

Appealing Avatars from 3D Body Scans

Perceptual Effects of Stylization

Reuben Fleming^{1,2}, Betty J. Mohler², Javier Romero³, Michael J. Black³, and
Martin Breidt²

¹ Sheffield Hallam University, Sheffield, UK

² Max Planck Institute for Biological Cybernetics, Tübingen, Germany

³ Max Planck Institute for Intelligent Systems, Tübingen, Germany

Abstract. Using styles derived from existing popular character designs, we present a novel automatic stylization technique for body shape and colour information based on a statistical 3D model of human bodies. We investigate whether such stylized body shapes result in increased perceived appeal with two different experiments: One focuses on body shape alone, the other investigates the additional role of surface colour and lighting. Our results consistently show that the most appealing avatar is a partially stylized one. Importantly, avatars with high stylization or no stylization at all were rated to have the least appeal. The inclusion of colour information and improvements to render quality had no significant effect on the overall perceived appeal of the avatars, and we observe that the body shape primarily drives the change in appeal ratings. For body scans with colour information, we found that a partially stylized avatar was perceived as most appealing.

Keywords: 3D Body Scan, Stylization, Avatar, Perception, Virtual Character

1 Introduction

Virtual avatars are frequently used in games, virtual worlds and for online communications. How an avatar is perceived by others is considered extremely important but creating a highly detailed and realistic avatar does not necessarily produce appealing results [1]. *Appeal*, one of the twelve principles of animation [8] is commonly used to describe well designed, interesting and engaging characters. In the same way animated movie and game characters require appeal in order for people to engage with them, virtual avatars also require appeal in order to engage others.

In this paper, we explore different ways in which we can increase the appeal of 3D body scans via stylization. In particular, we explore which styles make the virtual avatars most appealing, and what is the optimal amount of stylization for the most successful ones. We also examine the role of other factors in the appeal of the final stylized avatar, namely the realism of the renders and the appearance of the original subject. For this, we acquire 3D body scans of real

people using a state-of-the-art capturing system and automatically create virtual avatars with different styles and degrees of stylization. We render these avatars with and without colour information and also improve the render quality to simulate varying levels of realism. We then study, in multiple experiments, which factors affect the perceived appeal of these avatars.

2 Background

Human-like virtual characters are becoming more and more present in our lives. Since virtual environments often try to replicate the real world, a natural design choice for these virtual characters is to make them look as real as possible. This is becoming easier thanks to the improvements of 3D capturing systems in terms of speed, accuracy, quality and price. However, not all virtual characters are designed with reality in mind. Highly stylized avatars, such as the ones used by the Xbox and Wii game consoles, seem to suggest that people find the stylized look very appealing. The stylized approach to avatar creation has traditionally required the user to manually define and personalize their avatar; creating a true resemblance of that person is hard to achieve. Game studios are already attempting to cut out the manual work associated with this process by capturing gamers' features and applying them to in-game characters. For example *NBA 2K15* [18] attempts to capture the player's face while *Kinect Sports Rivals* [13] captures the entire body. Extracting information about the player is currently achieved using peripherals such as Microsoft's *Kinect 2* and Sony's *PS4* camera.

Despite the availability of both high quality, realistic body scans and increasingly convincing avatars from consumer peripherals, highly stylized characters are still extremely popular and can be seen in many animated movies, games and even web-based support services. Animation studios such as *Disney*, *Blue Sky* and *Pixar* tend to aim for a highly stylized look when designing characters. Attempts to increase realism in stylized characters have often led to negative reactions from viewers [12]. Examples of this include films such as *Polar Express* [21], *Mars Needs Moms* [19], and *The Adventures of Tintin* [17].

Furthermore, negative reactions to the almost, but not quite, photo-realistic characters seen in some film and game productions [3] and the human-like robots that inspired the "Uncanny Valley" hypothesis [14] suggest that stylizing (therefore decreasing realism), rather than attempting to improve realism, could actually produce more appealing results. For example, Inkpen and Cedlins [6] conducted a survey into peoples' comfort when communicating with avatars and found that although respondents were comfortable interacting with both realistic and cartoon looking avatars, some avatars rated highly realistic were also felt to be eerie or creepy. It is therefore possible that, even if one created a photo-real digital replica of a human, a stylized version could still be more appealing.

Experiments performed by Seyama and Nagayama [16], where they morphed between artificial and real faces, suggest that stylization should be done consistently in order to avoid viewer discomfort as abnormal features (e.g. cartoon eyes applied to a human head) produced the lowest pleasantness scores in their

experiments. These conclusions match similar findings by Macdorman et al. [9] who found that a mismatch between the size and texture of the eyes and face was especially prone to making a character look eerie. Interestingly, again for faces, results by McDonnell et al. [11] suggest that both highly realistic and highly abstract render styles were both considered appealing but render styles in-between the two were considered unappealing. Recently Zell et al. [20] found that while face shape was important for portraying realism, it was the material – specifically the albedo map – that played the key role in influencing the perceived appeal of CG characters and a stylized smooth looking skin texture (obtained via blurring) was perceived as being more appealing than a realistic skin texture.

To go beyond previous work on face appeal and investigate the perceptual effects of stylization on full-bodied avatars, we conducted two separate experiments. The aim of the first experiment was to investigate the effects of stylization when applied to average male and female body shapes containing no colour information. This allowed us to focus purely on body shape. In the second experiment we stylized 3D body scans of individual people. The aim here was to assess the effect of the subjects’ appearance, (both in terms of body shape and colour information) and render quality on perceived appeal and realism. Overall the aim of these experiments is to better understand how to create more appealing avatars from 3D body scans.

3 Stylization of 3D Body Scans

3.1 3D Body Scanning

To represent realistic human body shape and appearance we first captured detailed 3D body scan and colour information from 4 actors (2 female).

The system is composed of 22 measurement units, each including a pair of black and white cameras observing a projected speckle light pattern, and one 5 megapixel colour camera that captures the body appearance 2 ms after the speckle image, synchronized with white flash lights. The system provides very good coverage of the entire human body and can resolve the 3D locations of a point on the body to approximately 1 mm.

Raw scans contain noise, occasional holes and can be difficult to manipulate. Consequently we registered a common 3D body template to the scans in order to obtain a coherent mesh topology for each. For this we used a statistical 3D body model that compactly parametrises body deformations automatically in terms of shape changes due to subjects’ identity and body pose [4]. The model is based on approximately 2100 body scans of male and 1700 body scans of female subjects from the US and EU CAESAR dataset [15].

Through this process we obtained 3D geometry that resembles the original 3D scans but also shares a common layout. We call these meshes *registrations*. Since registrations share mesh topology, their vertex positions fully describe the body geometry. Therefore, morphing between two or more bodies can be performed as a weighted sum of the corresponding vertex positions across registrations.

Once registered, we extracted both albedo and pre-lit colour maps [2] from each scan (using the UV texture coordinate layout defined in the template). The albedo maps were illuminated using simple point lights within 3ds Max 2014 (Autodesk, USA). The pre-lit colour maps (Fig. 1 right), provided a more realistic look and, unlike the albedo maps (Fig. 1 left), required no CG lighting due to the lighting information already being embedded within the maps.

