Facial expression of emotion and perception of the Uncanny Valley in virtual characters
Tinwell, Angela; Grimshaw, Mark; Abdel Nabi, Debbie; Williams, Andrew

Published in:
Computers in Human Behavior

DOI (link to publication from Publisher):
10.1016/j.chb.2010.10.018

Publication date:
2010

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
? You may not further distribute the material or use it for any profit-making activity or commercial gain
? You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: January 15, 2016
Facial expression of emotion and perception of the uncanny valley in virtual characters

Angela Tinwell
*University of Bolton, A.Tinwell@bolton.ac.uk*

Mark Grimshaw
*University of Bolton, m.n.grimshaw@bolton.ac.uk*

D. Abdel Nabi
*University of Bolton, a.nabi@bolton.ac.uk*

Andrew Williams
*University of Bolton, A.Williams@bolton.ac.uk*

Digital Commons Citation
http://digitalcommons.bolton.ac.uk/gcct_journalspr/14

This Article is brought to you for free and open access by the School of Games Computing and Creative Technologies at UBIR: University of Bolton Institutional Repository. It has been accepted for inclusion in Games Computing and Creative Technologies: Journal Articles (Peer-Reviewed) by an authorized administrator of UBIR: University of Bolton Institutional Repository. For more information, please contact ubir@bolton.ac.uk.
Facial Expression of Emotion and Perception of the Uncanny Valley in Virtual Characters

Angela Tinwell\textsuperscript{a}, Mark Grimshaw\textsuperscript{a}, Debbie Abdel Nabi\textsuperscript{b} and Andrew Williams\textsuperscript{a}

\textsuperscript{a}School of Business and Creative Technologies, The University of Bolton, Deane Road, Bolton, BL3 5AB, United Kingdom

\textsuperscript{b}School of Health and Social Sciences, The University of Bolton, Deane Road, Bolton, BL3 5AB, United Kingdom

Correspondence and proofs should be addressed to Angela Tinwell, School of Business and Creative Technologies, The University of Bolton, Deane Road, Bolton, BL3 5AB, UK.

Tel: +44 (0)1204 903589, Fax: +44 (0)1204 903500, e-mail: A.Tinwell@bolton.ac.uk

Keywords: Uncanny Valley; Facial expression; Emotion; Characters; Video games; Realism
Abstract

With technology allowing for increased realism in video games, realistic, human-like characters risk falling into the Uncanny Valley. The Uncanny Valley phenomenon implies that virtual characters approaching full human-likeness will evoke a negative reaction from the viewer, due to aspects of the character’s appearance and behavior differing from the human norm. This study investigates if “uncanniness” is increased for a character with a perceived lack of facial expression in the upper parts of the face. More important, our study also investigates if the magnitude of this increased uncanniness varies depending on which emotion is being communicated. Individual parameters for each facial muscle in a 3D model were controlled for the six emotions: anger, disgust, fear, happiness, sadness and surprise in addition to a neutral expression. The results indicate that even fully and expertly animated characters are rated as more uncanny than humans and that, in virtual characters, a lack of facial expression in the upper parts of the face during speech exaggerates the uncanny by inhibiting effective communication of the perceived emotion, significantly so for fear, sadness, disgust, and surprise but not for anger and happiness. Based on our results, we consider the implications for virtual character design.
Facial expression of emotion and perception of the Uncanny Valley in virtual characters

1. Introduction

1.1 The Uncanny Valley

Jentsch (1906) first introduced the subject of “The Uncanny” into contemporary thought in an essay entitled “On the Psychology of the Uncanny”. The uncanny was described as a mental state where one cannot distinguish between what is real or unreal and which objects are alive or dead. Referring to Jentsch’s essay, Freud (1919) characterized the uncanny as a feeling caused when one cannot detect if an object is animate or inanimate upon encountering objects such as “waxwork figures, ingeniously constructed dolls and automata” (p. 226).

In 1970 (as translated by MacDorman and Minato, (2005)) the roboticist Masahiro Mori made associations of the uncanny with robot design. Mori observed that as a robot’s appearance became more human-like, a robot continued to be perceived as more familiar and likeable to a viewer, until a certain point was reached (between 80% and 85% human-likeness), where the robot was regarded as more strange than familiar. As the robot’s appearance reached a stage of being close to human, but not fully, it evoked a negative affective response from the viewer. Fig. 1 depicts a visualization of Mori’s theory showing familiarity increasing steadily as perceived human-likeness increases, then decreasing sharply, causing a valley-shaped dip.
The unpleasant feelings evoked by the uncanny have been attributed to it being a reminder of one’s own mortality (MacDorman, 2005; Mori, 1970). Kang (2009) however, suggested the negative impact of the uncanny is related to how much of a threat a character is perceived to be and how much control we have over the potentially threatening or dangerous interaction.

1.2 Explorations of cross-modal influence and the Uncanny Valley

Advances in technology have facilitated increased visual realism in video games and designers in some game genres are creating near-realistic, human-like characters. Contrary to Mori’s advice, these designers are aiming for the second peak as enhanced realism is believed to improve the player experience and sense of immersion (e.g. Ashcraft, 2008; Plantec, 2008). As characters approach high levels of human-likeness and exhibit human-like motor behavior, aspects of their appearance and behavior are being placed under greater scrutiny by the audience. Factors
such as facial expression may appear odd or unnatural and can adversely make a character appear life-less as opposed to life-like. As with robots, highly human-like video game characters may be subject to the Uncanny Valley phenomenon (e.g. Brenton, Gillies, Ballin, & Chatting, 2005; Gouskos, 2006; MacDorman, Green, Ho, & Koch, 2009; Pollick, 2009).

