# Warm flesh instead of cold steel: Can the Uncanny Valley effect be replicated with biological faces?

Faculty of Behavioural, Management and Social Sciences, University of Twente

Human Factors and Engineering Psychology

**Bachelor** Thesis

Marcel Pertenbreiter

s2151308

First Supervisor: Dr. Martin Schmettow

Second Supervisor: Dr. Peter de Vries

# **UNIVERSITY OF TWENTE.**

### Abstract

The uncanny valley effect describes the adverse reactions people experience when confronted with a highly human-like looking entity such as humanoid robots or virtual characters. The reason that almost human-like robots elicit a feeling of eeriness remained questionable as of now. Numerous research points to various explanations of the uncanny valley effect, including high-level cognitions, developmental factors, individual differences and evolution. As the existing research points towards an evolutionary origin of the UV effect, the present study investigated if the UV effect can be replicated with the biological faces of primates, human ancestors, and humans. The participants in this study performed an online survey in which they were presented with each stimulus for a brief time and were asked to rate them based on their perceived eeriness and likability. The results show that each participant experienced the uncanny valley effect when confronted with biological faces. Thus, this study delivered conclusive evidence that the uncanny valley effect is likely a universal phenomenon with its origin within evolution. Further speculations are developed about the benefit of the UV effect for the preservation of humans by arguing that it served as a drive for collective decision making or prevented interbreeding with other hominins in the past. The results of this study allow digging deeper into the view of the UV effect as an evolutionary mechanism.

### Introduction

### The Uncanny Valley

Imagine yourself in the future living in a retirement home. However, you are being nursed by humanoid robots that look almost identical to humans. How would you feel about this? The present research on highly human-like robots would say that emotions of eeriness and unease would dominate our reaction to these robots. At first glance, it would seem convenient that humans enjoy interacting with human-like robots, given that humans enjoy interacting with other humans that are similar to them (Paiva et al., 2005). However, should the humanoid robot become highly human-like and almost resemble a human in its appearance, then a person's response to this robot shifts from empathy to a feeling of adversity and eeriness (Mori, MacDorman, & Kageki, 2012). This effect is called the 'uncanny valley' and was first mentioned

40 years ago by the Japanese professor Masahiro Mori (Mori, 1970). This phenomenon displays a severe threat to the design of robots that are aimed at resembling humans.

If the robots are almost but not perfectly human-like, then people might be less willing to interact with them due to the eeriness they experience when looking at them. In a study by Smith, Sherrin, Fraune, and Šabanović (2020), they discovered that humans are less likely to engage and interact with a robot should this robot elicit feelings of disgust and eeriness. Furthermore, should people be less likely to interact with a highly human-like robot due to the uncanny valley effect, then the question arises whether robots should be designed simplistically to avoid the uncanny valley effect in the first place. Hence, the uncanny valley effect can be avoided by either designing robots with low similarity to humans or by designing robots that imitate them perfectly (Zhang et al., 2020). This shows that designing robots around the UV effect also has ecological benefits, with robots of low human likeness potentially performing better than their counterparts while also having less production and design expenses (Tung & Chang, 2013).

Alternatively, the uncanny valley effect might be avoided by knowing what features of a highly human-like character lead to this feeling of eeriness. This would enable the possibility of designing human-like characters that do not lead to an adverse reaction by minimising or excluding said features. In order to discover the features of the human-like character, in particular its face, that bring about this negative effect, it is essential to unravel the roots and causes of the uncanny valley effect. Multiple theories try to explain the origins of the uncanny valley effect. The two most researched theories among those are concerned with either cognitive or evolutionary mechanisms. Recent evidence points toward an evolutionary explanation that describes the UV effect as a means of self-preservation. Therefore, this study aims to investigate if the UV effect is evolutionary by presenting participants with various biological faces (primates, human ancestors/relatives, humans). If the UV can be found with biological faces, this would present conclusive evidence for an evolutionary origin. Nonetheless, it is essential to consider other explanations of the UV effect and see how they might influence each other in the expression of the UV effect. This study aimed to investigate the following research question: 'Do biological faces trigger the Uncanny Valley effect?'.

### **Theories on the Uncanny Valley**

Broadly summarised, the current research showed that two main theories aim to explain the UV effect. The first is the fast system theory (Haeske & Schmettow, 2016). This theory states that humans immediately experience a feeling of eeriness when looking at the faces of highly human-like characters without the need to consciously reflect on what they are observing. Thus, it is an automatic and unconscious brain process that occurs within a short amount of time in which the face is being observed (Slijkhuis, 2017).

On the contrary, the second theory, the slow system, states that the UV effect arises when the observer has enough time to consciously reflect on the human-like character and consequently, notice the discrepancy between the almost human-like character and the actual appearance of a real human (Cheetham, Pavlovic, Jordan, Suter, & Jancke, 2013). Within both theories, some multiple other theories and hypotheses try to discover the origins of the Uncanny Valley effect.

### Cognitive theories of the Uncanny Valley

Regarding the latter, more perception and cognition heavy theory, there is one hypothesis that has received increasing research attention. This is the category ambiguity hypothesis. Category ambiguity is supposed to occur when it becomes difficult to decide to which category an object belongs. Therefore, should a human-like character possess both traits that are commonly associated with, for example, robots, as well as attributes that are uniquely associated with humans, then this results in problems of determining if the presented character belongs to the category of humans or robots (Strait et al. 2017). Burleigh et al. (2013) argue that the ambiguity resembles a conflict that creates a feeling of discomfort and unease similar to that of the state of cognitive dissonance, in which an individual struggles to combine two or more conflicting ideas, attitudes, and beliefs (Elliot & Devine, 1994).

Support for this hypothesis comes from a study by Mitchell, Szerszen, Schermerhorn, Scheutz, and MacDorman (2011). They investigated to what extent robots and humans would be perceived as eerie if a robot speaks with the voice of a human and a human with the voice of a robot. Their results show that the congruent condition (human-figure with human voice and robot-figure with robotic voice) is rated significantly less eerie than the incongruent condition. Thus, the violation of an expectation (e.g. robotic appearance with human voice) might have made it troublesome to put the character (human and robot) into its supposed category resulting in an adverse reaction.

Another hypothesis that is similar to that of the category ambiguity is the violation of expectation hypothesis. This hypothesis claims that humans have developed throughout their life expectations about how specific entities are supposed to behave. A highly human-like robot would, due to its similarity to humans, be expected to also act like a human. If the robot should violate this expectation, then there is a discrepancy between expectation and reality, leading to a feeling of eeriness (MacDorman & Ishiguro, 2006).