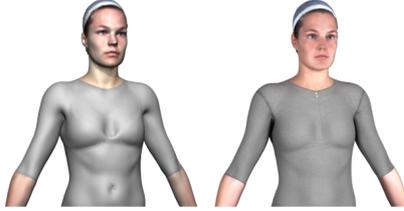


Fig. 1: Detail showing female actor (F_1) with CG lit albedo map (left), and pre-lit colour map (right).

3.2 Shape Stylization

For our initial experiment, 10 style templates (5 male and 5 female) were created in the digital sculpting software Zbrush (Pixologic, USA) (Fig. 2). The two most appealing styles for each gender were then used again for our second experiment.

The style templates were created by a trained 3D artist and were modeled to match the body proportions of existing stylized designs (see Tab. 1). The style templates used the same topology as the registrations, therefore direct morphing between registrations and style templates was made possible by simply interpolating vertex positions. While creating the styles is a time-consuming manual task that requires a skilled artist, it is worth noting that thanks to the shared topology, this only needs to be done once and then the templates can be reused for all future scans.

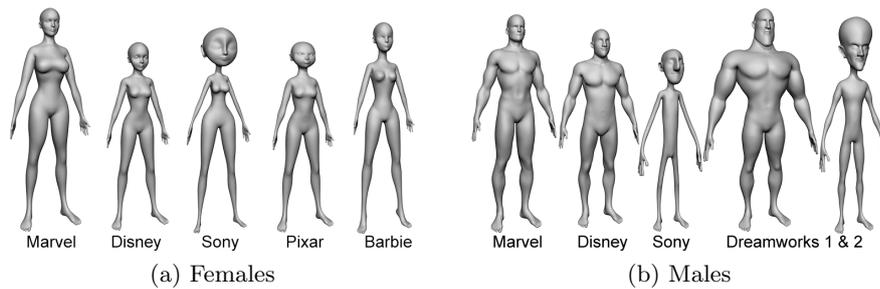


Fig. 2: Style templates created from character reference.

Table 1: Style template references

Gender	Style Name	Description
Female	Marvel	Typical female Marvel superheroine (Marvel Entertainment, USA)
Female	Disney	“Rapunzel” from the animated movie “Tangled” (Disney, 2010)
Female	Sony	“Sam Sparks” from the animated movie “Cloudy with a Chance of Meatballs” (Sony, 2009)
Female	Pixar	“Princess Merida” from the animated movie “Brave” (Pixar, 2012)
Female	Barbie	“Barbie” doll (Mattel, USA)
Male	Marvel	Typical male Marvel superhero (Marvel Entertainment, USA)
Male	Disney	“Kristoff” from the animated movie “Frozen” (Disney, 2013)
Male	Sony	“Flint Lockwood” from the animated movie “Cloudy with a Chance of Meatballs” (Sony, 2009)
Male	Dreamworks 1	“Metro Man” from the animated movie “Megamind” (Dreamworks Animation, 2010)
Male	Dreamworks 2	“Megamind” from the animated movie “Megamind” (Dreamworks Animation, 2010)

3.3 Body Shape Generation

For experiment 1 we used average male and female body shapes (Fig. 4). As the process for generating these shapes was identical, the following describes the process for generating the average female body M (Fig. 3 top left). The average shape was computed by averaging the $n \approx 2100$ female or the $n \approx 1700$ male original registrations O_i (as described in section 3.1): $M = \frac{1}{n} \sum_{i=1}^n O_i$.

In experiment 2, rather than stylize the average body shape, we stylized registrations F obtained from 3D body scans of individual people. In order to maintain the characteristic body features (even when fully stylized) we augmented the resulting morph between registration F_i and template S_j with data derived from the registration: Half of the Euclidean difference between the mean body shape M and the registration F_i was added back to the current body style template S_j (right term of Eq. 1). See Fig. 3 for an illustration of the process.

We obtain the final stylized avatar shape $N_{i,j|\alpha}$ used for one trial (with identity i , style j and stylization level $\alpha \in [0 \dots 1]$) by linear interpolation between the original registration F_i and the current style template S_j , augmented by the individual body characteristics. This helped to maintain the more subtle characteristics of the original body shape while keeping the overall proportions of the stylized models.

Eq. 1 was implemented using the *Morpher* modifier in 3ds Max, morphing the scans from their original state to the 100% stylized model (see Fig. 3 bottom for examples). Two intermediate states were also generated at 33% and 66% stylization.

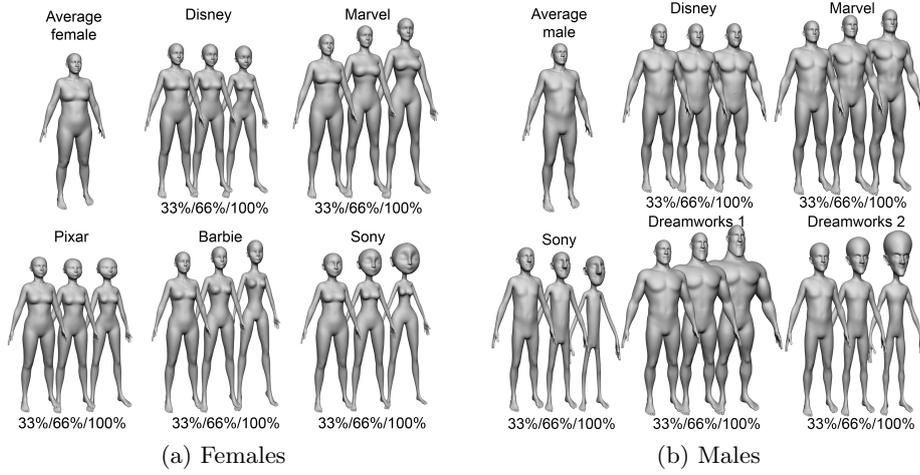


Fig. 4: Styles used in Experiment 1.

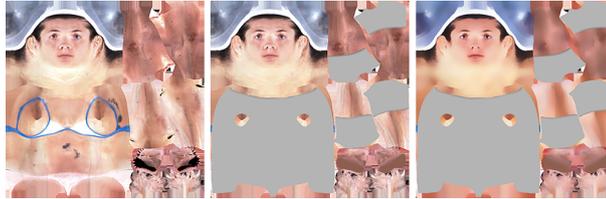


Fig. 5: Original albedo map (left), cleaned up and grey body suit added (center), stylized result (right).

Both the albedo and pre-lit colour maps had a resolution of 1024×1024 pixels. Parameter α in Eq. 1 was used to linearly blend between the normal and stylized versions of these maps.

4 Experiment 1

For our first experiment we rendered avatars with no colour information within the Unity 4 real-time game engine (Unity Technologies, Denmark). We displayed an average male and female body, 5 different style templates for each body and 2 in-between levels of stylization (33 and 66%) for each template (Fig. 4). This resulted in a total of 32 avatars (the average male and female body shapes plus 30 stylized bodies). Stimuli were presented in a random order.

4.1 Experimental Design

Male and female stimuli were shown to two separate groups of people. We tested 18 participants (10 female, mean age 26.5 yr, $SD = 7.5$) for the female stimuli, and 17 participants (8 female, mean age 29.5 yr, $SD = 5.9$) for the male

stimuli. Both were conducted in accordance with the Declaration of Helsinki. All participants signed informed consent and were financially compensated. We asked participants for their judgements of appeal on a 7-pt Likert scale when observing our stimuli. Participants were instructed to answer quickly and simply provide their first impression. Participants were also given an in-depth definition of appeal before they began:

High appeal: “you find the person extremely appealing. This may include finding the person extremely engaging, likeable and/or generally pleasing to the eye.”