Design guidelines have been authored to advise character designers on how to avoid the Uncanny Valley. Such guidelines have included factors such as facial features and proportion and level of detail in skin texture (e.g. Green et al., 2008; MacDorman, Green, Ho, & Koch, 2009; Seyama & Nagayama, 2007). Hanson, (2006) found that by changing a character’s features to a more cartoon-like style eliminated the uncanny. Schneider, Wang, & Yang, (2007) identified that character designs of a non-human appearance with the ability to emote like a human were regarded more positively. These authors acknowledge that the results from their experiments provide only a partial understanding of what a viewer perceives to be uncanny, based on “inert” (unresponsive) still images. The majority of characters featured in animation and video games do not remain still, and cross-modal factors such as motion, sound, timing and facial animation contribute to the Uncanny Valley (Richards, 2008; Weschler, 2002). When a human engages with an android, behavior that seems natural and appropriate from the android (referred to as “contingent interaction” by Ho, MacDorman & Pramono [2008, p. 170]) is important to obtain a positive response to that android (Bartneck, Kanda, Ishiguro, & Hagita, 2009; Kanda, Hirano, Eaton, & Ishiguro, 2004). Previous authors (such as Green et al., 2008; Hanson, 2006; MacDorman et al., 2009; Schneider et al., 2007) state that, had movement been included as a factor, the results and conclusions drawn from their experiments might have differed.

Previous attempts to recreate an Uncanny Valley shape do not comply with Mori’s (1970) diagram and suggest that it may be too simplistic with various factors (including dynamic facial expression) influencing how uncanny an object is perceived to be (see e.g. Bartneck, et al., 2009; Ho et al., 2008; MacDorman, 2006; Minato, Shimda, Ishiguro, & Itakura, 2004; Tinwell and Grimshaw, 2009; Tinwell, Grimshaw, & Williams, 2010).
1.3 Facial expression of emotion and the Uncanny Valley

It is well-documented that, in humans and animals, successful recognition of each type of the six universally recognized basic emotions, anger, disgust, fear, happiness, sadness and surprise, (Ekman, 1992a, 1992b) serves a different adaptive (survival or social interaction) function (Darwin, 1872; Ekman, 1979, 1992a, 1992b). For example, detection of fear and sadness in others may foretell potential harm or distress to self and humans react instinctively to such emotions to avoid a possible threat. As Blair, (2003) states:

Fearful faces have been seen as aversive unconditioned stimuli that rapidly convey information to others that a novel stimulus is aversive and should be avoided (Mineka & Cook, 1993). Similarly, it has been suggested that sad facial expressions also act as aversive unconditioned stimuli discouraging actions that caused the display of sadness in another individual and motivating reparatory behaviors (Blair, 1995 p. 561).

It has been suggested that disgust also serves the adaptive function of evoking a negative, aversive reaction from the viewer; it warns others to be concerned about potential infection or approaching a distasteful object (Blair, 2003). In contrast, displays of anger or embarrassment do not serve to act as unconditioned stimuli for instrumental learning. Instead, they are important signals to modulate current behavioral responses, particularly in social situations involving hierarchy interactions (Blair & Cipolotti, 2000; Keltner & Anderson, 2000).

Ekman and Friesen’s (1978) Facial Action Coding System (FACS) has been integrated within facial animation software to achieve authentic facial expression of emotion in realistic,
human-like video games characters. Dyck, Winbeck, Leiberg, Chen, Gur, and Mathiak (2008) conducted a study to investigate whether the facial emotional expressions of a virtual character could be recognized as easily as those produced on human faces. Still images of virtual characters expressing the emotions happiness, sadness, anger, fear, disgust, and neutral were compared to still images of humans expressing the same emotions at the same medium levels of intensity. Emotion recognition between the two groups indicated that while emotions expressed were, for the most part, recognized equally well in humans and virtual characters, the two emotions fear and sadness achieved better recognition rates when presented in the virtual character than when expressed on human faces; disgust was the only emotion that could not achieve an acceptable recognition rate in virtual characters when compared to humans and was mainly confused with anger. The emotions anger and happiness were recognized equally well for both groups. Again, the authors acknowledged that the results of their study might have been different had animated characters been used to assess how emotions can be interpreted with motion (and speech).

There have been no studies investigating if the type of emotion portrayed by a virtual character influences level of perceived uncanniness and if so which emotions are most significant in exaggerating the uncanny. Considering that anger, fear, sadness and disgust may be considered signals of a threat, harm or distress (Ekman, 1979), it would be reasonable to suggest that these survival-related emotions will be regarded as more uncanny in near human-like (but not quite fully authentic), virtual characters; especially when part of the facial expression of them is aberrant, blurring the clarity of their depiction. Emotions not associated with threat or distress (i.e. happiness and surprise) may be regarded as less important, formidable or essential for survival and therefore less noticeably strange or uncanny; even when the animation of facial features appears odd or wrong to the viewer.
1.4 The importance of upper face animation

During speech nonverbal signals are used to interpret the emotional state of a person. Nonverbal signals are largely conveyed by the upper part of the face; the lower region of the face constrained by the articulatory processes (Busso & Narayanan, 2006; Ekman, 2004, 1979; Ekman & Friesen, 1978, 1969). For example a narrowing of the eyes and a shaking fist show that a person is angry. Raised eyebrows typically demonstrate the emotion surprise (Ekman, 2004). Brow lowering and raising are used as “batons” (Ekman, 2004 p. 41), to provide additional emphasis for words or phrases. A lowered brow is associated with negative emotions such as “fear, sadness … anger” (Ekman, 2004, p. 42) to accentuate a negative word. A raised brow is likely to be associated with more positive emotions such as happiness and surprise to emphasize a more positive word such as “easy, light, good” (Ekman, 2004, p. 42). Communication is enriched by nonverbal signals generated by movement in the forehead and eyelids to display emotional content. As such, it has been recommended that when creating virtual characters, the upper face must be modeled correctly to avoid confusion as to the affective state of the character (Busso & Narayanan, 2006). The results from the previous study by Tinwell et al. (2010) support these findings. Strong relationships were identified between the uncanny in virtual characters and an awareness of a lack of expressivity in the upper face region, the forehead area being of particular significance. The current authors aimed to extend this by investigating whether inadequate movement in the upper facial areas may have differential effects depending on which emotion is being depicted.