One expectation people have about humans is that they are uniquely capable of experiencing high-level emotions and sensations compared to other animals and especially machines. In turn, this would mean that machines, including robots, are expected to express non or only a minimal amount of information that could suggest this capacity to feel and sense. Human-like characters/robots like androids, on the other hand, might indicate through, for instance, facial expressions that they possess said traits (Gray & Wegner, 2012). This particular violation of the expectation that robots should not have described traits defines the mind perception hypothesis (Zhang et al., 2020).

### Developmental factors of the Uncanny Valley

In addition to cognitive influences, other researchers proclaim that developmental factors play a role in the effect of the uncanny valley. Lewkowicz and Ghazanfar (2011) examined the theory that humans and monkeys both develop a prototype of the faces of their species during infancy. This prototype helps them to detect and identify slight anomalies that distinguish human faces from faces of other species and robots. In their study, they presented 6-, 8-, 10-, and 12-monthold infants with three different entities: a human, an uncanny avatar with increased eye size, and a realistically looking avatar. It is assumed that six-year-old infants, compared to older infants, did not yet develop a prototype of a human face and would, thus, not discriminate between the uncanny avatar face and the human face. This means, in turn, that six-month-old infants are narrower in the faces they can differentiate, being able to identify not only human faces but also that of different monkey species. However, the development of a human face prototype helps to improve the perceptual expertise for the specific stimuli features that belong to this prototype. As

a result, older infants have a more detailed stimulus structure that makes it easier to detect features in other faces that deviate from this structure.

In conclusion, these findings show that the six-year-old months would not look longer on the human face than on the uncanny avatar face with unnatural eye size. The experiment of their study confirmed the expectations with six to twelve months-old infants looking significantly more on the human face than on the uncanny character compared to the six-month-old infants. The six-month-old infants, on the other hand, even preferred the uncanny character over the human face. Thus, at six months and onwards, the infants gathered enough experience and information about human faces to detect even minor differences. The study demonstrated that developmental factors might play an essential role in the causality of the uncanny valley. Notably, the study's authors did not neglect the effect of evolutionary factors on the uncanny valley but instead highlighted that it is likely an interaction between development and evolutionary aspects.

### The influence of individuality on the Uncanny Valley

Different personality traits and facets have also been researched in relation to the uncanny valley effect. With the assumption that the UV effect is of evolutionary origin, it becomes a universal experience that almost everyone experiences. In accordance, the impact of individuality should be limited or non-existent. McDorman and Entezari (2015) examined to what extent nine unrelated traits influence the uncanny valley sensitivity (ratings of eeriness and warmth) of participants when looking at entities of different human likeness. Eight traits had a significant positive correlation with the eeriness rating. Regarding the dimension of warmth, seven out of nine traits had a significant negative correlation.

Overall, this shows how human individuality might impact the occurrence and intensity of the uncanny valley effect. For instance, individuals who score higher in perfectionism focus more on details to find mistakes and flaws that make them feel inadequate when they find them (Rice, Bair, Castro, Cohen, & Hood, 2003). Thus, the imperfections created by the mismatch between the human-like entity's face and the brain model of an actual human face are more easily detected by perfectionistic individuals, presumably enhancing the feeling of eeriness (MacDorman & Ishiguro, 2006). Although these results suggest that individuality asserts an effect on the UV, other studies provide different results concerning the four traits religious fundamentalism, negative attitude towards robots, human-robot uniqueness, and animal reminder sensitivity. Haeske and Schmettow (2016) researched if the UV effect can be found when the stimuli (robot faces) are presented to the participants only for a short duration of 100ms or for an unlimited duration. The ratings of the short presentation time were highly correlated with that of the long condition, demonstrating that it is likely a rapidly performing system that leads to the uncanny valley effect. There was still a discrepancy between the eeriness ratings of the short presentation time. This variance requires an explanation that is not concerned with individuality because none of the four traits included in the study had any predictive power neither in the short nor in the long condition (presentation time).

### Evolutionary theories of the Uncanny Valley

Evolutionary theories and hypotheses of the uncanny valley effect have been investigated and backed up by only a handful of studies. The first person to point out that the uncanny valley effect might have evolutionary roots was Mori himself, who argued that the feeling of eeriness serves as a form of self-preservation from proximal sources of danger (Mori et al., 2012). Now, the question arises from what kind of evolutionary threats this feeling of eeriness is supposed to protect us.

One possible threat is displayed by pathogens and the associated contamination (Zhang et al., 2020). Some pathogens can be detected through specific outward appearances. For instance, the skin rashes of people suffering from Hansen's disease might trigger a disgust response to avoid this individual and the associated possibility of contamination (Oum, Lieberman, & Aylward, 2011). Similarly, highly human-like entities could possess some facial cues that the human species has evolved to circumvent by experiencing eeriness when seeing them. Moreover, genetic closeness makes it more likely that the disease or illness the person is carrying is transmittable. Since human-like entities could be falsely perceived as being human and having similar genetic material, it is plausible for the human brain to think of them as probable pathogen transmitters (MacDorman & Ishiguro, 2006). Support for this theory has been only given indirectly by showing that individuals who are more sensitive to the emotion of disgust are also more susceptible towards the uncanny valley effect (MacDorman & Entezari, 2015).

Another evolutionary hypothesis termed mortality salience states that human-like entities elicit adverse reactions because they remind us of death and our mortality. For instance, an emotionless and pale face of a humanoid robot might be perceived as a corpse. McDorman and Ishiguro (2006) went a step further. They claimed that human-like robots might not only subconsciously create a fear of annihilation but also that of being replaced by a human-like entity and the fear that humans themselves can be reduced to and are nothing more than soulless machines or beings.

The mortality salience hypothesis is strongly connected to the terror management theory. According to this theory, unconscious thoughts of death and their corresponding anxiety are counteracted by highlighting and defending one's worldview and self-esteem (MacDorman, 2005). The dead appearance of some human-like entities represents one factor that could nonconsciously remind humans of their mortality. McDorman (2005) tried to provide evidence for the mortality salience hypothesis by investigating the extent to which participants defend and support their worldview when observing an uncanny robot compared to a control group that were shown an image of a human. The results favour the hypothesis showing that participants in the experimental group were more likely to prefer information that supported their worldview compared to the control group. However, caution is advised because the study was only conducted with one stimulus, and it is unknown if this exact stimulus triggers any confounding variables.

Moosa and Ud-Dean (2010) emphasise that the pathogen avoidance hypothesis is too specific and instead developed the danger avoidance hypothesis. Centuries ago, the human life span was about 25 years, implying that many deaths were premature being the result of "predators, invaders, disaster or disease". Corpses could be seen as indicators of potential dangers such as predators still lurking in the bushes or lethal gases that steadily spread out through the air. According to the danger avoidance hypothesis, a dead looking human-like entity would create a feeling of eeriness that signalises that there is danger around and elicits a response to be cautious (Moosa & Ud-Dean, 2010).