Average appeal: “you neither find the person appealing nor unappealing. They are acceptable but essentially you are indifferent, finding them neither engaging nor disengaging, likeable nor unlikeable.”

Low appeal: “you find the person extremely unappealing. This may include finding the person extremely disengaging, unlikeable and/or repellent.”

4.2 Experiment 1: Results for Female Avatars

A partially stylized avatar was generally rated as being more appealing than the average female body shape while the amount of stylization for optimal appeal depends on the style, (Fig. 6 and Tab. 2).

Table 2: Mean female appeal ratings (standard deviations in parentheses). The average female body shape was rated as 3.61 (SD = 0.92) on appeal.

Style	33% stylized	66% stylized	100% stylized
Marvel	4.94 (1.211)	5.00 (1.085)	4.17 (1.724)
Disney	4.72 (1.127)	4.83 (1.425)	3.67 (1.847)
Barbie	4.39 (1.195)	4.00 (1.572)	2.33 (1.138)
Pixar	4.39 (1.092)	3.44 (1.504)	2.17 (1.043)
Sony	2.22 (1.003)	1.56 (0.784)	1.28 (0.461)

We ran a repeated measures ANOVA on the rating of appeal for female avatars with two within subject factors: Style (5 levels: Marvel, Disney, Sony, Dreamworks 1, Dreamworks 2) and Percent Stylized (4 levels: average female, 33, 66, 100%) and one between subject factor: Participant Gender (2 levels: male, female). There was no significant effect of participant gender on rating of appeal on female avatars, $F(1,16) = 0.027$, $p = 0.871$, $\eta_p^2 = 0.002$.

There was a significant effect of percent stylized on appeal ratings, $F(3,48) = 8.485$, $p < 0.001$, $\eta_p^2 = 0.347$. 100% stylized had significantly lower appeal ratings than both 33% and 66% while both 33% and 66% stylized female avatars were found to be more appealing than the average female body (as revealed by post-hoc comparisons using Bonferroni adjustments for multiple comparisons, ($p < 0.001$)).

There was also a significant effect of style on appeal ratings, $F(4,64) = 63.03$, $p < 0.001$, $\eta_p^2 = 0.798$. The five styles did not have the same effect on appeal ratings across percent stylized (as revealed by post-hoc comparisons using

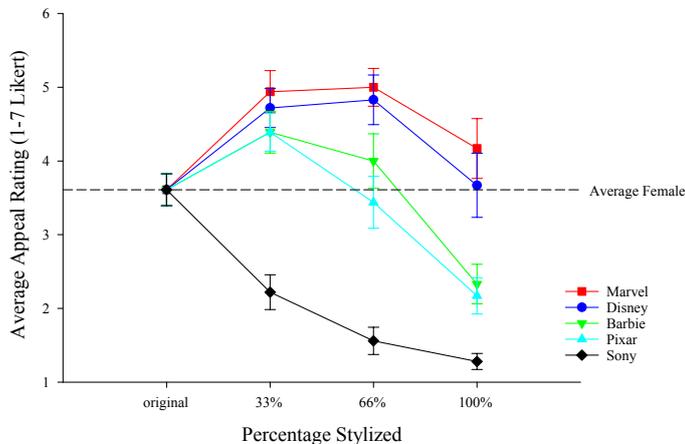


Fig. 6: Female appeal ratings on a 7-pt Likert scale ($n=18$). Ratings for the average female have been included in this graph as a reference. Error bars represent one standard error of the mean in all presented figures.

Bonferroni adjustments). Most notably, the Sony style had a significantly lower rating of appeal than all of the styles ($p < 0.001$) and was always rated as less appealing than the average female body. The appeal ratings for the Marvel and Disney styles did not significantly differ from each other ($p = 1.00$) and appeal ratings for the Barbie and Pixar styles also did not significantly differ from each other ($p = 0.907$). All other comparisons between styles resulted in significantly different appeal ratings ($p < 0.05$).

These main effects were conditioned upon a significant interaction between style and percent stylized, $F(12,192) = 8.154$, $p < 0.001$, $\eta_p^2 = 0.422$. We therefore tested for an effect of percent stylized for each individual style. The effects of percent stylized were significant for all styles ($p < 0.05$) except for the Marvel style ($p = 0.12$). Not surprisingly, there was a simple effect of style found when testing at each percent stylized, driven by the low ratings of appeal for Sony ($p < 0.05$). In addition, at 100% stylized the Marvel and Disney appeal ratings were significantly greater than both Barbie and Pixar appeal ratings (as revealed by post-hoc comparisons using Bonferroni adjustments, ($ps < 0.05$)).

These results show that for all styles except Sony, a partially stylized avatar leads to the greatest appeal ratings and depending on the style either 33% or 66% leads to optimal appeal. Excluding the Sony style, an improvement upon the original appeal ratings for the average female body of approximately 29% could be achieved by stylizing the body shape.

4.3 Experiment 1: Results for Male Avatars

A partially stylized avatar was often rated as more appealing than the average male body shape while the amount of stylization for optimal appeal depends on the style, (Fig. 7 and Tab. 3).

Table 3: Mean male appeal ratings (standard deviations in parentheses). The average male body shape was rated as 2.94 (SD = 1.30) on appeal.

Style	33% stylized	66% stylized	100% stylized
Marvel	4.00 (1.12)	5.12 (1.06)	5.59 (1.33)
Disney	4.06 (1.03)	4.35 (1.17)	4.71 (1.31)
Dreamworks 1	4.12 (1.76)	3.71 (1.86)	3.00 (1.73)
Sony	2.41 (1.33)	1.47 (0.62)	1.12 (0.49)

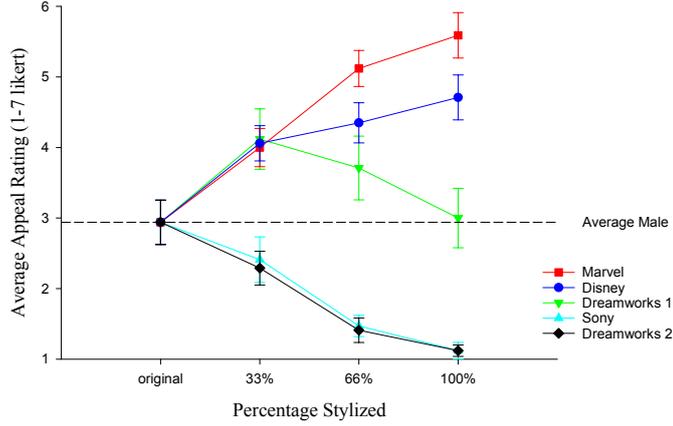


Fig. 7: Male appeal ratings on a 7-pt Likert scale ($n=17$). Ratings for the average male have been included in this graph as a reference.

We ran a repeated measures ANOVA on the rating of appeal for male avatars with two within subject factors: Style (5 levels: Marvel, Disney, Sony, Dreamworks 1, Dreamworks 2) and Percent Stylized (4 levels: average male, 33, 66, 100%) and one between subject factor: Participant Gender (2 levels: male, female). There was no significant effect of Participant Gender, $F(1,15) = 1.831$, $p = 0.196$, $\eta_p^2 = 0.109$.

