Render me Real? Investigating the Effect of Render Style on the Perception of Animated Virtual Humans

Rachel McDonnell*
*Trinity College Dublin

Martin Breidt†
†Max Planck Institute for Biological Cybernetics

Heinrich H. Bülthoff†‡
‡Korea University

Figure 1: *Male avatar rendered in different visual styles, ranging from realistic to abstract, based on the results in Section 5.*

Abstract

The realistic depiction of lifelike virtual humans has been the goal of many movie makers in the last decade. Recently, films such as *Tron: Legacy* and *The Curious Case of Benjamin Button* have produced highly realistic characters. In the real-time domain, there is also a need to deliver realistic virtual characters, with the increase in popularity of interactive drama video games (such as *L.A. Noire™* or *Heavy Rain™*). There have been mixed reactions from audiences to lifelike characters used in movies and games, with some saying that the increased realism highlights subtle imperfections, which can be disturbing. Some developers opt for a stylized rendering (such as cartoon-shading) to avoid a negative reaction [Thompson 2004]. In this paper, we investigate some of the consequences of choosing realistic or stylized rendering in order to provide guidelines for developers for creating appealing virtual characters. We conducted a series of psychophysical experiments to determine whether render style affects how virtual humans are perceived. Motion capture with synchronized eye-tracked data was used throughout to animate custom-made virtual model replicas of the captured actors.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—[Rendering]

Keywords: perception, uncanny valley, facial animation, motion capture

Links: DL PDF

1 Introduction

Choosing an appropriate rendering style for virtual humans in movies, games, and other domains can be a challenging task. Using near-human characters can be considered risky, as it is not only costly but can produce a negative audience reaction [Geller 2008; Levi 2004]. This effect may be one of the reasons why many studios use cartoon stylizations rather than photorealistic humans to depict their characters. Knowing the correlation between how realistic and how appealing characters of different render styles are perceived to be could help studios in making their decisions. There are many other applications besides the movie and game industry who could benefit from this knowledge. For example, in advertising it could be very advantageous to determine which human representations would best convince viewers to purchase their products. Furthermore, in virtual training applications, knowing which render style students would be more likely to trust could help with training success.

In this paper, we investigate if using realistic rendering does in fact produce a more negative response than using lower quality or stylized rendering. More specifically, we consider: Are there differences in how truthful we perceive real and virtual humans? Does rendering style influence trust? If we choose a stylized rendering, will this affect how appealing, friendly, or re-assuring the character appears? In order to do so, we perform a series of psychophysical experiments, where a range of 11 different render styles are applied to identical geometry and motion pairs.

*rachel.mcdonnell@cs.tcd.ie
†martin.breidt@tuebingen.mpg.de
‡heinrich.buelthoff@tuebingen.mpg.de

In our Explicit Experiments (Section 5), we allow participants to rate the realism of characters rendered using 10 different styles. We also collected participants' direct impressions of the characters, by asking them to rate how re-assuring, familiar, appealing, friendly, and trustworthy they found them to be. These ratings were collected separately on moving and still characters. This allowed us to determine the relationship between realism and appeal for virtual characters, and whether motion had an effect on the ratings. We found that highly realistic and highly abstract styles were considered appealing, whether still or moving. However, render styles that were in the middle of the range between abstract and realistic were considered unappealing, and the most unappealing character was considered even more so when motion was applied. Furthermore, a lack of appeal occurred for characters that were considered unfamiliar. We also determine whether or not animation artifacts are more forgiving on less realistic or stylized characters. We found that characters that were textured using human photographs appeared more unpleasant than stylized characters, when large motion anomalies were present.

In our Implicit Experiments (Section 6), we perform a more indirect test of character appeal using a lie detection paradigm. Using recorded truths and lies from two human actors and their virtual replicas, we determine if the style in which the characters are rendered has an effect on how truthful they are perceived to be. Also, if there are differences in ratings between real and virtual characters. Overall, we found that the audio track contained the most reliable cues for deception detection, and without this, participants were performing at chance level. With audio and visual cues, participants relied on visual more than auditory cues for the appealing virtual avatars. Furthermore, we found that rendering style did not bias 'truth' responses, and that lies and truths were perceived on real and virtual characters in the same way. In Section 8, we provide guidelines for developers, based on our results.

2 Related Work

The complex interaction between motion and appearance has been examined in previous work. Hodgins et al. [1998] found that viewers' perception of motion characteristics is affected by the geometric model used for rendering. They observed higher sensitivity to changes in motion when applied to a polygonal model, than a stick figure. Chaminade et al. [2007] investigated how the appearance of a range of characters influenced perception of their actions. They found that anthropomorphism decreased the tendency to report their motion as biological. In [McDonnell et al. 2009], we investigated the effect of body shape on the perception of bodily emotion. We found that perception of emotion is highly robust and mostly independent on the character's body appearance. More recently, Hodgins et al. [2010] conducted perceptual experiments to determine how degradation of human motion affects the emotional response of participants to animation. They found that removing facial animation and/or sound changed the emotional content that was communicated to their participants. The effect of render style was not investigated in their experiment.

The Uncanny Valley (UV) has become a standard term for the hypothesis that near-photorealistic virtual humans often appear unintentionally eerie or creepy. This UV phenomenon was first hypothesized by robotics professor Masahiro Mori [1970]. Mori predicted that as a robot looks more human it also looks more agreeable, until it appears so human that we begin to find subtle imperfections unsettling. He also hypothesized a stronger effect for moving versus still objects. This supposed negative response has been attributed to many causes such as motion artifacts or lack of familiarity. More recently, the UV hypothesis has been transfered to human avatars in computer graphics, and has been explored directly in some studies [MacDorman et al. 2009; Bartneck et al. 2007]. Hanson et al. [2005] conducted a web-based survey showing a range of humanoid depictions, from cartoon-like to realistic. Users in their study found each of the images to be positively acceptable, and their reaction never dipped into the negative region.

Previous work from the fields of psychology and neuroscience has shown that different regions of the brain are activated when presented with real and virtual stimuli. For example, Perani et al. [2001] conducted a Positron Emission Tomography scan study where participants viewed sequences of a real hand, a realistic virtual reproduction, and a low quality virtual hand. They provided evidence of the difference in neural processing between real and virtual stimuli. However, they found only a limited effect of the level of realism of the virtual hands. Han et al. [2005] measured brain activity using functional Magnetic Resonance Imaging (fMRI) while participants viewed cartoons or live action movies. The authors conclude that their findings could suggest that the human brain functions in a different way when interacting with real people than with artificial or cartoon characters. Mar et al. [2007] conducted an fMRI study using footage derived from the film *Waking Life* as stimuli. Participants saw identical biological movement for the stylized and live-action footage. They found that brain activation was higher in live-action portrayals of social interaction, which implied that participants were more engaged when viewing the live-action sequences.

More recently, Chen et al. [2010] found that adaptation to cartoon faces with large eyes shifted participants' preferences for human faces towards larger eyes. Also, Saygin et al. [2012] conducted an fMRI experiment where participants viewed a human and her robotic replica with and without silicon skin performing recognizable tasks. They found that brain activity was higher for the human-like robot and attributed this higher activation to the incongruence between motion and appearance. Other studies have shown that people can respond socially to human and nonhuman entities [Slater and Steed 2002; Reeves and Naas 1996] and that they can engage with virtual humans whether or not they look human [Nowak and Biocca 2003].

