Distribution Agreement

________________ ______________

Shensheng Wang Date

THE UNCANNY VALLEY: EXISTENCE, EMOTIONAL COMPONENTS AND POSSIBLE EXPLANATIONS

By

Shensheng Wang

Master of Arts

Developmental Psychology

____

PHILIPPE ROCHAT, Ph.D.

Advisor

____

PHILIPPE ROCHAT, Ph.D.

Committee Member

____

PHILLIP WOLFF, Ph.D.

Committee Member

____

SCOTT O. LILIENFELD, Ph.D.

Committee Member

Accepted:

____

Lisa A. Tedesco, Ph.D.

Dean of the James T. Laney School of Graduate Studies

____

Date

The Uncanny Valley: Existence, Emotional Components and Possible Explanations

By

Shensheng Wang

B.S., Nankai University, 2012

Advisor: Philippe Rochat, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of Master of Arts in

Developmental Psychology

2014

Abstract

The Uncanny Valley: Existence, Emotional Components and Possible Explanations

By Shensheng Wang

As robots gradually resemble humans, they become increasingly familiar to people, until a point at which their subtle imperfections render them uncanny, eliciting the so-called "Uncanny Valley" phenomenon (UVP). In the recent decades, the UVP has become a pressing issue that hinders developing more realistic androids and 3D animations. However, researchers disagree on whether the phenomenon is real or not, and if it is real, how to explain it. Specifically, the uncanny feeling linked to the faces of human replicas remains unexamined from a strictly psychological perspective. We addressed these issues with three studies: In the first study, we conducted a survey in which 62 adult participants were asked to rate the realism of 89 human and android faces and the emotional responses these faces elicited, such as eeriness, disgust and fear. In the second study, we tested another 62 adult participants with a visual looming task (Vagnoni, Lourenco, & Longo, 2012) to implicitly measure the emotional responses to the 89 faces. In the third study, we tested the hypothesis that the uncanny feeling is associated with the sensation of uncertainty (Jentsch, 1906) by conducting a reaction time-based sorting task, in which another 36 adult participants were asked to quickly sort faces as either real or unreal and their reaction time was recorded. In the first two studies, we plotted the emotional responses against the realism and showed that their relation resembled the uncanny valley in their trends and shapes, suggesting the existence of the UVP. The first study also indicated that the UVP may have the emotional components of eeriness and disgust. The third study confirmed the uncertainty hypothesis by showing that the faces that were sorted with uncertainty manifested by longer reaction time were associated with the strongest negative emotional responses. Possible explanations for the results and future directions were proposed.

The Uncanny Valley: Existence, Emotional Components and Possible Explanations

By

Shensheng Wang

B.S., Nankai University, 2012

Advisor: Philippe Rochat, Ph.D.

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Arts

in Developmental Psychology

2014

Acknowledgements

I would like to thank my advisor, Dr. Philippe Rochat, for his guidance and support throughout this entire research project. I am also thankful to my committee members, Dr. Phillip Wolff and Dr. Scott O. Lilienfeld, for their help and advice. Finally, I would like to express my gratitude to members of the Rochat Lab, family, and friends for encouraging me throughout this research.

Table of Contents

List of Tables and Figures

The Uncanny Valley: Existence, Emotional Components and Possible Explanations

Introduction

We live in an uncanny world. From the déjà vu experience to the prevalence of Zombie in a pop culture, the uncanny experience haunts us in various situations. Although the uncanny experience is not unfamiliar to people, it has not been well understood and lacks scientific investigation. The recent decades witness a growing interest in this mysterious feeling, provoked by androids and computer-generated imagery (CGI) that increasingly resemble humans but reach the stage of being "uncanny".

This specific problem of how realistic a human replica ought to be to appear likable is part of the Uncanny Valley phenomenon (UVP). The term "uncanny valley" was initially coined in 1970 by a Japanese roboticist Masahiro Mori, who predicted that as robots progressively resemble humans, they will become increasingly familiar to us; yet before they achieve 100% human resemblance to be acceptable, there will be a point where they appear unfamiliar to us and elicit an intense uncanny sensation (Mori, 1970).

Does the uncanny valley exist? If it exists, what is the uncanny feeling when we see a realistic face falling into the uncanny valley? Mori did not answer these questions, nor did previous empirical studies that attempted to reconstruct the uncanny valley show consistent evidence regarding its existence (Bartneck, Kanda, Ishiguro & Hagita, 2007; Hanson, 2005; MacDorman, 2006, MacDorman & Ishiguro, 2006). With two innovative paradigms, a visual looming task and a reaction time-based sorting task, in addition to a face rating survey, the present study is intended to contribute to the empirical investigation on the UVP, in the context of which we hope to answer the questions of 1) whether the uncanny valley exists, 2) what the "uncanny" feeling is, and 3) why this phenomenon occurs in androids, computer-generated characters and other human replicas.

The Uncanny experience

You may encounter the uncanny feeling many times in your personal experience but it is hard to recall vividly without being in the actual situations that elicit this feeling. To get a sense of the uncanny feeling, imagine that you shake a person's hands, whose cold temperature immediately reminds you that it is in fact a prosthetic limb. Maybe instead of feeling empathetic about the person who lost his arms, you experience an eerie sensation, presumably due to the fact that you mistake the prosthetic hands for real (Mori, 1970; Poliakof, Beach, Best, Howard, Gowen, 2013).

A prosthetic hand illustrates the phenomenon of the uncanny valley: the realism of a human replica increases its familiarity, yet it can be too realistic to be repulsive. Anecdotal evidence has shown that an android, the robot designed to be indistinguishable from humans in its appearance and behaviors falls into the "uncanny valley" (MacDorman & Ishiguro, 2006). In addition to that, the UVP has been reported in wax figures ("*House of Wax*", 1953, 2005), sculptures (Brenton, Gillies, Ballin, & Chatting, 2005), mannequins, dolls, and 3D animation characters (Kaba, 2013), many of which have been the themes of horror movies ("*The Hills Have Eyes*", 2006), or generate an unexpected horror ("*The Polar Express*", 2004). The uncanny valley has been found also in real people who undergo plastic surgery to resemble dolls, and the photographs of symmetrical faces.

Instead of being a newly-established concept brought by the androids and computer-generated characters, the discussion on the "uncanny" feeling associated with human replicas can be dated back to the early 20th century when Ernst Jentsch (1906) and Sigmund Freud (1919) both mentioned the "uncanny" experience in the context of Holffmann's story "Sandman" (1816) featuring a lifelike doll Olympia. Nevertheless, the uncanny experience is not limited to dolls and androids. As Freud noted, the uncanny feeling can be experienced in castration fear, involuntary repetition, coincidences, life after death, being buried alive and many other instances (Freud, 1919). However, most of these situations associated with "uncanniness" can hardly be tested empirically. Thanks to the UVP, the human-like artificial faces of androids and other human replicas may serve as a test bed for researchers to examine the uncanny feeling from an experiential perspective.

Theories on the uncanny valley phenomenon

1. <u>Imperfections in realism</u>

What makes an android uncanny? The notion that the uncanny feeling is due to the imperfections in the very realistic appearance and/or behavior of an android that abruptly render it unreal has been proposed by Mori and many other researchers either directly or indirectly (MacDorman & Ishiguro, 2006; Mori, 1970). In previous studies, researchers have studied various factors that make a robot unreal, including its physical appearance of the face (Hanson, 2005; MacDorman, Green, Ho, & Koch, 2009; Seyama & Nagayama, 2007), emotional expressions (Tinwell, Grimshaw, Nabi, & Willams, 2011), body movements (Chaminade et al, 2007; Groom, Nass, Chen, Nielsen, Scarborough, & Robles, 2009; Piwek, McKay, & Pollick, 2014; Saygin et al., 2012;

Thompson, Trafton, and McKnight, 2011), voice (Mitchell, Szerszen, Lu, Schermerhorn, Scheutz, & MacDorman, 2011) and a combination of multiple factors contributing to the realism (Tinwell, Grimshaw, & Williams, 2010)

Although the various aspects of imperfection in realism for a robot are linked to the UVP, the ways in which it is imperfect do not converge to offer a unified explanation for why the UVP occurs. Therefore, alternative theories that go beyond the imperfections in realism have been proposed to explain the UVP from the following perspectives.

2. Evolutionary/developmental origin

The evolutionary perspective attempts to explain the origin of the UVP in the context of human evolution by stating that the uncanny valley stems from self-preservation, a fundamental issue about humans ourselves.

2.1 Pathogen avoidance

The evolutionary perspective in the UVP originated from Mori who suggested that the uncanny response might be crucial for the self-preservation of human beings (Mori, 1970). Christian Keysers suggested a link between the uncanny feeling and Rozin's Theory of Disgust (Rozin & Fallon, 1987), and interpreted the uncanny feeling as disgust, resulting from an evolved cognitive mechanism for pathogen avoidance that serves adaptive functions (Ho, MacDorman, & Pramono, 2008; MacDorman & Ishiguro, 2006). According to this view, humans perceive the androids as carrying transmittable diseases that they are bound to avoid.

2.2 Mate selection

Evolutionary pressure shapes human preference for beauty and this aesthetic preference is hardwired into the nervous system of humans (Rhodes & Zebrowitz, 2001). An explanation based on evolutionary aesthetics is therefore proposed to account for the UVP (MacDorman & Ishiguro, 2006). In this view, humans have evolved under the selection pressure the sensitivity to the fitness of the mating partners and the fitness is often indicated by the physical appearance. Consequently, the uncanny feeling linked to the androids may stem from their unrealistic appearance that may suggest low fertility or weak immune system. Hanson (2005) showed that the UVP can be overcome by improving the aesthetic property in android design which, according to Hanson, illustrates the importance of attractiveness in the UVP.

2.3 Terror management theory and the fear of mortality

In addition to disgust and attractiveness, Sara Kiesler proposed that the uncanny feeling, according to terror management theory (Greenberg, Pyszczynski, Solomon, Simon, & Breus, 1994), may be associated with the fear of inevitable death of human beings and may be elicited by an android. To test this hypothesis, MacDorman and Ishiguro (2006) examined how viewing an android face versus an Asian girl would affect the worldview protection (a form of distal defense subconsciously modulated by death-related concepts) of the participants differently. Their findings confirmed the proposed hypothesis that the uncanny feeling may be due to the fear of death. Alternatively, the fear may be linked to the imagination of being replaced by the android Doppelganger, losing body control, or being deprived of soul (Ho et al., 2008).

2.4 Evolutionary versus Developmental origin

To demonstrate that the UVP has an evolutionary origin, Steckenfinger & Ghazanfar (2009) studied the preferential looking of monkeys among real money faces, unrealistic synthetic monkey faces, and realistic synthetic monkey faces. They found that monkeys preferred real monkey faces and unreal synthetic faces to realistic synthetic faces, indicating the uncanny valley of monkey. With a similar experimental paradigm, Lewkowicz and Ghazanfar (2012) tested the alternative explanation: the uncanny valley could be developmental in origin, where experience plays a crucial role. They found that not until 12 months of age did infants show a visual preference for human faces and unrealistic avatar faces to realistic avatar faces. Their findings suggested that the UVP may emerge from the early experience with real human faces through perceptual learning that biases our responses towards conspecific faces rather than faces of other species, or non-human beings. However, the nature-nurture debate on the origin of the UVP has not been resolved with research showing contrasting evidence to Lewkowicz & Ghazanfar (2012) that 7-month-old infants equally attended to the face of their own mother, a stranger but avoided the intermediate morphing of the two (Matsuda, Okamoto, Ida, Okannoya & Myowa-Yamakoshi, 2012). To what extent the UVP is evolutionary in origin as opposed to learned with experience requires further investigation.

3. Expectancy Violation, Categorical Uncertainty and Bayesian model

Other influential explanations for the UVP focus on the perceptual and cognitive mechanisms underlying the perception of human replicas. For example, the violation of expectation hypothesis suggests that very realistic human replicas elicit expectations for being real, but cannot measure up to these expectations due to the imperfections in their appearance and behaviors (MacDorman, 2006; Mitchell, Szerszen, Lu, Schermerhorn, Scheutz, & MacDorman, 2011). The violation of expectation may derive from a mismatch between appearance and behavior (Saygin, Chaminade, & Ishiguro, 2010; Saygin, Chaminade, Ishiguro, Driver, & Frith, 2012), among incongruent sensory inputs from multiple modalities (Mori, 1970; Steckenfinger & Ghazanfar, 2009), or result from conflicting cues of even a single modality. For example, Mori (1970) mentioned that prosthetic hands with hair and vein violated the expectations for being real due to its cold temperature and lack of soft tissue. He suggested that the mismatch between visual and tactical input created a sense of strangeness. Empirical evidence has confirmed this uncanny feeling generated by a cross-modal mismatch (Mitchell, Szerszen, Lu, Schermerhorn, Scheutz, & MacDorman, 2011; Tinwell, Grimshaw, & Williams, 2010). According to Mitchell et al (2011), a mismatch in the human realism of the appearance of a face and the voice was correlated with the perceived eeriness. In addition, motion exaggerated the uncanny phenomenon and magnified the perceived fear (Saygin et al., 2011; Tinwell et al., 2010).

In fact, a variety of situations in which people experience unease may be linked to the violation of expectation (Olson, Roese, & Zanna, 1996). Therefore, MacDorman (2006) warned that the uncanny feeling may be linked to the violation of expectation specifically in human forms. Regardless of the specific forms of expectancy violation, the underlying cognitive and affective mechanisms of the UVP remain unexplained: Why experiencing expectancy violation (e.g., a surprise) is necessarily linked to the uncanny feeling? To answer why the violation of expectation about a human replica may be linked to the uncanny feeling, researchers proposed the categorical uncertainty hypothesis.

The pioneering work on the categorical uncertainty hypothesis about the uncanny feeling was done by Jentsch (1906). In the article "On the psychology of the uncanny", Jentsch noticed, "Doubt and uncertainty, particularly whether an apparently animate being is really alive; or conversely, whether a lifeless object might not be in fact animate, are at the core of the uncanny". Although the uncertainty appears in many theoretical accounts for the UVP, it has not been empirically tested until Moore (2012). Using a Bayesian model, Moore (2012) extended the "perceptual magnet effect" (Feldman, Griffiths, & Morgan, 2009) from one to multiple perceptual cues to examine the hypothesis of categorical uncertainty. The "perceptual magnet effect" predicts that people are more sensitive to the distinctions at categorical boundaries than those far from the boundaries; thus, the perception at the categorical boundaries is distorted. According to Moore, when people categorize faces based on multiple perceptual cues (e.g., size of the eyes, skin color, etc.), the resulting perceptual distortion linked to each perceptual cue may indicate misaligned categorization, which creates a perceptual tension that in turn is experienced and labeled as the uncanny feeling.

4. Mind Perception/anthropomorphism

Instead of the attention on the imperfections of the human replicas, Gray and Wegner (2012) emphasize the extraordinary realism they have achieved to the extent that it prompts human observers to attribute mind to the replicas, that is, to anthropomorphize inanimate entities as humans. And the attributions of mind, particularly the human experience (the ability to feel and sense) rather than agency (the ability to act and do), is fundamental and unique for a real person but not an android (Carpenter, Eliot, & Schultheis, 2006; Gray & Wegner, 2012).

Defining the uncanny valley phenomenon

In recent years, studies have tested the previously mentioned theories with a wide variety of stimuli (for example: a continuum of morphed facial pictures from a doll to a person, Hanson, 2005; computer-generated avatars of incremental photorealism, MacDorman et al., 2009; pictures of existing humanlike robots of various forms, Ho & MacDorman, 2010; real androids in the lab, MacDorman, Minato, Shimada, Itakura, Cowley, & Ishiguro, 2005) and experimental paradigms (for example: subjective ratings; fMRI, Saygin et al., 2012; preferential looking, Steckenfinger & Ghazanfar, 2009), but only generate mixed support for its existence (Bartneck, Kanada, Ishiguro, & Hagita, 2009; Burleigh et al., 2013; Hanson, 2005; Pollick, 2010; Thompson, Trafton, & McKnight, 2011). It has raised fundamental theoretical questions about how to define the UVP.

1. <u>The interpretation of "Shinwa-kan"</u>

The discrepancy in the previous findings about the existence and explanations of the UVP is partially due to the confusion in translating and conceptualizing the dependent variable of "shinwa-kan", the y-axis in the hypothetical graph of the UVP. The original translation of shinwa-kan was familiarity (Mori, 1970). However, researchers have questioned this translation, because it is hard to imagine a negative familiarity (MacDorman, 2006). Alternative translations are proposed. In fact, "shinwa-kan" is a neologism in Japanese that does not have a direct equivalent in English. However, by separating the word into two components "shinwa" and "kan", which mean "mutually be friendly" or "having similar mind", and "the sense of ", Bartneck et al., (2005) proposed that “affinity” and in particular “likability” are more suitable translations than “familiarity”.