The outward appearance of an individual cannot only be representative of a potential pathogen the person is carrying, but it might entail information about the genetic material that is important for reproduction. Due to the evolutionary drive, ancestors sought to reproduce with individuals who were equipped with the correct genetic material and the corresponding traits,

which facilitated their adaptation to their environment and, consequently, making it more likely for them to survive and reproduce. Attractiveness serves as an indicator to choose individuals to breed with who have 'good' genetic material.

Attractiveness is associated with several traits important for reproduction, such as in the case of an individual's health, greater attractiveness comes with a decreased likelihood of possessing genetic anomalies that lead to disabilities, retardation and diseases (Voland & Grammer, 2003). The flaws and imperfections of human-like entities might make them less visually appealing, leading to an eerie feeling when being confronted with them. Hanson (2005) tested this so-called evolutionary aesthetics hypothesis by showing participants morphed faces that range on a continuum from robot to human faces. Besides, they artificially changed the appearance of the faces to make them more appealing regardless of whether the face is robotic, human-like, or completely human. Their results show that the visually pleasing faces all were consistently rated low in eeriness even when the face was nearly human-like (Zhang et al., 2020).

### Evidence for the evolutionary theories

All these evolutionary hypotheses have in common that they portray the uncanny valley effect as a mechanism that increases the organism's reproductive fitness. Moreover, should the uncanny valley effect at its core truly stem from evolution, then it would be a universal phenomenon that can be experienced by all humans and possibly even other species. Evidence for this claim and the evolutionary hypotheses in general stems from research by Steckenfinger and Ghazanfar (2009). They exposed macaque monkeys to three different facial stimuli of their species and measured the time they spent looking at each stimulus. The first stimulus was a real face of a monkey. The other two stimuli were synthetically created stimuli of monkey faces. One aimed to look realistically and the other unrealistically. They hypothesised that the monkeys will spend less time looking at the realistic, synthetic monkey face compared to the other two faces because it appears to be conspecific at first, which, however, cannot meet the expectation of the concept of a conspecific. The study results are in line with this hypothesis, showing that the uncanny valley effect can be reproduced with primates points toward an evolutionary framework.

Siebert et al. (2020) conducted a similar study in which they presented monkeys with visual stimuli of their species with varying degrees of realism. As expected, the monkeys showed

an avoidance reaction towards uncanny stimuli, but the stimuli eliciting this reaction were those of intermediate realism and not high in realism. Instead, it seems that abnormal features in appearance led to an adverse reaction. Arguably, this provides conclusive evidence for evolutionary theories and against the hypothesis of category ambiguity since the 'valley' of the UV was not present for the realistically looking monkey, meaning that they were not falsely mistaken for belonging to their species.

Moreover, the uncanny valley effect can be found for the majority of individuals even when the stimuli, in this case, robotic faces, are only presented for a maximum of 50 milliseconds (Moll and Schmettow, 2015). In other words, the participants would not have enough time to consciously reflect on what they are observing, which indicates that it is likely an automatic process having evolutionary origins. Nonetheless, the UV effect was found for everyone with a presentation time of two seconds, which allows for conscious reflection of the stimuli. Thus, it is likely that the UV effect has an evolutionary basis that might vary in multiple aspects (e.g., strength) depending on the magnitude of cognitive influences.

Although people might differ in their sensitivity to the UV and the exact human likeness score that leads to the effect, the UV can nonetheless be found to exist for almost every person (Koopman, 2019), marking it as a universal experience. The variations within the UV experience are presumably the result of individual differences and cognitive mechanisms that were already discussed.

#### The present study

Since the current research points toward an evolutionary framework of the UV, the present study aims to provide evidence that the UV effect is indeed evolutionary in nature. Therefore, an experiment was conducted that presented the participants with various biological faces of primates and human ancestors (such as the Australopithecus Africanus or Neanderthalensis) that range from low human likeness (e.g. rhesus monkey) to almost human-like (human ancestors) to completely human. One aspect shared by all evolutionary theories of the UV effect is that they are not necessarily restricted to robotic faces. For instance, the pathogen avoidance hypothesis speculates that our human ancestors have evolved mechanisms that detect facial cues indicative of transmittable diseases. Thus, this mechanism was not developed (or arose through selective drives) to protect us from off-looking, human-like, robotic faces in the first place but of biological faces such as that of humans. If the UV effect can be replicated with biological faces, this would prove that the evolutionary theories of the UV, which focus on our adaption to biological faces, hold true. As a result, the UV effect could be classified as a phenomenon of evolution that increased the reproductive fitness of humans and primates and maybe also did so for our human ancestors. The following research question is examined: *'Do biological faces trigger the Uncanny Valley effect?'*.

Since emotions are deeply connected to evolution, it could be of value for our understanding of the UV to investigate the extent to which the facial expressions of the stimuli faces affect the perceived eeriness. In a study by Tinwell, Grimshaw, Nabi, and Williams (2011) six different emotional expressions have been investigated regarding their relationship with the perceived human likeness and familiarity of virtual characters. These emotions were happiness, sadness, anger, fear, surprise, and disgust. Among these emotions, happiness was rated lowest in familiarity. This is surprising because happiness is usually associated more positively. The researcher reason that it results from the expression being perceived as artificial since it was modelled on the character through a human actor. Nonetheless, artificially created facial expressions and the resulting impressions should be evaluated with caution. On the other hand, sadness was rated the highest in familiarity presumably because the participants anthropomorphised the virtual characters (Gong, 2008) and felt empathetic towards them.

Facial expressions with the UV have also been researched regarding the mortality salience hypothesis by Koschate, Potter, Bremner, and Levine (2016). The expression of basic emotions (happiness and sadness) of human-like entities has strongly reduced the occurrence of thoughts of death in their study. Although the reduction in thoughts of death also reduced the UV effect, the UV effect was not eliminated completely, meaning that the mortality salience hypothesis might only be one of the multiple evolutionary factors explaining the UV.

As of now, the exact effect of the emotional expressions of human-like characters on the perceived eeriness remains questionable. For this reason, various facial expressions were included in the study as a confounding variable to test whether they exert a significant effect on the perceived eeriness. This would also grant certainty for future research into the UV effect of whether it is necessary or fruitful to include facial expressions.