On the basis of results from previous studies indicating features that contribute to the uncanny (e.g. Green et al., 2008; MacDorman et al., 2009; Tinwell, 2009; Tinwell & Grimshaw, 2009, Tinwell, Grimshaw & Williams, 2010), the current study examined the phenomenon of the Uncanny Valley using a male human (referred to as the group human) and a male virtual character (“Barney” from the video game Half-Life 2, (Valve, 2008)) who produced synchronized and emotionally congruent utterances. To investigate how an ambiguity of facial expression during
speech and how blurring the salience of an emotion may exaggerate the uncanny in virtual characters, the virtual character was used in two experimental conditions: one in which movement above the lower eyelids was disabled (a partially animated condition named *lack*) and one in which movement was not disabled (a fully animated condition named *full*).

To establish if the type of emotion conveyed by a character also affects perceived uncanniness for that character, all six basic emotions (anger, fear, disgust, happiness, sadness and surprise) were portrayed (facially and prosodically) by the human and virtual character. The experimental hypotheses were as follows:

**H1a. Effect of condition:** For all of the six basic emotions of anger, disgust, fear, happiness, sadness and surprise (Ekman, 1992a, 1992b): the group human will be rated as most familiar and human-like; the group full (facial animation) will be rated second highest for familiarity and human-likeness; and the group lack (of upper facial animation) will be rated as least familiar and human-like.

**H1b. Emotion type:** The degree of difference in perceived familiarity and human-likeness between full and lack will vary across the six emotions due to differences in the psychological significance of each emotion type and variations in the importance of upper facial cues in their successful detection. Negative valence emotions (aversive or warning stimuli), that is anger, fear, sadness and disgust will attract the lowest ratings of familiarity and human-likeness (i.e. be perceived as most uncanny) in human and in both animated conditions, especially so in lack.
2. Method

2.1 Design

A 3 x 6 repeated measures design was used in the study. The independent variables (IVs) were: (1) the type of character (a) human (a human actor), (b) full (a virtual character with full facial animation), and (c) lack (the virtual character with movement disabled in the upper part of the face); and (2) the type of emotion expressed facially and orally by the human or virtual character (six types; the six basic emotions: anger, fear, disgust, happiness, sadness, and surprise (Ekman, 1992a, 1992b)). A seventh level of this IV, neutral (tone and facial expression) was employed as a control measure only to ensure participants were (1) paying attention when judging emotions, (2) could recognize neutrality within a given face, and (3) that they could respond to the different emotions presented. Neutral data was not subjected to analysis except in terms of participant accuracy and selection.

The dependent variables were: (1) ratings of familiarity and; (2) ratings of human-likeness as the characters portrayed each of the six emotions (and neutral) in the video.

2.2 Participants

Conflicting views exist as to whether gender has a significant impact on perception of the uncanny (see e.g. Bartnek et al., 2009; Green et al., 2008; Ho et al., 2008). Potential gender disparities were controlled for this study by using male participants only. The participants consisted of 129 male, university students, with a mean age of 21.8 years ($SD = 2.44$ years). The students were studying in the areas of: video game art, video game design, video game software development, special effects development for television and film, and sound engineering. Students
were selected from these disciplines as they were expected to have a similar level of exposure to realistic, human-like virtual characters.

2.3 Human and virtual character stimuli

For the character type human, video recordings of an actor were made. The videos were filmed within a video production studio at the University of Bolton, using a Panasonic (AG-HMC/50P), portable, high definition video camera. The duration of each video was 4 seconds. For the recordings, the actor had been instructed to say the line, “The cat sat on the mat”, and to use appropriate facial expressions and intonation of speech for each of the six emotions (plus neutral) as he spoke. The actor was provided with a list of the seven emotions (anger, fear, disgust, happiness, sadness, surprise and neutral) and provided with a rehearsal time for each emotion so that he could practice the required facial expression and intonation of speech. The actor was instructed to convey moderate to high intensity for each emotion, apart from the neutral state where he was asked to show no facial expression or intonation of voice.

Full and lack: FacePoser software (Valve, 2008) was used to create footage of the virtual character. Facial animation was choreographed using Flex animation techniques, based on Ekman and Friesen’s (1978) FACS. The muscular activities used to generate changes in facial appearance for a particular emotive state have been assigned specific Action Units (AU). In the upper face region, three AUs were identified across the six basic emotions for movement of the eyebrows: AU1, Inner Brow Raiser; AU2, Outer Brow Raiser and AU4, Brow Lowerer; and two for the eyelids, AU5, Upper Lid Raiser and AU7, Lid Tightener. AUs can work unilaterally or bilaterally to create the appropriate action (Ekman, 1979; Ekman & Friesen, 1978). The facial expression tool FacePoser, assigns each AU to an interactive Flex Slider bar that allows a user to create and manipulate facial expressions for a character. The Flex Slider bar Lowerer is used to simulate AU4, Brow Lowerer. The slider bar can then be adjusted to a varying level of intensity to depict the
required intensity of expression (in this case a frown) for the character’s facial appearance. For example, to generate the emotion anger, the appropriate AUs are selected (as displayed in a human face) and altered to the required intensity by adjusting the Flex Slider bars. This technique was used to simulate expressions for the animated sequences.

Using Flex animation techniques based on FACS allowed control of individual parameters for each facial muscle in the lower, mid and upper part of the character’s face. For each of the six basic emotions Flex Slider bars were selected for each active AU (see Table 1).