2. The specification of human realism

Another source of disparity in previous findings is linked to the ambiguity in the definition of human realism. David Hanson (2005) pointed out the importance of defining human realism. He defined it as “being within the possible, naturally-occurring appearance of real human beings.” He then noted that realism can be defined across several domains: static, dynamic appearance and “contextual responsiveness” (contingent interaction). However, in order to obtain experimental control, researchers often systematically manipulated and examined only one aspect of realism in their studies, overlooking the multidimensionality nature of human realism. This methodological limitation often hinders the interpretation and integration of findings across different researchers and studies.

3. Undermined validity

Another problem of the inconsistency is ecological validity, particularly the difficulty in generalizing the empirical findings from one experiment to another, and from empirical settings to real life situations. Even among the studies that replicated the UVP (MacDorman, 2006; Seyama & Nagayama, 2007; Saygin et al., 2011), researchers still face the challenge of generalizing their explanations to a different set of experimental stimuli. For example, with a doll-to-person continuum of morphed pictures, Seyama & Nagayama (2007) reconstructed the uncanny valley by showing the pleasantness level of the participants varies along human realism as predicted in the UVP. However, their hypothesis that the uncanny valley is due to perception of abnormal facial features cannot account for the UVP observed in Saygin (2011), where an android robot performed human behavior without showing the previously-tested facial abnormalities. Even using the same technique of morphing but a different pair of doll and real human faces to create a realism continuum, Hanson (2005) in his experiments showed that the UVP could be reconstructed, surpassed, and even inverted, depending on the stimuli to be morphed. Therefore, a serious experimental concern arises regarding whether the so-called uncanny feeling reported in these experiments was essentially the same across studies, which significantly undermines the validity of these findings.

The emotional components of the uncanny valley phenomenon

The clarification of the meanings of “shinwa-kan”, “uncanny” and “human realism” is one crucial step leading to a consensus among researchers on defining the UVP. In addition to the clarification of these concepts, research questions need to address the fundamental issues in the UVP.

Since Mori proposed the UVP, most research questions have focused on explaining how this phenomenon occurs, without first showing strong empirical evidence concerning its existence. On top of that, the present studies lack in a clear definition of the uncanny feeling. Whether the uncanny feeling is primarily about fear (mortality salience hypothesis), disgust (pathogen avoidance hypothesis), or a sense of unease associated with human replicas (uncertainty hypothesis) is unknown.

In response to these questions in the UVP, Ho et al. (2008) analyzed the subjective ratings of animate robots that varied in appearance and behavior from mechanical to realistic. They found the term fear highly predictive of attributions of eerie or creepy to the robots. Disgust, shock, and nervousness were also significant predictors. Instead of reducing eerie or creepy to a single emotional term, the findings suggested that "the uncanny valley may not be a single phenomenon but rather a nexus of phenomena with disparate causes." Interestingly, the findings from subjective ratings coincide with Freud's interpretation of uncanny feeling by examining the meanings of the term "uncanny" in multiple languages. In his essay *The Uncanny*, Freud mentioned that

"The subject of the 'uncanny' is a province of this kind. It undoubtedly belongs to all that is terrible—to all that arouses dread and creeping horror; it is equally certain, too, that the word is not always used in a clearly definable sense, so that it tends to coincide with whatever excites dread. Yet we may expect that it implies some intrinsic quality which justifies the use of a special name. One is curious to know what this peculiar quality is which allows us to distinguish as "uncanny" certain things within the boundaries of what is "fearful." (Freud, "The Uncanny", 1919)

Has Freud's "peculiar type of fear" description of the uncanny been empirically tested? If so, is the meaning of "uncanny" clear enough for researchers to reexamine the mechanisms underlying the UVP? YES and NO. On one hand, the findings in Ho et al. (2008) have indicated that fear may be a major emotional component of the uncanny feeling. On the other hand, however, a sample of 14 animate androids varying in both their appearance and behavior realism may not be able to represent the full range of human realism in the UVP. Furthermore, the researchers failed to distinguish between the appearance and behavior aspects of realism, leaving the independent variable human realism not defined. The present study is intended to remedy these caveats while testing whether the UVP exists in Study 1 and 2, using static images of 89 faces of human replicas and real people of various forms. Furthermore, we conducted Study 3 to examine a possible explanation for this phenomenon known as the categorical uncertainty hypothesis.

Current Studies

We conducted three studies to answer whether the UVP exists, and if it exists, what are the possible explanations. In Study 1, we conducted a face rating survey with 89 static facial images to test if the uncanny valley exists. We asked participants to rate 8 statements about the realism of each face and their emotional responses to it. Then we plotted ratings on each emotional response against the ratings on realism.

In Study 2, we conducted a computer-based visual looming task to confirm our previous findings. The hypothesis was: if the UVP exists, in addition to the subjective ratings, the behavioral responses will also form a relation with the realism that resembles the uncanny valley function. In the visual looming task, we presented the same 89 faces on a looming trajectory towards the face of the participant from a computer screen. Then we asked participants to estimate the speed of movement of each face by pressing a button when they feel the looming face makes contact with their own faces. Inspired by Vagnoni, Lourenco, &Longo (2012), we associated the reaction time of estimated time-to-collision to the emotional responses to the faces, with which we plotted against the ratings on realism from Study 1 in order to reconstruct the uncanny valley function.

In Study 3, we examined whether the UVP is due to the sense of uncertainty using a reaction time-based sorting task. In the sorting task, we presented each of the 89 faces on a computer screen and asked the participant to sort them as either "real" or "unreal" as quickly as possible. We analyzed the choice as well as the reaction time of the participants in sorting each face to a designated category. Finally, we tested the uncertainty hypothesis in explaining the UVP by using sorting data to predict the various ratings of the faces.

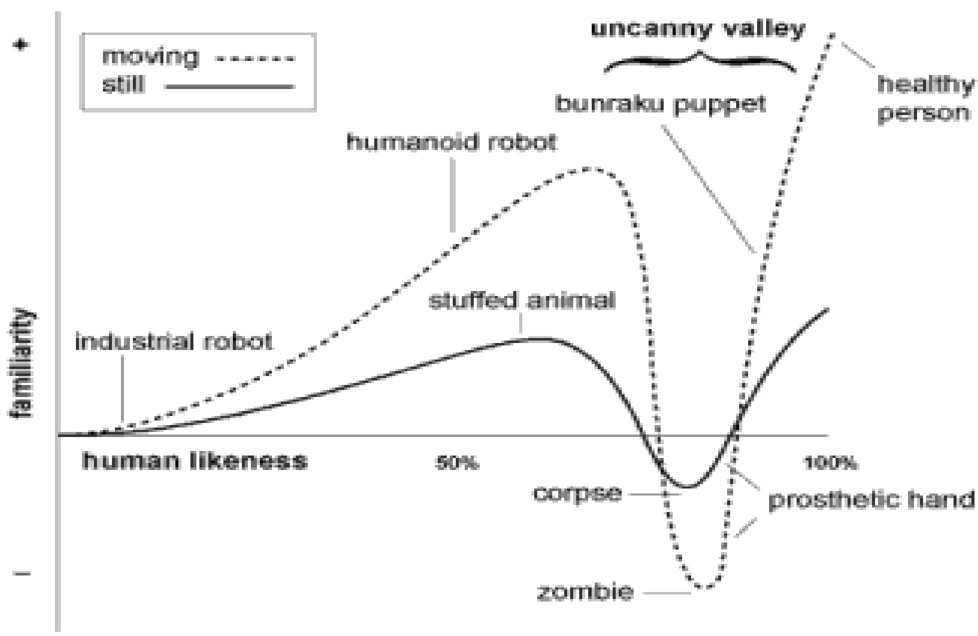


Fig. 1 Mori's 1970 graph of the uncanny valley

Study one

In this study, we examined whether the UVP exists or not by plotting emotional ratings against ratings on realism. If it exists, we asked which emotional responses form a function with the facial realism that characterizes the UVP.

Method

*Participants***.** Sixty-two undergraduate psychology students (39 females; mean age = 24.18 years, SD = 9.01) were recruited from Georgia Highlands College and participated in return for course credit. All participants provided oral informed consent.

Stimuli and procedure

Static facial images

Face rating survey included 40 “uncanny” faces resulting from searching “Uncanny Valley” in Google Images, and 49 normal faces. These uncanny images consisted of human replicas (androids, dolls, wax figures, mannequins, computer-generated imagery (CGI), Zombies) and faces from real people (human Barbie dolls, plastic surgery, faces with Bell’s palsy and symmetrical faces created by mirroring one half of the face). The normal faces were selected both from the Google Images and from college student photographs we took for previous studies, and were matched onto the uncanny faces individually in gender, age and facial features (e.g., mustache). All images were resized (300 px width 380 px height) and their external contour was cropped in an oval shape in Photoshop CS6 software.

Procedure

In a 30-person classroom, an experimenter conducted a paper-based survey among the students (i.e., participants) to obtain face ratings. Before the survey, the experimenter informed the participants that they were going to see some faces, and their task was to rate each face on 8 statements about it from a seven-point Likert scale (1-Not at all, 7-Very much, see Appendix A). During the survey, the experimenter presented a PowerPoint slide consisting of 89 faces (one face per slide) on a projector screen (170 cm width 127 cm height) located about 4 meters in front of the participants. The experimenter presented one slide at a time for proximately 45-60 seconds, during which all the participants had sufficient time to rate a face on the 8 statements. Then the experimenter moved to the next slide. The whole survey took about an hour.

The 8 statements were of the form either “This figure makes me feel_____,” or “This figure looks _____,” and the blank was filled with one of 8 terms. The 8 terms consisted of “eerie”, “disgusted”, “unsettling”, “attractive”, “threatening”, “likable”, “real”, and “alive”. Among these terms, real rating was included to measure the facial realism. Disgusted, threatening, alive and attractive were included to directly test the theoretical explanations of the UVP as pathogen avoidance mechanism (Rozin & Fallon, 1987), fear of mortality (MacDorman & Ishiguro, 2006) and aesthetic concern (Hanson, 2005). In addition, eerie was included because Ho et al (2008) indicated eerie as best capturing the visceral reaction in the uncanny valley, yet other studies have shown mixed evidence regarding whether the uncanny valley exists if eerie rating is plotted against real rating (MacDorman & Ishiguro, 2006; MacDorman, 2006; MacDorman et al., 2009; Poliakoff et al., 2013; Thompson, et al., 2011). Finally, unsettling was included because it was frequently used to describe the feeling people experience in the UVP in addition to eerie (Hanson, 2005; MacDorman, 2005; Steckenfinger & Ghazanfar, 2009) but was not directly tested.

We hypothesized that if the phenomenon exists, we would expect graphs depicting the emotional responses as a function of facial realism to exhibit a similar pattern predicted by the UVP (Fig. 1): as the facial realism increases, there will be an initial increase to a peak, then a decrease to the valley, finally is followed by a rebound exceeding the first peak in the emotional response.

Results

Descriptive statistics of the rating data revealed substantial variability across participants, which indicated that although participants used the same Likert scale to rate faces, they differed significantly in how they associated the rating scores with their own evaluation of the facial realism. Therefore, the ratings of a face were not informative unless they were compared to the ratings of other faces from the same participant. And by evaluating one face in comparison to the other faces, we were able to compare the ratings across the participants. In other words, although each individual participant differed in their overall inclination in rating higher or lower scores, we assumed they maintained a consistency in their own rating, particularly on a single statement across the 89 faces. Therefore, on each of the 8 statements per participant, the original rating scores were standardized by taking the mean and standard deviation of rating scores across the 89 faces, resulting in 89 Z scores. For example, Participant 1 rated 89 faces on eeriness with a mean of 1.90 and standard deviation of 1.91. As a result, the original rating score of the participant on Face 1 as 5 was transformed to (5-1.90)/1.91 = 1.62. In light of this standardization, a Z score on eeriness indicated the eeriness of a face in relation to the other 88 faces for a given participant. Furthermore, after standardization the Z scores were also comparable across the participants. The following analysis was all based on the standardized rating scores.

First, for each face, we obtained a mean Z score on each of the 8 statements by averaging across the 62 participants. Then, we plotted the mean scores of “eerie”, “disgusted”, “unsettling”, “attractive”, “threatening”, “likable” and “alive” against “real” rating, resulting in 7 line graphs (see Appendix B).

However, the relationship between the realism of the 89 faces and their induced emotional responses was not clearly visualized in the scatter plots, mainly due to the fact that the 89 faces were not evenly distributed on a continuum of realism. To remedy this problem, we divided the 89 faces into 10 bins (groups) of equal with a gradation of realism and then plotted the graphs using group means instead of the means of the individual faces. Since the realism of the 89 faces was not evenly distributed, the number of faces falling into each subgroup differed among the 10 subgroups. A line graph depicting the relationship between ratings on disgusted and ratings on real is shown below (for graphs of other relationships, see Appendix C).

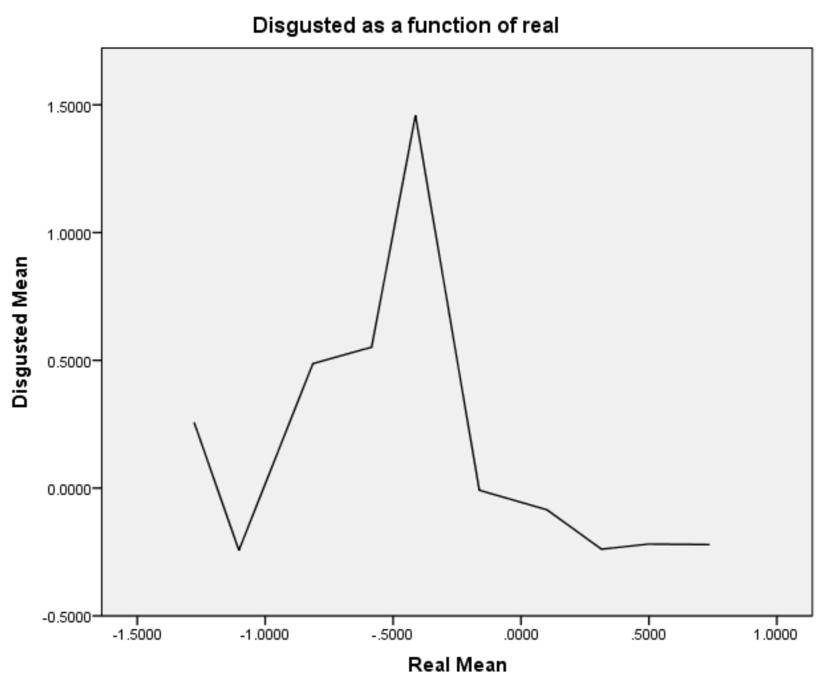


Fig. 2 Mean ratings on disgusted as a function of mean ratings on real for10 face groups.

Fig. 2 showed the mean ratings on disgusted were the highest for faces that were medium realistic compared to the others. The results were in accordance with the UVP in that as faces became more realistic, the emotional responses of the participants underwent a sudden drop (in the case of disgusted rating, it was reflected as an abrupt increase in Fig. 1). A one-way repeated measures ANOVA confirmed that participants rated the 20% real (corresponding to the first valley in the graph) faces significantly less disgusted than 10% real (corresponding to the point at far left in the graph) faces ($p < 0.05$), rated 50% real (corresponding to the peak in the graph) faces most disgusted compared to other faces ($p < 0.05$), and 100% real (corresponding to the point at far right in the graph) faces as the same as the 20% real faces ($p > 0.05$). Similar results were found in the ratings on eeriness.

Discussion

The trend in the UVP (up-down-up) was reconstructed among all 7 attributes of a face, among which ratings on eeriness and disgust formed a relationship with facial realism that most resembled the function predicted by the UVP (Fig. 2). The dip identified in the present data indicated that faces rated as medium realism to fall into the uncanny valley. The relationships between other emotional responses and facial realism showed similar trends but did not always map onto the uncanny valley function phenomenon precisely as Mori predicted (see Appendix C).

The graphs need to be interpreted with caution: First, the resemblance in the shapes of the graphs may indicate that the uncanny feeling is about eeriness and disgust. Yet, alternative explanations could be the 89 faces varied in human realism in different forms, eliciting some but not all of the emotional responses that did not change and potentially not to the same degree (The variation in the emotional components represented in each face may explain why the shapes of the graphs look so different) . Given the 89 faces we used, it was possible that the eerie and disgusted feelings were more salient emotional responses than the others for our specific set of stimuli in this experiment. With another set of stimuli, the exact emotional responses that form the uncanny valley function with facial realism may change.