### Methods

### Procedure

First, the participants received a short overview of the study procedure and were told that the aim of the study is to assess their emotional response to different faces. The Uncanny Valley effect was not mentioned in order to prevent a possible response bias. If the participants signed the Informed Consent, they were presented with an overview image consisting of 16 different faces to get them acquainted with the presented stimuli. After they have been familiarised with the type of faces, the participants could begin the study.

The participants were presented with each stimulus for a maximum of 2 seconds. Then, they rated all stimuli based on their eeriness on two different scales, which were shown after each stimulus. The stimuli were divided into four blocks with each block consisting of 25 stimuli. Within each block, the order of the stimuli has been randomised. Between these blocks, the participants had the chance to take a break.

The first scale consisted of only one item, whereas the second scale used five different items to measure the construct of eeriness. The five items were equally randomised among all stimuli for every individual participant. After rating the stimuli, the participants were asked to fill out three different personality questionnaires. In the end, the participants were debriefed about the study and were given the opportunity to write a comment on the survey.

### Stimuli

In total, 100 stimuli were used in the study. The stimuli included faces of primates, human relatives and ancestors, humans and robotic faces. Displayed in numbers, 12 faces presented the group of Homo Sapiens, 27 faces belonged to various extinct species of the Genus Homo, three faces to the species of Neanderthalensis, 47 faces to various primates, and 11 stimuli showed robotic faces. The biological faces were chosen in a way that they cover the range of human likeness from low human likeness (primates), to almost human-like (human ancestors/relatives), to human. The robotic faces stemmed from a previous study by Koopman (2019) and were included to compare the results of biological faces with that of robotic faces.

Moreover, to filter the most suitable stimuli, the current study used criteria established by Marthur & Reichling (2016) that were partly changed and modified to suit the context of the study.

Inclusion criteria:

1. The full face is shown from top of head to chin.

2. The face is shown in frontal to 3/4 aspect (both eyes visible).

3. The individual belongs to a species that has lived or is still living (no fictitious faces)

4. The image depicts:

a. A living individual

b. Real remainders of the animal that have been taxidermied to closely resemble the way it looked when it was alive

c. A realistic hominid bust (no artificially created CG or morphed face)

5. In case of b) and c), it is shown as it looked when it was alive (not missing any hair, facial features or skin)

6. The resolution of the photo is sufficient to yield a final cropped image with a resolution of at least 450x450 pixel

Exclusion criteria:

1. The individual represents a famous person (for the human image)

2. The image shows other faces or body parts that would appear in the final image.

3. Objects or text overlap the face.

The majority of stimuli were collected from the catalogue of hominid busts of John Gurche (http://gurche.com/) and the open-access databases Global Biodiversity Information Facility (https://www.gbif.org/) and PrimFace (https://visiome.neuroinf.jp/primface/). The remaining stimuli were collected through targeted google searches to include faces of modern humans of different ethnicities and human and non-human primates displaying various emotions. Altogether 111 biological faces were collected. Each stimulus was rated by the four researchers based on its human likeness (from 0 - 100), the perceived emotional valence (from -100 - 100) and the perceived emotional expression based on seven categories (neutral, happiness, sadness, anger, fear, surprise, disgust). The interrater reliability for the human likeness was excellent (>0.92). The biological stimuli with the lowest inter-rater agreement regarding human likeness were removed from the set ending up with the final set of 90 biological stimuli. All stimuli except for some robotic faces were placed on a white background to eliminate the possibility of the background asserting an effect on the participants' emotional response.

### Measures

To measure the eeriness elicited by the stimuli, two scales were used. The first measure was used in a study by Marthur and Reichling (2016). It entails one scale with two opposing items, namely "less friendly, more unpleasant, creepy" vs "friendly and pleasant, less creepy", that was displayed on a continuous visual analogue scale (VAS) ranging from -100 to +100. The scale measures the likability of the facial stimuli.

The second scale was developed by Ho and Macdorman (2017) and consists of five different item pairs: uninspiring–spine-tingling, boring–shocking, predictable–thrilling, bland– uncanny, and unemotional–hair-raising. These items were presented on a VAS ranging from 0 to 100 and not -100 to 100, since the left side of the scale (e.g the concept uninspiring) describes how a face is normally perceived: with no great emotional reaction or, in other words, neutrally. All five items demonstrated excellent psychometric properties. Both measures were translated into Dutch and German and were used in the present study. Moreover, the Very Short Authoritarianism Scale (VSA), the Short Form Need For Closure Scale (NFC-SF), the short form of the Big Five Inventory (BFI-S) were included in the study.

### Materials

The survey was created on the website Qualtrics. An integrated feature of Qualtrics was used to give the participants the opportunity to change the language of the eeriness scales. The Dutch and German translations for the five-item eeriness scale were created by Koopman (2019). The respective translations for the likability scale have been developed specifically for this study.

### **Participants**

In total, 84 participants were included in the study. Of those 84 participants, 2 denied informed consent. The recruitment of participants was conducted via the University of Twente's own subject pool (SONA), consisting mainly of University students. Further participants were gathered through social media platforms as well as through connections to close relatives and friends of the researchers. Personal information such as demographic data was not collected because it is investigated if the UV effect is innate and universal, making the effect of individual differences such as age and gender irrelevant and implausible.

### Design

The study used a repeated measure design, with each participant rating all 100 stimuli based on their eeriness. Regarding the research question, the independent variable in this experiment was the human likeness score of the stimuli and the dependent variable the perceived eeriness.

### **Data Analysis**

### Polynomial regression

The first step was to analyse the results on a population level. Therefore, we took the eeriness ratings of the participants and extracted the mean rating for each stimulus. These ratings were then plotted in four different polynomial models with eeriness as the dependent variable and human likeness as the predictor variable. These four models range from a grand mean function to the third-degree polynomial function. The depicted emotional valence was added as a control variable. The predictive accuracy of the four models was estimated using the leave-one-out approximation (LOO).

It was predicted that a third-degree polynomial graph most accurately describes the relationship between eeriness and human likeness since in comparison to a first and second-degree graph, it is the only model that has both a peak and a trough in its curve (Schmettow, 2021, s. 5.5). This is also a characteristic of the uncanny valley curve. The second-degree curve has either a peak or trough, and the first-degree model describes a curve with a linear pattern. Mathur and Reichling (2016) compared the first-, second-, and third-degree polynomial models on their data and discovered that the third-degree function is the most suitable. Since the cubic model is the only model among the four that can potentially generate an uncanny curve, it was used as a test statistic for the MCMC sampling to identify the probability of an uncanny valley curve.