Table 1

<p>| Active Action Units (AU) used for the Six Basic Emotions in the Two Virtual Character Conditions. |
|-------------------------------------------------|-------------------------------------------------|</p>
<table>
<thead>
<tr>
<th>Condition</th>
<th>Full</th>
<th>Lack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>AU4, 7, 17, 24</td>
<td>AU17, 24</td>
</tr>
<tr>
<td>Disgust</td>
<td>AU7, 9, 10, 15, 17</td>
<td>AU9, 10, 15, 17</td>
</tr>
<tr>
<td>Fear</td>
<td>AU1, 2, 4, 5, 15, 20, 26</td>
<td>AU1, 2, 5</td>
</tr>
<tr>
<td>Happiness</td>
<td>AU1, 6, 7, 12</td>
<td>AU6, 12</td>
</tr>
<tr>
<td>Sadness</td>
<td>AU1, 4, 15, 23, 38</td>
<td>AU15, 23, 38</td>
</tr>
<tr>
<td>Surprise</td>
<td>AU1, 2, 5, 26</td>
<td>AU26</td>
</tr>
</tbody>
</table>
Further animation was then created by manipulating the Flex Sliders, copying the human actor’s facial expression, during the line of speech. The intensity of each Flex Slider bar was adjusted at points throughout the animation timeline to mimic the intensity of facial movements made by the actor, for example, for different words or pauses. For the emotion anger, the eyebrow movement of the human to cause a frown intensified as the word “mat” was pronounced. Thus the intensity of the corresponding Flex Slider bar, Lowerer was increased at this point for the virtual character to replicate the facial movement intensity made by the actor as he pronounced this word. The intensity was then reduced as the actor’s face began to relax at the very end of the sentence.

Two animations were created for each of the six basic emotions: The first animation with a fully animated facial expression for the virtual character (full); and the second where all movement in the top part of the face had been disabled (lack). In the latter case, all Flex Sliders used to control movement in the character’s eyebrows and eyelids were deselected for each of the six emotions (see Table 1). The Flex Sliders controlling facial muscles in the mid and lower part of the face were left fully animated. For all videos, headshots were placed against a dark background.¹

Although the purpose of this study was not to investigate emotion recognition, the recognition rates of perceived emotion for both the human and virtual stimuli were compared to ensure that the movement in the virtual characters should replicate that of the human footage. The human stimuli achieved a mean recognition rate of 79.81% ($SD = 12.61\%$) and the fully animated virtual characters a mean recognition rate of 73.75% ($SD = 20.08\%$) (see Table 2). Participants achieved high recognition rates for the control state neutral in both the human (89.92%) and virtual character (91.47%), implying that the participants could recognize when no emotion was being presented. The virtual character was expected overall to have a slightly lower mean value due to restrictions of facial movement with current facial animation software.

¹ Note: To create a neutral facial expression in FacePoser all Action Units are set to zero. Therefore, only one facial stimulus was created for the facial expression neutral for the virtual character; presented to participants once. The expression neutral was used as a control for this study to analyse user response for judging emotion.
### Table 2

**Percentage Accuracy Ratings for the Six Basic Emotions and a Neutral Expression, in the Three Conditions (N = 129).**

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Anger</th>
<th>Disgust</th>
<th>Fear</th>
<th>Happiness</th>
<th>Sadness</th>
<th>Surprise</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Human</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>82.95</td>
<td>6.98</td>
<td>1.55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Disgust</td>
<td>10.85</td>
<td>72.87</td>
<td>2.33</td>
<td>0</td>
<td>7.75</td>
<td>0.78</td>
<td>1.55</td>
</tr>
<tr>
<td>Fear</td>
<td>0.78</td>
<td>4.65</td>
<td>75.97</td>
<td>0</td>
<td>10.85</td>
<td>6.20</td>
<td>0</td>
</tr>
<tr>
<td>Happiness</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>86.82</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sadness</td>
<td>0</td>
<td>1.55</td>
<td>0</td>
<td>0.78</td>
<td>55.04</td>
<td>0.78</td>
<td>7.75</td>
</tr>
<tr>
<td>Surprise</td>
<td>1.55</td>
<td>12.40</td>
<td>18.60</td>
<td>12.40</td>
<td>1.55</td>
<td>90.70</td>
<td>0.78</td>
</tr>
<tr>
<td>Neutral</td>
<td>3.88</td>
<td>1.55</td>
<td>1.55</td>
<td>0</td>
<td>24.81</td>
<td>1.55</td>
<td>89.92</td>
</tr>
<tr>
<td><strong>Virtual Character-Full</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>79.84</td>
<td>40.31</td>
<td>0.78</td>
<td>3.10</td>
<td>0.78</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Disgust</td>
<td>14.73</td>
<td>39.53</td>
<td>2.33</td>
<td>2.33</td>
<td>2.33</td>
<td>3.10</td>
<td>1.55</td>
</tr>
<tr>
<td>Fear</td>
<td>0.78</td>
<td>3.10</td>
<td>51.94</td>
<td>1.55</td>
<td>2.33</td>
<td>3.10</td>
<td>0</td>
</tr>
<tr>
<td>Happiness</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>77.52</td>
<td>0</td>
<td>1.55</td>
<td>0.78</td>
</tr>
<tr>
<td>Sadness</td>
<td>1.55</td>
<td>12.40</td>
<td>6.98</td>
<td>0.78</td>
<td>86.82</td>
<td>2.33</td>
<td>4.65</td>
</tr>
<tr>
<td>Surprise</td>
<td>0.78</td>
<td>2.33</td>
<td>36.43</td>
<td>13.18</td>
<td>0</td>
<td>89.15</td>
<td>1.55</td>
</tr>
<tr>
<td>Neutral</td>
<td>2.33</td>
<td>2.33</td>
<td>1.55</td>
<td>1.55</td>
<td>7.75</td>
<td>0.78</td>
<td>91.47</td>
</tr>
<tr>
<td><strong>Virtual Character-Lack</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>66.67</td>
<td>24.03</td>
<td>0.78</td>
<td>0</td>
<td>3.10</td>
<td>0.78</td>
<td>0</td>
</tr>
<tr>
<td>Disgust</td>
<td>27.13</td>
<td>54.26</td>
<td>6.98</td>
<td>2.33</td>
<td>6.98</td>
<td>3.88</td>
<td>1.55</td>
</tr>
<tr>
<td>Fear</td>
<td>2.33</td>
<td>0.78</td>
<td>16.28</td>
<td>1.55</td>
<td>6.20</td>
<td>3.10</td>
<td>0</td>
</tr>
<tr>
<td>Happiness</td>
<td>1.55</td>
<td>0</td>
<td>0</td>
<td>83.72</td>
<td>0</td>
<td>1.55</td>
<td>0.78</td>
</tr>
<tr>
<td>Sadness</td>
<td>0</td>
<td>10.85</td>
<td>3.10</td>
<td>0</td>
<td>37.21</td>
<td>0</td>
<td>4.65</td>
</tr>
<tr>
<td>Surprise</td>
<td>0</td>
<td>0.78</td>
<td>57.36</td>
<td>5.43</td>
<td>2.33</td>
<td>56.59</td>
<td>1.55</td>
</tr>
<tr>
<td>Neutral</td>
<td>2.33</td>
<td>9.30</td>
<td>15.50</td>
<td>6.98</td>
<td>44.19</td>
<td>34.11</td>
<td>91.47</td>
</tr>
</tbody>
</table>
2.4 Apparatus