It has been shown that certain psychological responses, including emotional arousal, are commonly generated by deceptive situations [DePaulo et al. 2003]. Experimental literature indicates that people can discriminate lies from truths with above-chance levels of accuracy [DePaulo et al. 2003]. For virtual avatars, Steptoe et al. [2010] assessed the impact of the addition of realistic eye motion in avatar mediated communication. They found that the addition of eye movement increases participant accuracy in detecting truth and deception when interacting with virtual avatars.

Figure 2: *Performance Capture with synchronized body, face, eye motion, audio, and video reference.*

3 Performance Capture

We decided to use motion-capture technology to animate the faces of our virtual characters, since we wanted to precisely match the timing and intensity of the real motion. Furthermore, we did not want the interpretation of the action by an animator to be an extraneous factor. The first step was to conduct a series of performance capture sessions to collect motion and audio data from real people, and to create virtual models to apply the motion to. We opted to record interview style sequences, where the person answering the questions was recorded, while the interviewer was out of the capture volume.

Two actors participated in each recording session. One took the role of the "interviewer" whose voice alone was recorded. The second was the "interviewee" whose voice, face, body and eye movements were captured (Figure 2). Even though we only intended to view the characters from the shoulders up, we included full body motion, as we felt the shifting of position in the pelvis and the arm, neck and head motion would be important to realistically portray the performance of the actor. Furthermore, as in [Hodgins et al. 2010], we wanted to ensure that we had natural, ecologically valid stimuli that would be similar to those found in real applications. Eye-capture was also included in our dataset as we felt it would increase the fidelity of the recorded performance.

All actors were non-professionals and accustomed to the motion capture setup and environment. A series of questions were asked at random by the interviewer. The questions were selected such that a short answer response would occur, in order to get an immediate reaction from the interviewee. We ensured that questions did not relate to the physical appearance of the character (e.g., "what color are your eyes?"), and could not have multiple answers (e.g., "what is your favorite film?"). The interviewee was told in advance whether to lie or tell the truth to the question. The answers were not rehearsed and the questions were not known in advance to ensure a natural reaction.

Motion capture was conducted using a 13 camera Vicon optical system, where 52 markers were placed on the body and 36 markers on the face. Typical motion-capture artifacts were avoided by ensuring accurate capture with minimal trajectory re-construction. A head-mounted eye-tracking device (SMI Eyelink II) was used to capture the movements of the left eye relative to the head at 250Hz. The eye-tracker is accurate to within 0.5 degrees of visual angle. Before every capture take, calibration was performed using nine fixation targets on a screen that was positioned in the capture area. Once calibration was complete, the actor was free to move away from the screen as we recorded the rotation angle of the eye relative to the head, not to the screen. Two studio condenser microphones were placed near the actors and recorded their voices on two separate tracks.

3.1 Motion Mapping

The body motion (captured at 120Hz) was mapped onto a skeleton, where joint angles were computed and used to drive the virtual character in Autodesk 3ds Max 2012. The facial motion was directly exported as 3D marker motion. In order to avoid any retargeting errors due to slight differences in the head bone animation (as approximated from the joint angles), the facial marker motion was not directly applied to the bones of the characters. The rigid head motion was first removed, using three markers that were mounted on the eye-tracking rig, and moved only with the actor's head. Subsequently, the markers were linked to the head bone in order to follow the motion of the character's head exactly. A bone-based approach that used linear blend skinning was used to drive the facial geometry. We chose this type of deformation (over more sophisticated methods such as [Curio et al. 2006]) as it is the predominant animation method used in computer games. Twenty of the characters' facial bones were then constrained to their corresponding optical markers to produce the animation. Since motion was only mapped from the actor to his/her virtual replica, retargeting was not necessary, and the markers could drive the bones directly.

Finally, the eye rotations were computed and applied directly on to the bones driving the eye balls in 3ds Max. Since we only tracked the motion of the left eye, an additional rotational offset was added to its motion before applying it to the right eye ball. The offset was chosen such that the eye convergence point was approximately at the position of the interviewer's face when the actor looked at her. This approximation worked successfully for our recordings, since the actors kept eye contact with the interviewer during the recording and did not focus on nearby objects.

4 Models and Render Styles

Two virtual replicas were made of the motion captured actors (1 male and 1 female). A 3D scan of each actor's face was taken using a structured light 3D scanner (ABW). This scan was used, along with a series of photographs taken from a series of different angles by an artist to create virtual models of the actors. The virtual models were typical "next gen" game characters, with both facial and body rigs, and high quality diffuse, opacity and normal-map textures.

We created 11 rendering styles in total for use in our experiments. In order to ensure high quality rendering, we commissioned some of the styles from a professional artist working in the games industry. Other styles were created with the guidance of an artist and with plugins for 3ds Max. We differentiate in notation between characters that used photographs of real humans for their diffuse textures (Human) and those that used more cartoon-like shaders and textures (Toon). Figure 3 depicts the first 10 styles, and Figure 9 shows an example of the eleventh style under motion.

- *ToonPencil*: was included as a highly abstract style, in order to determine if a large loss in visual information would affect perception. A high quality line-drawing plugin (Illustrate 5.6) for 3ds Max was used to create the style. There was no lighting/shading information present for this style, just black lines with a white background.
- *ToonFlat*: was a flat-shaded typical cartoon style, generated using Illustrate 5.6. Colors were chosen to match the human diffuse texture maps as closely as possible. No light source was used and no shading information was present.
- *ToonShaded*: a two-color shaded version of ToonFlat, to determine if the extra shading detail affected perception.
- *ToonBare*: this style was constructed using standard Phong shading on a character with diffuse color only and no texture maps.
- *ToonCG*: This style was created using a two pass render by an artist working in the games industry. We asked the artist for a highly appealing style, similar to that used in computer graphics cartoons in the industry, in order to examine participant reactions to a highly appealing character. The style used subsurface scattering and painted textures to create a soft skin. A blue rim light was composited to add depth and appeal to the character, a commonly used technique in the industry.
- *HumanIll*: Negative audience reaction to characters has often been attributed to the appearance of disease or death. Therefore, a character that we felt looked ill was included, in or-

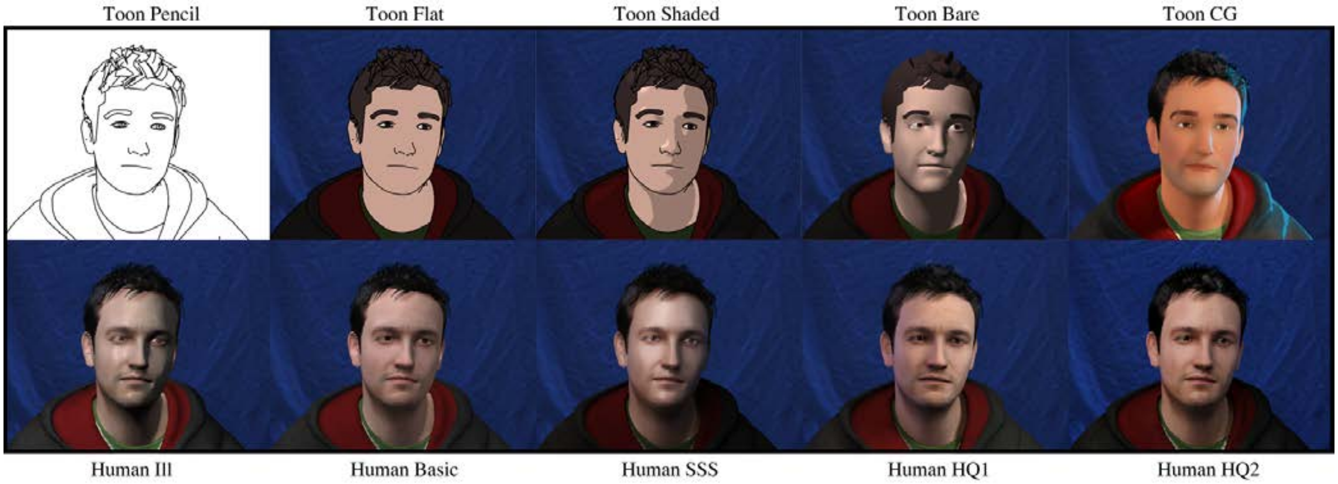

Figure 3: *The ten render styles used in our experiments, ranging from abstract to realistic.*

der to determine participant reactions to a highly unappealing character. We desaturated and added a yellow hue to the diffuse map along with shading the skin of the character using a waxy material. The eyes were glazed over using a semi-transparent glass shader on the cornea.