### Calculating the position of the trough

Multiple steps must be performed to estimate the local minimum of a third-degree polynomial graph (Schmettow, 2021, s. 5.5). The first step is to find the stationary points, which are the points at which the function is neither increasing nor decreasing anymore, meaning that the slope is equal to zero. If the slope is equal to zero, then this means that there is no gradient. The derivative of the function indicates the gradient of the graph. Thus, the stationary points can be

found by setting the derivative of the third-degree polynomial function equal to zero and solving the equation:

$$f'(x) = \beta_1 + 2\beta_2 x + 3\beta_3 x^2 = 0$$

The third polynomial degree can represent a parabolic form that hits point zero once at a minimum and once at a maximum. In a similar notion, these can be found by setting the second derivative of the third-degree polynomial function equal to zero and solving the equation:

$$f''(x) = 2\beta_2 + 6\beta_3 x = 0$$

This could lead to two solutions, with one indicating a minimum and the other a maximum. The slope of a minimum, or a trough, switches from negative to positive, hitting x = 0 while rising. Thus, a local minimum can be identified if f''(x) > 0 and vice versa for a maximum point (Schmettow, 2021, s. 5.5).

### Multilevel model

Based on this research, it is assumed that the uncanny valley has an evolutionary origin making it a universal experience. Analysing the results on a population level might demonstrate that an uncanny valley effect can be identified on average. It does not, however, guarantee that each participant experienced the UV effect. One could consider a situation in which a few participants did not experience the UV effect, but their effect size was not strong enough to render the uncanny valley effect as not present on a population level (Schmettow, 2021, s. 6). Therefore, the next step was to analyse the results on a participant level to see how each participant reacted to the stimuli. Moreover, the overall variation of the participant population was introduced as a random factor (standard deviation) alongside the population level (fixed effect) in a multilevel model analysis (Schmettow, 2021, s. 6). This analysis is indicative of the variation that can be explained by individual differences compared to that of the fixed effect/population level.

#### *Psychometrics vs Designometrics*

In the introduction it was mentioned that the present research shows contradictory findings concerning the influence of individual traits on the UV effect. One plausible reason that McDorman and Entezari (2015) found an effect of individuality on the eeriness rating is because their eeriness rating scale did not measure the eeriness of the stimuli itself, as it should have, but instead measured the perceived eeriness of the participants in response to the stimuli. Whereas

the qualities and attributes of individuals, such as intelligence, can be measured directly by letting the participants rate items (person by items), the attributes of designs, for example, faces, can only be measured through ratings of persons on items in response to multiple designs. That is the key difference between a two-dimensional psychometric scale (measuring person by items) and a three-dimensional designometric scale (measuring designs by items by person) as termed and developed by Schmettow and Bosrci (2021). A psychometric scale needs a large sample size of participants to effectively differentiate between persons. In a similar notion, a designometric scale with its focus on discerning between designs needs a large sample of designs to account for the variance between the designs.

McDorman and Entezari (2015) used too few designs. In other words, they used only six different stimuli in total. Hence, the variance in terms of eeriness between the designs might have been too low, skewing the results. Moreover, it is not enough to merely use a large number of designs to measure designs. The data analysis needs to be conducted from a designometric perspective too (Schmettow & Bosrci, 2021). Unfortunately, this was not considered by McDorman and Entezari (2015) either. The present study analysed the results in a designometric fashion by averaging the participants' ratings on the items over the designs. The participants' ratings were added as the random participant level effects in the multilevel model.

### Beta regression

The rating scales used to measure the perceived eeriness have a set range (one from 0-100, the other from -100-100). These fixed boundaries needed to be considered when analysing the results. Hence, a beta regression was applied that uses a logit link function and a double bounded error distribution (Schmettow, 2021, s. 7.4.2). Since the boundaries of a beta distribution range from 0 to 1, the ratings needed to be rescaled accordingly in order to lie within this range (excluding the boundaries) (Schmettow, 2021, s. 7.4.2).

### Distributional model

In another data set using rating scales to let participants assess robots of different human likeness, Schmettow (2021) observed that the individual participants' ratings varied greatly around the population level. This variation differed from the regular mean-variance relationship captured by the usual Beta model (Schmettow, 2021, s. 7.5). The variation was likely the result

of the process of anchoring. Anchoring describes how participants themselves define the endpoints of a bounded rating scale (Schmettow, 2021, s. 7.4). Whereas one participant would think that an (extremely) shocking face (boring-shocking item of the eeriness scale) would look like that of a zombie in a horror movie, another participant might set the boundary for a shocking face much lower. Ultimately, these imaginary boundaries affect to what extent the participants utilise the full visual analogue scale. Anchoring can be partly controlled by applying a distributional model to the multilevel model. The distributional model can account for broader and narrower distributions of the participant variation (Schmettow, 2021, s. 7.5). In other words, the distributional model allows putting a response variance parameter on the participant level variation, which can be used to account for the different response patterns of the individual participants (Schmettow, 2021, s. 7.5). Thus, a distributional model was used to account for the effects of anchoring that the subjective rating scales used in the study are susceptible for.

### Universality

At last, the probability of finding an uncanny valley curve on a participant level will be investigated. The theory that the ucanny valley effect is a universal phenomenon experienced by everyone is a universal statement in itself (Schmettow, 2021, s. 6.4). Concerning absolute evidence, universal theories or statements cannot be proven, but they can be rejected. One single piece of evidence that contradicts the universal statement suffices to abandon the theory and its' pursue entirely (Schmettow, 2021, s. 6.4). Thus, to declare the UV effect as a universal phenomenon, it needs to be demonstrated that every participant undergoes the typical uncanny valley curve that consists of two stationary points with one indicating a minimum and the other a maximum. Another condition is that the maximum point, marking the shoulder, precedes the minimum, the trough. At last, using the MsCMC sampling method, the probability of finding a UV curve estimated on a participant level. The probability of finding such a curve needs to be high for every participant to prevent rejecting the theory.

### Results

In this section, the obtained data of the participants was analysed based on the procedure given in the data analysis. First, the relation between human likeness and the perceived eeriness was

investigated on a population-level model. Next, the participant level was included in the model, making it a multilevel model. The participant-level model allowed to determine the probability of the presence of the UV effect for the individual participants. Moreover, a distributional beta regression was employed to account for the effect of anchoring. At last, the effect of emotional valence on the eeriness ratings was considered.

### Polynomial regression on population-level

Investigating the relationship between human likeness and eeriness with the leave-one-out method showed that the cubic model does not have the best predictive accuracy on a population level. The model with the best predictive accuracy was the linear model with an *IC* of -4.559, followed by the quadratic model, then comes the cubic model and at last the grand mean model (see *Table 1*). As can be seen later, differences on the participant level, also in terms of the response style, significantly affected the population level.