Video footage of each stimulus was presented to participants via an online questionnaire. The 20 videos (human = 7, full = 7, and lack = 6) (see Footnote 1) were played in a random order. Participants each had an individual computer station within a computer lab and used a “Dell E207WFPc 20.1 inch Widescreen” monitor to observe the video footage and a set of “Speed Link, Ares² Stereo PC Headphones” for the sound.

2.5 Procedure

Participants were presented with 20 video clips showing the six basic emotions: anger, disgust, fear, happiness, sadness and surprise (Ekman, 1992a, 1992b) in addition to a neutral expression for both a virtual character and a human.

After watching each video clip, participants were required to rate the character on a 9-point scale in terms of how human-like they perceived it to be (1 indicating they thought it to be non-human-like and 9 conveying that they thought it very human-like) and how strange or familiar they judged it to be from 1 (very strange) to 9 (very familiar). Nine-point scales were used to allow future direct comparison with previous experiments investigating the Uncanny Valley in characters (see e.g. MacDorman, 2006; Tinwell, 2009; Tinwell and Grimshaw, 2009; Tinwell et al., 2010) where 9-point scales were also used. For each video, participants were also asked to select which facial expression best described the character from: (a) anger, (b) disgust, (c) fear, (d) happiness, (e) sadness, (f) surprise, and (g) neutral. They were then asked to rate their level of experience both
playing video games and using 3D modeling software from the options, (1) none, (2) basic, or (3) advanced.  

3. Results

The first hypothesis (H1A) proposed that, for both familiarity and human-likeness, there would be a main effect of character type (human videos versus full animation versus lack of upper face animation). Specifically, regardless of emotion type, human videos would be rated highest on both DV measures followed by those in the full condition with lack videos rated lowest. Table 3 (and Figures 2 & 3) shows the mean ratings for familiarity and human-likeness associated with each emotion for each condition indicating support for the hypothesis. Consistently, videos of humans were rated as more familiar and human-like than those with full animation stimuli attracting higher ratings than those with no upper facial animation.  

---  

2 86.05% of participants had an advanced level of playing video games, with 13.95% a basic level of experience. For experience of using 3D modeling software: 46.51% had a basic level, 35.66% an advanced level, and 17.83% no experience.
3.1 Effect of stimulus condition on familiarity ratings

Table 3

Mean Ratings (and SD) for Familiarity and Human-Likeness in the Three Conditions across All Six Emotions (N = 129).

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Condition</th>
<th>Familiarity</th>
<th>Human-lokeness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Human</td>
<td>Full</td>
<td>Lack</td>
</tr>
<tr>
<td></td>
<td>M  SD</td>
<td>M  SD</td>
<td>M  SD</td>
</tr>
<tr>
<td>Anger</td>
<td>6.92 1.90</td>
<td>4.36 1.94</td>
<td>4.18 1.84</td>
</tr>
<tr>
<td>Disgust</td>
<td>6.44 2.23</td>
<td>4.62 1.87</td>
<td>4.30 1.76</td>
</tr>
<tr>
<td>Fear</td>
<td>6.03 2.22</td>
<td>4.30 1.95</td>
<td>3.60 1.86</td>
</tr>
<tr>
<td>Happiness</td>
<td>6.01 2.70</td>
<td>4.03 2.04</td>
<td>3.93 2.08</td>
</tr>
<tr>
<td>Sadness</td>
<td>6.54 2.01</td>
<td>5.05 1.90</td>
<td>4.57 1.83</td>
</tr>
<tr>
<td>Surprise</td>
<td>6.46 2.11</td>
<td>4.66 1.93</td>
<td>4.29 1.86</td>
</tr>
</tbody>
</table>
To assess the significance of these results a 3 x 6 repeated measures analysis of variance (ANOVA) and planned comparisons were applied to the data. The results of the ANOVA revealed a significant main effect of character type, $F (2, 256) = 30.436, p < .001$. Planned comparisons showed that, overall, there were significant differences between the ratings for human versus full animation characters, $F (2, 256) = 115.311, p < .001$; human versus lack of upper facial animation, $F (2, 256) = 155.687, p < .001$ and; full animation versus lack of upper facial animation, $F (2, 256) = 18.966, p < 0.001$, thus supporting Hypothesis H1A.
3.2 Effect of stimulus condition on human-likeness ratings

Figure 3 demonstrates that, once again, for all emotions, human videos attracted the highest ratings of human-likeness, a naturally anticipated result. This was followed by full then lack. However, unlike familiarity judgments, for this measure all emotions were rated very similarly in the human condition; very little variation in mean ratings of human-likeness can be observed across the six emotions for this condition.

