A perceptual study on the manipulation of facial features for trait portrayal in virtual agents

Ylva Ferstl∗
Graphics, Vision and Visualisation Group
Trinity College Dublin
yferstl@tcd.ie

Rachel McDonnell
Graphics, Vision and Visualisation Group
Trinity College Dublin
ramcdonn@tcd.ie

Figure 1: Examples of the faces used as stimuli. a) Wide face exp. 2, b) standard face with large eyes exp. 1, c) narrow face exp. 2, d) narrow face with small eyes exp. 1, e) standard face with medium eye size exp. 1, f) wide face exp. 2.

ABSTRACT

Human perceptual studies have shown that facial characteristics affect judgments about the personality of a person. For example, larger facial width has been associated with judgments of aggressiveness, dominance, and untrustworthiness. Previous studies of virtual faces have not been able to reflect the same perceptual rules, but have used characters with unrealistic feature sizes or highly abstract characters. For this study, we created virtual characters with realistic feature dimensions and investigated the effects of facial width and eye size on personality perception. Our results indicate that virtual characters may indeed follow different perceptual rules for facial width, and care must be taken when manipulating eye size. These findings are useful for effective character design for video games, movies, and embodied virtual agents.

CCS CONCEPTS
• Applied Computing → Psychology; Computer Games; • Computing methodologies → Perception;

KEYWORDS
character design, physiognomy, face perception

ACM Reference format:

∗Corresponding author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IVA ’18, November 5–8, 2018, Sydney, NSW, Australia
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-6013-5/18/11...$15.00
https://doi.org/10.1145/3267851.3267891

1 INTRODUCTION

Virtual humans are finding a growing number of applications, such as in the social media apps Spaces by Facebook, Bitmoji and Genies, as well as computer games and human-computer interfaces. With the increasing usage of virtual humans, their visual design becomes a more important factor. Finding rules relating to the perception of a virtual character’s appeal and/or personality could make networking in virtual spaces more effective or provide guidelines for artists when creating digital models for games and movies. Research shows that personality perception of human faces is strongly influenced by specific facial features [2, 11, 13, 14, 18]. Among the features influencing judgments of personality is the facial width-to-height ratio (FWHR), found to influence perception of dominance, honesty, and aggressiveness [3, 8, 19], and eye size, found to influence perception of dominance and honesty [10, 24].

More recently, these findings have been tested for virtual characters, on humanoid [22] and non-humanoid avatars [6]. In both cases, facial width seemed to have the opposite effect on personality perception than observed for humans, while eye size manipulations suffered from inducing a perception of eeriness. It is therefore unclear if and how the rules found for human faces can be applied to virtual characters.

In this study, we sought to build upon the limited studies of perception of personality traits of virtual faces. We used more realistic-looking characters than in previous studies in order to be able to better compare the results to findings for real human faces. In experiment 1, we manipulated FWHR as well as perceived eye size. In experiment 2, we further explored the effects of FWHR. In each case, we aimed to preserve a level of naturalness of the manipulations of facial features. We applied all manipulations to
both a male and a female character. Our results support previous findings for virtual characters that some facial features affect personality perception differently in virtual characters [6, 22]. Our findings have implications for effective character design in virtual environments.

2 BACKGROUND

Studies have associated the facial width-to-height ratio (FWHR), the width of the face divided by the height of the upper face, with perception of threat, dominance, as well as trustworthiness. A meta-analysis study concluded that individuals with larger FWHR were judged by observers as more threatening, more dominant and less attractive, compared to those with smaller FWHRs, especially for male faces [8]. Male faces with larger FWHR have furthermore been found to be judged as less trustworthy [15, 19].

Eye size has similarly been associated with personality judgments of human faces. Specifically, faces with larger eyes are perceived as more honest [24] and less dominant [10] than faces with small eyes.