Although the cubic model was not the preferred model on the population level, *Figure 1*. demonstrates that the cubic model draws a similar curve through the relationship of eeriness as does the LOESS model. Since the LOESS creates a smooth curve between the two variables, it can be said that the cubic model accurately describes the relationship between the two variables as well.

Moreover, applying the Monte-Carlo Markov Chain (MCMC) method on the cubic model showed that the probability of the population level is an uncanny valley curve equals 72.5%. Due to these, in part, opposing results, it is necessary to consider the participant level.

### Table 1

Model	IC	Estimate	SE	diff_IC
M_poly_1	looic	-4.559410	9.863243	0.000000
M_poly_2	looic	-3.651307	9.411036	0.908103
M_poly_3	looic	-1.364203	9.659988	3.195206
M_poly_0	looic	-1.119816	9.272131	3.439594

Model ranking by predictive accuracy

### Figure. 1



Graph of the eeriness ratings of the participant averaged over each stimulus

Note. The observed values each display one of the stimuli. The x-axis represents the human likeness ranging from 0 to 1. The y-axis represents the rated eeriness (inversed) ranging from 0 to 1.

### Multilevel model with distributional beta regression

When the participant level is taken into the equation, the picture becomes more apparent. *Table* 2. summarises the results of the multilevel model with a beta distribution. The mean level of the eeriness ratings (intercept) of all stimuli roughly amounts to 25. Strikingly, the participant-level standard deviation of the intercept is 17.25. Thus, participants differ considerably when it comes to rating the stimuli based on their perceived eeriness. These differences can be best described by the individual differences in the utilisation of the rating scales. As the individual curves in *Figure*. 2 all resemble an uncanny valley curve, the participant of the first curve mainly used eeriness ratings ranging from 0.6 to 0.8, whereas the participant of the third graph (starting from the left), applied ratings ranging from 0.5 to 0.65.

### Table. 2

Population-level coefficients with random effects standard deviations

fixef	center	lower	upper	SD_Part	
Intercept	0.2423457	0.1349185	0.3537239	0.1725790	NA
NA	1.2259843	1.0541065	1.4012498	NA	0.5545595
valence	0.0035227	0.0026571	0.0044081	0.0012479	NA
humLike	2.1916328	1.4888111	2.8686954	0.1213779	NA
humLike_2	-5.0889238	-6.6040510	-3.5435657	0.1514233	NA
humLike_3	3.3125477	2.3226758	4.2920712	0.1318868	NA

### Figure 2.

Seven graphs each illustrating the responses of a different participant plotted as a third-degree polynomial



Note. The x-axis presents the human likeness score of the stimuli. The y-axis presents the

### Figure. 3

Spaghetti plot of the responses on participant level computed as a third-degree polynomial graph



Since these variations mainly stem from the different response styles regarding anchoring (see Method section on *distributional model*), a distributional model was applied with a response variance parameter that adjusts the ratings accordingly. The spaghetti plot in *Figure. 3* shows the individual curves of all participants comprised in one graph and controlled for the effects of divergent response styles. What can be seen is that all the curves have the typical shape of an uncanny valley curve with a maximum point and a minimum point that is located on the right of the maximum.

A narrower look at the positions of the trough reveals that the troughs range from 0.625 *(human-likeness)*, describing the participant with the lowest trough position to 0.829 *(human-likeness)*, the participant with the highest trough position. All other troughs position themselves between these two values.

Moreover, *Figure*. 5 shows the positions of the trough and the shoulder for every participant. The figure clearly illustrates that there is little variation in the positions of both the

shoulder and trough between the participants. This suggests that the human likeness of the stimuli as a predictor sufficiently explains the consistent eeriness ratings and positions of both shoulder and trough. This also means that there is not much room left for other predictors or factors such as emotional valence or individuality to explain these ratings.

Finally, the probability was calculated that the individual curves represent an uncanny valley curve. As depicted by *Figure. 4*, most participants were highly likely of experiencing the uncanny valley effect with a likelihood of more than 91%. Only four participants scored below 91%, but their probability was still relatively high ranging from 79% to 86%. This shows that all participants were likely of experiencing the uncanny valley effect.

### Figure 4.





### Figure. 5

### Individual position of shoulder and trough



### **Emotional Valence**

The emotional valence affected the eeriness ratings only slightly by an increase in the eeriness ratings by a factor of 0.0035 with each increase in emotional valence (regardless of positive or negative) (*Table. 2*). On the other hand, the certainty is good as there is a probability of 95% that the true value lies between 0.00271 and 0.00437.

Moreover, as illustrated by *Figure. 6*, the faces were perceived as eliciting the most eeriness, the closer the emotional valence reached the zero mark with a slight downward trend as the emotional valence becomes highly positive. With rising emotional valence (regardless of whether the valence was positive or negative), the faces were increasingly rated as less eerie.

### Figure. 6.

Correlation between the predictor variable 'emotional valence' and the dependent variable 'eeriness ratings' for each stimulus



### Discussion

### **Conclusion of results**

This study delivered conclusive evidence that the uncanny valley effect is not solely bound to the negative experience of human-like, artificial entities like robots or animated characters. The evidence is conclusive since the fitted curves in their appearance all resembled an uncanny valley curve with a shoulder and a trough on participant level. Moreover, it has been shown that there is a high probability that all these individual curves indeed present an uncanny valley curve. Thus, the uncanny valley can be found when the stimuli consist of human-like, biological faces such as primates or human ancestors. The research question: *'Do biological faces trigger the Uncanny Valley effect?'* can be affirmed. Should the UV effect indeed have evolutionary origins, then it is a universal trait that is shared among (almost) all individuals. In line with this condition, the results show that every participant was highly likely of undergoing the effect. These results advocate evolutionary explanations and theories, including the danger- and pathogen-avoidance hypotheses, the evolutionary aesthetic hypothesis, and the mortality salience hypothesis.

Next to the research question, the emotional valence was included as a control variable to test whether it can impact the eeriness ratings next to the factor of the human likeness. The results show that the faces were perceived as less eerie the closer the emotional valence reached the value of zero, which describes a neutral facial expression. However, the effect size of the emotional valence was small, rendering its explanatory power relatively insignificant.

### Limitations of the study

The first limitation concerns the background that the stimuli faces were placed on. The initial background of the stimuli was removed, and the stimuli were all placed on a blank, white background. Many of the stimuli were not cut out perfectly or smoothly meaning that they might have appeared edged or odd. Moreover, seeing biological faces without any neck or body could give the impression of looking at a corpse, which could interfere with the eeriness ratings. At last, a person looking at a primate might strongly expect to observe this stimulus in its natural habitat and not on a white background. However, all stimuli received the same treatment so it would affect all stimuli equally making the effect less significant if it asserted an effect in the first place. Besides, placing the faces on a white background controlled for the influence of the background.