- *HumanBasic*: This style used only diffuse human textures, using Phong shading with default point light sources in 3ds Max. This style was a typical game-style which produced no shadows or reflections.
- *HumanSSS*: Sub surface scattering was used to shade the skin of the character with area lights to create reflections and shadows. The cornea of the eyes reflected light using a solid glass shader.
- *HumanHQ1*: An artist working in the games industry created this style using three area spotlights, and ray-traced shadows.
- *HumanHQ2*: Characters using this style were shaded using ray-traced area lights, indirect illumination, soft shadows, and physically accurate reflections with a Fresnel-like falloff.
- *HumanWrinkle*: This style was the same as HumanHQ2 but with additional normal maps for producing dynamic wrinkle effects on the skin, driven automatically by the compression of the geometry edges. Wrinkle normal maps were created by an artist using photographs of the actors' faces in poses that created laughter lines, crows feet, and forehead wrinkles (Figure 9). This wrinkle style was included as this feature could be used to identify facial deceit [Ekman and Friesen 2003].

5 Explicit Experiments

There has been much speculation about whether or not it is profitable to invest large resources into trying to achieve photorealistic characters, as the audience reaction to these characters can often be negative. For example in the movie *Polar Express*, characters were realistically rendered yet were frequently considered off-putting when compared to the stylized characters used in movies such as *The Incredibles* [Levi 2004]. In this first set of experiments, we explicitly ask participants to rate different render styles applied to the same character, in order to determine the direct effect of render style on their impression of the character. The first block used only still images, while characters in the second block had motion applied. A between-groups design was used for this experiment where participants in the first and second blocks differed. Based on Mori's theory [1970], we hypothesized that participants would give abstract characters more positive ratings than characters that they considered to have almost photo-realistic appearance. Also, that this effect would be amplified when motion was present.

5.1 Exp 1: Still vs. Moving

In psychology, rating scales are commonly used to gather information on subjective conditions such as attitudes and emotions of participants, which could not be collected directly. We chose this technique in order to collect the subjective judgments of participants on a range of scales. In order to avoid ambiguity in the terms used in the scales, we presented participants with a full list of definitions before the experiment. This list was visible on a monitor next to the experiment, so that participants could refer to it when necessary.

- *Extremely abstract - Extremely realistic*: This scale was used to determine if participants found the image to be photorealistic or abstract.
- *Extremely unappealing - Extremely appealing*: This scale told us if participants would be captivated by a movie with the character as a lead actor, or if they would not find the character appealing in any way.
- *Extremely unfamiliar - Extremely familiar*: We wished to determine if participants had seen something like the images before, or if they were totally unfamiliar with them. This scale was included as negative reactions to characters are often attributed to lack of familiarity.
- *Extremely eerie - Extremely re-assuring*: We based this scale on MacDorman et al. [2009] in their study on uncanny facial images. The scale was included to compare our results with previous studies.
- *Very unfriendly - Very friendly*: We wished to determine if render style could make a character appear more friendly.
- *Very untrustworthy - Very trustworthy*: When asked directly, we wanted to know if participants would rate the different render styles with different levels of trust.

5.1.1 Block 1 – Still Images

For this block, we used the male model rendered in 10 styles, as described in Section 4. The eleventh style with animated wrinkles was not included since the wrinkles would not be visible for a neutral expression without motion. A still image of each of the 10 render styles in three different neutral poses was rendered at a resolution of 720×576 pixels. We chose neutral poses in order to avoid any emotional bias due to the facial expression. In total, 30 still images were used as stimuli in this experiment, and the poses served as repetitions of the styles.

Seventeen volunteers (7F, 10M) took part in this experiment. As with all subsequent experiments, participants were naïve to the purpose of the experiment and from different educational backgrounds. Also, University ethical approval was granted for all experiments, and participants received a book voucher to compensate for their time. Participants were presented with a questionnaire where they rated each image from 1–7 on each of the 6 rating scales. Each stimulus remained on the screen until participants completed the 6 questions and pressed the space-bar to move to the next image. In order to avoid participants seeing the same image one after the other, stimuli were presented in pseudo-random order. This meant that participants viewed a full sequence of all 10 styles in random order before being presented with the next repetition of the 10 styles, etc.

5.1.2 Block 2 – Movies

In this block, motion was applied to the character, in order to determine if motion affected participant ratings. We chose three motion clips from the male actor recorded in the motion capture session described in Section 3. The clips ranged in length from 6–10 seconds. Clips were selected such that the motion looked natural and contained some visible eye, eyelid, and lip motion. The motion was applied to the corresponding character model and rendered in the full 11 styles described in Section 4.

Seventeen (7F, 10M) new volunteers that did not take part in Block 1 were recruited for this experiment. Participants were told that the characters would be speaking in the motion clips, but that they would not hear the audio, to ensure that they based their decisions on visual information alone. As before, they were presented with a questionnaire where they rated each movie from 1–7 on each of the 6 rating scales. The movie looped on the screen until participants completed the 6 questions and pressed the space-bar to move to the next clip. As before, stimuli were presented in pseudo-random order.

5.1.3 Results

The ratings for each participant were averaged repetitions (image or movie clip). We first tested if there was a difference between HumanWrinkle and HumanHQ2 in Block 2, and found no significant difference for any of the rating scales. We did not include HumanWrinkle in the subsequent analysis since there was no corresponding style in Block 1. For each of the rating scales, a two-way repeated measures ANalysis Of VAriance (ANOVA) was conducted on the data from both blocks with within-subjects factor *render style* (10) and between-subjects factor *movement* (2 - still or moving).

Realism

We first analyzed the data for the Realism ratings to determine if our stimuli covered a large range. A main effect of render style was found ($F_{9,288} = 170.94, p \approx 0$). Post-hoc analysis on this and all subsequent experiments was conducted using Newman-Keuls tests for comparison of means, and only significant results at the 95% level are reported. Figure 4 shows that the ToonPencil style was rated as the most abstract (1.5 on the scale), and HumanHQ2 was considered very realistic (6 on the scale). Posthoc analysis showed that six distinct levels of realism were found: 1. ToonPencil; 2. ToonFlat, ToonBare, ToonShaded; 3. ToonCG; 4. HumanIll; 5. HumanBasic, HumanSSS; 6. HumanHQ1, HumanHQ2. No main effect of the presence or absence of motion was found, indicating that the appearance dominated and adding motion did not increase the realism ratings of the characters. Finally, no interaction occurred, implying that each of the render styles were rated in the same manner whether moving or still.