Second, the limitations of the face rating survey prevented us from drawing a confident conclusion about the emotional components of the uncanny feeling based on our data. One limitation of the current rating included that the faces were always presented in a fixed order, which might cause adjacent faces in the presentation to influence the ratings of one another. In addition, the order of statements was also fixed across participants, which caused high correlations among the 7 subjective ratings.

Last but not least, neither Mori nor other researchers including ourselves have provided a definite and absolute function of the "uncanny valley", that is, what does the "uncanny valley" look like? For example: What are the height of the peaks and the depth of the valley? Along the x-axis, where is the point at which (e.g., degree of realism) the valley occurs? Without answering these questions, comparing the shapes of graphs plotted with data to the hypothetical graph may be incomplete in concluding which emotional response best capture the uncanny feeling.

However, ratings from all the emotional responses agreed on showing a non-monotonous relation with facial realism, which supported the existence of the UVP. In addition to the subjective ratings, we hypothesized that if the UVP exists the behavioral responses will also form a relation with the realism of the faces that resembles the uncanny valley function. We tested this hypothesis in Study 2.

Study two

In the second study, we tested if the UVP exists from a new perspective by introducing a behavioral test-visual looming task-to implicitly evaluate the emotional responses, particularly linked to fear, disgust and avoidance behavior.

Objects that systematically expand in size create a visual equivalent of impending collision in the field of view of the participants, thus eliciting avoidance responses. This is known as the visual looming effect manifested in human infants as well as adult rhesus monkeys (Schiff, Caviness, Gibson, 1962). A recent study showed that the looming effect was modulated by the emotional components of the looming stimuli. Specifically, people perceived threatening stimuli as moving faster than the non-threatening stimuli, showing stronger avoidance behavior exemplified by the underestimation of imagined time-to-collision. The magnitude of underestimation in perceived time-to-collision was proportion to the self-reported degree of fear (Vagnoni, Lourenco, &Longo, 2012).

As for the UVP, if the emotional responses elicited by the human replicas modulate the estimation of time-to-collision, we expect a difference in the estimated time-to-collision among faces differing in their uncanniness. Since the UVP has been linked to the negative emotional responses of fear, disgust and the sensation of uncertainty, we expected shorter estimated time-to-collision for faces that fell into the uncanny valley than those that did not. Therefore, we hypothesized that if the UVP exists, as faces increase in realism, the corresponding estimated time-to-collision will change in the form of the uncanny valley function.

Method

*Participants***.** Sixty-two undergraduate psychology students (43 females, mean age = 19.60 years, SD = 1.89) were recruited from Emory University and participated in return for course credit. All participants provided written informed consent.

Stimuli and procedure. Participants sat at a table about 40 cm in front of a 19-inch monitor (75 Hz refresh rate). The same 89 faces from Study 1 were presented in a program run by MATLAB (Mathworks, Natick, MA). On each trial, a face gradually enlarged for a second at a given rate, corresponding to 2.5, 3.5 and 4.5 s of time-to-collision. On the first frame, the width of the face was either 400 or 500 pixels (15.1° or 18.9° visual angle). The size of image differed so that estimation of time-to-collision involved both the speed and the initial size of an image.

90 images were presented twice and were randomly assigned to each of 6 (3 time-to-collisions x 2 initial sizes) conditions, resulting in 180 trials separated in 30 blocks (each block consisted of 6 trials corresponding to 6 conditions). Since it was impossible to evenly distribute 89 faces to 6 presentation conditions, one face (Face 89) was duplicated to measure up to the demand of the experimental design. The order of trials was randomized. After the participant responded to on one trial, the next trial began after a random interval of 300-800 milliseconds.

Participants were all instructed that they would see faces expanding in size as if they were approaching and that the faces would disappear after a second. From the moment the face disappeared, the task of the participants was to imagine it continuing to approach at the same speed and to press the space bar when they judged the face would have collided with their own face. Reaction time from the moment each face disappeared to the moment of button pressing was recorded as the measurement of estimated time-to-collision. Each participant ran a practice trial with looming geometric shapes before the test with faces. As we discussed previously that shorter estimated time-to-collision was associated with faces that elicited stronger negative emotional responses, our hypothesis was that as faces increased in realism, the corresponding estimated time-to-collision would first increased, then decreased, finally rebound to a higher level than its initial stage to form a trend predicted by the UVP.

Results

Visualization of the data indicated the existence of extreme observations of reaction time in the responses, which may be due to fatigue or failure of the participants in following the instruction to respond properly (The computer will not register a button pressing as a valid response until the face disappeared. Failing to follow the instruction to press the button results in extremely long reaction time) Since the extreme reaction time may bias the mean estimated time-to-collision, reaction time from each participant was standardized to detect outliers. And the reaction time was obtained from 6 presentation conditions differing in the actual time-to-collision and initial size; it rendered the reaction time between different conditions incomparable. Therefore, the standardization was performed separately on each of the 6 presentation conditions. Cases with a Z score above 3 or below -3 were excluded as outliers (109 trials were excluded, consisting of 0.99% of all trials).

A repeated measures ANOVA examined the two presentation factors, actual time-to-collision (2.5, 3.5, 4.5s) and initial size (400, 500 pixels wide), on the estimated time-to-collision (in milliseconds). Results showed main effects of actual time-to-collision and initial size: participants perceived faces that started closer to them [$F(1, 61) = 41.37$, $p<0.05$] and those presented with shorter actual time-to-collision as moving faster [$F(2, 122) = 48.017$, $p<0.05$]. No interaction between the two factors was found [$F(2, 122) = 1.562$, $p=0.218$]. The results indicated that most participants estimated the time-to-collision consistently with the actual speed and initial size in the presentation conditions. 6 participants showed the reversed pattern (e.g., 2.5s actual time-to-collision was estimated as longer than 4.5s actual time-to-collision); therefore they were excluded in later data analysis.

We examined whether the UVP exists by plotting the estimation time-to-collision against the realism for all 89 faces. Since there was no interaction between the actual time-to-collision and the initial size, we collapsed trials across initial size levels.

For each participant, each of the 89 faces was presented twice and assigned to two of the six presentation conditions; therefore we obtained reaction time from either one time-to-collision level or two. In order to obtain an index that reflected the emotional modulating effect of each face, we averaged the reaction time across 56 participants separately on three actual time-to-collision levels. Then we plotted the 3 means of estimated time-to-collision on the y-axis and the 3 corresponding actual time-to-collision on the x-axis of a Cartesian coordinate plane, resulting in three points in the plane. Thereby a line through the two points at both ends was drawn and its slope was calculated. The resulting slopes captured the overall trend of estimation for each face.

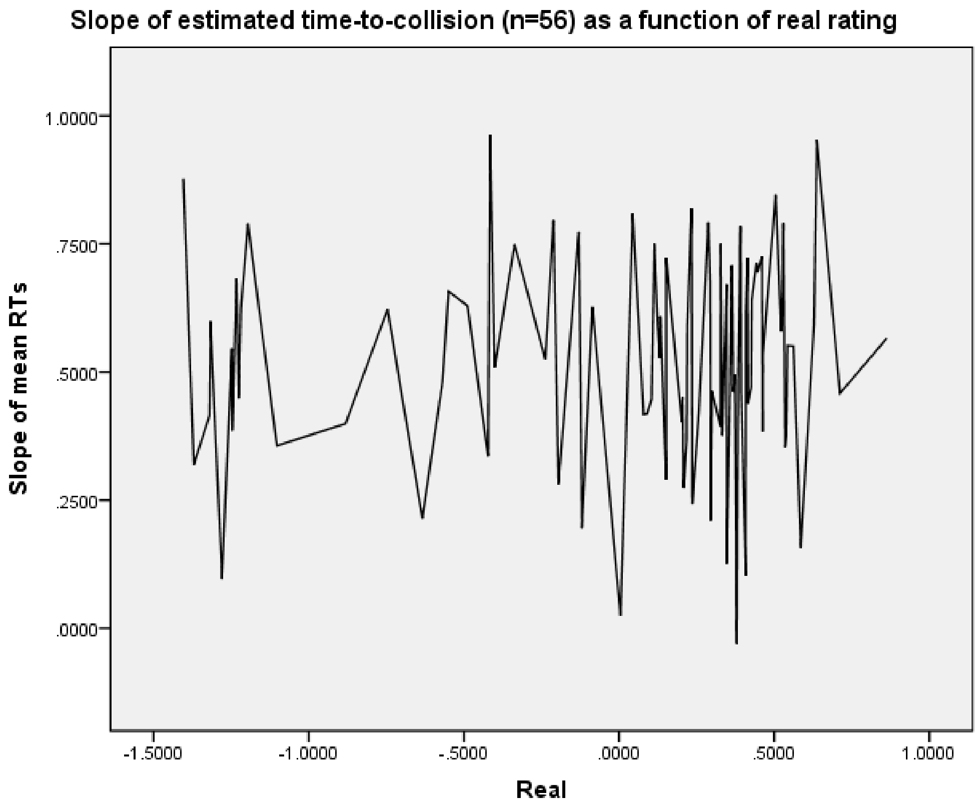


Fig. 3 Slope of estimated time-to-collision as a function of rating on realism.

In our previous attempt to summarize the emotional modulating effect on estimated time-to-collision, we assumed the least individual difference among participants in their responses to the looming faces so that the reaction time of a participant on one level was comparable to that of another participant on a different level. However, substantial individual difference in reaction time may exist, rendering the direct averaging across participants subject to bias of idiosyncratic responses. To remedy this, we converted the reaction time to Z scores on three actual time-to-collision levels for each participant separately. For example, a Z score of -1 on 2.5s time-to-collision level indicated that its original reaction time was 1 standard deviation below the mean reaction time of all faces presented at the actual time-to-collision condition for a given participant.

In order to be compared with estimated time-to-collision, the actual time-to-collision of 2.5, 3.5 and 4.5s were likewise standardized. Since the means and standard deviations of reaction time varied among participants, each of the three time-to-collisions 2.5, 3.5 and 4.5s was converted into 56 different Z scores accordingly, and averaged into three mean Z scores. With standardized reaction time of both estimated and actual time-to-collisions, slope for each face was calculated based on Z scores. The resulting slope was plotted against rating on real, as shown in the line graph below (Fig. 4).

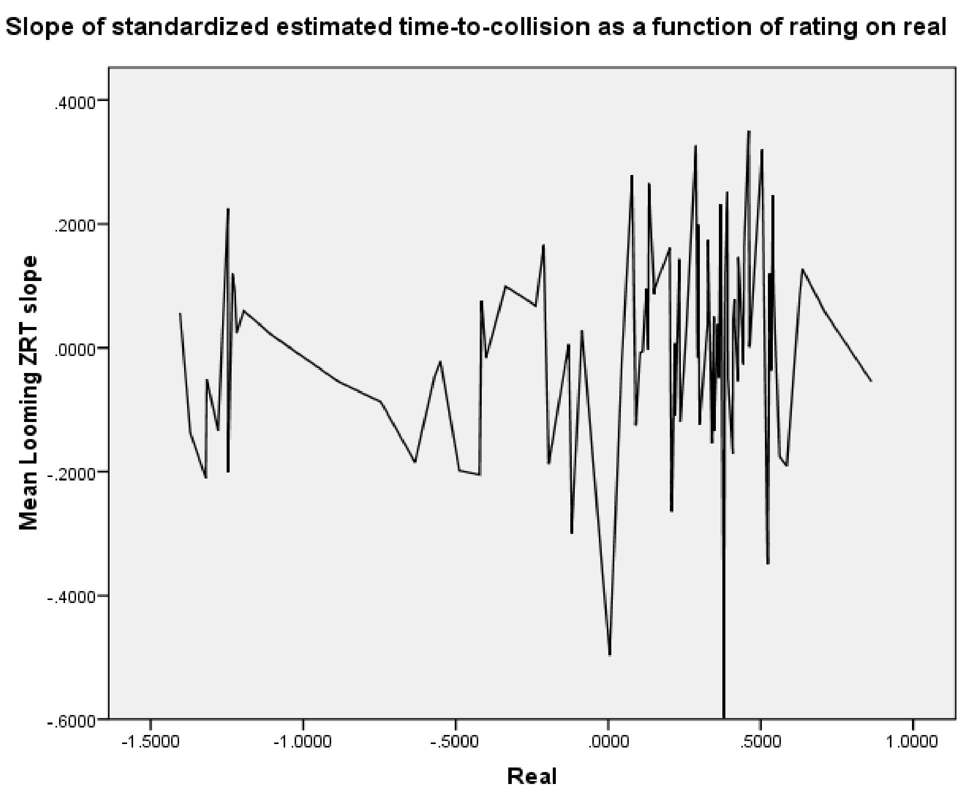


Fig. 4 Slope of standardized estimated time-to-collision as a function of rating on realism.

As we discussed previously, the relationship between the realism of a face and its induced emotional responses was not clearly visualized from the graph above. Therefore, we employed the same 10 face groups that were defined in the first study to examine the relationship between the realism of face and the corresponding slopes by comparing the groups instead of comparing the individual faces. A slope was derived for each group of faces, using the mean RTs of all the in-group faces presented on each of the three time-to-collision levels across 56 participants. Slopes derived from standardized RT were plotted against the ratings on realism as shown in the graphs below (Fig. 5), in which a function resembling the uncanny valley appeared, suggesting that the UVP may exist.

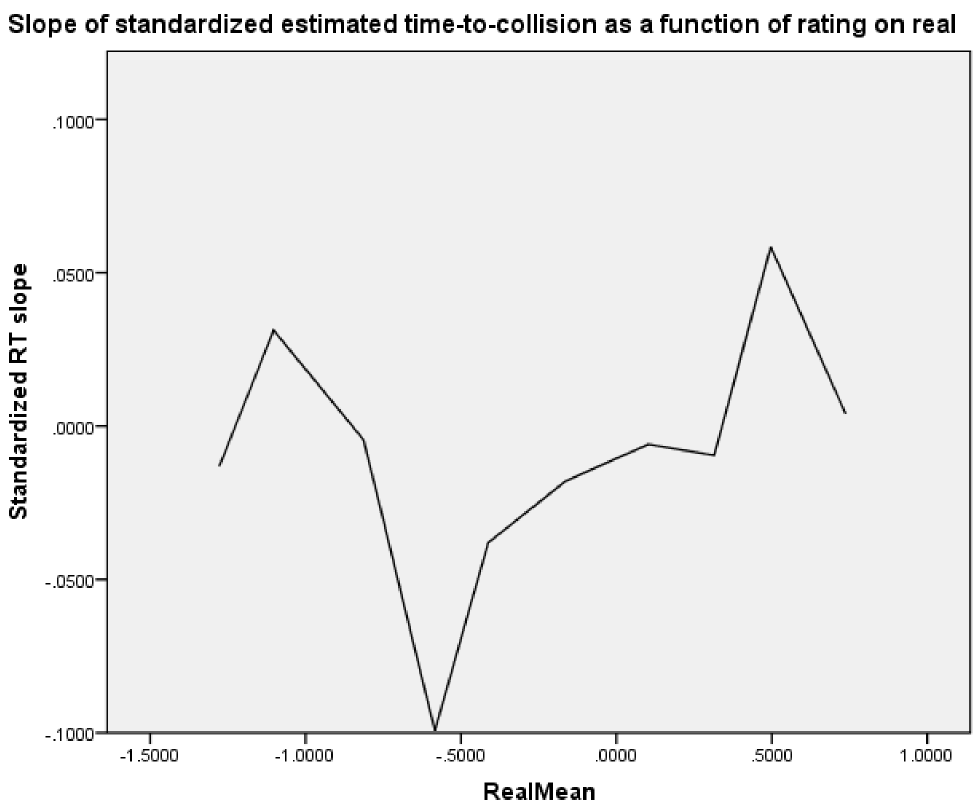


Fig. 5 Slope of standardized estimated time-to-collision as a function of rating on realism for 10 face groups

Discussion

In Study 2, we examined the emotional responses to faces in relation to the realism using a visual looming task. First, an index of "slope" characteristic of the emotional modulating effect in the looming task of each face was obtained by integrating the estimated time-to-collision on multiple presentation conditions for each face. Subsequently, combined with Study 1, the slopes were plotted against the ratings on realism for the 89 faces, which formed a pattern that resembled the uncanny valley. However, the reaction time in the visual looming task may not exclusively measured the exact "uncanny" feeling, because the estimated time-to-collision is equally subject to various other emotional modulation (e.g., threating) that might not necessarily be the "uncanny" feeling, and it is possible that these other emotional components might generate the current findings.