In total, the participants have had to rate 100 stimuli. This might have been exhaustive for some of the participants. This concern is also reflected in a few comments the participants could give at the end of the study. If the participants felt exhausted, they might not have rated the stimuli based on their feeling but asserted random numbers to the stimuli. Attention checks could have given information if the participant paid attention to the study, but they were not used in this study. To counteract the effect of the burden of too many stimuli, the stimuli were divided into four blocks with breaks in-between that gave the participants the chance to take time off if needed. Within the blocks, the stimuli were presented in random order. Consequently, a fatigue effect would have affected all stimuli relatively equally. Moreover, the process of perceiving and judging the stimuli is to a large extent unconscious and fast, requiring only a minimal amount of effort.

The survey was distributed to people that might have not possessed adequate English language skills and only the rating skills and a few descriptions were translated into German and Dutch. Nonetheless, understanding the meaning of the rating scales and having a basic understanding of what to do, was probably enough to complete the study appropriately.

At last, the eeriness rating scale used in the study that was developed by Ho and Macdorman (2017) initially consisted of two factors that in combination, assess eeriness. This study only used the factor '*Spine-tingling*' without the '*Eerie*' factor. The *Spine-tingling* factor only explains 39.5% of the total variance of the eeriness construct. Eeriness could have been assessed more accurately if the items of the *Eerie* factor were included in the study. This does not necessarily have to be a downside. Replications of this study (or similar studies) could use the *Eerie* factor to investigate the UV effect. Finding differences in the results between these two factors would give a clearer picture of the attributes that contribute to the experience of eeriness in the UV effect.

#### Speculating about the Origins of the Uncanny Valley within Evolution

The aforementioned evolutionary hypotheses of the UV already provide multiple explanations for its causality. Knowing through the results of this study that the UV is likely an evolutionary development gives more room for speculating about other possible theories and causalities.

### Eeriness as directing collective decisions about fight- or flight-responses

The stimuli that fell into the trough of the uncanny valley consisted mainly of faces depicting ancestors or relatives of the homo sapiens such as the Homo Erectus or Neanderthalensis. It can be valuable to investigate the past relationships between the homo sapiens and other species of the genus homo. Travelling into the past and looking at how the modern human, the homo sapiens, emerged, leads to two different theories. The first theory, the "out of Africa" (OOA) theory states that homo sapiens have evolved into the modern human within Africa between 200.000 and 60.000 years ago. From there one, it is assumed that they moved to Asia and/or Europe where they replaced other hominins including the homo Erectus, homo Neanderthalensis and the Denisovans (Ko, 2016).

The second hypothesis, the multiregional evolution hypothesis, assumes that through interbreeding, the various hominins evolved together into the subtypes of the homo sapiens, which in turn, make up a continuous gradient human species instead of one that could be defined in categorical terms (Ko, 2016). The most recent view on the OOA hypothesis keeps its focus on the evolution of the homo sapiens with its origin in Africa but allows for the possibility of interbreeding between humans and other hominins, namely Neanderthalensis, H. Heidelbergensis, and Denisovans. Thus, it combines the two theories (Ko, 2016). What both of these theories have in common is that they describe how the modern human conquered regions and led to the extinction of other hominins either through sheer violence (OOA theory) or through a combination of violence and interbreeding (multiregional evolution hypothesis). Furthermore, genetic studies have shown that in comparison to other hominins, the homo sapiens was extremely aggressive and hostile (Ko, 2016). This hostility is presumably the result of the dangerous environment that Africa displayed at that time due to the multitude of large predators that inhabited the land. The aggressive tendencies of humans facilitated their survival (Ko, 2016).

Within this framework, the eeriness that the humans at that time experienced might have served the purpose of directing coordinated aggression towards the other hominins. Research into the mechanism of disgust has shown that besides protecting the organism from disease-transmitting entities, it can also be seen as a form of warning other individuals within one's group to avoid organisms that are potentially dangerous (Tybur, Lieberman, Kurzban, & DeScioli, 2013). The expression of disgust or eeriness provides the organism, in this case, the homo

sapiens, with information on how to effectively deal with the opposing threat of other hominins, for example, the homo Neanderthalensis. This would enable them to make a collective decision of either engaging in a fight with the enemy or to decide to flee whereas the sole feeling of anger could lead to irrational decisions causing the death of the organism (Scarpa & Raine, 2000). For instance, a member of the group homo sapiens might make an irrational decision that is based solely on anger by rushing towards multiple enemies with the high likelihood of being killed or severely hurt. The action of warning others within the group could be regarded as being altruistic. Altruistic behaviour does not fare well with evolution as entities that are overly concerned for others are less likely to survive and/or reproduce leading to negative selection (Sesardic, 1995). The act of warning others in this framework, however, is not purely altruistic. The warning helps to gather the support of the other group members and would thus, serve the purpose of protecting the individual as its first priority.

The notion that it is safer to attack other communities or groups is strengthened by observation of primates that engage in between-community raiding as long as one's own group is superior in number and strength (Pandit, Pradhan, & Balashov, 2016). Under this line of reasoning, experiencing eeriness fulfilled the purpose of informing others by expressing the emotions of disgust that the opposing group or individual might be dangerous and consequently, helped to make a collective decision of either fighting or fleeing.

### Eeriness as a mechanism that prevents interbreeding

In the previous section, the possibility was addressed that the homo sapiens partially emerged from interbreeding between the various genera of homo (Ko, 2016). Evidence for these interbreeding events comes from studies that have shown that the DNA of neanderthalensis can be recognised to an extent in the Asian population (Vernot & Akey, 2014). On the other hand, the fact that the homo sapiens interbred with the other genera and that the resulting genetic effects are still present today does not infer that interbreeding between genera was favoured and increased by evolutionary mechanism. Rather it is the opposite. Cheetham, Pavlovic, Jordan, Suter, and Jancke (2013) have shown that there is a negative or purifying selection force working against the archaic DNA within the human genome. This is presumably the result of absent or reduced fertility in male hybrid offspring. The interbreeding between the two genera of Papio

and Manaca illustrates this problem with the majority of offspring of these crossings not surviving and the single one that survived being infertile (Laas, 2001).