Figure 3. Line graph of mean ratings for human-likeness in the three conditions across all six emotions.

ANOVA analysis of human-likeness data yielded a significant main effect of character type, $F (2, 256) = 10.944, p < .001$. Planned comparisons showed significant differences between: the
ratings of human versus full animation videos, $F(1, 128) = 333.314, p < .001$; human versus lack of upper facial animation, $F(1, 128) = 488.367, p < .001$; and full animation versus lack of upper facial animation, $F(1, 128) = 39.349, p < .001$, again supporting Hypothesis H1A.

### 3.3 Effect of emotion type on ratings of familiarity and human-likeness in each condition

Hypothesis H1B stated that perception of uncanniness (indexed by ratings of familiarity and human-likeness) would, in each condition, vary across the six emotions portrayed in the videos. Scrutiny of Figures 2 & 3 and Table 3 offer support for H1B. The emotion rated as the most familiar in the human condition was anger, whereas sadness was rated as the most familiar in the full and lack conditions. Happiness was rated lowest for familiarity both in the human and full conditions but not for lack. Disgust and surprise were rated almost identically in all three conditions. An interesting finding was that familiarity ratings of fear plummeted most dramatically when upper facial animation was lost. Consistent with the pattern of results for familiarity for measures of human-likeness, sadness emerged as the receiving the highest ratings in full and lack conditions, while fear ratings were substantially reduced when a lack of upper facial animation was introduced. Human-likeness ratings of happiness increased (marginally) between full and lack.

Two 3 x 6 repeated measures ANOVA and planned comparisons were conducted to assess the significance of the above. The results confirmed some support for Hypothesis H1B. A significant interaction between character type x emotion for both familiarity, $F(10, 1280) = 3.281, p < .001$, and human-likeness ratings, $F(10, 1280) = 6.065, p < .001$, were observed. Looking at the means (see Table 3), it appears that the transition between human to full to lack affects ratings of some emotions more than others. Specifically:

i) Expression of anger is rated highest on familiarity when viewing human stimuli, but this rating sinks comparatively dramatically when the participant is viewing a fully or partially animated character.
ii) Sadness makes a relatively more favorable transition from human to full animation conditions but, ratings for familiarity and human-likeness decrease substantially when there is no upper face animation.

iii) For fear, the shift from human to fully animated character conforms to that of other emotion types portrayed with a reduction in perceived familiarity and human-likeness. However, when upper facial animation is removed, the effect is more deleterious, (in terms of uncanniness) for fear than any other emotion. Familiarity and human-likeness ratings plunge lower than ratings for the other five emotions in Condition 3.

The results of planned comparisons supported the above observations. For every emotion, ratings of familiarity and human-likeness differed significantly between human and full and human and lack ($p < .001$). In terms of familiarity, significant differences between full and lack were observed for: fear, $F = (1, 128) = 16.362, p < .001$; sadness, $F (1, 128) = 9.876, p < .01$; disgust, $F (1, 128) = 4.686, p < .05$; and surprise, $F (1, 128) = 4.568; p < 0.05$. There were no significant differences observed between full and lack ratings of familiarity relating to anger ($p > .05$) or happiness ($p > .05$). For human-likeness, significant differences between full and lack were found for: fear, $F (1, 128) = 36.151, p < .001$; sadness, $F (1, 128) = 15.028, p < .001$; surprise, $F (1, 128) = 13.173, p < .001$; disgust, $F (1,128) = 9.667, p < .01$; and anger, $F (1, 128) = 5.181, p < .05$. No significant difference was found with regard to happiness ratings from full and lack ($p > .05$).

A clear pattern of results can be seen. For both familiarity and human-likeness; going from full animation to loss of movement in the upper face has the greatest adverse effect, in terms of perceived uncanniness, when fear is being portrayed in the video. This is followed by the negatively valenced sadness, surprise then disgust. Anger videos were only judged significantly lower in lack than full in terms of human-likeness; familiarity ratings were unaffected. Ratings of
the portrayal of happiness do not appear to have been affected by loss of upper facial animation. Collectively, the results offer substantial support for Hypothesis H1B.

4. Discussion

4.1 Effect of inadequate upper facial animation on the Uncanny Valley

A primary aim of the current research was to investigate the effect of inadequate facial movement (leading to an ambiguity in the emotional expression being portrayed) on perceived uncanniness in virtual characters. As was predicted, humans were rated the most familiar and human-like, followed by fully animated characters and then partially animated characters with movement disabled in the upper face region. These results may be accounted for in the following ways.