To investigate whether faces of virtual humans are judged similarly to human faces, Wang et al. [22] used Second Life characters for comparison tests of personality. In contrast to human face studies, they did not find wider faces to be judged as less trustworthy, and wider faces were perceived as less aggressive compared to narrow faces. The authors found no link between eye size and ratings of dominance, while big eyes were linked to increased honesty ratings only for female avatars. However, their large-eyed avatars had high ratings of eeriness, possibly affecting other ratings. Additionally, the virtual characters used in their study, the second life avatars, were rather unrealistic and may not be directly comparable to human faces.

Further exploring the transferability of findings for the perception of personality of human faces, Ferstl et al. [6] investigated effects of facial features of very abstract looking virtual characters. They showed that the perceptual rules found for real human faces do not apply to abstract faces, and in some cases are the reverse. Similar to the findings of Wang et al. [22], they found narrow faces to be perceived as more aggressive. Additionally, narrow faces were perceived as more dominant than round or wide faces.

In this study, we assess whether virtual faces with more realistic feature sizes (compared to Wang et al. [22] and Ferstl et al. [6]) could reproduce the rules of personality perception found for human faces, or if perception of virtual faces indeed follows different rules.

3 EXPERIMENT 1

We created our stimuli with the Morph Character System for the Unity3D game engine. First, we created a standard facial blendshape configuration for both the female and the male character. We aimed to get the blendshape configurations as close as possible to the average faces found by Gruendl [9], which are based on 64 and 32 photographs of female and male faces, respectively (see Figure 2). Our resulting facial configurations served as the standard from which we derived all other configurations, namely narrower and wider faces, as well as eyes that would be perceived smaller and larger. We created combinations of all three facial widths and eye sizes, leading to a total of 9 configurations per gender, i.e., 18 in total.

Figure 2: Our standard faces (left image sides) compared to the average female and male face found by Gruendl [9] (right image sides).

Actual eye size does not seem to vary greatly between individuals [1], and perceived eye size may chiefly depend on eye shape and setting in the skull. Round eyes, for example, reveal more of the white of the eye and might appear larger. Protruding eyes may also appear large, whereas, hooded eyes and monolid eyes may appear smaller. For manipulating perceived eye size, we hence did not change the actual size determined by creating our standard faces, but instead varied eyelid size and shape, and the proportion of eyeball covered by both the upper and the lower eyelids. We kept FWHR within realistic ranges [8]. Still images of all characters with FWHR measurements are included in Appendix (section 1).

The experiment was designed with the Unity3D game engine and it was run online. Each of the characters were presented once and in random order, but male and female characters were always alternated to avoid direct comparisons. The characters were presented with a looped animation comprised of a slight idle head movement and regular eye blinks in order to increase naturalness. Eye blinks occurred 12 times per minute, guided by previous findings for average blink rates [5]. Each character was displayed in an uninterrupted animation loop until the participant finished rating it on 5 measures of personality and affinity, presented below the character, one at a time. Participants were asked to rate perceived dominance, trustworthiness, aggressiveness, appeal and eeriness of the presented character on a 7-point Likert scale (see Appendix, section 3). Before starting the experiment, participants completed two training trials with characters not included in the actual experiment. The training trials were identical to the experimental trials but the training characters were different from the experimental characters, with a different face and hair style.

The online experiment was distributed via university mailing lists. We collected data from 33 participants (12 females, 1 other gender, ages 20-48 years, M = 32.5, SD = 7.7), all of whom gave informed consent regarding their participation.

3.1 Results

We examined the effect of the factors facial width-to-height ratio (FWHR) and and eye size on the ratings given for each of the 5 questions (aggressiveness, dominance, trustworthiness, appeal, eeriness). We treated the questionnaire scores as ordinal data and analyzed the data by fitting a cumulative link model, using clm from the R ordinal package [4]. Results are visualized by diverging stacked bar charts of the rating scores in Figure 3.
Perception of aggressiveness was significantly influenced by FWHR ($\chi^2(2) = 12.73, p < .01$) as well as eye size ($\chi^2(2) = 9.91, p < .01$). A narrow face was perceived as significantly more aggressive than both standard FWHR ($p < .001$) and wide faces ($p < .01$), and small eyes were likewise perceived as more aggressive than both medium ($p < .01$) and large eyes ($p < .01$).