With both studies, we confirmed that the responses of the participants toward the 89 faces were not a monotonous function of the realism of the faces, suggesting the existence of the UVP. However, two questions remained unanswered: First, what is the mechanism that underlies the UVP? Second, given the discussion on the relation among the uncanny feeling, multiple emotional responses and the behavioral responses indicated by the estimated time-to-collision, to what extent does the uncanny feeling distinguish itself from other emotions? To answer these questions, we introduced a new behavioral test called sorting task to test 1) if the categorical uncertainty about whether a replica is real or unreal can account for the UVP and 2) whether the uncanny feeling associated with the faces falling into the "valley" is different from other negative feelings elicited by faces outside the valley.

Study three

From Study 1 and Study 2, we observed a relationship between the facial realism and the emotional responses that resembled the uncanny valley function in its trend and shape, examining both the subjective ratings and behavioral measures modulated by the emotional responses. In the following study, we attempted to examine the cognitive mechanism that underlies this relationship, by testing the hypothesis that the UVP is due to the uncertainty (Jentsch, 1906), particularly of identifying a face as either real or unreal (Moore, 2012). In previous studies this uncertainty has been linked to either a cross-modal mismatch (Tinwell et al., 2010) or mismatch of conflicting visual cues (MacDorman et al., 2009; Mitchell, Szerszen, Lu, Schermerhorn, Scheutz, MacDorman, 2011; Seyama & Nagayama, 2007). However, none of these studies directly measured the subjective uncertainty, but assuming that a mismatch would necessarily cause a sense of uncertainty. To test the uncertainty hypothesis, we introduced a sorting task, in which participants were asked to identify each face as either a real person or not a person. Meanwhile, their reaction time was recorded as the measurement of uncertainty. According to the uncertainty hypothesis, we predicted that faces sorted as neither real nor unreal (the faces in-between) will 1) elicit the longest reaction time and 2) in turn be associated with the strongest negative emotional (the so-called "uncanny "feeling). 3) We asked if the sense of the "uncanny" feeling can be predicted by the magnitude of the reaction time.

Method

Participants. 36 undergraduate students (20 females, mean age = 20.36 years, SD=2.29) from Emory University were recruited to participate in the sorting task, among who 11 first participated in the visual looming task.

Stimuli and procedure. In the sorting task, each participant sat in front of a 19-inch monitor (75 Hz refresh rate), where the 89 faces were presented one by one in a random order. Each trial began with a fixation cross at the center of the screen for 800 milliseconds. One of the 89 faces was presented successively at the same place. The task of a participant was to sort the face as either real or unreal as quickly as they can by pressing two buttons: a green button on the left indicated "real" and a red button on the right indicated "unreal". At the moment the participant pressed the button, the face disappeared and was followed by a blank screen of 500 milliseconds before the next trial began. The reaction time between the onset of the face presentation and the response of the participant was recorded as a measure of the uncertainty in sorting.

Results

Before analyzing the data, we compared the data from participants who only participated in the sorting task to those who were also tested in the looming task (n=11). If viewing the 89 faces in looming task did not influence the sorting (both in terms of the outcome of sorting and the reaction time), we could combine the data from both groups of participants to increase the sample size. A correlation analysis indicated high correlation between the two groups of participants, both in mean reaction time (Pearson $r = 0.611$, $p < 0.01$, 2 tailed) and in the percentage of participants that sorted a face as real

(Pearson r = 0.954, $p<0.01$, 2 tailed). We therefore were able to combine data from two groups of participants.

Among the 36 participant, 18 faces were consensually sorted as real and 7 faces as unreal. Although their sorting of the other faces was inconsistent, a binomial test indicated that if over 23 (including 23) participants sorted a face as real (or unreal), the agreement reached a significant level ($p <0.05$). In other words, faces that did not reach this degree of agreement were sorted by chance. We called these faces the in-between.

We categorized the 89 faces into three groups: faces that were sorted as significantly real, significantly unreal and the faces in between. We first confirmed with a one-way repeated measures ANOVA that the in-between faces elicited significantly longer reaction time than the real and unreal categories, [$F (2, 70) = 9.341$, $p <0.05$], whereas the reaction time did not differ between the real and unreal faces ($p=0.125$). The relationship between reaction time and the real category was illustrated in the line graph below. As the faces ranged from unreal, in-between to real, the corresponding reaction time in sorting they elicited first increased and then decreased.

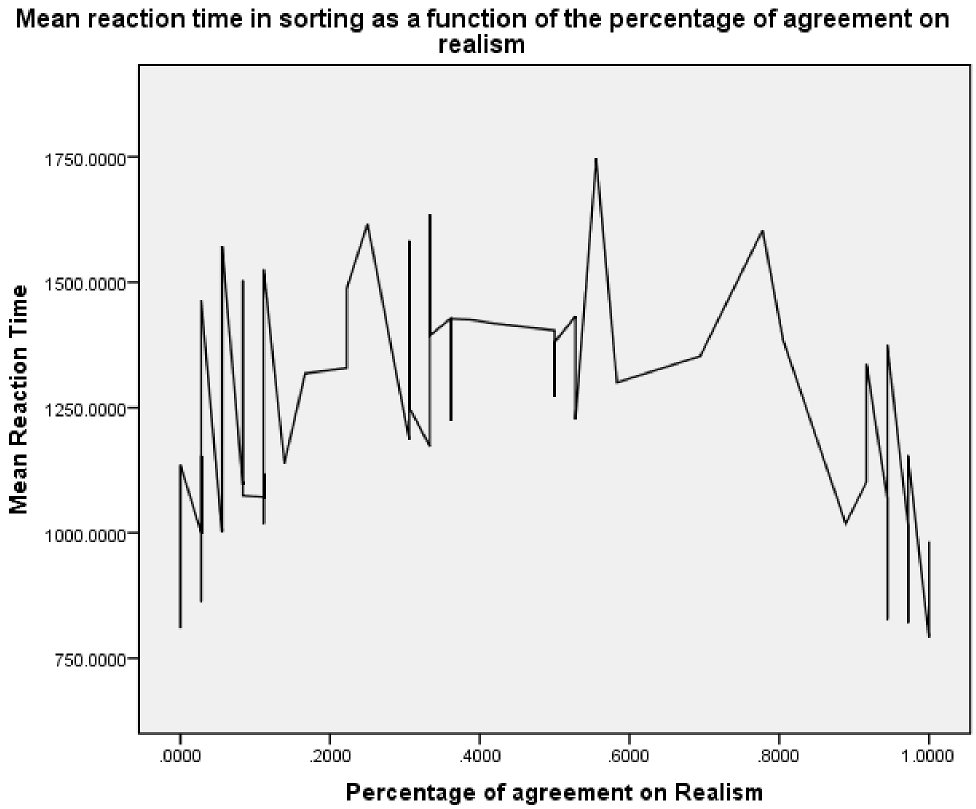


Fig. 6 Reaction time of sorting as a function of percentage of agreement on face realism

Using ratings on emotional responses from Study 1 as the measurement of dependent variable, we then tested our second hypothesis about whether the in-between faces that aroused the longest reaction time were also associated with stronger negative emotional responses. One-way repeated ANOVA confirmed that the in-between faces elicited significantly higher ratings on eerie, disgusted, unsettling, and threatening. In contrast, they also elicited significantly lower ratings in positive emotional responses such as attractive and likable.

Specifically, ratings on eerie, disgusted, and threatening, the in-between faces received the highest ratings, followed by the unreal faces and the real faces. A similar pattern was found in ratings on unsettling; however, participants did not experience more negative emotions for the unreal faces than for the real faces. On the contrary, the in-between faces were found less attractive and less likable than the unreal faces, and the real faces in turn, with the difference among all pairs of comparison reaching significant level ($p<0.05$). Interestingly, a pairwise comparison showed that the in-between faces were significantly more real than the unreal faces ($p<0.05$) but did not significantly differ from the real faces ($p=0.067$). It suggested that the faces that were in between were rated as closer to the real than the unreal faces, yet this subtle difference between faces that are real and in-between may be crucial for inducing the uncertainty and subsequently be associated with stronger negative emotional responses. Finally, a comparison among the three categories of faces in the ratings on alive showed that both unreal faces and those were in-between were rated as not alive compared the real faces. Noted that the in-between faces were rated as less alive than the unreal faces, though the difference of which was marginally significant ($p=0.052$). In other words, although the in-between faces looked very realistic, they were rated as more “dead” than the unreal faces that were completely inanimate. This finding was consistent with Gray, Knickman, & Wegner (2011), which found that patients in persistent vegetative state (PVS) were considered more dead than dead. Graphs illustrated the comparison among unreal, in-between and real faces on multiple emotional responses ratings in Appendix D.

Discussion

We confirmed our first hypothesis by showing that the faces in-between elicited a sense of uncertainty, evidenced by the significantly longer reaction time in sorting. The second hypothesis of categorical uncertainty as an explanation of the UVP was also confirmed by showing that the in-between faces were associated with the strongest negative feelings and the least positive emotional responses in the ratings. By linking the sense of uncertainty to stronger negative emotional responses, we provided further evidence that the UVP exists and that the uncertainty may be associated with the uncanny feeling.

In terms of the psychology of the uncanny feeling to human replicas, the present findings suggested that the “uncanny” feeling about faces at the boundaries between being real and unreal distinguishes itself from the negative emotional responses towards unreal faces, both qualitatively and quantitatively.

First, the negative emotional responses to unreal faces were different from that to in-between faces qualitatively. Taken the ratings on disgusted for example, the real faces to the unreal faces were significantly different in their ratings [$F(2, 122) = 93.463$, $p < 0.05$; pairwise comparison showed $p < 0.05$] (Fig. 7), whereas no difference was found in the corresponding reaction time [$F(1, 35) = 2.472$, $p = 0.125$; pairwise comparison showed $p = 0.125$] (Fig. 8). On the contrary, the comparison between real and in-between faces indicated significant difference both in ratings [pairwise comparison showed $p < 0.05$] and in reaction time [pairwise comparison showed $p < 0.001$]. By comparing the negative emotional responses towards unreal versus those directed to in-between faces, it became clear that the sense of uncertainty, exemplified by the long reaction time in sorting, differentiated the two types of negative emotional responses.

Second, the uncanny feeling also distinguished itself from other negative feelings towards unreal faces by showing significantly stronger negative emotional responses (pairwise comparison between unreal and in-between faces in ratings on disgusted indicated p <0.05). The negative feeling about the in-between faces was significantly more intense than the unreal face.

Taken together the difference in both aspects, it suggested that the uncertainty could not account for the difference in emotional responses between unreal faces and real faces, yet may be explanatory of the significant increase in the magnitude of negative responses towards the in-between faces, which might define the “uncanny”.

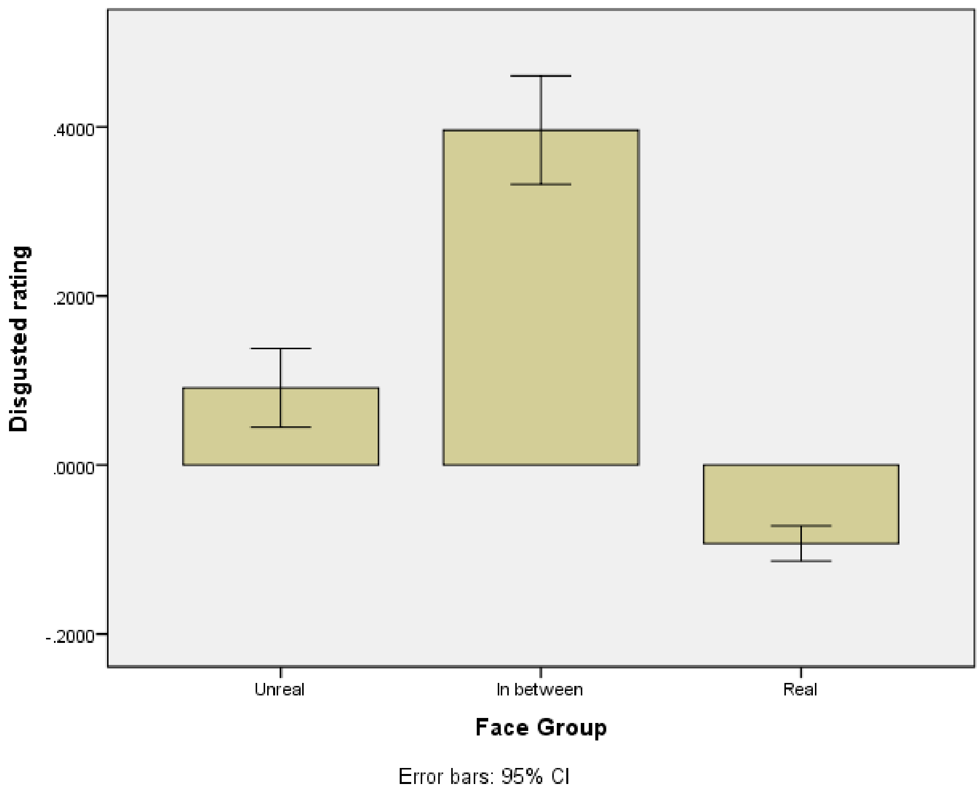


Fig. 7 Ratings on disgusted of unreal, in-between, and real faces

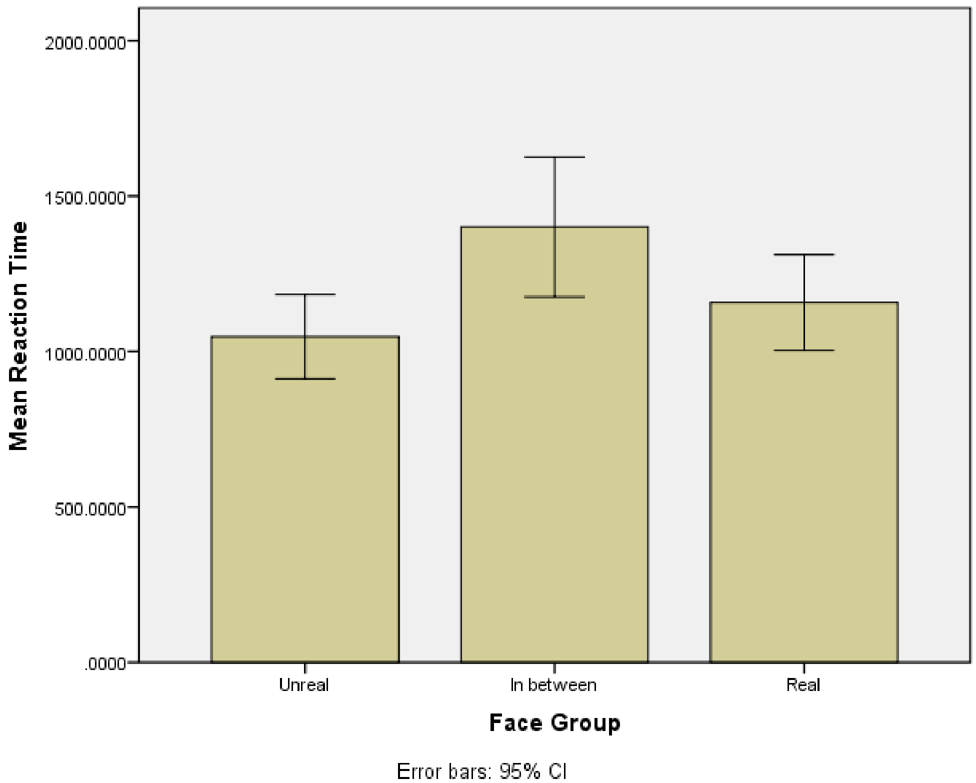


Fig. 8 Reaction time of unreal, in-between, and real faces in the sorting task

Conclusion and future directions

Since Mori coined the term “uncanny valley” in 1970, androids, computer-generated characters and other human replicas have become increasingly realistic and entered various aspects of social activities. The question of how to improve the likability of human replicas by increasing their physical and behavioral realism is central to the research on the UVP. Several hypotheses have been proposed to explain this phenomenon by focusing on its evolutionary/developmental origin, cognitive mechanism, and the fundamental distinctions between human and non-human beings in possessing the “mind”. However, empirical investigation on the UVP has not provided a clear picture about 1) if the uncanny valley exists, 2) what the uncanny feeling stands for and 3) what may account for s of this phenomenon, evidenced by the accumulating inconsistent findings and a lack of unified methodology in addressing these fundamental questions.