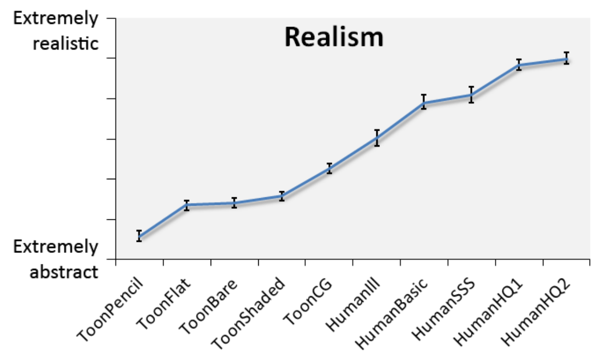

Figure 4: *Main effect of render style on Realism ratings. Error bars show standard error of the mean (as in all other graphs).*

Appeal

The ratings for Appeal were then analyzed using a two-way ANOVA, as before. A main effect of render style was found ($F_{9,288} = 26.54, p \approx 0$). Post-hoc analysis showed that four of the most abstract and the two most realistic styles were rated as the most appealing (ToonPencil, ToonFlat, ToonShaded, ToonCG, HumanHQ1, HumanHQ2) and significantly more appealing than all other styles $p < 0.05$ in all cases). Characters that were rated halfway between abstract and realistic in the Realism ratings, were considered the least appealing, with HumanIll being rated as significantly less appealing than all other styles ($p < 0.0005$ in all cases). HumanBasic and ToonBare were the next least appealing ($p < 0.04$ in all cases), followed by HumanSSS ($p < 0.05$ in all cases). These findings imply that characters that appear highly abstract or highly realistic are considered equally appealing. Furthermore, that *a drop in appeal occurs for characters that are neither abstract nor realistic, and therefore may be difficult to categorize*.

No main effect of movement was found as before, however there was an interaction between movement and render style ($F_{9,288} = 2.4, p < 0.02$). Post-hoc analysis showed that this was due to the fact that HumanIll was rated as significantly less appealing when moving than when still ($p < 0.009$). This implies that *extremely unappealing characters appear even less appealing when moving than when still*. This effect did not occur for any of the other characters, as there was no significant difference between their ratings when still or moving (Figure 5, left). On observing the average ratings in Figure 5, the appealing styles appear to be rated as more appealing with motion applied, but these differences were not significant in our study.

Re-assuring

Our statistical analysis showed a main effect of render style ($F_{9,288} = 38.2, p \approx 0$). As with the Appeal ratings, the most realistic and most abstract styles were considered the most re-assuring (ToonPencil, ToonFlat, ToonShaded, ToonCG, HumanHQ1, Hu-

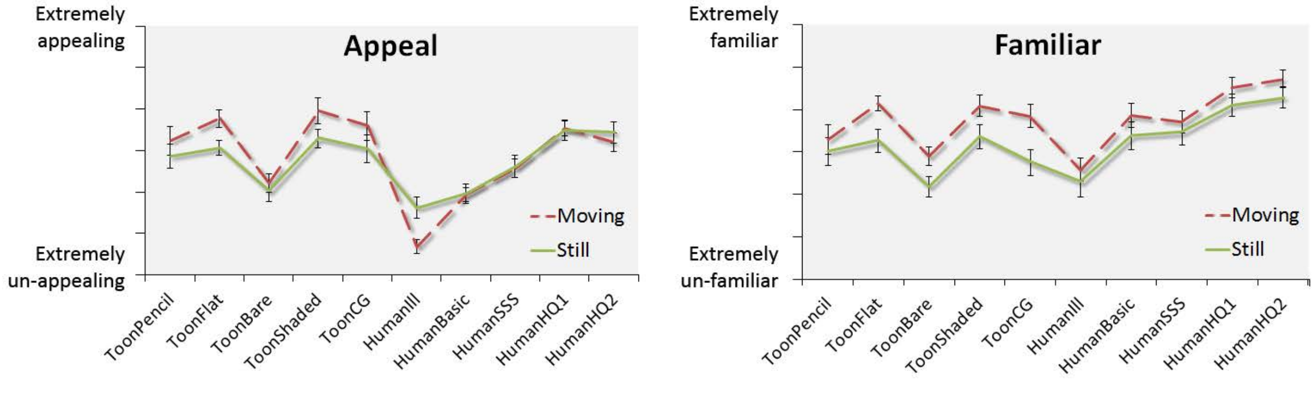

Figure 5: *Interaction between render style and movement on (left) Appeal and (right) Familiarity ratings. Graphs are ordered from the most abstract to the most realistic, as rated in the Realism ratings.*

manHQ2) and equally re-assuring except for ToonFlat and ToonShaded which were considered more re-assuring than HumanHQ2 ($p < 0.05$ in both cases). HumanIll was rated as the most eerie, and significantly more so than all other styles ($p < 0.00004$ in all cases). Finally, HumanBasic, ToonBare, and HumanSSS were rated as the next most eerie ($p < 0.00004$ in all cases). No significant main effect of movement, and no interaction was found.

Familiar

Participants found it difficult to rate the stimuli on the familiarity scale, as evidenced by the narrow range of average ratings in the graph (Figure 5, right). However, an overall main effect of render style was found ($F_{9,288} = 18.9, p \approx 0$). Post-hoc analysis revealed that the most realistic styles (HumanHQ1 and HumanHQ2) were rated as the most familiar to participants ($p < 0.002$ in all cases). HumanIll and ToonBare were rated equally as the least familiar ($p < 0.003$ in all cases). All other styles were rated as just above halfway between familiar and unfamiliar. Unlike the other scales, a main effect of movement was found for familiarly ratings ($F_{1,32} = 4.72, p < 0.04$), where moving stimuli were rated overall as significantly more familiar to participants than still images (Figure 5, right). No interaction occurred between movement and render style.

Friendly

A main effect of render style was observed ($F_{9,288} = 30.45, p \approx 0$). Post-hoc analysis showed that HumanIll and HumanBasic were rated as equally unfriendly and more unfriendly than any of the other render styles ($p < 0.00005$ in all cases). HumanSSS and ToonBare were rated as the next least friendly ($p < 0.009$ in all cases). The most realistic render styles were rated as more friendly than HumanSSS and ToonBare almost as friendly as the rest of the Toons. Finally, the majority of the Toons (ToonPencil, ToonFlat, ToonShaded, and ToonCG) were rated as the most friendly ($p < 0.009$ in all cases). No main effect of movement was found, which shows that the motion did not increase how friendly the character appeared. No interaction was found which implied that participants found the characters equally friendly regardless of whether they were moving or still.

Trustworthy

For the Trustworthy ratings, a main effect of style was also observed ($F_{9,288} = 35.53, p \approx 0$). HumanIll was considered significantly less trustworthy than all other styles ($p < 0.004$ in all cases). HumanBasic was the next least trustworthy ($p < 0.004$ in all cases). ToonBare and HumanSSS were next ($p < 0.002$ in all cases). Finally, the most realistic Humans, and all other Toons were rated as equally trustworthy, and more trustworthy than all other styles ($p < 0.004$ in all cases). There was no main effect of motion, and no interaction.