There was no significant effect of percent stylized on appeal ratings, $F(3,45) = 2.745$, $p = 0.429$, $\eta_p^2 = 0.059$. There was however a significant effect of style on appeal ratings, $F(4,60) = 67.743$, $p < 0.001$, $\eta_p^2 = 0.819$. The five styles did not have the same effect on appeal ratings across percent stylized (as revealed by post-hoc comparisons using Bonferroni adjustments). Most notably the Sony and Dreamworks 2 styles had significantly lower appeal ratings than the other styles at all percentage stylized ($p < 0.001$). These two styles were always rated as less appealing than the average male body shape and did not significantly differ from each other ($p < 0.001$). Two other style comparisons were not significantly different in appeal ratings, specifically Marvel and Disney ($p = 0.084$) and Disney and Dreamworks I ($p = 0.434$). All other comparisons between styles resulted in significantly different appeal ratings ($p < 0.05$). Marvel clearly re-

ceived the highest appeal rating (4.90), followed by Disney (4.37) and then finally Dreamworks I (3.61).

These main effects were conditioned upon a significant interaction between style and percent stylized, $F(12,180) = 21.056$, $p < 0.001$, $\eta_p^2 = 0.584$. We therefore tested for an effect of percent stylized for each individual style. The effects of percent stylized were significant for all styles ($p < 0.05$) except for the Disney style ($p = 0.238$). Not surprisingly, there was a simple effect of style found when testing at each percent stylized driven by the low ratings of appeal for Sony and Dreamworks 2 ($p < 0.05$). In addition, only at 100% stylized were the Marvel and Disney appeal ratings significantly greater than Dreamworks I (as revealed by post-hoc comparison using Bonferroni adjustments, ($p < 0.05$)). For Disney and Marvel, appeal ratings were significantly greater than the average male avatar at all percent stylized ($p < 0.05$), with Bonferroni adjustments for multiple comparisons). For Dreamworks 1 only 33% stylized was rated as significantly more appealing than the average male body shape ($p = 0.01$).

These results show that for all styles except Sony and Dreamworks 2, a stylized avatar leads to the greatest appeal ratings and depending on the style, either 33% (Dreamworks 1) or 100% (Marvel and Disney) leads to optimal appeal. Excluding the Sony and Dreamworks 2 styles, an improvement upon the original appeal ratings for the average male body of approximately 63% could be achieved by stylizing the body shape.

5 Experiment 2

Both the Marvel and Disney styles provided the most significant improvements for appeal ratings in experiment 1 so were therefore used to stylize the body scans of real people in our second experiment. For this experiment, 28 images were rendered in 3ds Max at 500×1000 pixels using our CG lit albedo maps, 7 rendered images per character: three percentages at two styles each, plus the original body scans (2 male and 2 female). Another 28 images were rendered in 3ds Max using our pre-lit colour maps at the same resolution (Fig. 8).

5.1 Experimental Design

There were 61 Participants (35 female) with an average age of 28.08 yr (SD = 7.6) and none of the participants had been involved in experiment 1. The experiment was conducted in accordance with the Declaration of Helsinki. All participants signed informed consent and were compensated. There was no time limit assigned to the sessions but participants were instructed to answer quickly and provide their first impression. The experiment lasted approximately one hour.

There were two tasks in this experiment. First, participants provided appeal ratings as in experiment 1. Second, participants were asked to dynamically create the most appealing and repelling avatars.

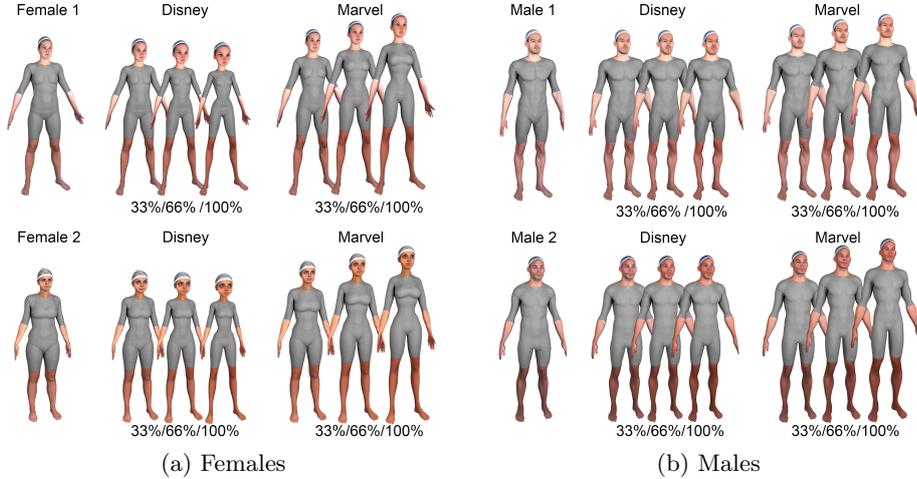


Fig. 8: Styles rendered using pre-lit colour maps.

Task 1: Likert Scale Ratings

Experiment 2 was a repeated measures design and involved rating male and female avatars on a 5-pt Likert scale and had three within subject factors for female avatars: Actor Identity (2 levels: F_1 , F_2), Style (2 levels: Marvel, Disney), Percent Stylized (3 levels: 33, 66, 100%) and three within subject factors for male avatars: Actor Identity (2 levels: M_1 , M_2), Style (2 levels: Marvel, Disney), Percent Stylized (3 levels: 33, 66, 100%) and one between subject factor of Render Quality (2 levels: albedo maps with CG lighting (42 participants), pre-lit colour maps without CG lighting (19 participants)).

In addition to appeal, we asked our participants six further questions (Tab. 4) to see whether our stylization technique had a significant effect on other character traits⁴.

This resulted in a total of 7 questions being asked for 28 avatars (2 styles \times 3 stylization levels) + 4 original body scans (2 male and 2 female). We additionally asked the participants who saw the avatars with pre-lit colour maps to rate realism. The avatars were presented in a random order. Questions were asked for each avatar in the same order as in Table 4.

Task 2: Interactive Creation of Appealing Characters

In a second task, 37 participants were asked to specify the exact level of stylization they found to be most appealing/repelling by interactively adjusting a slider in real-time within the Unity 4 game engine.

In addition to the stylized result controlled by the slider, the participants were also shown the original 3D body scan next to the avatar they were manipulating,

⁴ Our questions were inspired by existing research, e.g. *appeal* [11, 10] and *sympathy* [5], *neuroticism* inspired by the Big Five personality traits [7].

Table 4: Avatar questionnaire

1	Do you feel as if you can sympathise with this person?
2	Are you repelled by this person?
3	Would you accept help from this person?
4	Do you find this person engaging?
5	How appealing do you find this person?
6	How charming do you find this person?
7	Does this person come across as being neurotic?

providing them with a non-stylized reference. Albedo maps and CG lighting were used for this task. The dependent variable was the chosen percentage of stylization and the fixed factor was the style (Marvel and Disney) and the actor identity (F_1 , F_2 , M_1 and M_2). Each trial always began with the slider value set to 0% (i.e. showing the original body shape). This resulted in eight trials (two styles for each of the four actor identities). They were presented in random order and the instructions were as follows:

“Make this person as *appealing* as possible.”, and “Make this person as *repelling* as possible.”