With the importance of studying the uncanny feeling and its lack of empirical evidence being said, the present study attempted to examine the uncanny experience in the context of the UVP. As an empirical examination of the UVP, our study was new in 3 aspects: First, we obtained diversity in the forms of human realism by including not only androids but also wax figures, dolls, mannequins, CGIs and real people as stimuli in all three studies. By doing so, our findings were more likely to be generalized to the UVP encountered in everyday experience than some of the previous studies that created artificial "uncanny" stimuli that were absent in real world (MacDorman et al., 2009; Seyama & Nagayama, 2007), thus improving the ecological validity of the present study. Second, we proposed new methods in testing the UVP and measuring variables. For example, two novel experimental paradigms (the visual looming task and sorting task) provided us with reaction time for more reliable psychophysical measurements of the facial realism and the emotional responses of the participants. The combinations of the rating and behavioral tests compensated for the subjectivity of self-report in previous studies. Finally, Study 3 is the first study in applying nonparametric statistics to data analysis in the study on the UVP, resulting in quantitative measurement of the sense of uncertainty. Study 3 is also the first research in measuring the subjective sensation of uncertainty instead of unwarrantedly assuming faces at the boundaries will necessarily cause stronger uncertainty.

One of the main goals of the present study was to test if the UVP occurs among a variety of human replicas whose faces approach realism from multiple dimensions to different extent. By doing so, we generalized the findings from previous studies using morph images to reconstruct the UVP. Morphing is a methodology that is often used in previous studies but is limited in two aspects: first, morphing only included a limit number of examples of human replicas; second, morphing introduced additional imperfections in appearance (e.g., the blurring effect of a morphed image) that may confound the original imperfections in realism.

In Study 1, we confirmed the existence of the UVP by showing that the relation between the emotional responses and the realism of faces followed a function predicted by the UVP. We also draw our tentative conclusion that the emotional components of the UVP include eeriness and disgust, which requires further examination. In Study 2, we utilized a visual looming task to probe the responses of the participants towards various looming faces that differed in their realism. With reaction time data, we showed that participants estimated faces of medium human realism to appear looming towards them faster than the faces that were more or less realistic, again indicating the existence of the UVP. In Study 3 we tested the hypothesis that the UVP may be due to the categorical uncertainty we experience when faces locate at the boundary between real and unreal. Our data suggested that the in-between faces elicited longer reaction time in sorting and were associated with stronger negative emotional responses compared to both unreal and real faces. We also showed strong evidence that the uncanny feeling linked to very realistic human replicas was a unique emotional response that differed from those towards unreal faces both in its magnitude and form.

However, across three studies the “uncanny” faces did not overlap with each other much. One explanation for the inconsistency in the “uncanny” faces may be due to the different paradigms we used to examine the UVP: In Study 1, the emotional responses of the participants to the 89 faces were obtained via subjective ratings, which involved a conscious and in-depth reflection about one's own feelings. In Study 2, the responses of another group of participants were implicitly measured by virtue of the estimated time-to-collision of looming faces. To what extent the reaction time in the visual looming task measured the so-called "uncanny" feeling needs further investigation, yet it provided researchers a novel tool to visualize and examine the subjective feeling about faces which is often obscure and mysterious. Therefore, it is not surprising to have the faces identified as most unpleasant (e.g., eerie, disgusted) not identical to those that elicited the shortest estimated time-to-collision, because with different methods, Study 1 and Study 2 may test the different perspectives of uncanny feeling.

In fact, people may argue that the conclusion we made in both studies that the UVP exists is arbitrary, the same as previous studies did to interpret their findings, because the evidence we all relied on was the extent to which the graphs we plotted resemble that predicted by Mori in his initial article. However, neither Mori nor any other researchers so far have found the point at which the "valley" and "peak" on the graph are located. Whether it is 75% real or 95% real largely depend on the two ends of faces defining the continuum of realism. As a result, we did not emphasize on the exact shape of the uncanny valley in our interpretation of the data, but highlighted the fact that all the graphs showed a non-monotonous trend in the relation between emotional responses and the realism of faces, which is a rare consensus researchers have reached in verifying the existence of the UVP.

Future directions

In all three studies, we examined the most basic form of the UVP using static facial images only, instead of full body with motion and voice; but we still found evidence that suggested the existence of the UVP. It implied that although the uncanny valley occurred when there was a mismatch between appearance and voice or body movement, the perception of biological motion and auditory perception were not necessities for the UVP. Given the diversity of the stimuli in the three studies, it is unlikely that the emotional responses could be accounted by one dimension of realism. Instead, there may be profound mechanism(s) beyond the multidimensionality of realism that underlie the UVP, which according to the findings from Study 3, may point to the sense of uncertainty about whether a face is a real person or not.

The diversity of stimuli in the present studies reflected the multidimensionality of realism in real world examples of human replicas. It has both its pros and cons: it improved the ecological validity, whereas undermined the experimental control. Further studies may benefit from systematical manipulation of realism on multiple dimensions in search of UVP. Subsequently, researchers may empirically examine Moore's categorical uncertainty explanation of the uncanny valley by showing that the uncanny feeling is due to the differential perceptual distortion from multiple perceptual cues that creates a perceptual "tension" (Moore, 2012).

According to the mind perception approach proposed by Gray and Wegner (2012), another direction of future research may be to focus on the individual difference in attributing human mind to inanimate human replicas and other factors that may affect the android-human interaction by changing the propensity of a person to perceive an android as possessing human mind (Waytz, Cacioppo, Epley, 2010).

Reference

Bartneck, C., Takayuki, K., Ishiguro, H., & Hagita, N. (2007). Is the uncanny valley an uncanny cliff? *Proceedings of the 16 th IEEE International Symposium on Robot and Human Interactive Communication*, RO-MAN 2007, Jeju, Korea, pp. 368-373.

Bartneck, C., Kanda, T., Ishiguro, H., & Hagita, N. (2009). My robotic doppelganger – A critical look at the uncanny valley. *The 18th IEEE International Symposium on Robot and Human Interactive Communication*, Toyama, Japan, Sept. 27-Oct.2, 2009

Brenton, H., Gillies, M., Ballin, D. & Chatting, D. (2005). The uncanny valley: does it exist? *19th British HCI Group Annual Conference: workshop on human-animated character interaction*.

Burleigh, T. J., Schoenherr, J. R., & Lacroix, G. L. (2013). Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in human behavior*, http://dx.doi.org/10.1016/j.chb.2012.11.021.

Carpenter, J., Eliot, M., & Schultheis, D. (2006). The Uncanny Valley: Making Human-Nonhuman Distinctions. In *Proceedings of the 5th International Conference on Cognitive Science* (pp. 81-82).

Chaminade, T., Hodgins, J., & Kawato, M. (2007). Anthropomorphism influences perception of computer-animated characters' actions. *Social cognitive and affective neuroscience, 2*(3), 206-216.

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological review*, *116*(4), 752.

Freud, S. (1919). The Uncanny. In j. Strachey & A. Freud (Eds.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud* (pp. 217-256). London: Hogarth Press.

Gray, K., Knickman, T. A., & Wegner, D. M. (2011). More dead than dead: Perceptions of persons in the persistent vegetative state. *Cognition, 121,* 275-280

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition, 125,* 125-130.

Greenberg, J., Pyszczynski, T., Solomon, S., Simon, L., & Breus, M. (1994). Role of consciousness and accessibility of death-related thoughts in mortality salience effects. *Journal of Personality and Social Psychology, 67* (4): 627–637.

Groom, V., Nass, C., Chen, T., Nielsen, A., Scarborough, J. K., & Robles, E. (2009). Evaluating the effects of behavioral realism in embodied agents. *International Journal of Human-Computer Studies*, *67*(10), 842-849.

Hanson, D. (2005). Expanding the aesthetic possibilities for humanoid robots. *In IEEE-RAS*

international conference on humanoid robots, Tsukuba, Japan.

Ho, C. -C., MacDorman, K. F., & Pramono, Z. A. D. D. (2008). Human emotion and the uncanny valley: a GLM, MDS, and Isomap analysis of robot video ratings. *In*

Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction, HRI '08 (pp. 169–176). Presented at the HRI '08, New York, NY. doi:10.1145/1349822.1349845.

Ho, C. C., & MacDorman, K. F. (2010). Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior*, *26*(6), 1508-1518.

House of Wax (Warner Brothers Pictures 1953, 2005)

Jentsch, E. (1906). On the psychology of the uncanny, *Angelaki, 2* (1): 7-16.

Kaba,F. (2013). Hyper-realistic characters and the existence of the uncanny valley in animation films. *International review of social sciences and humanities, 4* (2): 188-195.

Lewkowicz, D. J., & Ghazanfar, A. A. (2012). The development of the uncanny valley in infants. *Developmental Psychology, 54* (2): 124-132.

MacDorman, K. F., Minato, T. Shimada, M. Itakura, S., Cowley, S. & Ishiguro, H. (2005). Assessing human likeness by eye contact in an android test bed. CogSci2005 – XXVII Annual Conference of the Cognitive Science Society.

MacDorman, K. (2006). Subjective Ratings of Robot Video Clips for Human Likeness, Familiarity, and Eeriness: An Exploration of the Uncanny Valley. *ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science* (pp. 26–29).

MacDorman, K., & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies 7* (3): 297-337.

MacDorman, K. F., Minato, T. Shimada, M. Itakura, S., Cowley, S. & Ishiguro, H. (2005). Assessing human likeness by eye contact in an android test bed. CogSci2005 – XXVII Annual Conference of the Cognitive Science Society.

MacDorman, K. F., Green, R., Ho, C., & Koch, C. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior, 25* (3): 695-710.

Matsuda, Y. T., Okamoto, Y., Ida, M., Okanoya, K., & Myowa-Yamakoshi, M. (2012). Infants prefer the faces of strangers or mothers to morphed faces: an uncanny valley between social novelty and familiarity. *Biology Letters, 8*(5), 725-728.

Mitchell, W. J., Szerszen, K. A., Lu, A. S., Schermerhorn, P. W., & MacDorman, K. F. (2011). A mismatch in the human realism of face and voice produces an uncanny valley. *i-Perception, 2,* 10-12.

Moore, R. (2012). A Bayesian explanation of the 'Uncanny Valley' effect and related psychological phenomena. *Scientific reports, 2,* 864.

Mori, M. (1970). The Uncanny Valley. *Energy, 7* (4), pp. 33-35.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *Robotics & Automation Magazine, IEEE, 19*(2), 98-100.

Olson, J.M., Roese, N.J., & Zanna, M.P. (1996). Expectancies. In E.T. Higgins & A.W.

Piwek, L., McKay, L. S., & Pollick, F. E. (2014). Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition*, *130*(3), 271-277.

Poliakoff, E., Beach, N., Best, R., Howard, T., & Gowen, E. (2013). Can looking at a hand make your skin crawl? Peering into the uncanny valley for hands. *Perception*, *42*(9), 998-1000.

Pollick, F. E. (2010). In search of the uncanny valley. In *User centric media* (pp. 69-78). Springer Berlin Heidelberg.

Rhodes, G., Zebrowitz, L. A., Clark, A., Kalick, S. M., Hightower, A., & McKay, R. (2001). Do facial averageness and symmetry signal health?. *Evolution and Human Behavior*, *22*(1), 31-46.

Rozin, P., & Fallon, A. E. (1987). A perspective on disgust. *Psychological review*, *94*(1), 23.

Saygin, A. P., Chaminade, T., & Ishiguro, H. (2010, August). The perception of humans and robots: Uncanny hills in parietal cortex. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2716-2720).

Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J., & 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): 413-422.

Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence*, *16* (4), 337–351.

Steckenfinger, S. & Ghazanfar, A. (2009). Monkey visual behavior falls into the uncanny valley. *Proceedings of the National Academy of Sciences*, 106, 18362-18366

The Hills Have Eyes (Fox Searchlight Pictures 2006)

The Polar Express (Warner Brothers Pictures 2004).

Thompson, J. C., Trafton, J. G. & McKnight, P. (2011). The perception of humanness from the movements of synthetic agents. *Perception 40* (6): 695–704.

Tinwell, A., Grimshaw, M., & Williams, A. (2010). Uncanny behavior in survival horror games. *Journal of Gaming and Virtual Worlds*, *2* (1), 3-25.

Tinwell, A., Grimshaw, M., Nabi, D. A., & Williams, A. (2011). Facial expression of emotion and perception of the uncanny valley in virtual characters. *Computers in human behavior, 27* (2): 741-749.

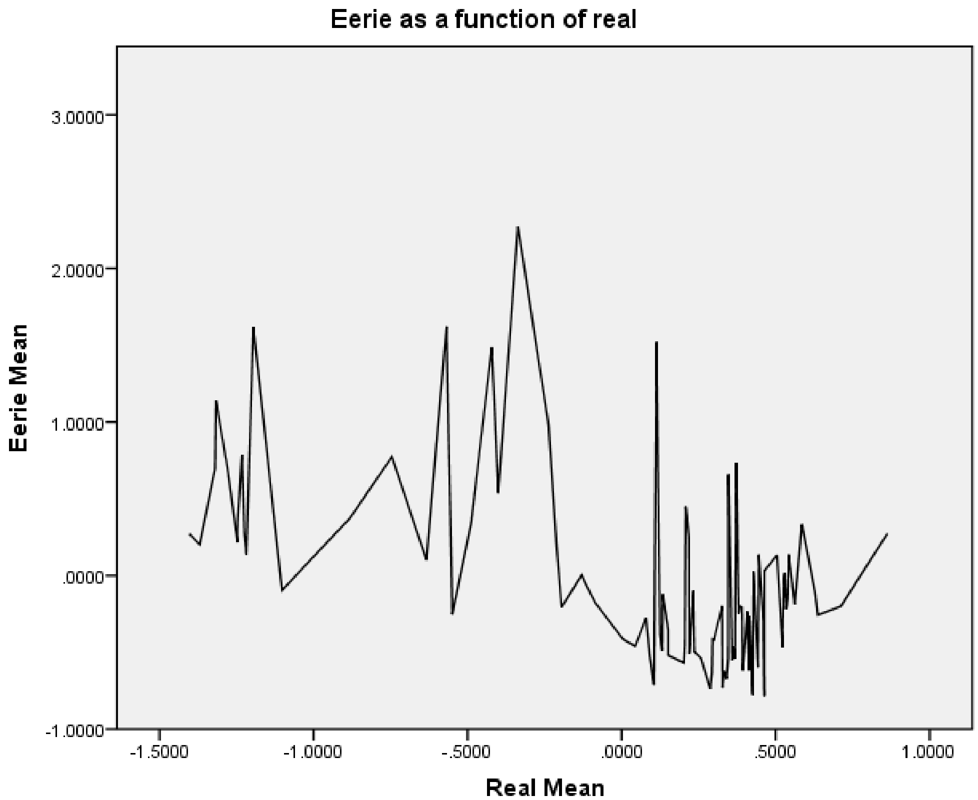
Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on psychological science, 5* (3): 219-232.

