of the avatar. We found an offset of appealing characteristics between heroes and villains but a similar development with increased age of the participants. Most feminine faces of villains, heroes, and arbitrary faces were designed by the 25-39 y/o age group, who gave their avatars strong feminine traits maximizing their perceived likeability and physical attractiveness. With increased age, male participants designed more prominent facial features with clear gender cues for their stereotypes, except for the repulsive faces, which were designed less exaggerated and more reasonable when the participants were older. We conclude that stereotypical avatars are designed to match individual concepts and that these concepts evolve with the participants' age.

Analyses of the parameter usage showed that female participants put more time and effort in face customization of their avatars than male ones. This is confirmed by previous work, showing that they they are more concerned about their virtual appearance [56, 19]. Considering such metadata also enables us to predict the participants' gender with an accuracy of 91.1% by using the SVM classification method. Age (44.2%) and stereotype (54.9%) of a face could be predicted correctly with less probability. However, using classification to determine additional demographics of the users enables new use cases in RPGs, for example, and a wide range of further applications supporting avatar customization. Based on the demographics of the players or the appearance of their avatars, games can adapt their stories and content in order to provide an individual experience.

DESIGN IMPLICATIONS

Implications for designers and developers who aim to consider the demographics of their users is two-fold: (1) We strongly recommend considering the target groups to increase users' affinity for their virtual avatar. Men and women have different preferences for their avatars and change over time. Younger participants prefer less realistic and more stylized characters, while older target groups prefer faces with more details and less stylization. Younger participants desired more androgynous avatars with less masculine or feminine features than older age groups. While adult males increase the masculinity and femininity of their avatars, older female participants designed their avatars more masculine. This could have far reaching consequences for the design of the players' character and the design of games. (2) We recommend avatar customization with a wide variety of parameters and additional metadata when the participants' gender, age, or stereotypical face should be determined. In particular, gender can be accurately determined using SVM classification. Additional stereotypical features (tattoos, piercings, hairstyles, etc.) could potentially further increase the classification accuracy. Such classification could be used to design adapted story plots in RPGs or in life simulations such as *The Sims* series [36].

FUTURE WORK

We contribute with the full data set of our study[5], which enables researchers to further explore the characteristics of different avatars. We recommend exploring additional classification algorithms such as neural networks and clustering to predict the appearance of the avatar and demographics of the user. In line with Schwind et al. [51], we recommend comparing the appearance of the avatar with the users themselves. The *faceMaker* is an interactive system, however, no further interaction with the characters is foreseen, which could be investigated by future studies. Furthermore, the rendered face in *faceMaker* can be blended between the male and female average face, however, not between different ages. As participants used the skin detail parameter to equip their avatars with more skin structures, we assume that older participants potentially increase the age of their avatars systematically. This could be confirmed by a further study.

Further research could also investigate why it is more likely that women and men design female avatar faces. Future investigations could also investigate a potential relation of our results to the phenomenon of gender swapping and how people do associate physical attractiveness with gaining benefits in games. We also recommend a comparison of our data set with users of other cultures and different ethnical backgrounds, as the present study only examined Caucasian faces and the population in Central Europe. Cultural or ethnic differences of the preferred characteristics could further increase our knowledge of virtual avatars and our understanding of players.

ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (DFG) within project C04 of SFB/Transregio 161, the German Academic Exchange Service (DAAD) RISE Research Internship Program 2018, and by the graduate program *Digital Media* of the Universities of Stuttgart and Tübingen, and the Stuttgart Media University. Thanks go to Solveigh Jäger, Katrin Wolf, Katharina Leicht, Verena Dengler, Julia Sandrock, and Brighten Jelke.