Uncanniness ratings were low in the human condition. The accurate facial expression of emotion (and expressions congruent with the intonation of speech, thus elaborating or confirming the tone used by the character) resulted in no sense of strangeness for the viewer. Conversely, in the lack condition, the total absence of upper facial animation accompanying speech may have evoked a sense of the virtual characters resembling the “waxwork figures or automata” referred to as characterizing the uncanny by Freud (1919, p. 226); the frozen upper face may have made it difficult to distinguish whether the characters were real (authentically human) or unreal or alive or dead, another catalyst for the sensation of uncanniness (Jentsch, 1906). The ambiguity of emotional expression portrayed by lack characters would have (also), unavoidably, presented an incongruity with the intonation of speech used by the character and resulted in a further sense of unease or uncanniness. Even in the full condition, the facial movements inevitably demonstrated subtle deviations from the human norm (Mori, 1970) and may have been perceived slightly at odds with the tone of speech used, leading to a similar, but less intense uncanny reaction.
These results add to our understanding of the possible functional significance of the experience of the uncanny. The uncanny may be related to the importance of being able to swiftly and accurately detect the emotion being expressed by another as it helps us predict their likely behavior. When a combination of tone of voice and facial expression is used to indicate a person’s affective state, the recipient expects a confirmatory congruence between the two. Any observed incongruence alerts people to oddness and the possibility of unpredictability of behavior which is alarming (even distressing and scary) as it may present a potential threat to personal safety. Hence, the sensation of uncanniness may serve to act as a sign of unpredictability and danger.

4.2 Effect of emotion-type on perceived uncanniness

The results of the current study also provide fairly compelling evidence that perception of the uncanny in virtual characters displaying inadequate facial animation is greatly influenced by the type of emotion the character is portraying. Those partially animated virtual characters expressing the emotions fear, sadness, surprise and disgust were judged significantly more uncanny than fully animated characters expressing the same emotions. In contrast, lack of upper facial animation had a lesser effect on perceived uncanniness of the emotions anger and happiness.

4.2.1 Fear, sadness and the uncanny

It is possible that the reason why the emotion fear (and to a lesser extent, sadness) was perceived as significantly more uncanny when emotional expressivity had been removed from the upper part of the face is that the partially animated stimuli represents a corpse or zombie-like state; a reminder of death’s omen (MacDorman, 2005; Mori, 1970). Both fear and sadness require the smallest movement of both upper and lower facial features (Pollick, Hillô, Calder & Paterso, 2003) and when animation is removed from the upper part of the face, the character’s physiognomy may
then appear paralysed, reminiscent of a rigamortis state. Furthermore, as Spadoni (2000) suggested, any incongruence in the upper facial movement and tone of speech representing an emotional state (as in the lack condition of this study) leads to a, “Frankenstein’s monster effect” (Spadoni, 2000, pp. 102-3).

An alternative explanation is that facial expression of these “aversion-reinforcement” emotions are regarded as comparatively more uncanny when features of them are somehow unexpectedly aberrant as in the lack condition. The possibility exists that uncanniness becomes ever more salient when it is difficult for the viewer to interpret these particular emotional states because people are acutely aware of the potential negative consequence of misinterpretation. The ambiguity or uncertainty presented by inadequately animated virtual characters attempting to convey fear and sadness may be a hair-trigger for elicitation of a sense of the uncanny, as a precursor to a fight-or-flight response. There are parallels here to discussions on the relationship between uncertainty and sound (Ekman & Kajastila, 2009; Kromand 2008) and sound, fear and the Uncanny Valley (Grimshaw, 2009) which likewise identify a fearful response to sonic ambiguity in computer games of the survival horror genre.

The emotion fear is, inevitably, strongly associated with the uncanny (Ho et al., 2008). Previous conceptualizations of what leads to a sense of the uncanny may help in understanding why this is. Freud defined the uncanny as an object that appears unfamiliar (Freud, 1919) and the unfamiliar naturally evokes fear. For example, in children, the unfamiliar, such as their mother (uncharacteristically) wearing a bathing cap, or deviations in expectations of how an object should and will behave and how it actually behaves, elicits fear (Charlesworth & Kreutzer, 1973). Given this, the observed relatively stronger sense of the uncanny in odd virtual characters attempting to portray fear, is, perhaps, predictable.

With regards to sadness, despite the fact that removing upper facial animation in the virtual character led to lower familiarity and human-likeness ratings, participants rated both fully and partially animated virtual characters expressing sadness as comparatively less uncanny than when
exhibiting any other emotion (see Figures 2 & 3). An explanation for this may be due to the human tendency to anthropomorphize virtual characters (Gong, 2008), a tendency, it is proposed, that meets two needs; social interaction and control of the environment (Epley, Waytz, Akalis & Cacioppo, 2008). However, a consequence of this may be that when we encounter a virtual character expressing sadness we may experience a level of empathy with them or sympathy towards them. Even when that character’s face might be considered lacking full communicatory value by having little or no movement in the upper face region, the distinctive lower face expression of sadness (droopy mouth) may be sufficient to compensate. Any sense of unfamiliarity of lack of human-likeness may, in these circumstances, be tempered or obscured by a sense of sympathy.

4.2.2 Surprise and the uncanny: the perils of ambiguity

Given that the ability to accurately detect surprise in others is not necessarily important to one’s personal survival (certainly not as critical as the ability to detect fear or anger in others), it is a little more unexpected that shifting from fully to partially animated virtual characters who were conveying surprise was accompanied by an increased sense of uncanniness. However, the similarities in facial expression for the emotions surprise and fear make them the most difficult emotions to distinguish between (Ekman, 2003). This is corroborated by data from the current study. Although high emotion recognition ratings were achieved for surprise in both the human and animated characters with over half of the participants identifying the emotion surprise correctly even in the partially animated state (see Table 2), errors in recognition of surprise, in both virtual character conditions, were almost exclusively due to the participant mistaking surprise for fear (see Table 2). In everyday life there is, of course, less ambiguity as the surprise is affectively contextualized by antecedent events. However, this result suggests that at least some of the characteristics that make fear significantly more uncanny than others in the lack condition may also be present for the emotion surprise. This may account for the greater than anticipated uncanniness
ratings to surprise witnessed for conditions full and lack and data more closely resembling the negative emotions (e.g. fear, sadness, disgust) than the positively valenced happiness.