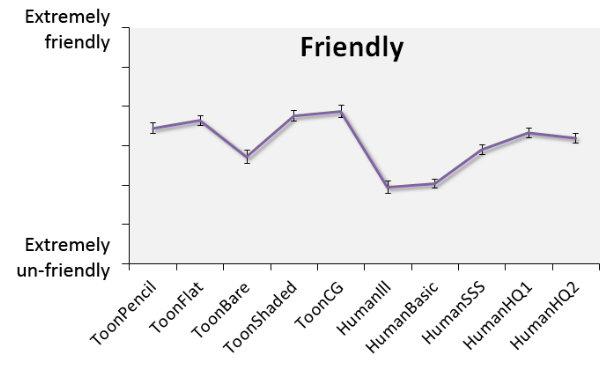

Figure 6: *Main effect of style for Friendly ratings.*

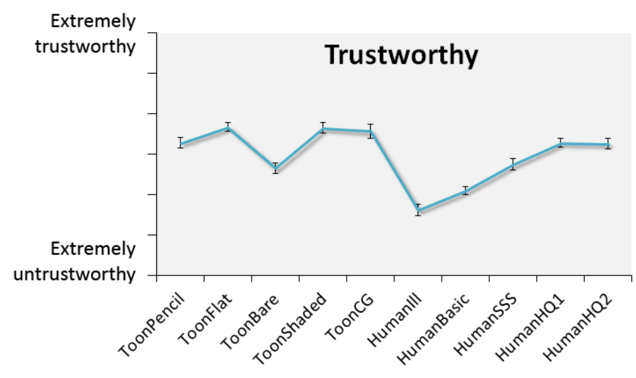

Figure 7: *Main effect of style for Trustworthy ratings.*

5.2 Exp 2: Motion Anomalies

It has often been speculated that the more realistic a character appears, the more likely people are to notice subtle problems in its animation. In this experiment, we aim to test how unpleasant animation artifacts are perceived to be, and if they appear more unpleasant on realistically rendered characters.

Hodgins et al. [2010] investigated the effect of animation artifacts on the ability of a character to convincingly act out a scene. They tested a range of artifacts and found that when the facial animation from one half of the face was removed, participants found this extremely noticeable. They also found that removing all eye motion was not considered a disturbing artifact. We chose these two artifacts in our experiment in order to determine if the effect of render style changed the ratings for a very severe (half face) and a barely noticeable (static eyes) artifact. We hypothesized that the half face condition would be the most unpleasant, and that motion alterations would be perceived as more unpleasant on the realistically rendered model.

The motion alterations were applied to the character in 3ds Max and rendered out in five styles: HumanWrinkle, HumanIll, HumanBasic, ToonCG, and ToonPencil. The styles were chosen as a sample of styles of different appeal, based on the ratings in Exp 1. The same three clips as in Section 5.1.2 were selected, which ranged in length from 6–10 seconds.

Twelve volunteers (6F, 6M) took part in this experiment. Participants were asked to view each motion and then to rate it using the keyboard as input, on a scale from 1–7, where 1 indicated "extremely unpleasant" and 7 "extremely pleasant". They viewed 45 movies in pseudo random order (5 render styles $\times$ 3 motion alterations (no alteration, half face, static eyes) $\times$ 3 motion clips).

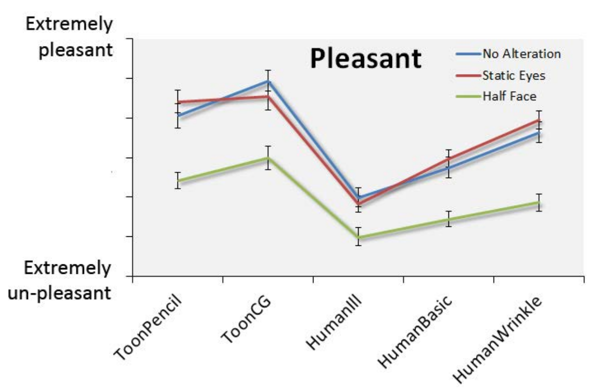

Figure 8: *Interaction between render style and motion alteration.*

5.2.1 Results

The ratings for each motion alteration, for each render style were averaged per participant over repetitions. A 2-way repeated measures ANOVA was conducted on the results, with factors *motion alteration* (3 - no alteration, static eyes, half face) and *render style* (5). A main effect of motion alteration was found ($F_{2,22} = 37.28, p \approx 0$), and post-hoc analysis revealed that this was due to the half face alteration being rated as significantly more unpleasant than either of the other two ($p < 0.0002$ in both cases). There was no significant difference between ratings for movies with static eyes or no alteration, implying that participants did not find this artifact disturbing (as in Hodgins et al. [2010]). The character's head did not move on the horizontal plane very much in our movies, which may account for the fact that static eyes were not perceived as unpleasant. A main effect of render style was also found ($F_{4,44} = 29.96, p \approx 0$). Post-hoc analysis showed that HumanIll was rated as the least pleasant ($p < 0.005$ in all cases), HumanBasic was the next least pleasant ($p < 0.005$ in all cases). HumanWrinkle and ToonPencil were rated as equally pleasant ($p < 0.05$ when compared to all other styles). Finally, ToonCG was rated as the most pleasant, and significantly more so than all others ($p < 0.05$ in all cases).

An interaction also occurred ($F_{8,88} = 4.06, p < 0.0004$). Figure 8 shows that with no alteration and with half face alteration, ToonCG was perceived as more pleasant than ToonPencil, but with static eyes, there was no difference. For the half face condition, pleasantness ratings were significantly lower for the Human than for the Toon characters. This implies that *large motion anomalies are much more acceptable on Toon characters, than on characters with human textures applied*. The Toon characters with the extreme half face alteration were in fact rated as more pleasant than the HumanIll character with no alteration, and equally pleasant to the HumanBasic character with no alteration. Contrary to popular belief, the artifacts applied to the most realistic model were not considered as the most disturbing overall.

Relative differences

Analyzing the absolute difference in ratings gives us an overall picture of how pleasant participants found the various styles, with different anomalies applied. However, it does not tell us which of the styles was most affected by the anomalies. Therefore, we calculated the differences in rating between the no alteration condition and the anomaly conditions, per render style, per participant. We then conducted an ANOVA on the values and found no main effect of style, but a main effect of motion alteration ($F_{1,11} = 47.64, p < 0.00004$) which showed that render styles with the half face anomaly applied had a greater drop in appeal than when they were displayed with static eyes. An interaction between alteration and render style also occurred ($F_{4,44} = 5.94, p < 0.0007$). For the static eyes condition, the differences were very low and hovered around zero, implying that there was very little difference between ratings for the no alteration and static eyes conditions. For the half face condition, there was less of a drop in pleasantness ratings for character's that were already considered unpleasant with no alteration (HumanIll and HumanBasic) than those that were considered pleasant with no alteration (ToonPencil, ToonCG and HumanWrinkles).

6 Implicit Experiments

Social interaction with virtual humans is becoming popular in interactive drama video games, where players decisions and actions during the game affect the narrative. For example, in the video game *L.A. Noire*, gamers are asked to interact on a higher level with characters than ever before, by trying to determine if a virtual character is lying to them or not. So far, these types of games have chosen human-like characters, presumably to allow for more believable performances and interaction. In this experiment, we investigate if the choice of rendering style is important for the display of subtle cues such as those that occur when lying, and if different rendering styles will evoke different reactions to the same performance.

We chose a deception task for this experiment, where participants were asked to detect if a virtual character was telling the truth or lying. Previous work using videos of real people has shown that deception detection is difficult, with participants usually being able to detect a lie at just above chance level. We postulated that rendering style would unconsciously influence their decisions, where a negative reaction towards a character could increase the 'lie' responses. In the Audio/Visual Experiment, we tested the ability of participants to detect lies on three virtual styles along with two control cases. The first control was *audio only*, where participants made judgments based on the audio track alone. The second control condition was the digital video recorded during the capture session (*real*). Finally, the Visual Only Experiment tested participants ability to detect lies using only visual cues (i.e., no audio).