For this task participants were instructed to take their time and be as accurate as possible. Participants were also instructed to adjust the slider once more after deciding on a level of stylization to verify their decision before submitting a rating.

5.2 Experiment 2: Task 1: Results

For the most part responses to our questions show a similar pattern across percent stylized for all actors: an increase in rating at 33% and a decrease in rating at 100% (as shown in Fig. 9). We decided to analyze only appeal and realism ratings as appeal seems to be representative of the majority of the questions in Table 4. Indeed the mean appeal rating did not significantly differ from all other questions ($p < 0.05$) (with the exceptions only being neurotic and repelling that were inverted in scale). Realism appears to be the only question that demonstrates a unique pattern across percent stylized so we analyzed realism ratings separately.

Appeal Rating Results

Average appeal rating for the original female body scans was 2.69. The average appeal ratings for female style templates was 3.49 for Marvel and 3.19 for Disney. The average appeal ratings for the original male body scans was 3.08 (slightly higher than the female original body scans). The average appeal ratings for the male style templates was slightly lower than the female styles at 3.2 for Marvel and 2.89 for Disney (see Tab. 5 and Fig. 10). In relative terms, our method achieves approximately 34% improvement in appeal ratings for the original female body scans by stylizing at 33% or 66% towards Marvel or Disney styles and approximately 8% improvement in appeal ratings for the original male body scans by stylizing 33% towards Disney or 66% towards Marvel styles.

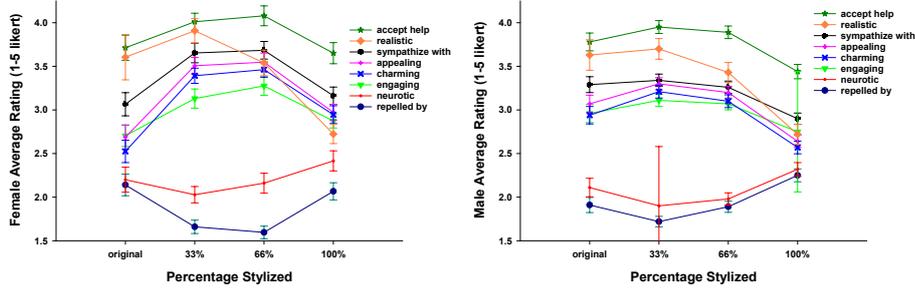


Fig. 9: Female (left) and male (right) ratings on a 5-pt Likert scale for all questions, collapsed across all styles, identities and render quality (n=61, n=19 for realism). Note: Repelling and neurotic ask for negative traits and therefore have inverted values.

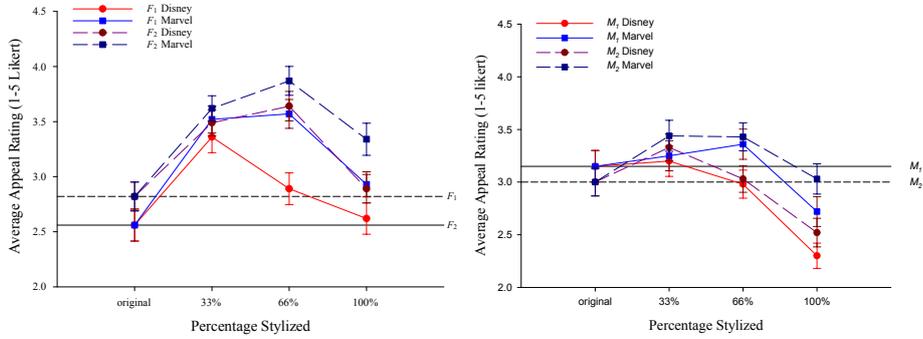


Fig. 10: Female (left) and male (right) appeal ratings on a 5-pt Likert Scale, collapsed across both render styles (n=61).

Since the style templates and actors are unique for female and males we ran separate repeated measures ANOVA on the rating of appeal for male and female avatars with three within-subject factors: Actor Identity (2 levels, F_1 , F_2/M_1 , M_2), Style (2 levels: Marvel, Disney) and Percent Stylized (4 levels: original, 33, 66, 100%) and two between subject factors of Participant Gender (2 levels: male, female) and of Render Quality (2 levels: albedo maps with CG lighting, pre-lit colour maps without CG lighting).

Surprisingly, for both male and female avatars we found no significant effect of render style on appeal ratings (the albedo maps with CG lighting versus the pre-lit colour maps without CG lighting stimuli), Female: $F(1,57) = 0.833, p = 0.365, \eta_p^2 = 0.014$, Male: $F(1,57) = 0.933, p = 0.338, \eta_p^2 = 0.02$. We also found no effect of participant gender on appeal ratings for both the female avatars: $F(1,57) = 0.152, p = 0.698, \eta_p^2 = 0.003$ and the male avatars: $F(1,57) = 0.022, p = 0.883, \eta_p^2 = 0.000$.

Table 5: Female and male appeal ratings (standard deviations in parentheses).

Level	Actor F_1		Actor F_2		Actor M_1		Actor M_2	
	Marvel	Disney	Marvel	Disney	Marvel	Disney	Marvel	Disney
original	2.56(1.13)	2.56(1.13)	2.82(1.02)	2.82(1.02)	3.15(1.18)	3.15(1.18)	3.00(1.03)	3.00(1.03)
33%	3.52(0.96)	3.36(1.11)	3.62(0.90)	3.49(0.92)	3.25(1.12)	3.20(1.14)	3.44(1.15)	3.33(1.00)
66%	3.57(1.02)	2.89(1.13)	3.87(1.02)	3.64(1.05)	3.36(1.24)	2.98(1.04)	3.43(1.04)	3.03(1.10)
100%	2.93(0.89)	2.62(1.11)	3.34(1.14)	2.89(1.02)	2.72(1.11)	2.30(0.94)	3.03(1.13)	2.52(1.06)

Similar to experiment 1, we found a significant effect of percent stylized on appeal ratings for female avatars: $F(3,171) = 22.044, p < 0.001, \eta_p^2 = 0.279$ and male avatars: $F(3,171) = 15.267, p < 0.001, \eta_p^2 = 0.211$. Post-hoc comparisons using Bonferroni adjustments revealed that the original female body scans are rated significantly lower than 33% and 66% ($ps < 0.001$), but not 100% stylized female avatars ($p = 0.291$). Also 33% and 66% stylized female avatars do not significantly differ from each other on appeal ratings ($p = 1.0$), while 100% stylized female avatars are significantly lower rated on appeal than both 33% and 66% female stylized avatars ($p < 0.001$, as revealed through post-hoc comparisons using Bonferroni adjustments). The difference in appeal can be seen as an increase in appeal from the original to 33% or 66% stylized female avatars, followed by a reduction in appeal at 100% to similar appeal ratings as the original female body scans. Post-hoc comparisons using Bonferroni adjustments revealed that 33% stylized male avatars are rated significantly higher on appeal than the original male body scans ($p = 0.041$) and 100% stylized is significantly lower rated on appeal than the original body scans, 33% and 66% stylized male avatars ($p < 0.05$). The difference in appeal can be seen as a slight increase in appeal from the original to 33% stylized male avatars, followed by a slight decline at 66% and a steeper decline in appeal with 100

For males and females, we also found a significant effect of style on appeal ratings, females: $F(1,57) = 15.885, p < 0.001, \eta_p^2 = 0.218$ and males: $F(1,57) = 15.184, p < 0.001, \eta_p^2 = 0.210$. For both males and females, the gender respective Marvel template is significantly higher rated on appeal than the gender specific Disney template (as revealed through post-hoc comparisons using Bonferroni adjustments, $p < 0.001$). Only for the females did we find a significant effect of actor identity, females: $F(1,57) = 8.237, p < 0.05, \eta_p^2 = 0.126$ and males: $F(1,57) = 0.253, p = 0.617, \eta_p^2 = 0.004$. A stylized female F_1 is significantly lower rated on appeal than a stylized female F_2 (as revealed through post-hoc comparisons using Bonferroni adjustments, $p < 0.001$).