[5] https://github.com/valentin-schwind/facemaker

REFERENCES

1. K. C. Apostolakis and P. Daras. 2013. RAAT - The reverie avatar authoring tool. In *2013 18th International Conference on Digital Signal Processing (DSP)*. 1–6. DOI:http://dx.doi.org/10.1109/ICDSP.2013.6622788
2. James Batchelor. 2017. *Uncharted: Drake's Fortune*. Website. (11 December 2017). https://www.gamesindustry.biz/articles/2017-12-11-uncharted-series-sales-passes-41-million
3. Elizabeth Behm-Morawitz and Dana Mastro. 2009. The Effects of the Sexualization of Female Video Game Characters on Gender Stereotyping and Female Self-Concept. *Sex Roles* 61, 11 (01 Dec 2009), 808–823. DOI:http://dx.doi.org/10.1007/s11199-009-9683-8
4. Julian Benson. 2015. *The Tomb Raider reboot is the most successful game in the series' history*. Website. (7 April 2015). https://www.pcgamesn.com/tomb-raider/the-tomb-raider-reboot-is-the-most-successful-game-in-the-series-history
5. BioWare. 2017. *Mass Effect: Andromeda*. Game [PC,PS4,XBOXONE]. (21 March 2017).
6. Max V. Birk, Cheralyn Atkins, Jason T. Bowey, and Regan L. Mandryk. 2016. Fostering Intrinsic Motivation Through Avatar Identification in Digital Games. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2982–2995. DOI: http://dx.doi.org/10.1145/2858036.2858062
7. J. Bovet, M. Barket-Defradas, V. Durand, C. Faurie, and M. Raymond. 2017. Women's attractiveness is linked to expected age at menopause. *Journal of Evolutionary Biology* 31, 2 (2017), 229–238. DOI: http://dx.doi.org/10.1111/jeb.13214
8. Bram P. Buunk, Pieternel Dijkstra, Detlef Fetchenhauer, and Douglas T. Kenrick. 2002. Age and Gender Differences in Mate Selection Criteria for Various Involvement Levels. *Personal Relationships* 9, 3 (2002), 271–278. DOI:http://dx.doi.org/10.1111/1475-6811.00018
9. Romina Carrasco, Steven Baker, Jenny Waycott, and Frank Vetere. 2017. Negotiating Stereotypes of Older Adults Through Avatars. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction (OZCHI '17)*. ACM, New York, NY, USA, 218–227. DOI: http://dx.doi.org/10.1145/3152771.3152795
10. Debaleena Chattopadhyay and Karl F. MacDorman. 2016. Familiar faces rendered strange: Why inconsistent realism drives characters into the uncanny valley. *Journal of Vision* 16, 11 (sep 2016), 7. DOI: http://dx.doi.org/10.1167/16.11.7
11. Donghun Chung, Brahm Daniel deBuys, and Chang S. Nam. 2007. Influence of Avatar Creation on Attitude, Empathy, Presence, and Para-Social Interaction. In *Human-Computer Interaction. Interaction Design and Usability*, Julie A. Jacko (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 711–720.
12. Vinet Coetzee, Jaco M. Greeff, Ian D. Stephen, and David I. Perrett. 2014. Cross-Cultural Agreement in Facial Attractiveness Preferences: The Role of Ethnicity and Gender. *PLOS ONE* 9, 7 (07 2014), 1–8. DOI: http://dx.doi.org/10.1371/journal.pone.0099629
13. Core Design. 1996-2013. *Tomb Raider*. Game [PC]. (March 1996-2013).
14. John F. Cross and Jane Cross. 1971. Age, sex, race, and the perception of facial beauty. *Developmental Psychology* 5, 3 (1971), 433–439. DOI: http://dx.doi.org/10.1037/h0031591
15. Karen Dion, Ellen Berscheid, and Elaine Walster. 1972. What is beautiful is good. *Journal of Personality and Social Psychology* 24, 3 (1972), 285–290. DOI: http://dx.doi.org/10.1037/h0033731
16. Nicolas Ducheneaut, Ming-Hui Wen, Nicholas Yee, and Greg Wadley. 2009. Body and Mind: A Study of Avatar Personalization in Three Virtual Worlds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 1151–1160. DOI: http://dx.doi.org/10.1145/1518701.1518877
17. Alice H. Eagly, Richard D. Ashmore, Mona G. Makhijani, and Laura C. Longo. 1991. What is beautiful is good, but . . .: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin* 110, 1 (1991), 109–128. DOI: http://dx.doi.org/10.1037/0033-2909.110.1.109
18. Christina R Glaubke, Patti Miller, McCrae A Parker, and Eileen Espejo. 2001. Fair Play? Violence, Gender and Race in Video Games. (2001).
19. Rosanna E. Guadagno, Nicole L. Muscanell, Bradley M. Okdie, Nanci M. Burk, and Thomas B. Ward. 2011. Even in Virtual Environments Women Shop and Men Build: A Social Role Perspective on Second Life. *Comput. Hum. Behav.* 27, 1 (Jan. 2011), 304–308. DOI: http://dx.doi.org/10.1016/j.chb.2010.08.008
20. S. Hagerman, Z. Woolard, K. Anderson, B. W. Tatler, and F. R. Moore. 2017. Women's self-rated attraction to male faces does not correspond with physiological arousal. *Scientific Reports* 7, 1 (2017), 1–6. DOI: http://dx.doi.org/10.1038/s41598-017-13812-3
21. Carrie Heeter, Rhonda Egidio, Punya Mishra, Brian Winn, and Jillian Winn. 2008. Alien Games: Do Girls Prefer Games Designed by Girls? *Games and Culture* 4, 1 (dec 2008), 74–100. DOI: http://dx.doi.org/10.1177/1555412008325481
22. Garrett Hulfish. 2016. The uncanny valley effect is strong with 'L.A. Noire' in 4K Ultra HD. Digital Trends. (Oct 2016). https://www.digitaltrends.com/gaming/la-noire-4k-trailer/