Dominance ratings were significantly influenced by both FWHR ($\chi^2(2) = 19.37, p < .001$) and eye size ($\chi^2(2) = 18.01, p < .001$). Narrow faces were perceived as more dominant than both standard ($p < .01$) and wide faces ($p < .001$). Small eyes were also perceived as more dominant than both medium ($p < .01$) and large eyes ($p < .001$).

Perceived trustworthiness was significantly influenced by both FWHR ($\chi^2(2) = 8.33, p < .05$) and eye size ($\chi^2(2) = 10.33, p < .01$). Narrow ($p < .01$) as well as wide faces ($p < .05$) were perceived as less trustworthy than standard width. Small eyes were rated less trustworthy than medium eyes ($p < .01$).

Appeal was significantly influenced by both FWHR ($\chi^2(2) = 92.29, p < .001$) and eye size ($\chi^2(2) = 9.92, p < .01$), where wide faces were perceived as less appealing than standard ($p < .001$) and narrow faces ($p < .001$), as were small eyes compared to medium eyes ($p < .01$).

Eeriness was influenced by both FWHR ($\chi^2(2) = 13.72, p < .01$) and eye size ($\chi^2(2) = 20.10, p < .001$). Perceived eeriness was higher for narrow ($p < .01$) and wide faces ($p < .001$), compared to standard faces. Large eyes were perceived as most eerie (compared to medium eyes: $p < .001$; compared to small eyes: $p < .05$), and small eyes were perceived as more eerie than medium eyes ($p < .05$).

We used the likelihood ratio test to assess whether the interaction between the two factors are supported by the data for each question. The likelihood ratio statistic was small in all cases, indicating that interactions are not supported by the data.

3.2 Discussion

Eye size has previously been linked to perceptions of trustworthiness and dominance for human faces, with large eyes rated as less dominant [10] and more trustworthy [24]. In line with this, we found small eyes to be perceived as more dominant than medium and large eyes, and less trustworthy than medium eyes.

Narrow faces were perceived as both more aggressive and more dominant, while wide faces did not differ significantly from standard width faces. Both narrow and wide faces were perceived as less trustworthy than faces with standard width. Wide faces were perceived as less appealing than both narrow faces and faces with standard facial width. Standard facial width was perceived as less eerie than both narrow and wide faces.

Wide faces were rated as rather unappealing, potentially influencing other ratings. We hypothesized that some of the male faces, especially wide faces, had a feminine look. The wide face may have induced perceptions of babyfacedness, as mentioned as a possible effect in [22]. We therefore ran a second experiment exploring the effects of facial width before drawing conclusions. In this second experiment, we aimed to make the narrow and wide faces more appealing, the male faces more masculine (especially the wide face), and the female faces more feminine.

4 EXPERIMENT 2

In experiment 2, we only manipulated facial width, while keeping eye size fixed. We aimed to make the narrow and wide facial shapes more appealing and natural. We changed the appearance of the male character to be more masculine, including the addition of a beard, and we aimed to make the female character more feminine by adding make-up and bigger lips. Still images of the new characters are shown in the Appendix, section 2. We tested the perception of these new faces along with the six standard eye-size faces from experiment 1, resulting in 12 test characters. We added a rating of masculinity/femininity to the question set.

The questionnaire was distributed via university mailing lists. We collected data from 27 participants (7 females, 1 other gender, ages 20–47 years, $M = 29.0, SD = 6.5$), all of whom gave informed consent regarding their participation. The experiment was run in the way previously described for experiment 1.

4.1 Results

We analyzed the data with the same statistical methods described for experiment 1, but with the factors FWHR, gender, and character (old versus new). We did not find any interactions with character and the other two factors, indicating that the direction of effects was not influenced by the factor character.