These effects suggest that both percent stylized and style influence the ratings of appeal of male and female avatars and for females there is an additional influence of actor identity. For the female avatars, these main effects were conditioned upon a significant interaction between actor identity, style and percent stylized, $F(3,171) = 3.45, p = 0.035, \eta_p^2 = 0.055$ and a for both male and female avatars there was a significant interaction between style and percent stylized, Female: $F(2,118) = 5.34, p < 0.01, \eta_p^2 = 0.083$ and Male: $F(2,118) = 4.27, p = 0.016, \eta_p^2 = 0.067$. The Disney F_1 style at 66% seems to drive the interaction between actor identity, style and percent stylized. This one data point deviates from the typi-

cal pattern of improved appeal at 33%, greater improvement at 66% and then a decrease in appeal at 100% stylized, as revealed through post-hoc comparisons using Bonferroni adjustments, $p < 0.001$. The male Disney avatars stylized at 100% have significantly lower appeal ratings than all other data points, where Marvel does not always decrease in appeal at 100% stylized.

Realism Rating Results

Realism ratings for the original female body scans was on average 3.61. Average realism for the stylized female avatar at 33% was 3.91, at 66% was 3.54, and at 100% was 2.72.

Realism ratings for the original male body scans was on average 3.64. Realism ratings were quite similar for the male avatars as compared to the female avatars. Specifically, average realism for the stylized male avatars at 33% was 3.7, at 66% was 3.44, and at 100% was 2.73 (see Fig. 11 and Tab. 6).

Interestingly for the all avatars at 33% stylized there was no decrease in ratings of realism as compared to the original scans, at 66% only the female F_1 Disney style and the male Disney style received a decreased realism rating. All 100% stylized body scans (both male and female) received a significantly lower realism rating.

We ran two separate repeated measures ANOVA on the rating of realism for both male and female avatars with three within subject factors: Actor Identity (2 levels, F_1 , F_2/M_1 , M_2), Style (2 Levels: Marvel, Disney) and Percent Stylized (3 levels: 33, 66, 100) and one between subject factor: Participant Gender (2 levels: male, female). Since we had not asked for realism ratings for the albedo colour maps there is no between subject factor for ratings of realism.

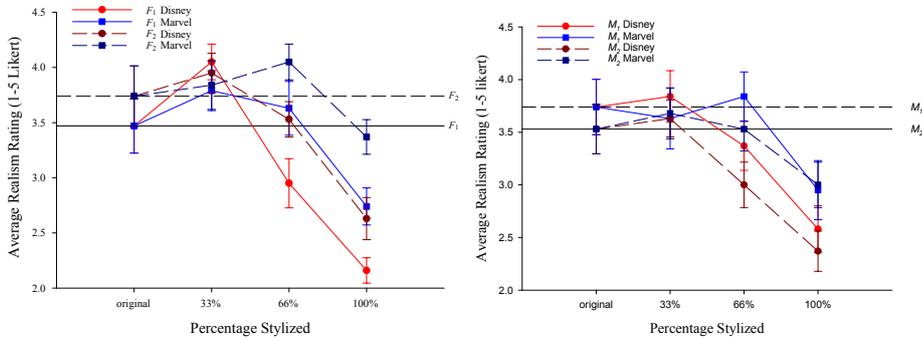


Fig. 11: Female (left) and male (right) realism ratings on a 5-pt Likert scale (n=19,17 respectively).

We found no effect of participant gender on realism ratings for both the female avatars: $F(1,17) = 0.196$, $p = 0.664$, $\eta_p^2 = 0.011$ and the male avatars: $F(1,17) = 0.004$, $p = 0.949$, $\eta_p^2 = 0.000$.

Table 6: Female and male realism ratings (standard deviations in parentheses).

Level	Actor F_1		Actor F_2		Actor M_1		Actor M_2	
	Marvel	Disney	Marvel	Disney	Marvel	Disney	Marvel	Disney
original	3.47(1.07)	3.47(1.07)	3.74(1.09)	3.74(1.09)	3.75(1.15)	3.75(1.15)	3.53(1.02)	3.53(1.02)
33%	3.79(0.79)	4.05(0.71)	3.84(0.96)	3.95(0.78)	3.63(1.26)	3.84(1.07)	3.68(1.06)	3.63(0.76)
66%	3.63(1.07)	2.95(0.97)	4.05(0.71)	3.53(0.70)	3.84(1.02)	3.37(0.01)	3.53(0.91)	3.00(0.94)
100%	2.74(0.73)	2.16(0.50)	3.37(0.68)	2.63(0.83)	2.95(1.22)	2.58(0.96)	3.00(0.94)	2.37(0.83)

We found a significant effect of percent stylized on realism ratings for the female avatars: $F(3,51) = 9.917$, $p < 0.001$, $\eta_p^2 = 0.368$ and the male avatars: $F(3,51) = 14.846$, $p < 0.001$, $\eta_p^2 = 0.5466$. Post-hoc comparisons using Bonferroni adjustments revealed that there was a significant decrease in realism between 100% stylized and the 33% and 66% stylized female avatars ($p < 0.001$), and a significant decrease in realism between 100% stylized and the original, 33% and 66% stylized male avatars ($p < 0.01$). All other comparisons were not significantly different. For both males and females, partially stylized avatars had quite similar ratings for realism as compared to the original body scans, while for the male avatars realism was significantly lower at 100% and for the female avatars even 100% stylized did not have significantly lower average realism than the original female body scans ($p = 0.07$).

We found a significant effect of style on realism ratings for the female avatars, $F(1,17) = 35.898$, $p < 0.000$, $\eta_p^2 = 0.679$ and the male avatars, $F(1,17) = 6.045$, $p < 0.05$, $\eta_p^2 = 0.262$. For both male and female avatars the gender respective Marvel template is rated to have a significantly higher realism rating compared to the gender respective Disney template (as revealed through post-hoc comparisons using Bonferroni adjustments, $p < 0.05$).

Again similar to the appeal ratings, only for the females did we find a significant effect of actor identity, female avatars: $F(1,17) = 19.242$, $p < 0.001$, $\eta_p^2 = 0.531$, male avatars: $F(1,17) = 3.259$, $p = 0.089$, $\eta_p^2 = 0.161$. Consistent with appeal ratings, a stylized female F_1 is rated significantly lower for realism than a stylized female F_2 (as revealed through post-hoc comparisons using Bonferroni adjustments, $p < 0.001$).