23. Zaheer Hussain and Mark D Griffiths. 2008. Gender swapping and socializing in cyberspace: an exploratory study. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society* 11, 1 (2008), 47–53. DOI: http://dx.doi.org/10.1089/cpb.2007.0020

24. Katherine Isbister. 2006. *Better Game Characters by Design: A Psychological Approach (The Morgan Kaufmann Series in Interactive 3D Technology)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

25. James D. Ivory. 2006. Still a Man's Game: Gender Representation in Online Reviews of Video Games. *Mass Communication and Society* 9, 1 (2006), 103–114. DOI: http://dx.doi.org/10.1207/s15327825mcs0901_6

26. Sara John. 2006. Un / Realistically Embodied : The Gendered Conceptions of Realistic Game Design. In *Advanced Visual Interfaces AVI, Workshop on Gender & Interaction, Real & Virtual Women in a Male World*. 5.

27. Jari Kätsyri, Klaus Förger, Meeri Mäkäräinen, and Tapio Takala. 2015. A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in psychology* 6, MAR (apr 2015), 390. DOI: http://dx.doi.org/10.3389/fpsyg.2015.00390

28. David J. Kelly, Paul C. Quinn, Alan M. Slater, Kang Lee, Liezhong Ge, and Olivier Pascalis. 2007. The Other-Race Effect Develops During Infancy: Evidence of Perceptual Narrowing. *Psychological Science* 18, 12 (2007), 1084–1089. DOI: http://dx.doi.org/10.1111/j.1467-9280.2007.02029.x PMID: 18031416.

29. Judith H. Langlois and Lori A. Roggman. 1990. Attractive Faces Are Only Average. *Psychological Science* 1 (1990), 115–121. DOI: http://dx.doi.org/10.1111/j.1467-9280.1990.tb00079.x

30. Anna Larsson and Carina Nerén. 2005. Gender Aspects on Computer Game Avatars. *Computer* November 2004 (2005).

31. David J. Lewkowicz and Asif A. Ghazanfar. 2009. The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences* 13, 11 (2009), 470 – 478. DOI: http://dx.doi.org/10.1016/j.tics.2009.08.004

32. Stacy Tessler Lindau, L. Philip Schumm, Edward O. Laumann, Wendy Levinson, Colm A. O'Muircheartaigh, and Linda J. Waite. 2007. A Study of Sexuality and Health among Older Adults in the United States. *New England Journal of Medicine* 357, 8 (2007), 762–774. DOI: http://dx.doi.org/10.1056/NEJMoa067423 PMID: 17715410.

33. Viola Macchi Cassia, Hermann Bulf, Ermanno Quadrelli, and Valentina Proietti. 2014. Age-related face processing bias in infancy: Evidence of perceptual narrowing for adult faces. *Developmental Psychobiology* 56, 2 (2014), 238–248. DOI: http://dx.doi.org/10.1002/dev.21191

34. Karl F. MacDorman and Debaleena Chattopadhyay. 2016. Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition* 146 (jan 2016), 190–205. DOI: http://dx.doi.org/10.1016/j.cognition.2015.09.019

35. Rosa Mikeal Martey, Jennifer Stromer-Galley, Jaime Banks, Jingsi Wu, and Mia Consalvo. 2014. The strategic female: gender-switching and player behavior in online games. *Information, Communication & Society* 17, 3 (2014), 286–300. DOI: http://dx.doi.org/10.1080/1369118X.2013.874493

36. Maxis/EA Games. 2000. *The Sims*. Game [PC]. (February 2000).

37. Christian A. Meissner and John C. Brigham. 2001. Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law* 7, 1 (2001), 3–35. DOI: http://dx.doi.org/10.1037/1076-8971.7.1.3

38. Monica K. Miller and Alicia Summers. 2007. Gender Differences in Video Game Characters' Roles, Appearances, and Attire as Portrayed in Video Game Magazines. *Sex Roles* 57, 9 (01 Nov 2007), 733–742. DOI: http://dx.doi.org/10.1007/s11199-007-9307-0

39. Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The uncanny valley. *IEEE Robotics and Automation Magazine* 19, 2 (jun 2012), 98–100. DOI: http://dx.doi.org/10.1109/MRA.2012.2192811