6.1 Exp 1: Audio/Visual

In this experiment, we tested participants between-groups in order to avoid them seeing the same truth and lie sequences on different render styles. Furthermore, we recorded controlled video data during the capture session, so that we could compare our virtual characters to their real counterparts. In this experiment, we test the ability of participants to detect lies using different visual and auditory cues. The five cases that we tested were: real, audio only, HumanIll, HumanWrinkle, and ToonCG. HumanWrinkle was chosen as the most realistic style that also had animated wrinkles which we felt would aid lie detection (Figure 9). ToonCG was chosen as an appealing style that we postulated might evoke positive reactions. Finally, HumanIll was included to determine the effect of a highly unappealing style. We hypothesized that lie detection would be best when viewing the real video, and worst with the audio alone. Also, that appealing characters would bias participants into believing more of their answers than unappealing syles.

Figure 9: *Female character rendered in HumanWrinkle style. (left) character without wrinkle maps, (right) character during eye-brow lift, displaying blended wrinkle maps.*

In order to create variation in the stimuli, we recruited two actors for this experiment - one male and one female. Two new capture sessions were conducted, as described in Section 3. Since we were recording video that would be used as stimuli in the experiment, we used the smallest possible markers that our system would allow (2mm). Furthermore, we placed the eye-tracker camera as far from the face as possible without affecting the tracking accuracy, in order to prevent occlusion of the face in the video stimuli (Figure 10, left). A blue backdrop was placed behind the actor, outside the capture volume of the motion capture system. This was to ensure that we could maintain a consistent background in all videos, and replicate it in the virtual environment.

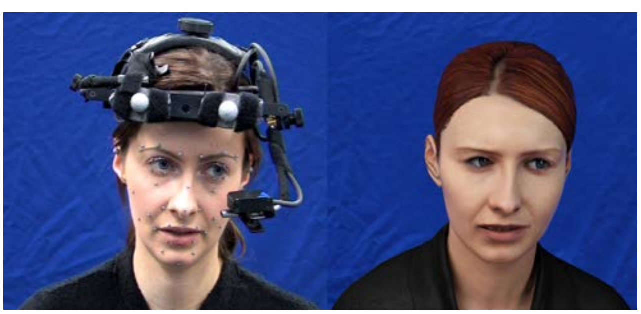

Figure 10: *Corresponding frames from the real video and the virtual character rendered in HumanWrinkle style, for the female actor.*

The video stimuli were recorded during the motion capture session, to ensure that identical video and virtual stimuli could be achieved. In order to match the camera in the virtual environment, the camera position and orientation relative to the actor was recovered using manually labeled marker locations in the image and their reconstructed 3D positions.

Seven truth and seven lie sequences were chosen from each actor's performance capture session. Clips were chosen such that the motion had minimal artifacts. The sequences ranged in length from 2 to 7 seconds. The eye, body, and face motions of the actors were applied to their custom-made virtual replicas. These sequences were then rendered in the three render styles (HumanIll, ToonCG, and HumanWrinkle).

Audio recordings from the interviewer and interviewee microphones were merged to a mono audio stream. The video stimuli were divided to precisely match the virtual sequences. One hundred and twelve movies were created in total (2 actors $\times$ 14 sequences (7 truths, 7 lies) $\times$ 4 representations (3 virtual, 1 real)). A further matching 28 audio tracks were also created (2 actors $\times$ 14 sequences). The character was displayed from the shoulders up and facing the participant, with a blue cloth background to match the video backdrop (Figure 10).

Sixty volunteers took part in this experiment (36M, 24F). The volunteers were split into five groups of 12, where each group was presented with one of the five representations (audio, real, HumanIll, HumanWrinkle, ToonCG). Stimuli were displayed on a 24" LCD monitor at a distance of 60cm and participants used closed-back headphones to listen to the audio. Movies were displayed at 720 $\times$ 576 uncompressed (to avoid compression artifacts). Participants viewed the entire clip and were then asked to decide whether the actor in the clip was telling a "lie" or the "truth", and indicated their response using a left or right mouse button click. Participants were instructed to try not to use special strategies for detecting lies and to follow their intuition each time. Participants viewed four repetitions of each block in pseudo-random order, to avoid exact repetitions being played in sequence. No feedback was given to indicate if their answers were correct.

6.1.1 Results

Lie Detection: Sensitivity

We first analyzed the ability of our participants to detect lies. We wanted to determine if this was affected by the representation (whether real or virtual), or the style used for rendering on the virtual models. Also, if participants were better at detecting the lies of either of the two actors. Treating the experiment as a signal detection task, we calculated sensitivity (d') for each actor, for each character representation using Equation 1 where HR is Hit Rate, FAR is False Alarm Rate and z(p) $\in [0, 1]$ is the z-score of p. High d' values indicate that participants are accurate at the task of detecting lies.

$$d' = z(\mathrm{HR}) - z(\mathrm{FAR}) \quad (1)$$

A two-way ANOVA was conducted with between groups factor *representation* (5 - real, audio, HumanIll, HumanWrinkle, ToonCG) and within factor *actor* (2 - male and female). We found that signal detection was low overall, but above chance (which is consistent with previous work [DePaulo et al. 2003]). We found a main effect of actor ($F_{1,55} = 17.78, p < 0.0001$) which showed that participants were much better at the task when viewing the female actor. Participants were performing close to chance level for the male actor. A main effect of representation was also found ($F_{4,55} = 2.66, p < 0.05$). Figure 11 shows the average d' scores for each of the representations. Contrary to our hypothesis, the audio track alone appeared to be the best representation for detecting lies, and was significantly better than audio/visual pairs rendered in HumanWrinkle and ToonCG ($p < 0.05$ in both cases) and close to significantly better than the Real videos ($p = 0.08$).

These results imply that the visual information was being used by participants, but was less reliable than the audio signal. A possible interpretation for the low ratings of the HumanWrinkle and ToonCG could be the fact that participants found these styles appealing (as rated in Section 5) and focused more on the visual rather than the auditory information, when viewing these styles. It is unlikely to be due to the motion capture accuracy, as ratings for all of the virtual styles were not significantly different from the real video stimuli. No interaction was found between representation and actor.