For the both the female and male avatars, these main effects were conditioned upon a significant interaction between style and percent stylized, female avatars: $F(3,51) = 13.751$, $p < 0.001$, $\eta_p^2 = 0.447$; male avatars: $F(3,51) = 7.67$, $p < 0.001$, $\eta_p^2 = 0.311$ and for the female avatars a significant interaction between actor identity and percent stylized $F(2,36) = 6.610$, $p < 0.005$, $\eta_p^2 = 0.269$. For the male avatars, this can be summarized by Disney having a more negative effect on realism than Marvel at increasing percent stylized. For the female avatars, these interactions can be summarized by the styles not changing realism ratings at 33% stylized, but as stylization increases, the styles and the actors are being perceived with a variety of different ratings of realism. High levels of Marvel stylization result in decreased realism ratings for F_1 but not so for F_2 , where the Marvel stylization seems to have the same influence on realism ratings across both female actors. These results are consistent with our appeal ratings and may

shed some light on why 100% stylization is seen as less appealing, because it is also seen as less realistic.

5.3 Experiment 2: Task 2: Interactive Creation of Avatars Results

When asked to interactively create the most appealing female avatar using sliders it was extremely rare that participants chose either the original body scan (3%) or the fully stylized avatar (5% of the trials). The average stylization amount chosen was 40.9% across actors and styles for an appealing avatar (see Fig. 13 and Tab. 7).

When asked to interactively create the most appealing male avatar using sliders the most common response for participants was to choose the original male body scan (15.5%) while the fully stylized avatar was rarely chosen (4.7% of the trials). The average stylization amount chosen was 35.6% across actors and styles for an appealing male character.

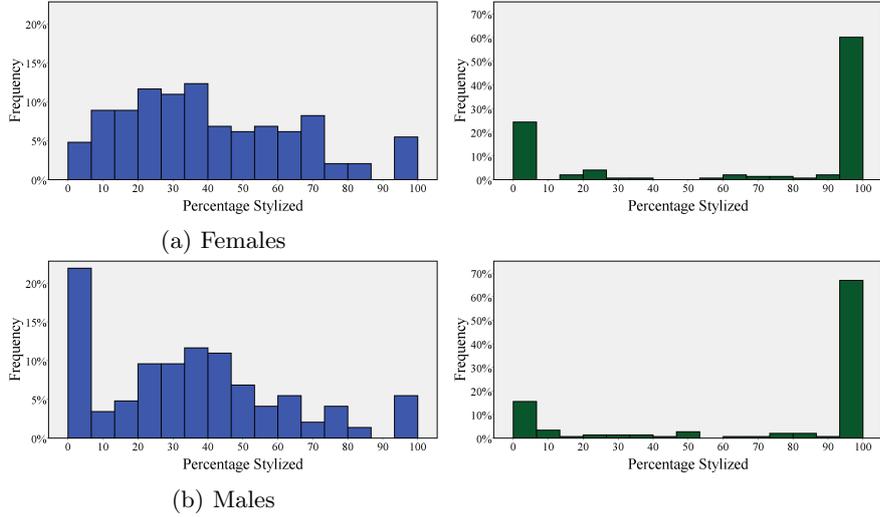


Fig. 12: Frequency of stylization level chosen interactively for the most appealing avatar (left) and most repelling avatar (right), collapsed across actor identity and style (n=37).

Table 7: Average stylization level in the method of adjustment task for most appealing character per actor and style (standard deviations in parentheses).

Style	F_1	F_2	M_1	M_2
Marvel	50.77% (26.3)	45.56% (24.90)	44.38% (27.30)	51.03% (36.93)
Disney	37.65% (23.3)	29.71% (22.02)	16.96% (17.60)	30.20% (23.04)

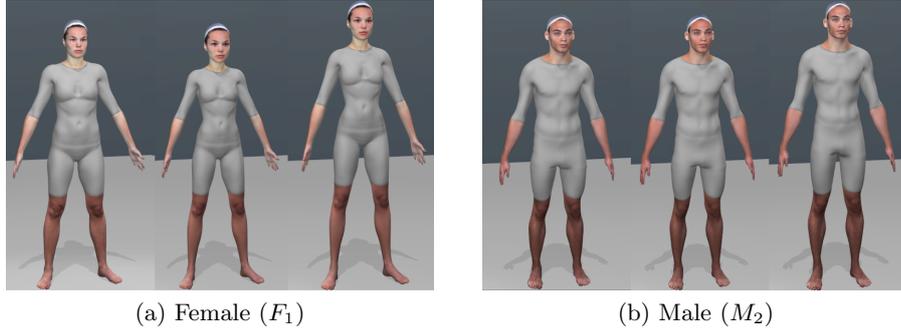


Fig. 13: Female avatars: Original body scan F_1 (left), 38% Disney (center), 51% Marvel (right). Male avatars: Original body scan M_2 (left), 30% Disney (center) 51% Marvel (right).

Due to bimodal distributions for the repelling avatars for both male and female, we decided to present the histograms for the percent stylized chosen (see Fig. 12). When participants were asked to create the most repelling female avatar they chose the 100% stylized version for 58% of the total trials and the original scan for 24% of the total trials. When asked to create the most repelling male avatar they chose the 100% stylized version for 64.2% of the total trials and the original scan for 13.5% of the total trials. This further confirms that the most appealing avatar is a partially, and specifically not a fully, stylized one. Overall, these results are also consistent with the repelling ratings from task 1. However, the degree to which the original body scans and 100% stylized male and female avatars were rated as repelling was at most only just above 2 on the 5-pt Likert scale (see Fig. 9). Important therefore to note is that original body scans and 100% stylized avatars were not seen as extremely repelling, but rather more repelling as compared to partially stylized avatars.

6 Discussion

For experiment 1, appeal ratings for the average male and female body shapes were highest when some level of stylization was applied (usually 33% or 66% but in some cases 100%). However, not all styles affect the resulting appeal score in a similar manner. For females, the Disney and Marvel styles were found to be consistently more appealing than the average body shapes (peaking at 66%). Styles with exaggerated proportions tended to be rated lower than styles with more human-like proportions, with the Barbie and Pixar styles producing lower appeal ratings than the Disney and Marvel styles. The Sony style (featuring the most exaggerated proportions) was rated as being consistently less appealing than the average female body shape regardless of the level of stylization applied. Male results show a slightly different pattern with the Marvel and Disney styles being considered significantly more appealing than the average male body shape and receiving the highest appeal ratings at 100%. However, the exaggerated

muscular style (Dreamworks 1), while most appealing at 33%, was perceived as being only slightly more appealing than the average male body shape at 100%. The other exaggerated male styles (Sony and Dreamworks 2) were both rated to be consistently less appealing than the average male body shape. For both males and females, our results seem to suggest a preference for less exaggerated body proportions.

For our second experiment involving the stylization of body scans featuring colour information, our stylization technique proved to be less effective for males. While a partially stylized avatar was rated as being most appealing for both males and females, the increase in appeal ratings was highest for females with even the fully stylized avatars receiving similar, or in some cases higher appeal ratings than the original actors. Interestingly, the male Marvel and Disney styles originally rated as being most appealing at 100% in experiment 1 had the opposite effect when being applied to our male actors and produced the least appealing avatars at 100%. This could be due to the already muscular physique of actors M1 and M2 compared to the average body shape used in experiment 1, resulting in a higher initial appeal rating and a less obvious change in body shape when being stylized.