Overall, narrow faces were perceived as significantly more aggressive ($\chi^2(2) = 19.72, p < .001$), more dominant ($\chi^2(2) = 21.17, p < .001$), and less trustworthy ($\chi^2(2) = 21.65, p < .001$) than faces with standard width. Narrow faces were rated more eerie and less appealing than faces with standard width ($\chi^2(2) = 20.28, p < .001$ and $\chi^2(2) = 20.09, p < .001$, respectively). Male faces were perceived as more aggressive ($\chi^2(2) = 34.36, p < .001$), more dominant ($\chi^2(2) = 19.72, p <
34.29, \( p < .001 \) and less trustworthy (\( \chi^2(2) = 21.65, p < .001 \)). The new male faces were perceived as more masculine than the old male faces (\( \chi^2(2) = 10.42, p < .01 \)). For the female characters, narrow and wide faces were perceived as less feminine than standard facial width (\( \chi^2(2) = 10.66, p < .01 \) and \( p < .05 \), respectively).

The new male faces were perceived as more masculine than the old male faces (\( \chi^2(2) = 21.65, p < .001 \)). For the female characters, narrow and wide faces were perceived as less feminine than standard facial width (\( \chi^2(2) = 10.66, p < .01 \) and \( p < .05 \), respectively).

Figure 4: Stacked bar charts for the results of experiment 2. Plotted is the frequency of responses for each of the 7 rating scores, for all facial width-to-height ratios (FWHR). (The y-axis represents the frequency of responses).

### Table 1: Comparison of results for facial width, for studies on real and on virtual faces.

<table>
<thead>
<tr>
<th>Researchers</th>
<th>Type of face</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geniole et al. [8]</td>
<td>Real</td>
<td>Wide faces more threatening, more dominant, and less attractive</td>
</tr>
<tr>
<td>Ormaiston et al. [15], Stirrat &amp; Perret [19]</td>
<td>Real</td>
<td>Wide faces less trustworthy</td>
</tr>
<tr>
<td>Wang et al. [22]</td>
<td>Virtual</td>
<td>Wide faces less aggressive</td>
</tr>
<tr>
<td>Ferstl et al. [6]</td>
<td>Abstract virtual</td>
<td>Narrow faces more aggressive and more dominant</td>
</tr>
<tr>
<td>Our findings</td>
<td>Virtual</td>
<td>Narrow faces more aggressive, more dominant, and less trustworthy. Wide faces less trustworthy and less attractive.</td>
</tr>
</tbody>
</table>

Our main findings of experiment 1 were reproduced; narrow faces were again rated as more aggressive, more dominant, and less trustworthy than standard faces. Male faces were perceived as more aggressive, more dominant, and less trustworthy than female faces. The fact that we did not find this in experiment 1 might be due to many of the male characters in experiment 1 looking somewhat feminine, through a round face or large eyes.

5 GENERAL DISCUSSION

Perceptual studies of human faces have found robust rules for the influence of facial features on perception of personality. Studies of virtual characters have not reflected these perceptual rules, but have been limited in number and restricted to unrealistic characters [6, 22]. We created characters with more realistic facial feature proportions than in previous studies.

In line with findings for human faces, we found small eyes to be perceived as more dominant than medium and large eyes, as well as less trustworthy than medium eyes. Large eyes yielded increased ratings of eeriness, emphasizing the need for careful feature design. In humans, large eyes are often considered attractive, especially in females (e.g., Mila Kunis, Winona Rider), and large-eyed cartoon characters may look appealing (e.g., animated Disney characters such as in Frozen, Tangled, etc.). However, more realistic-looking virtual characters may have to follow much more strict guidelines, as even our careful design of large (and small) eyes increased perceptions of eeriness compared to medium eyes.

Wide faces were perceived as less trustworthy than standard width, in line with previous findings for human faces [19, 23].