40. Naughty Dog. 2017. *Uncharted: Drake's Fortune*. Game [PS3,PS4]. (15 November 2017).

41. Kristine L Nowak and Christian Rauh. 2005. The Influence of the Avatar on Online Perceptions of Anthropomorphism, Androgyny, Credibility, Homophily, and Attraction. *Journal of Computer-Mediated Communication* 32 (2005), 178–197. DOI: http://dx.doi.org/10.1111/j.1083-6101.2006.tb00308.x

42. David I Perrett, Michael D Burt, Ian S Penton-Voak, Kieran J Lee, Duncan a Rowland, and Rachel Edwards. 1999. Symmetry and Human Facial Attractiveness. *Evolution and Human Behavior* 20 (1999), 295–307. DOI: http://dx.doi.org/10.1016/S1090-5138(99)00014-8

43. D. I. Perrett, K. J. Lee, I. Penton-Voak, D. Rowland, S. Yoshikawa, D. M. Burt, S. P. Henzi, D. L. Castles, and S. Akamatsu. 1998. Effects of sexual dimorphism on facial attractiveness. *Nature* 394, 6696 (aug 1998), 884–887. DOI: http://dx.doi.org/10.1038/29772

44. Azalea Reyes-Aguilar and Fernando A. Barrios. 2016. A Preliminary Study of Sex Differences in Emotional Experience. *Psychological Reports* 118, 2 (apr 2016), 337–352. DOI: http://dx.doi.org/10.1177/0033294116633350

45. Mark Rice, Ranieri Koh, Quintessence Liu, Qixiang He, Marcus Wan, Vanessa Yeo, Jamie Ng, and Wah Pheow Tan. 2013. Comparing Avatar Game Representation Preferences Across Three Age Groups. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 1161–1166. DOI: http://dx.doi.org/10.1145/2468356.2468564

46. Adam Rubenstein, Judith H. Langlois, and Lori A. Roggman. 2002. What Makes a Face Attractive and Why: The Role of Averageness in Defining Facial Beauty. (2002).

47. Edward Schneider and Shanshan Yang. 2007. Exploring the Uncanny Valley with Japanese Video Game Characters. *Intelligence* (2007), 546–549. http://www.digra.org/dl/db/07312.11004.pdf.

48. Sascha Schwarz and Manfred Hassebrauck. 2012. Sex and Age Differences in Mate-Selection Preferences. *Human Nature* 23, 4 (01 Dec 2012), 447–466. DOI: http://dx.doi.org/10.1007/s12110-012-9152-x

49. Valentin Schwind, Pascal Knierim, Cagri Tasci, Patrick Franczak, Nico Haas, and Niels Henze. 2017a. "These Are Not My Hands!": Effect of Gender on the Perception of Avatar Hands in Virtual Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1577–1582. DOI: http://dx.doi.org/10.1145/3025453.3025602

50. Valentin Schwind, Katrin Wolf, and Niels Henze. 2017b. FaceMaker—A Procedural Face Generator to Foster Character Design Research. In *Game Dynamics*. Springer International Publishing, Cham, 95–113. DOI: http://dx.doi.org/10.1007/978-3-319-53088-8_6

51. Valentin Schwind, Katrin Wolf, Niels Henze, and Oliver Korn. 2015. Determining the Characteristics of Preferred Virtual Faces Using an Avatar Generator. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '15)*. ACM, New York, NY, USA, 221–230. DOI: http://dx.doi.org/10.1145/2793107.2793116

52. Haeyeop Song and Jaemin Jung. 2015. Retracted: Antecedents and Consequences of Gender Swapping in Online Games. *Journal of Computer-Mediated Communication* 20, 4 (jul 2015), 434–449. DOI: http://dx.doi.org/10.1111/jcc4.12119

53. John P Swaddle and I. C. Cuthill. 1995. Asymmetry and Human Facial Attractiveness: Symmetry May not Always be Beautiful. *Proceedings of the Royal Society B: Biological Sciences* 261, 1360 (jul 1995), 111–116. DOI: http://dx.doi.org/10.1098/rspb.1995.0124

54. Paul Tassi. 2017. What BioWare Should Learn From The Reaction To 'Mass Effect: Andromeda'. Website. (March 2017). https://www.forbes.com/sites/insertcoin/2017/03/22/what-bioware-should-learn-from-the-reaction-to-mass-effect-andromeda/

55. Team Bondi. 2011. *L.A. Noire*. Game [PC,PS3,PS4,XBOX360]. (11 November 2011).

56. Natasha Veltri, Hanna Krasnova, Annika Baumann, and Neena Kalayamthanam. 2014. Gender differences in online gaming: a literature review. In *Twentieth Americas Conference on Information Systems*.