We found that efforts to make the avatars appear more realistic (through improvements in render quality) had no significant impact on appeal ratings. Considering our results, it could be that in contradiction to previous work [20], body shape is playing a larger role in the influence of perceived appeal than the inclusion (and modification) of surface colour, however further testing is needed to better understand this. Realism ratings displayed a similar pattern to appeal ratings and may provide some insight as to why higher levels of stylization are less desirable than partial stylization, as the highest level of stylization was rated lowest in realism. Interestingly, 33% and in some cases even 66% stylized body scans were not seen as less realistic as compared to the original body scans. This could be due to the stylization of the colour map drawing attention away from imperfections in the surface colour. The fact that there is not an immediate decrease in realism with stylization also is a result that further supports the use of stylization to create more appealing avatars.

We hope to further investigate the perception of realism in more depth to better understand the role realism plays in the appeal of our avatars and specifically further investigate why partial stylized avatars receive similar realism ratings as our original body scans.

A limitation of our findings is the small number of templates and body registrations presented in this paper and is also considered important for future research.

7 Summary

The stylization technique presented in this paper allows for the semi-automatic creation of more appealing personalized avatars from 3D body scans for use in games, virtual environments and online communications.

For stylizing 3D body scans, the main steps are the automatic registration of a statistical 3D body model to the body scan data, followed by the stylization through feature-preserving morphing with colour filtering and blending. Our technique can be automatically applied to any human body scan after registration of the scan to our 3D body template. The only manual process involved is the initial sculpting of the style templates. However, once created, these can be applied to an infinite number of body scans and based on the individual templates, can be applied at the optimal level of stylization required to maximize the appeal of these body scans. The rest of the process is fully automated. Based on the results presented in this paper, the proposed system can increase appeal ratings by approximately 34% for females (ranging from 29-39%) and 8.5% for males (ranging from 1.6% to 14.7%) depending on the actor and style.

We conducted multiple experiments in which participants were asked to rate the appeal of different avatars. Our results demonstrate that some stylization (approximately 48% for female and 44% for male) is perceived as most appealing on average across all styles and actors. This suggests that 3D body scans can be made to look more appealing with some level of stylization applied. Finally, when we asked our participants to interactively create the most appealing and repelling avatars from 3D body scans, we found again that a partially stylized avatar is the most appealing at approximately 41% for females and 36% for males. In contrast, when considering the percent of stylization people used to create repelling avatars, we found that 82% of the chosen female avatars and 78% of the chosen male avatars were made to be either 100% stylized or the original body scans themselves. For both male and female, a highly stylized avatar was most often chosen as the most repelling avatar.

Our results show that partially stylized male and female 3D body scans are perceived as most appealing as compared to the original and fully stylized body scans.

Acknowledgments We would like to thank Chris Ferguson, Joe Smallwood, Anna Wellerdiek, Alex Holmes and Michael Geuss as well as all the participants of our experiments.

References

1. Boberg, M., Piippo, P., Ollila, E.: Designing avatars. In: Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts. pp. 232–239. DIMEA '08, ACM, New York, USA (2008), <http://doi.acm.org/10.1145/1413634.1413679>
2. Bogo, F., Romero, J., Loper, M., Black, M.J.: FAUST: Dataset and evaluation for 3D mesh registration. In: Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). IEEE, Piscataway, NJ, USA (Jun 2014)
3. Chaminade, T., Hodgins, J., Kawato, M.: Anthropomorphism influences perception of computer-animated characters actions. *Social Cognitive and Affective Neuroscience* 2(3), 206–216 (2007), <http://scan.oxfordjournals.org/content/2/3/206.abstract>

4. Hirshberg, D., Loper, M., Rachlin, E., Black, M.: Coregistration: Simultaneous alignment and modeling of articulated 3d shape. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) *Computer Vision – ECCV 2012*, Lecture Notes in Computer Science, vol. 7577, pp. 242–255. Springer Berlin Heidelberg (2012), http://dx.doi.org/10.1007/978-3-642-33783-3_18
5. Hodgins, J., Jrg, S., O’Sullivan, C., Park, S.I., Mahler, M.: The saliency of anomalies in animated human characters. *ACM Trans. Appl. Percept.* 7(4), 22:122:14 (2010), <http://doi.acm.org/10.1145/1823738.1823740>
6. Inkpen, K.M., Sedlins, M.: Me and my avatar: Exploring users’ comfort with avatars for workplace communication. In: *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*. pp. 383–386. CSCW ’11, ACM, New York, USA (2011), <http://doi.acm.org/10.1145/1958824.1958883>
7. Johnson, J.A., Ostendorf, F., Johnson, J.A., Psycholog, D., State, P.: Clarification of the five-factor model with the abridged big five dimensional circumplex. *Journal of Personality and Social Psychology* pp. 563–576 (1993)
8. Johnston, O., Thomas, F.: *Disney Animation: The Illusion of Life*. Abbeville Press (1981)
9. MacDorman, K.F., Green, R.D., Ho, C.C., Koch, C.T.: Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior* 25(3), 695–710 (2009), <http://www.sciencedirect.com/science/article/pii/S0747563208002379>
10. McDonnell, R.: Appealing virtual humans. In: Kallmann, M., Bekris, K. (eds.) *Motion in Games*, Lecture Notes in Computer Science, vol. 7660, pp. 102–111. Springer Berlin Heidelberg (2012), http://dx.doi.org/10.1007/978-3-642-34710-8_10
11. McDonnell, R., Breidt, M., Bülthoff, H.H.: Render me real: Investigating the effect of render style on the perception of animated virtual humans. *ACM Trans. Graph.* 31(4), 91:1–91:11 (Jul 2012), <http://doi.acm.org/10.1145/2185520.2185587>
12. Melina, R.: Is ‘Mars Needs Moms’ Too Realistic? (2011), <http://www.livescience.com/33198-is-mars-needs-moms-too-realistic.html>, accessed: 2014-06-18
13. Microsoft: *Kinect Sports Rivals*. Computer Game (2014)
14. Mori, M., MacDorman, K., Kageki, N.: The Uncanny Valley [from the field]. *IEEE Robotics Automation Magazine* 19(2), 98–100 (2012)
15. Robinette, K.M., Blackwell, S., Daanen, H., Boehmer, M., Fleming, S.: *Civilian american and european surface anthropometry resource (CAESAR)*, final report. volume 1. summary. Tech. rep., US Air Force (2002)
16. Seyama, J., Nagayama, R.S.: The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments* 16(4), 337–351 (2007), <http://dx.doi.org/10.1162/pres.16.4.337>
17. Spielberg, S.: *The Adventures of Tintin*. Animated Feature Film (2011)
18. *Take-Two Interactive: NBA 2K15*. Computer Game (2014)
19. Wells, S.: *Mars Needs Moms*. Animated Feature Film (2011)
20. Zell, E., Aliaga, C., Jarabo, A., Zibrek, K., Gutierrez, D., McDonnell, R., Botsch, M.: To stylize or not to stylize?: The effect of shape and material stylization on the perception of computer-generated faces. *ACM Trans. Graph.* 34(6), 184:1–184:12 (Oct 2015), <http://doi.acm.org/10.1145/2816795.2818126>
21. Zemeckis, R.: *The Polar Express*. Animated Feature Film (2004)