We found narrow faces to be perceived as both more aggressive and more dominant, while wide faces did not differ significantly from standard width. Narrow faces were also rated less trustworthy. In light of the consistency of the results for the three studies of virtual faces known to us, this supports the notion that virtual faces are perceived differently from real human faces. A reason for this could be the tendency of villains in animated movies to be portrayed with sharp, long facial features (e.g., Captain Hook in Peter Pan, Scar in Lion King, Maleficent). This tendency could influence the perception of computer-generated characters towards automatic association of narrow faces with dangerous characters.
We considered that there may be a difference between real and virtual faces through other subtle cues. For example, wide faces in human males are associated with higher testosterone [8], which might in turn influence other facial features. However, studies on manipulated images of real human faces (that presumably also lack other subtle facial features associated with hormonal differences) reflected the same perceptual rules found for threats and dominance for human faces that were not manipulated [8].

In our study, narrow faces were also associated with increased ratings of eeriness and decreased trustworthiness which could have led to a general association with more negative traits, however, wide faces showed these same effects without an increase in aggressiveness and dominance ratings. Therefore, we do not believe that the perceived increased aggressiveness and dominance of narrow faces is due to increased eeriness or decreased trustworthiness.

Standard width tended to be rated more appealing and less eerie than narrow and wide faces, possibly showing a general preference for average faces, as previously found by [12, 16]. Average facial width was preferred even compared to narrow or wide faces that were made overall more appealing. Wide faces were perceived as least appealing, in line with previous findings for human faces [8].

In summary, our results support the notion that the appearance of virtual characters is judged differently from real human faces. Whereas wide human faces are perceived as more aggressive and dominant, portraying these same characteristic with virtual characters might be more effectively done with narrow faces. (See table 1 for comparison of results.)

Our results may be helpful for portraying personality more intentionally and effectively with virtual humans. For example, one could consider making subtle changes to the facial width and eye size of a social media avatar in order to appear more trustworthy. Providers of human-computer interfaces could benefit from using avatars that elicit more positive personality perceptions. A virtual agent that is perceived as friendly and trustworthy might increase acceptance and engagement of human users. Our findings could also be relevant for the design of virtual instructors or teachers, whose appearance seems to influence students’ learning outcomes [20]. For example, designing the virtual instructor to appear more trustworthy could improve a student’s perception of the instructor and positively impact their learning. Applying our findings for virtual characters in games by using a facial appearance consistent with the character’s intended traits could increase player engagement, similarly to how consistency of the character’s behavior and visual realism enhances player experience [21]. Previous work has also shown that personality perceptions of non-player characters in games can influence behavior [7], suggesting an application of our findings in game design.

Future work could explore whether matched virtual and real faces (i.e., a real face versus its virtual version) are also perceived differently. It would furthermore be interesting to investigate if and when increasingly realistic virtual humans would start following the perceptual patterns found for humans. Including a rating of naturalness of the virtual face could further help understand the perceptual patterns of human faces. Dynamic features of appearance may also be a subject of future work, extending and complementing the work of Ruhlund et al. [17]. Going forward, we would also like to investigate the effects of a virtual agent’s voice on user experience.

ACKNOWLEDGEMENTS

This research was funded by Science Foundation Ireland under the ADAPT Centre for Digital Content Technology (Grant 13/RC/2016) and the Game Face (13/DA/2135) project.

REFERENCES

APPENDIX
1  Still images of character stimuli of experiment 1 with FWHR measurements
2 Still images of new character stimuli of experiment 2 with FWHR measurements

3 Personality and affinity rating questionnaire

(1) "I find the character dominant and forceful."
   ('Not at all' - 'Extremely')
(2) "I find the character appealing."
   ('Not at all' - 'Extremely')
(3) "I find the character trustworthy."
   ('Not at all' - 'Extremely')
(4) "I believe this character would react aggressively if provoked."
   ('Not at all' - 'Extremely')
(5) "I find the character eerie."
   ('Not at all (= The character restores a sense of security, confidence, calm in me.)' - 'Extremely (= The character is gloomy and leaves me with a sense of fear.)')
(6) *Only in experiment 2* "I think this character looks ..."
   ('Very male' - 'Very female')