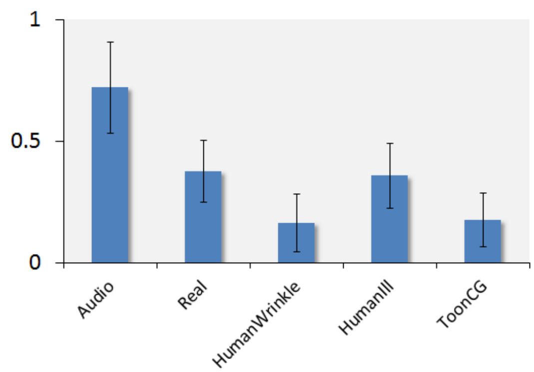

Figure 11: *Average d' sensitivity showing a main effect of character representation. Higher d' scores indicate that participants are better at the task of detecting lies.*

Lie Detection: Bias

We wished to determine if appealing render styles would implicitly bias participants towards rating the sequences as "truth", and vice versa for unappealing styles. Informed by Signal Detection Theory, we calculated the response bias *c* for each participant, per actor and representation using Equation 2. As before, a two-way ANOVA was conducted with between groups factor *representation* and within factor *actor*. No main effects and no interaction was found. This implied that *using appealing render styles does not bias participants towards believing that the character was telling the truth*. An overall truth bias was observed, however. This truth bias is consistent with literature [Buller and Burgoon 1996; Köhnken et al. 1989] and has been attributed to the fact that people are confronted with truthful statements more often than deceptive ones in daily life, and therefore expect statements to be truthful in an experimental situation.

$$c = -\frac{1}{2}(\mathrm{z}(\mathrm{HR}) + \mathrm{z}(\mathrm{FAR})) \quad (2)$$

Sequence Responses

The ability of participants to detect lies (d') was not the only measure that our experiment allowed us to test. Regardless of accuracy, we were also interested in the whether the question/answer sequence responses were consistent across render style (e.g., if a lie sequence was consistently responded to as 'truth' on one render style and 'lie' on another). In order to get a clearer picture of the responses for each of the 28 question/answer sequences over the four repetitions, we first calculated the average responses per sequence, per character representation. Using this data, we conducted a two-way ANOVA with within groups factor *question/answer sequence* (28 – 7 truths, 7 lies per actor) and between-groups factor *representation* (5 – real, audio, HumanWrinkle, HumanIll, ToonCG). No main effect of character representation was found, but a main effect of question/answer sequence was ($F_{27,1431} = 15.406, p \approx 0$). As can be seen by the spread of responses in Figure 12, participants were not performing homogeneously, despite the underlying "truth bias". Participants were not very accurate at determining which clips contained lies (low d'), but they were consistent in their responses for each sequence across render style (and both control cases), which provides evidence towards the fact that the performance dominated and therefore the render style did not influence their decisions. Finally, a significant interaction was not found, which implies that *responses for each of the 28 clips were consistent across representation*.

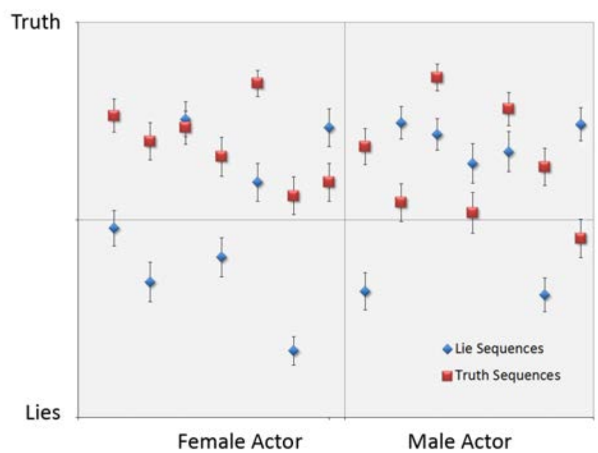

Figure 12: *Main effect of question/answer sequence. This graph shows the average ratings for the 7 truth and 7 lie sequences for each actor. Some lie sequences received high 'truth' ratings, but no truths had high 'lie' ratings. An overall 'truth' bias is also present.*

6.2 Exp 2: Visual Only

Auditory information strongly influenced the ratings of participant in the previous experiment. In this experiment, we tested participants ability to detect lies while using only visual information. This allowed us to determine if a difference in bias occurs for different render styles, when focusing solely on the appearance and motion of the character.

Since there was no audio to guide participants in this experiment, we used a within-subjects design. Three render styles were tested: the most realistic (HumanWrinkle), one of the most appealing (ToonCG), and the least appealing style (HumanIll). Seventeen participants who had not taken part in any of the audio/visual experiments took part in this experiment (12M, 5F). Eighty four movies were shown in random order (3 render styles × 2 actors (male and female) × 14 sequences (7 truth, 7 lies)). Using only visual cues, participants were asked to indicate for each movie whether they thought that the character was telling a lie or the truth.

6.2.1 Results

For each participant, d' and bias values were calculated per actor per render style. A two-way repeated measures ANOVA was conducted on the d' values and we found no main effect of style, actor and no interaction. An independent samples t-test was conducted on the sensitivity scores, comparing it to chance level ($d' = 0$) and there was no significant difference. This implied that participants were unable to do the task and were guessing their answers. A two-way repeated measures ANOVA was conducted on the bias scores and no main effects and no interaction was found. This shows that *even when presented with only visual cues, this still did not bias participants towards believing appealing characters over unappealing ones*.

7 Discussion

In this paper, we have investigated the effect of render style on the perception of virtual humans. A series of psychophysical experiments were conducted which explored both explicit and implicit measures on perception over a range of render styles, while keeping all other conditions constant. Eleven different render styles were rated by the participants for realism and this scale was subsequently used to compare against all other effects. We found that much of the information that we use to rate virtual characters is available in a still image. Movement changes only how familiar we find the characters, and also how appealing or pleasant they are considered. Highly unappealing characters are considered more so when movement is applied, and motion anomalies are considered more unpleasant on human than on cartoon render styles. We believe that this was due to the fact that humans are inherently conditioned to analyze human faces, and are therefore less forgiving of anomalies when a human photograph is applied to the model.

These results are very much in line with the Uncanny Valley theory. Contrary to the theory, however, was the fact that our most realistic character was often rated as appealing or pleasant as our cartoon characters. The drop in appeal for our stimuli came from characters that were rated as unfamiliar or those that were in the middle of our 'abstract to realistic' scale. We feel this could be attributed to the fact that these characters were difficult for the brain to categorize due to their uncommon appearance, as suggested in [Saygin et al. 2012].

In our implicit lie-detection experiments, we tested if the style of the character could go as far as to change our ability to perform a task, or bias us towards believing what they say. We found no difference in ratings or bias between any of our virtual characters or the corresponding real video sequences. This suggests that people can interact with virtual characters in the same way as real humans, primarily due to the content of the audio track. When asked directly, participants rated some characters as more trustworthy than others. However, the trustworthy characters did not bias participants unconsciously into believing more of their lies.

Our result is interesting since it shows that participants were so focused on the task, that the appearance of the character did not sway them. We found that the audio and animation contributed to the interpretation of the characters' intention rather than the render style. A similar result occurred in a previous study [McDonnell et al. 2009] where we found that realistic, motion-captured animations were enough to convince participants of the emotion of a character, regardless of the body that displayed that emotion. In the entertainment industry, one could speculate, based on our results, that blaming the render style for the box-office failure of photorealistic CG movies is not a valid argument. One could argue that it was the content (animation, audio, modeling) and not the render style that contributed towards the failure, as we found that render style does not alter the interpretation of content in a positive or negative manner.

To our knowledge, our lie detection experiment represents the first attempt in the literature at analyzing the effect of the "uncanny valley" implicitly. The fact that the render style did not bias participants into saying "lie" was not the answer that we had hypothesized, but is nonetheless an interesting finding which helps improve our understanding of the perception of virtual characters. We believe that implicit testing could prove very useful and practical to investigate the much discussed issue of the "uncanny valley".

Based on the results of our experiments, we can reassure developers that selecting realistic virtual characters is not as risky as often discussed (provided that the motion does not contain artifacts). Negative reactions occurred mainly for characters that used human texture maps, but that were not rendered with realistic eye and skin shaders. Cartoon characters were considered highly appealing, and were rated as more pleasant than characters with human appearance, when large motion artifacts were present. They were rated as more friendly than realistic styles and therefore might be more appropriate for certain virtual interactions (e.g., in rehabilitation simulations). However, we cannot say that *all* abstract cartoon characters are considered appealing. One of our render styles (ToonBare) was considered quite unappealing and evoked negative reactions from participants across most of the scales tested. We believe that this is related to the lack of familiarity of this style, as participants gave it a low rating on the familiarity scale.