4.2.3 Disgust

Blair (2003) states that expression of the emotion disgust serves the adaptive function of warning others to be concerned about a potentially dangerous situation or distasteful object. While this unsettling reaction may provide an explanation as to why disgust was regarded as significantly more uncanny in the full virtual characters and even more so for the lack characters, the observed effect is probably more attributable to the current limitations in facial animation software. Previous studies have shown an inadequate representation of the AU9, Nose Wrinkler in realistic, human-like characters due to low-polygon counts in this area (Dyck et al., 2008). As this is the most distinctive facial movement involved in expression of disgust, any perceived mismatch between this and the other facial movements involved in its expression (those in keeping with the human-like portrayal of disgust) may have exaggerated the uncanny. The expression appeared indistinct and wrong. Until modifications are made to improve the graphical realism for this facial movement, disgust may continue to be perceived as uncanny regardless of modifications made to other parts of the face.

4.2.4 Happiness

A lack of upper facial cues during speech did not influence perceived uncanniness for the emotion happiness. Indeed, the virtual character was regarded as slightly more human-like in the lack condition (see Figure 3). Explanations for this might include the possibility that the Uncanny Valley phenomenon is primarily driven by fear: a fear of one’s own death (Mori, 1970; MacDorman 2005); a fear of being over-powered by a man-made machine (Kang, 2009); a fear of being replaced by a doppelganger, and a fear of the unfamiliar (Freud, 1919). Unlike those emotions likely to cause
fear and alarm, an expression of happiness is more likely to evoke a more positive reaction from the viewer. Happiness is predominantly expressed by the mouth in the lower face region (Busso & Narayanan, 2006; Ekman, 1979; Ekman & Friesen, 1978). Hence, it may be that by disabling or reducing movement in the upper face region, the character is regarded more favorably. Viewers may concentrate in greater detail on the smile shape of the mouth, associated with evoking the positive reaction in this emotion.

Happiness was rated lowest for familiarity in the human and fully animated virtual character and lowest for human-likeness in the fully animated character. This result was unexpected, due to the more positive connotations associated with happiness. However, given that the smile produced by the actor was choreographed (not spontaneous) and that the smile of the virtual character was modeled on the actor’s, the low familiarity and human-likeness ratings may be attributed to participants being wary of being presented with a “false smile” (Ekman & Friesen, 1982, pp. 244-248) and equating this with strangeness.

4.2.5 Anger

A surprising result of this study was that anger was found to be no more uncanny with a lack of movement in the eyebrows and lids. As a lowered brow is such an important feature of this emotion, one may have expected results contrary to this study (Ekman, 2004). The perception of anger is, like fear crucial in one’s survival instincts, in that people have primal impulses to combat or flee in the presence of anger. Participants’ attained highest scores for familiarity and human-likeness for the emotion anger in the human, highlighting our inherent ability to recognize this emotion (see Figures 2 & 3). Yet, when displayed in a more primitive state in the virtual character, anger was regarded as no more uncanny with a lack of facial expressivity in the upper face. This may be attributable to other factors (such as prosody or appropriate lower facial animation).
eliminating or tempering a sense of the uncanny in characters attempting to portray anger or, indeed, happiness.

4.3 Limitations and future work

The current project used subjective ratings of stimulus uncanniness to assess the effect of strategic character design manipulations. Although this approach is well-established (e.g. Bartnek et al., 2009; Hanson, 2006; MacDorman, 2006; Seyama & Nagayama, 2007; Tinwell et al., 2010), it inevitably allows for the possibility of a number of variables contaminating the data (poor participant attention and motivation, individual differences in understanding of or defining the DV measures, and variability in scaling benchmarks). To help overcome this, future work will use the same stimuli and conditions (human plus fully or partially animated virtual characters) while measuring Autonomic Nervous System (ANS) electrophysiological responses (e.g. skin conductance and heart rate) to depictions of the six basic emotions. Participant ratings of the DV measures will also be taken. In this way, information can be gained on both the physiological and psychological impact of the uncanny in virtual characters and, importantly, the relationship between the two.

4.4 Conclusion

Data from this study imply that animated, high-fidelity, human-like, talking-head, virtual characters are rated by users as uncanny (less familiar and human-like) but significantly more so when movement, and therefore emotional expressivity, is limited in the upper face. More important, the magnitude of this increased uncanniness varies depending on which emotion is being communicated. Under these conditions, the emotions fear, sadness, disgust, and surprise, evoke a
strong sense of the uncanny but, despite aberrant upper facial movement, uncanniness is less noticeable for the emotions anger and happiness.

Overall, the results indicate that attempts to embed truly authentic and convincing human-like affective signals in video game characters still has some way to go. However, as previous authors suggest (McDorman, 2006; Tinwell et al., 2010), games designers may be able to make informed decisions as to how to control the perceived level of uncanniness in virtual characters. Modifications can be made to facial expressivity in the upper face, bespoke to each different emotion, allowing the designer to design for or against the uncanny as required. To avoid the uncanny in characters, designers should pay particular attention to upper facial expressivity when attempting to portray fear, sadness, and to a lesser extent disgust and surprise. With such little facial movement required to adequately convey the emotions fear and sadness any lack of movement in the upper face for these two emotions may evoke a perception of those characteristics typical of a zombie or corpse, increasing uncanniness and fear in the user. Similar design principles may be applied to the emotion surprise given the common misconception between this emotion and fear. Such modifications may be preferable for characters intended to be frightening depending on the game context. In contrast, designers wishing to successfully depict anger or happiness in their character need not invest substantial amounts of time in embedding upper facial expressivity, especially so if lower facial features and prosody support recognition.

Acknowledgments

For this study thanks is given to Daniel Whitehead for acting as the human stimulus, and Brennan Tighe and Roy Attwood at the University of Bolton for providing technical support with recording the video footage of the human. Thanks are also due to our colleague John Charlton for his expert advice and opinion.
References