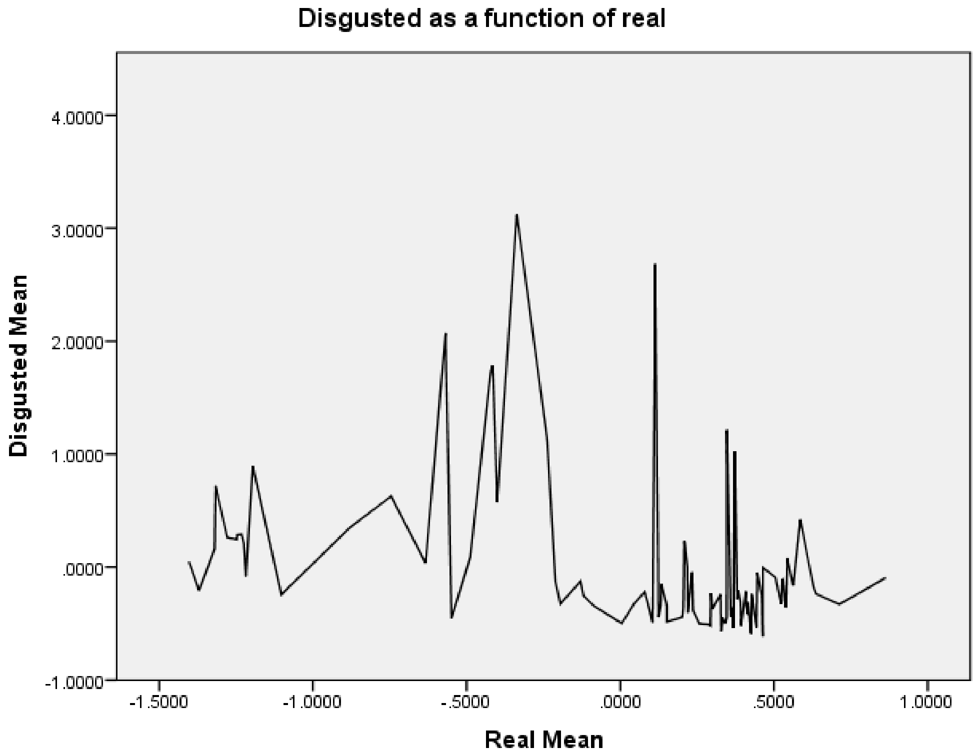
Appendix

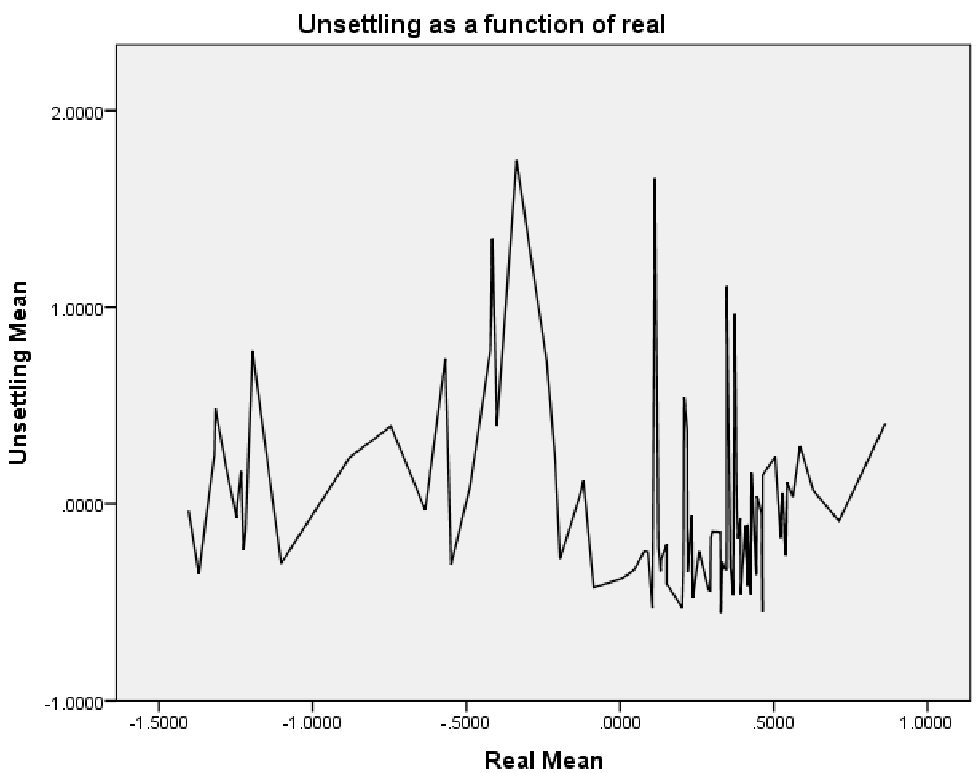
Appendix A

Slide	**Slide**
This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much	This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much
This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much	This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much
This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much	This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much
This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much	This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much
This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much	This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much
This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much	This figure makes me feel eerie? 1 2 3 4 5 6 7 Not at all very much

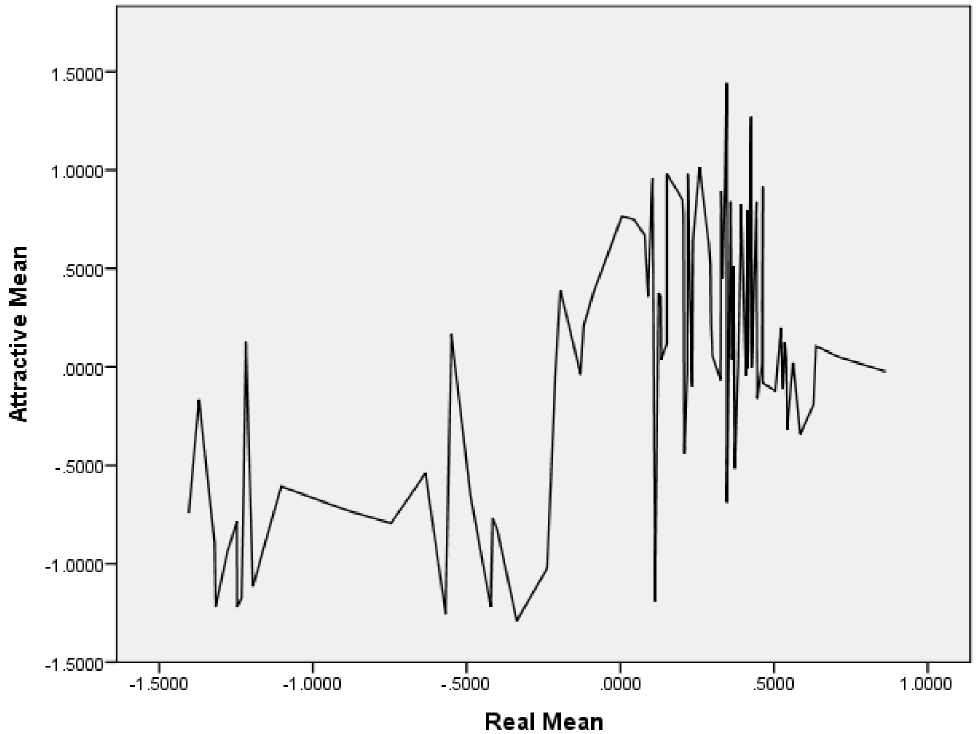
Appendix B


Eerie as a function of real
Eerie Mean
3.0000
2.0000
1.0000
.0000
-1.0000
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean

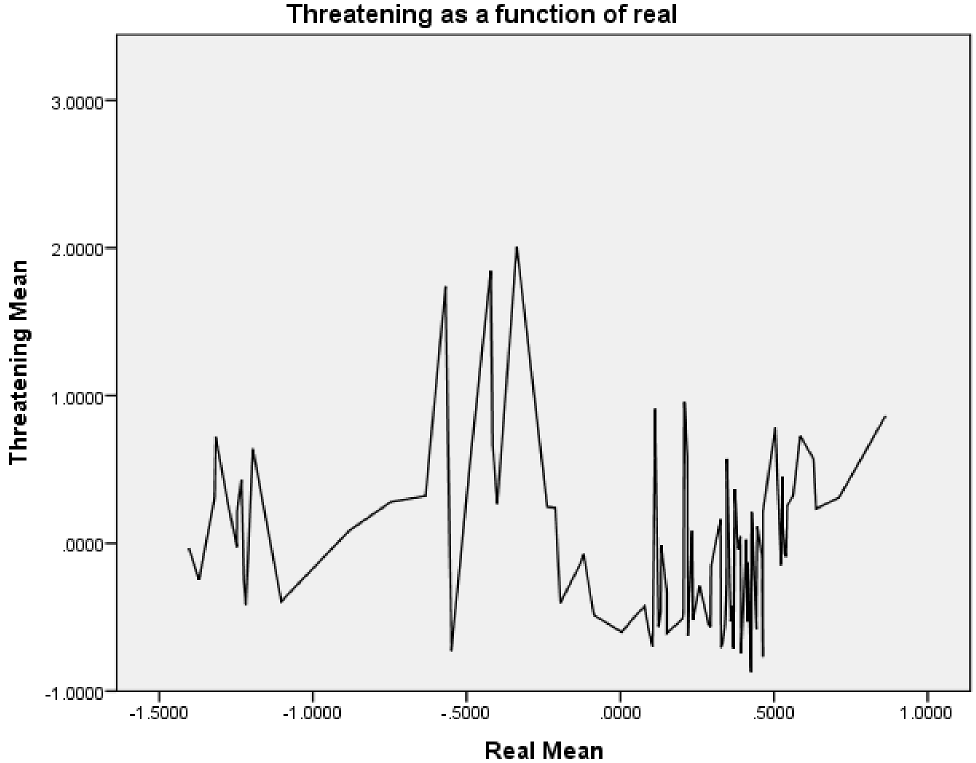

Disgusted as a function of real
4.0000
3.0000
2.0000
1.0000
.0000
-1.0000
Disgusted Mean
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean

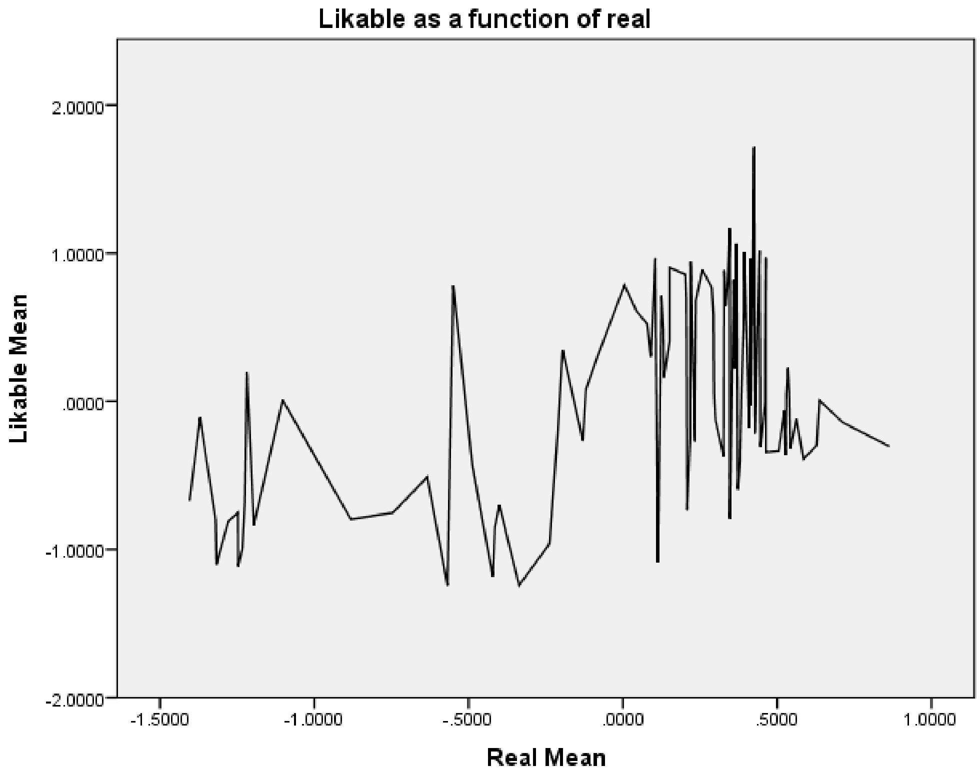

Unsettling as a function of real
2.0000
1.0000
.0000
-1.0000
Unsettling Mean
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean

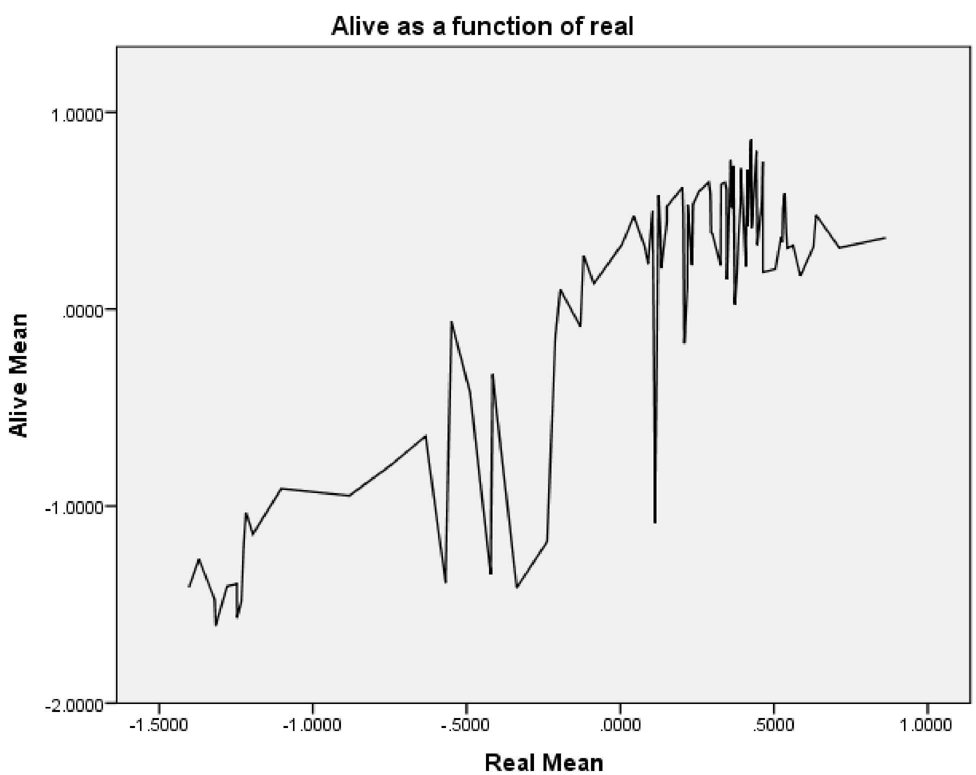
Attractive as a function of real



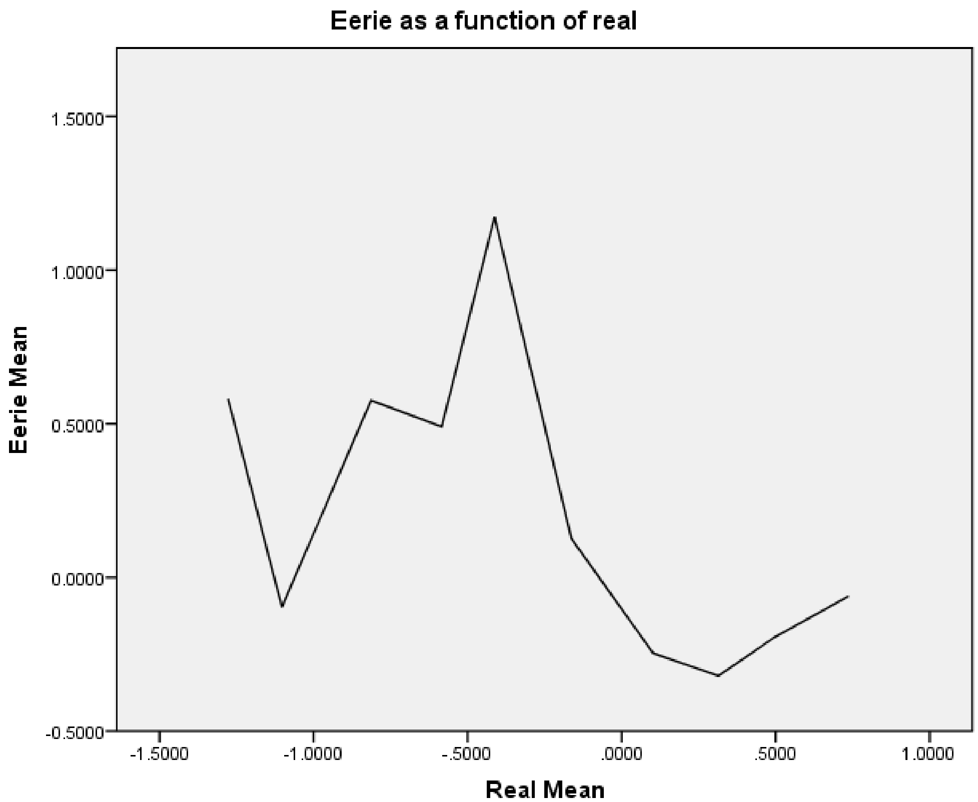
Threatening as a function of real

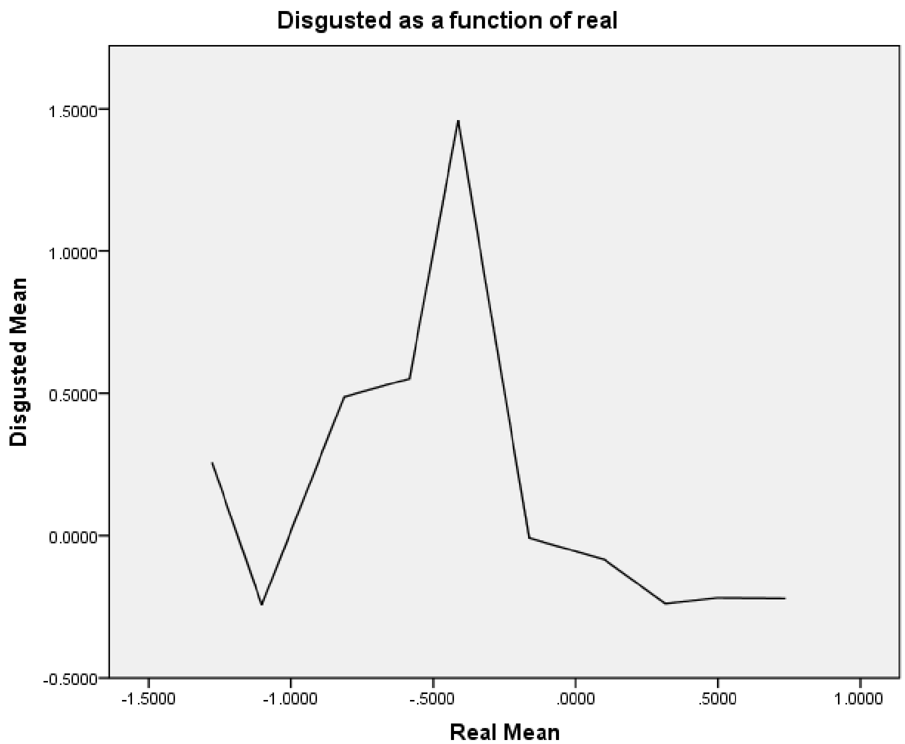


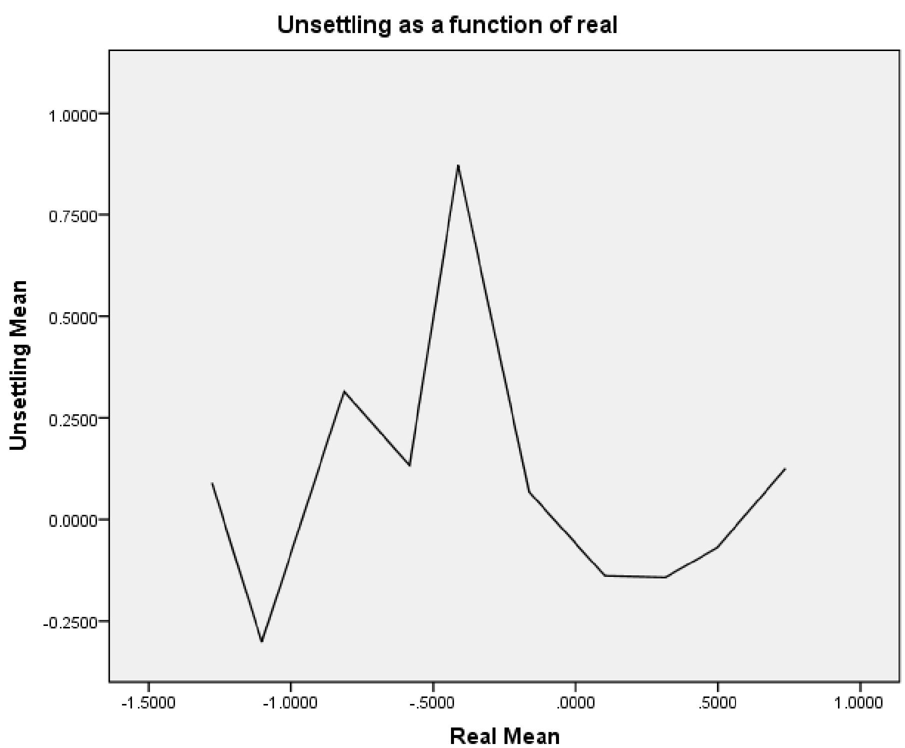

Likable as a function of real
2.0000
1.0000
.0000
-1.0000
-2.0000
Likable Mean
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean

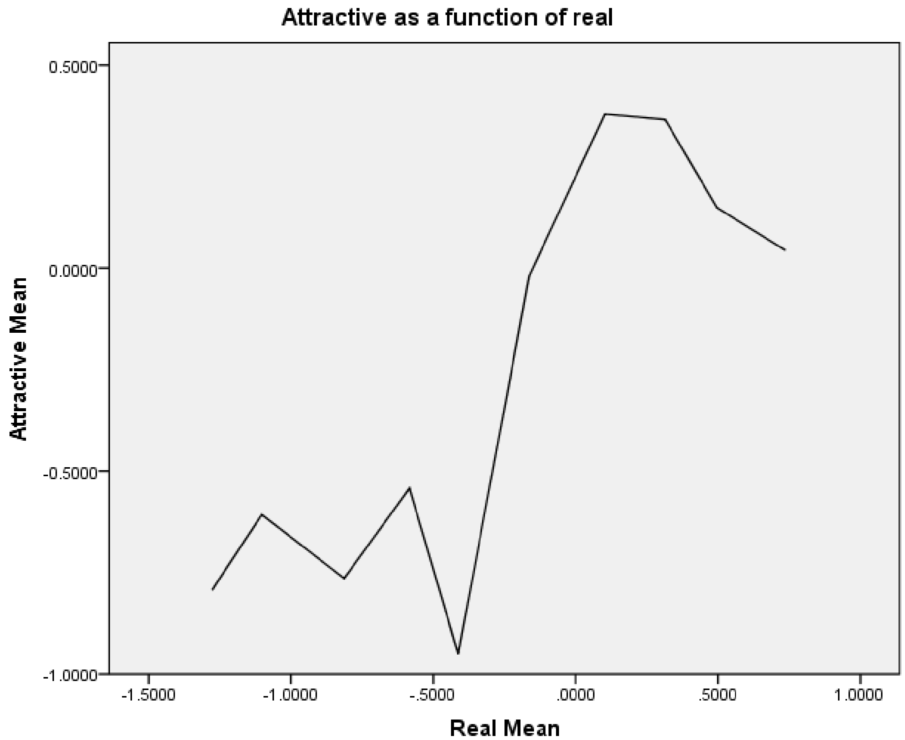

Alive as a function of real
1.0000
.0000
-1.0000
-2.0000
Alive Mean
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean

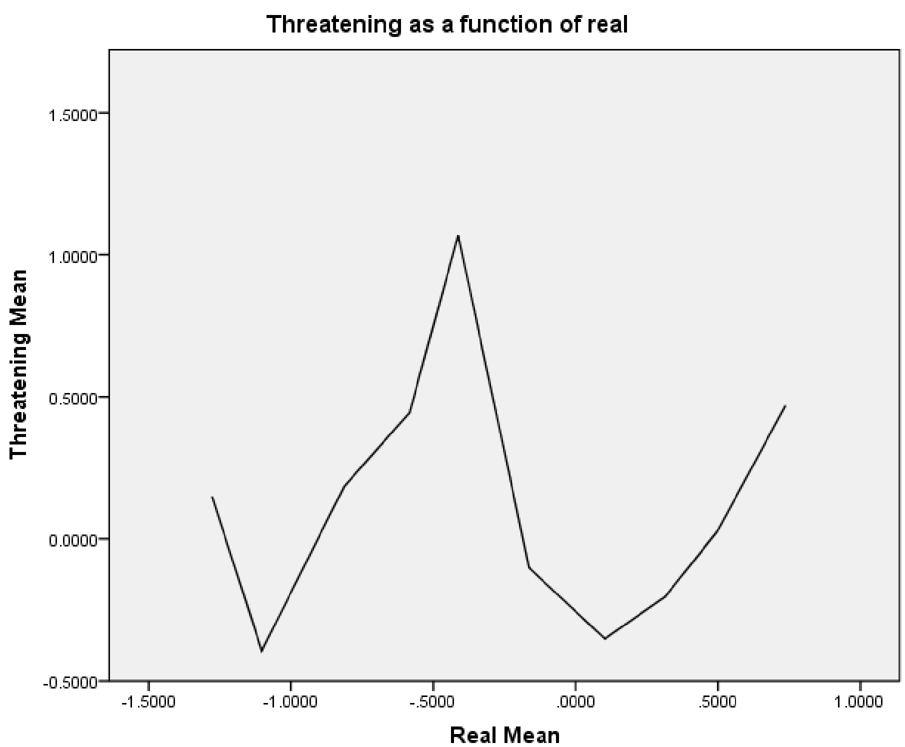
Appendix C

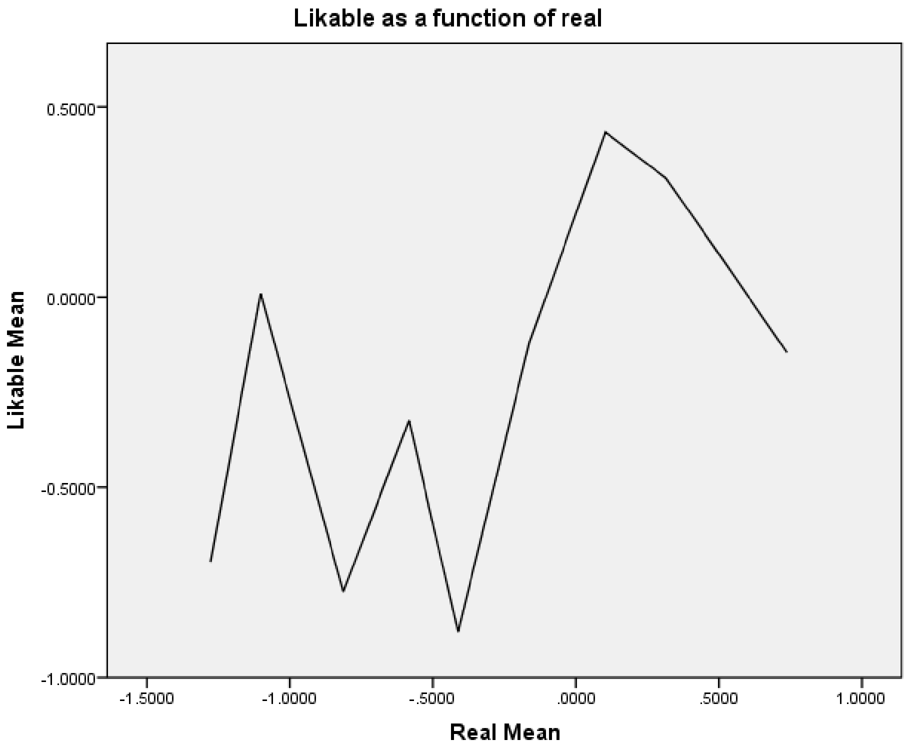

Eerie as a function of real
1.5000
1.0000
0.5000
0.0000
-0.5000
Eerie Mean
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean

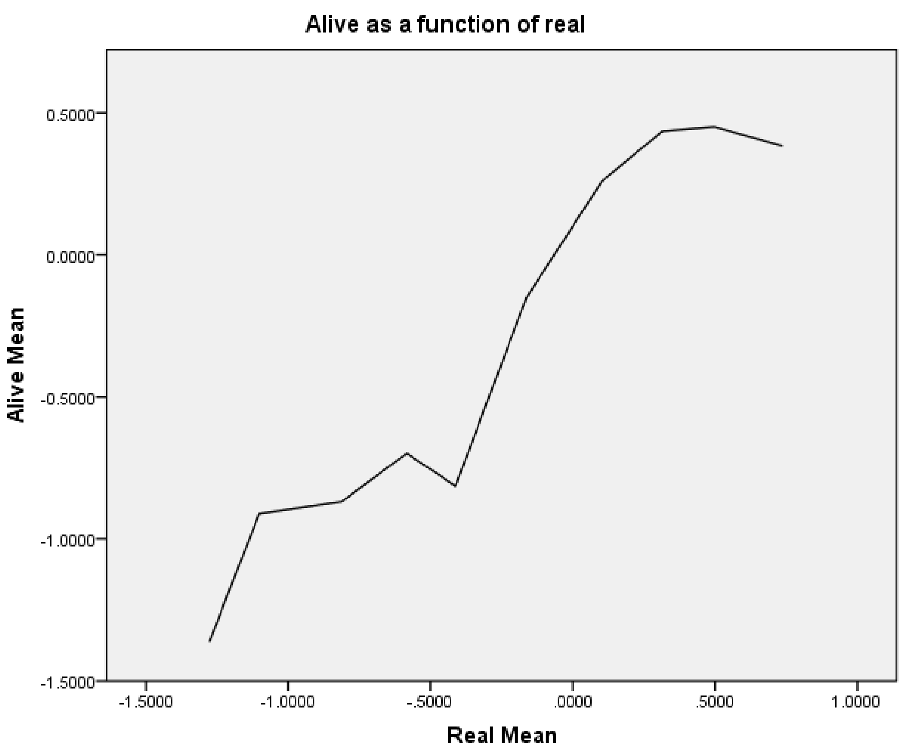

Disgusted as a function of real
Disgusted Mean
1.5000
1.0000
0.5000
0.0000
-0.5000
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean


Unsettling as a function of real
Unsettling Mean
1.0000
0.7500
0.5000
0.2500
0.0000
-0.2500
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean

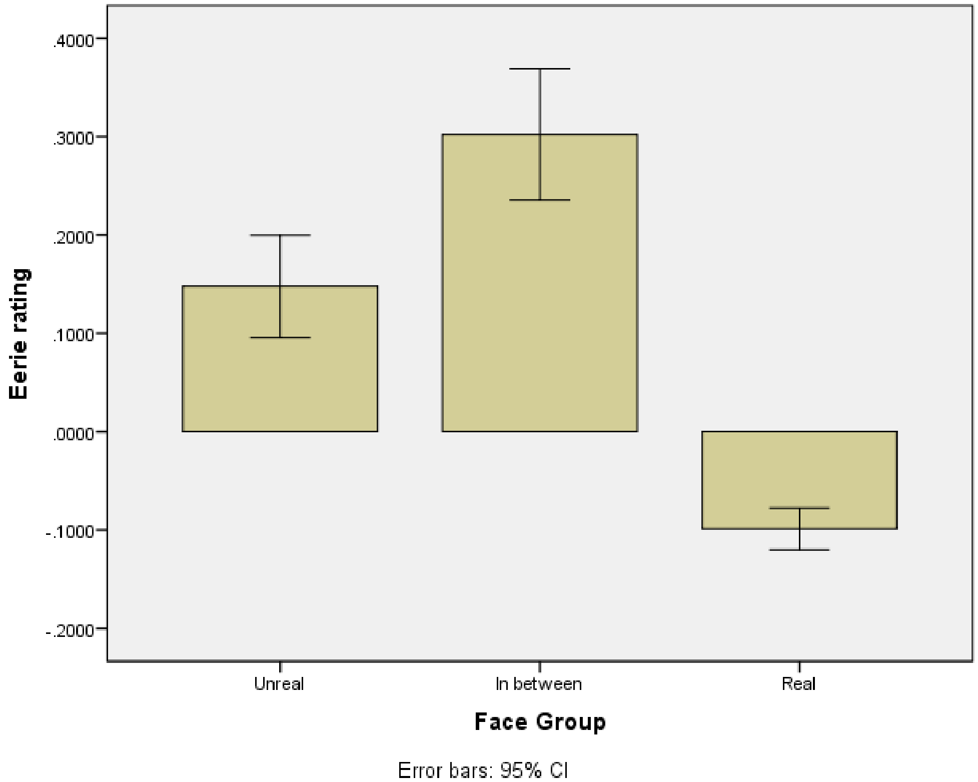

Attractive as a function of real
0.5000
0.0000
-0.5000
-1.0000
Attractive Mean
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean

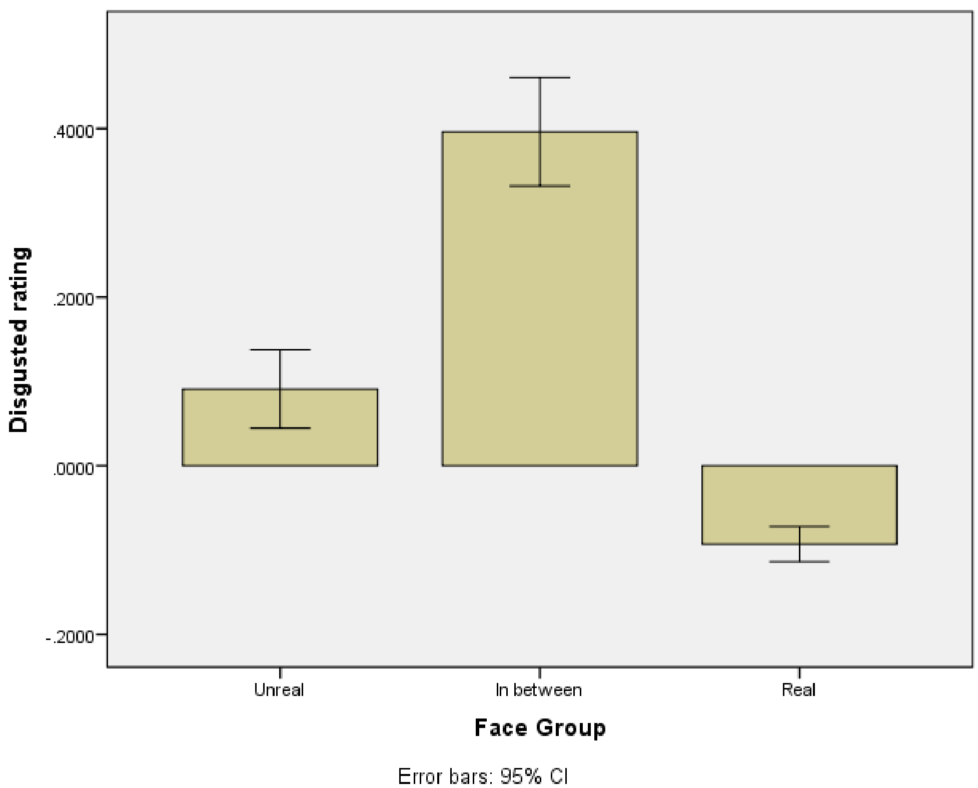

Threatening as a function of real
1.5000
1.0000
0.5000
0.0000
-0.5000
Threatening Mean
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean

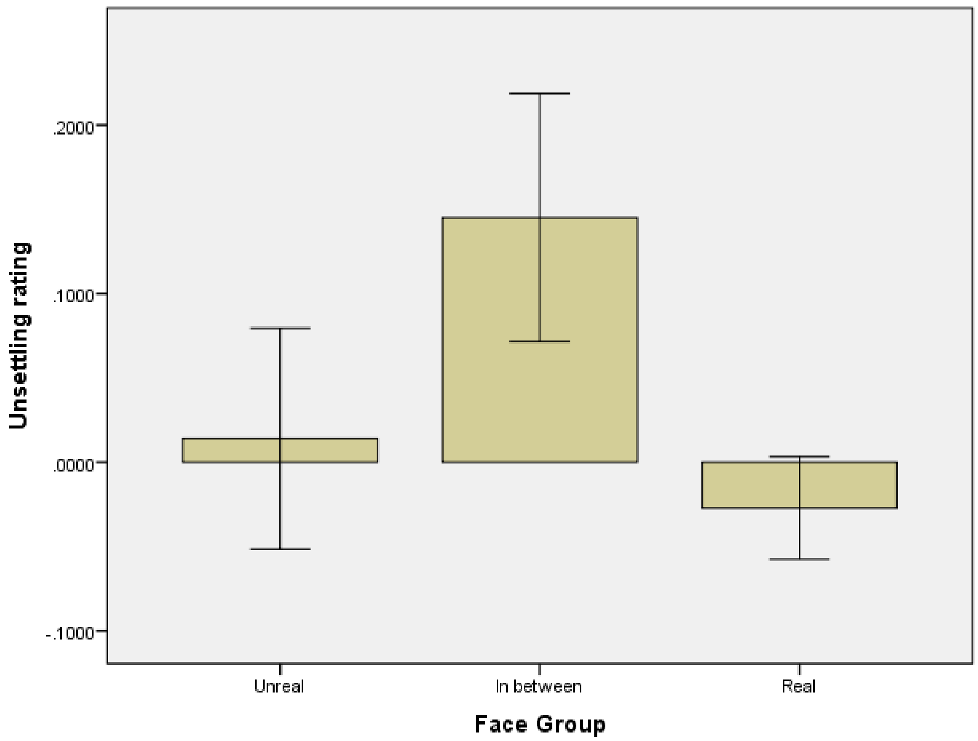

Likable as a function of real
Likable Mean
0.5000
0.0000
-0.5000
-1.0000
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean


Alive as a function of real
Alive Mean
0.5000
0.0000
-0.5000
-1.0000
-1.5000
-1.5000
-1.0000
-.5000
.0000
.5000
1.0000
Real Mean

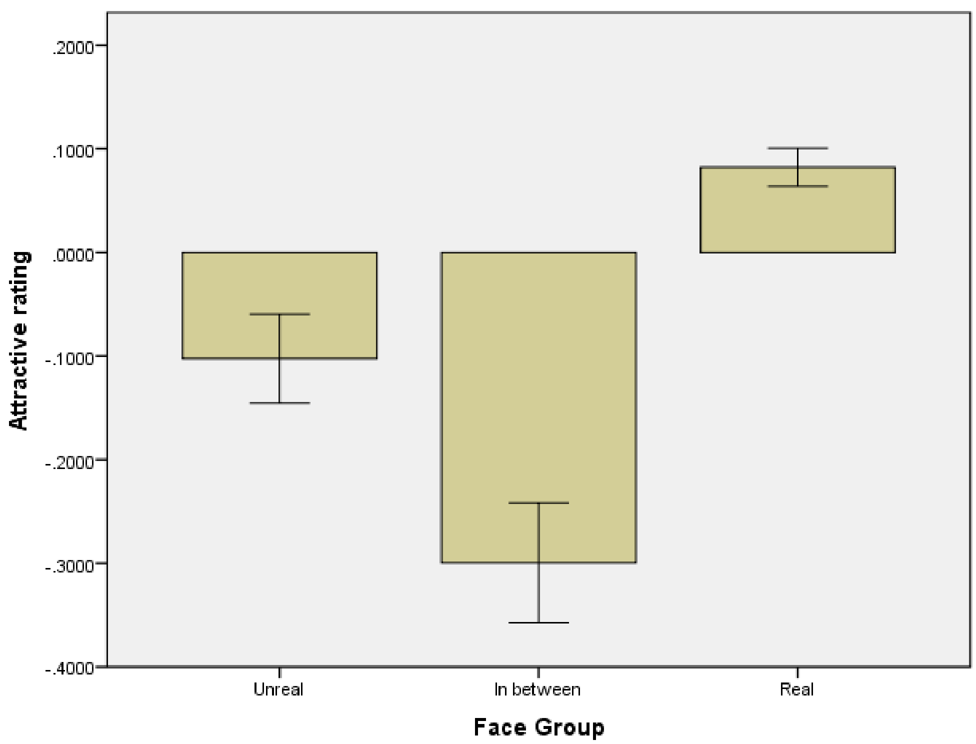
Appendix D


.4000
.3000
.2000
.1000
.0000
-.1000
-.2000
Eerie rating
Unreal
In between
Real
Face Group
Error bars: 95% CI

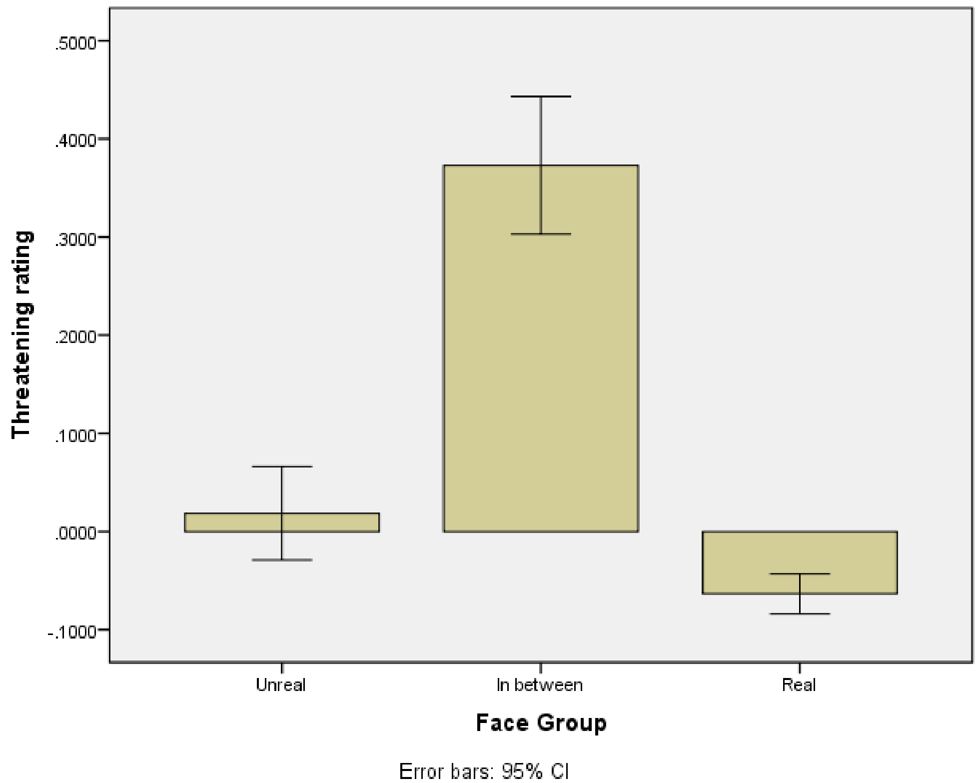

.4000
.2000
.0000
-.2000
Disgusted rating
Unreal
In between
Real
Face Group
Error bars: 95% CI

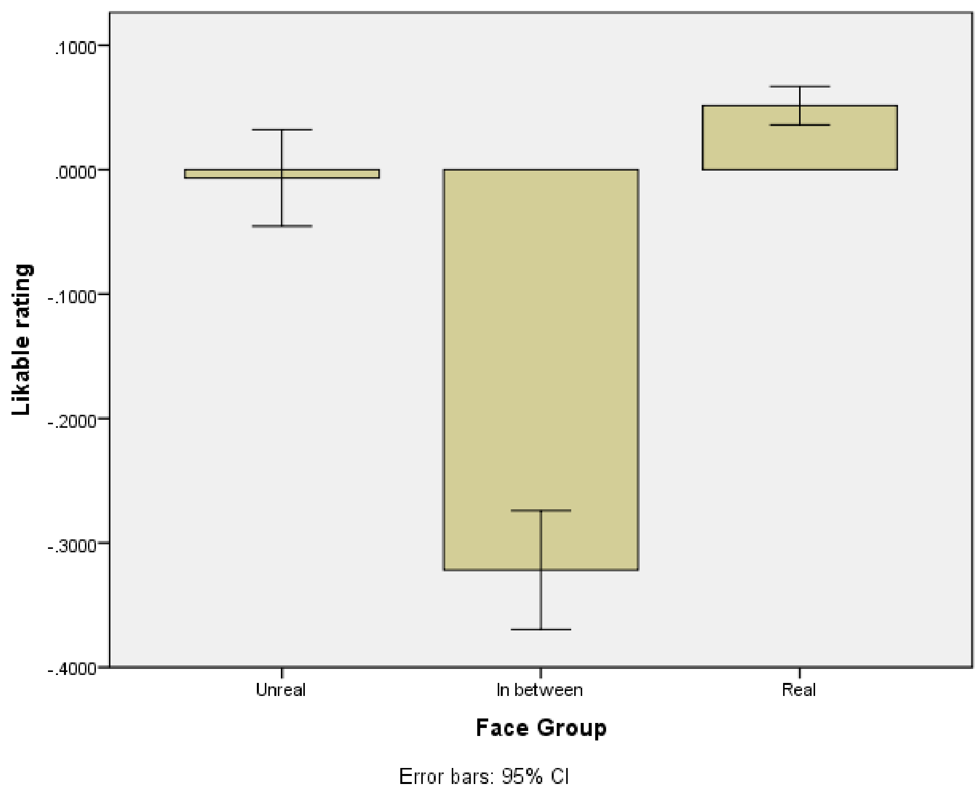


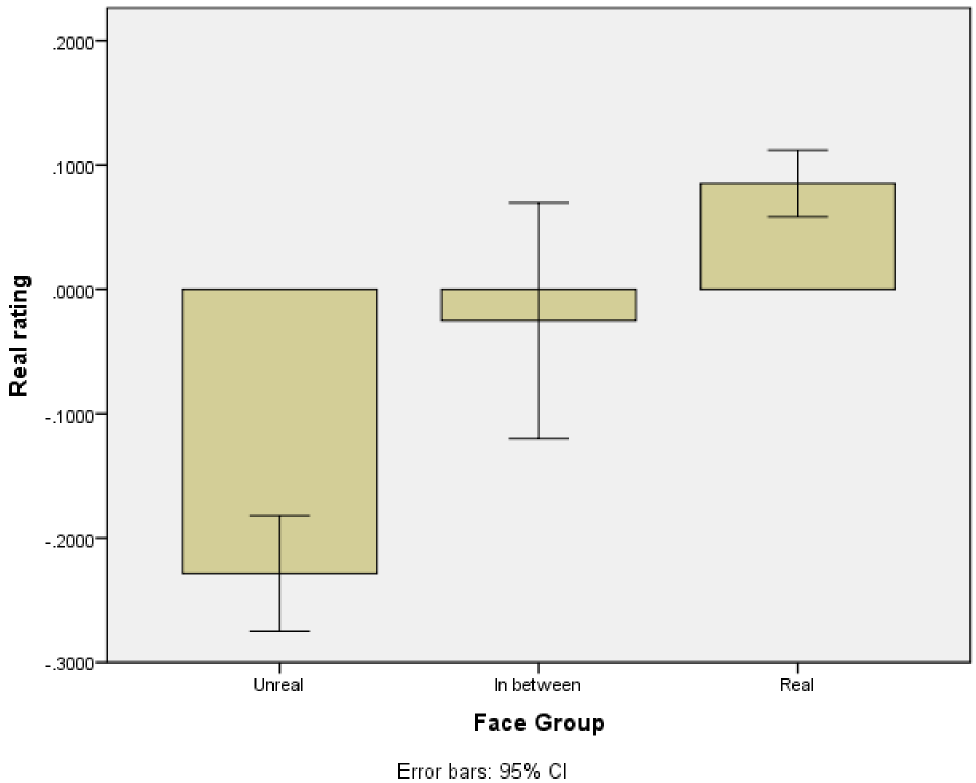
Error bars: 95% CI

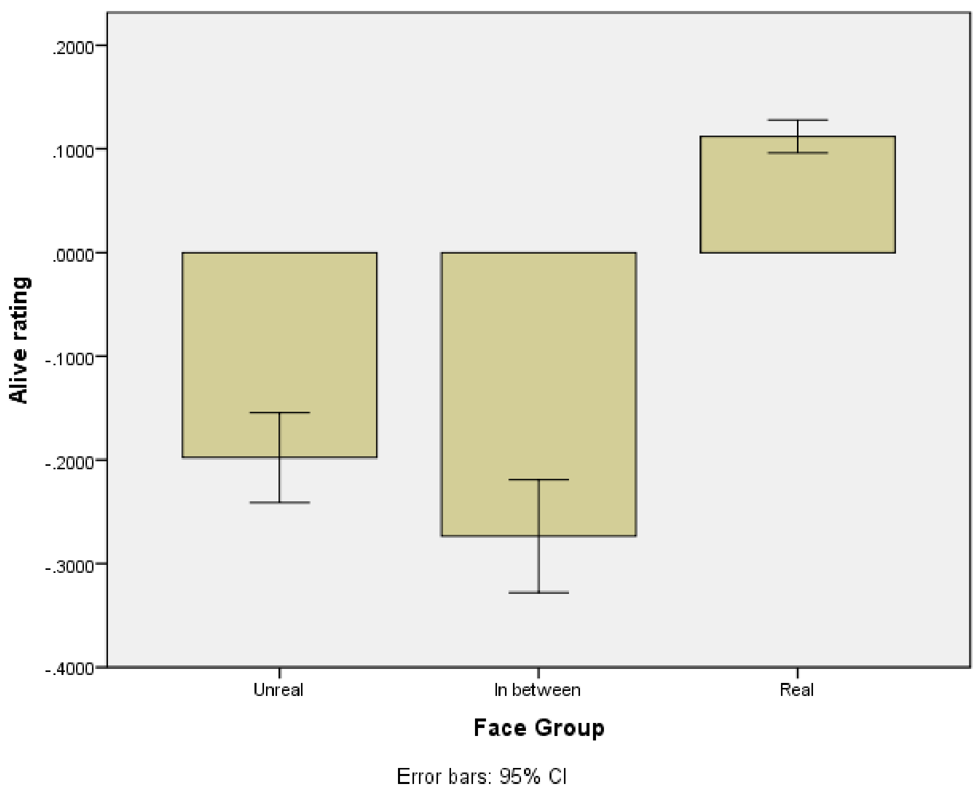


Error bars: 95% CI


.5000
.4000
.3000
.2000
.1000
.0000
-.1000
Threatening rating
Unreal
In between
Real
Face Group
Error bars: 95% CI


.1000
.0000
-.1000
-.2000
-.3000
-.4000
Likable rating
Unreal
In between
Real
Face Group
Error bars: 95% CI


.2000
.1000
.0000
-.1000
-.2000
-.3000
Real rating
Unreal
In between
Real
Face Group
Error bars: 95% CI


.2000
.1000
.0000
-.1000
-.2000
-.3000
-.4000
Alive rating
Unreal
In between
Real
Face Group
Error bars: 95% CI