8 Limitations and Future Work

In this paper, our aim was to investigate the effect of the typical render styles used in CG productions on the perception of avatars. One limitation is that we did not fully control for the effect of lighting and allowed the artist to have creative freedom. Future experiments using controlled lighting would be interesting in order to investigate the specific effects of the rendering parameters. Another limitation was that our animation clips did not contain emotional content. Testing the effect of emotion could be very interesting, to determine if performances with higher emotions elicit different responses. Furthermore, the actors' answers to the truth and lie questions were very short and to the point. Perhaps with longer exposure times per sequence, this could sway the participants' answers in a different direction. Future studies will investigate the implicit effect of render style on longer sequences or vignettes, such as those used in Hodgins et al. [2010]. One implicit effect that we did find was that participants were worse at the task of lie detection when viewing appealing virtual characters than when they were presented with audio alone. This could be due to the fact that they used more visual than auditory cues in making their decisions in these cases. Future experiments will probe further into the effects of auditory and visual cues using different render styles, and attempt to qualify the importance of each of the cues.

The video stimuli in our experiments showed actors with a range of motion capture markers on their faces, along with an eye tracking device on their head. We cannot be sure if ratings would have differed if the actors were shown without these. Using marker-less face and eye tracking could prove useful for future studies, to avoid this issue.

Eye-tracking to determine which parts of the faces were being attended to while performing the tasks might help us to further understand the difference in perception between different render styles. A bone-based deformation system was used to drive the faces of our characters. This computationally efficient method (which is standard practice in most video games) produced animation that was rated in the same way as real video, for lie detection. However, in future work, we would like to investigate the effect of higher quality deformation (e.g., including physically accurate wrinkles and muscles) on the perception of render style.

Acknowledgements

The authors would like to thank Cristóbal Curio for valuable discussions, artists Peter Lozsek and Pete McNally, and actor Simon Dobbyn. This work was sponsored by Science Foundation Ireland as part of the Captavatar and Natural Movers projects. This research was also supported by DFG grant Perceptual Graphics PAK 38 CU 149/1-2, EU Project "Tango" (ICT-2009-C 249858), and WCU (World Class University) program funded by the Ministry of Education, Science and Technology through the National Research Foundation of Korea (R31-10008).

References

BARTNECK, C., KANDA, T., ISHIGURO, H., AND HAGITA, N. 2007. Is the uncanny valley an uncanny cliff? In *Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN07)*, 368–373.

BULLER, D., AND BURGOON, J. K. 1996. Interpersonal deception theory. *Communication Theory 6*, 203–242.

CHAMINADE, T., HODGINS, J., AND KAWATO, M. 2007. Anthropomorphism influences perception of computer-animated characters' actions. *Social Cognitive and Affective Neuroscience 2*, 206–216.

CHEN, H., RUSSELL, R., NAKAYAMA, K., AND LIVINGSTONE, M. 2010. Crossing the 'uncanny valley': adaptation to cartoon faces can influence perception of human faces. *Perception 39*, 3, 378–386.

CURIO, C., BREIDT, M., KLEINER, M., VUONG, Q. C., GIESE, M. A., AND BÜLTHOFF, H. H. 2006. Semantic 3d motion retargeting for facial animation. In *APGV '06: Proceedings of the 3nd symposium on Applied perception in graphics and visualization*, 77–84.

DEPAULO, B. M., LINDSAY, J. J., MALONE, B. E., MUHLENBRUCK, L., CHARLTON, K., AND COOPER, H. 2003. Cues to deception. *Psychological Bulletin 129*, 74–118.

EKMAN, P., AND FRIESEN, W. V. 2003. *Unmasking the face: a guide to recognizing emotions from facial clues*. Malor Books.

GELLER, T. 2008. Overcoming the uncanny valley. In *IEEE Computer Graphics and Applications*, vol. 28, 11–17.

HAN, S., JIANG, Y., HUMPHREYS, G. W., ZHOU, T., AND CAI, P. 2005. Distinct neural substrates for the perception of real and virtual visual worlds. *NeuroImage 24*, 928–935.

HANSON, D., OLNEY, A., PRILLIMAN, S., MATHEWS, E., ZIELKE, M., HAMMONS, D., FERNANDEZ, R., AND STEPHANOU, H. 2005. Upending the uncanny valley. In *Proceedings of the National Conference on Artificial Intelligence (AAI05)*, 24–41.

HODGINS, J., O'BRIEN, J., AND TUMBLIN, J. 1998. Perception of human motion with different geometric models. *IEEE Transactions on Visualization and Computer Graphics 4*, 4, 307–316.

HODGINS, J. K., JÖRG, S., O'SULLIVAN, C., PARK, S. I., AND MAHLER, M. 2010. The saliency of anomalies in animated human characters. *ACM Transactions on Applied Perception 7*, 4, 1–14.

KÖHNKEN, H., GLEICHER, M., AND PIGHIN, F. 1989. Behavioral correlates of statement credibility: theories, paradigms and results. *Criminal behavior and the justice system: Psychological perspectives*, 271–289.

LEVI, S. 2004. Why Tom Hanks is less than human; while sensors cannot capture how humans act, humans can give life to digital characters. In *Newsweek 650*, 305–306.

MACDORMAN, K. F., GREEN, R. D., HO, C. H., AND CLINTON, T. K. 2009. Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior 25*, 695–710.

MAR, R. A., KELLEY, W. M., HEATHERTON, T. F., AND MACRAE, C. N. 2007. Detecting agency from the biological motion of veridical vs animated agents. *Social Cognitive and Affective Neuroscience Adcance Access 2*, 3, 199–205.

MCDONNELL, R., JÖRG, S., MCHUGH, J., NEWELL, F., AND O'SULLIVAN, C. 2009. Investigating the role of body shape on the perception of emotion. *ACM Transactions on Applied Perception 6*, 3, 14:1–14:11.

MORI, M. 1970. The uncanny valley. *Energy 7*, 4, 33–35.

NOWAK, K. L., AND BIOCCA, F. 2003. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence 12*, 5, 481–494.

PERANI, D., FAZIO, F., BORGHESE, N. A., TETTAMANTI, M., FERRARI, S., DECETY, J., AND GILARDI, M. C. 2001. Different brain correlates for watching real and virtual hand actions. *NeuroImage 14*, 749–758.

REEVES, B., AND NAAS, C. 1996. The media equation: How people treat computers, television, and new media like real people and places. In *Stanford, CA*, CSLI Publications.

SAYGIN, A. P., CHAMINADE, T., ISHIGURO, H., DRIVER, J., AND FRITH, C. 2012. The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social Cognitive and Affective Neuroscience 7*, 4, 413–442.

SLATER, M., AND STEED, A. 2002. Meeting people virtually: Experiments in shared virtual environments. In *The social life of avatars: Presence and interaction in shared virtual environments*, 146–171.

STEPTOE, W., STEED, A., ROVIRA, A., AND RAE, J. 2010. Lie tracking: social presence, truth and deception in avatar-mediated telecommunication. In *CHI '10: Proceedings of the 28th international conference on Human factors in computing systems*, 1039–1048.

THOMPSON, C. 2004. The undead zone: Why realistic graphics make humans look creepy. *Slate*.