Another explanation for the negative selection of archaic DNA focuses on the deleterious alleles and the smaller population sizes of other homo population such as Neanderthalensis. Neanderthalensis that possessed mildly deleterious alleles were not affected by selection due to the smaller population size in comparison to the modern human population size. As soon as these alleles were introgressed into humans, they were under the influence of negative selection (Cheetham, Pavlovic, Jordan, Suter, & Jancke, 2013). Despite the potential benefits of the homo sapiens to crossbreed with the Neanderthalensis enabling them to better respond to environmental conditions such as colder climate, the negative effects outweigh.

With this in mind, the UV effect can be seen as a mechanism that emerged through selective forces and resulted in humans avoiding interbreeding with other genera, which would have ultimately resulted in less reproductive success for the species as a whole. This means that this effect possibly originated in primates or other non-human animal species and was adapted to humans accordingly. The research into the development of a face prototype of Lewkowicz & Ghazanfar (2011) is important to consider here. The face prototype developed during infancy would allow the humans to differentiate (presumably on a subconscious level) between their own species and other genera and as a result, avoid crossbreeding.

### **Future research**

The present evidence that the UV effect is likely of evolutionary origin and can be found with biological faces, opens doors for new research. Regarding the first speculation that the UV effect had the purpose of protecting our ancestor from other hominins and helping them to make a collective decision, one could investigate how the UV effects unfold itself when multiple people as a group are confronted with human-like, biological faces. In case that they adopt each other's response, this could show that they try to find a mutual decision or response. However, one has to be careful with its interpretation since it could also simply be the result of social mechanism or phenomena such as the tendency to conform to the group. Another problem is that in research environments, the participants might try to repress their emotional expressions for various reason such as feeling embarrassed. The best solution would be to explore the group's reaction in a

'natural environment' without their awareness of being observed. It remains questionable whether this would be applicable concerning both its practicality and ethical considerations.

Proving that the uncanny valley serves as a buffer against interbreeding might be more challenging to investigate. One option could be to change the context of the situation by presenting the participants with a scenario or story of crossbreeding between species and compare the results with a control group. If the trough of the curve is deeper or more pronounced than that of the control group, then the (unconscious) reminder/story of interbreeding might have resulted in a greater protection mechanism of experiencing eeriness.

### References

Cheetham, M., Pavlovic, I., Jordan, N., Suter, P., & Jancke, L. (2013). Category processing and the human likeness dimension of the uncanny valley hypothesis: eye-tracking data. *Frontiers in psychology*, *4*, 108.

Dannemann, M., & Racimo, F. (2018). Something old, something borrowed: admixture and adaptation in human evolution. *Current opinion in genetics & development*, 53, 1-8.

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. Journal of Personality and Social Psychology, 67(3), 382–394. http://dx.doi.org/10.1037/0022-3514. 67.3.382.

Gong, Li. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. Computers in Human Behavior, 24, 1494-1509.

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, *125*(1), 125-130.

Haeske, A. B. (2016). The Uncanny Valley : Involvement of Fast and Slow Evaluation Systems. University of Twente.

Hanson, D. (2005, December). Expanding the aesthetic possibilities for humanoid robots. In *IEEE*-RAS international conference on humanoid robots (pp. 24-31).

Hare, B. (2017). Survival of the friendliest: Homo sapiens evolved via selection for prosociality. *Annual review of psychology*, *68*, 155-186.

Koopman, R. (2019). *The Uncanny Valley as a universal experience: a replication study using multilevel modelling* (Bachelor's thesis, University of Twente).

Ko, K. H. (2016). Hominin interbreeding and the evolution of human variation. *Journal of Biological Research-Thessaloniki*, 23(1), 1-9.

Koschate, M., Potter, R., Bremner, P., & Levine, M. (2016, March). Overcoming the uncanny valley: Displays of emotions reduce the uncanniness of human-like robots. In 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 359-366). IEEE.

Laas, B. C. (2001). " The Effects of Interbreeding on the Morphological and Genetic Attributes of Non-human Primates.

MacDorman, K. F. (2005, December). Mortality salience and the uncanny valley. In 5th IEEE-RAS International Conference on Humanoid Robots, 2005. (pp. 399-405). IEEE.

MacDorman, K. F., & Entezari, S. O. (2015). Individual differences predict sensitivity to the uncanny valley. Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 16, 141–172.

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

Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. Cognition, 146, 22–32. https://doi.org/10.1016/j.cognition.2015.09.008

Mitchell, W. J., Szerszen Sr, K. A., Lu, A. S., Schermerhorn, P. W., Scheutz, M., & MacDorman, K. F. (2011). A mismatch in the human realism of face and voice produces an uncanny valley. *i*-*Perception*, 2(1), 10-12.

Moll, B., & Schmettow, M. (2015). Investigating the origins of the uncanny valley : The effect of presentation time on ratings of uncanniness.

Mori, M. (1970). Bukimi no tani [the uncanny valley]. Energy, 7, 33-35.

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

Moosa, M. M., & Ud-Dean, S. M. (2010). Danger avoidance: An evolutionary explanation of uncanny valley. *Biological Theory*, 5(1), 12-14

Oum, R. E., Lieberman, D., & Aylward, A. (2011). A feel for disgust: tactile cues to pathogen presence. *Cognition and Emotion*, *25*(4), 717-725.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Woods, S., Hall, L., & Zoll, C. (2005). Learning by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence*, *19*(3-4), 235-266.

Pandit, S. A., Pradhan, G. R., Balashov, H., & Van Schaik, C. P. (2016). The conditions favoring between-community raiding in chimpanzees, bonobos, and human foragers. *Human Nature*, *27*(2), 141-159.

Perea-García, J. O., Kret, M. E., Monteiro, A., & Hobaiter, C. (2019). Scleral pigmentation leads to conspicuous, not cryptic, eye morphology in chimpanzees. *Proceedings of the National Academy of Sciences*, *116*(39), 19248-19250.

Scarpa, A., & Raine, A. (2000). Violence associated with anger and impulsivity. *The neuropsychology of emotion*, 320-339.

Schmettow, M. (2021). *New statistics for design researchers*. Springer International Publishing.https://doi.org/10.1007/978-3-030-46380-9

Sesardic, N. (1995). Recent work on human altruism and evolution. *Ethics*, 106(1), 128-157.

Siebert, R., Taubert, N., Spadacenta, S., Dicke, P. W., Giese, M. A., & Thier, P. (2020). A naturalistic dynamic monkey head avatar elicits species-typical reactions and overcomes the uncanny valley. *Eneuro*, 7(4).

Slijkhuis, P. J. H., & Schmettow, M. (2017). The Uncanny Valley Phenomenon A replication with short exposure times. University of Twente.

Schmettow, M. and Borsci, S., 2020. *RPubs - Designometrics is a cube, stupid!*. [online] Rpubs.com. Available at: <a href="https://www.rpubs.com/schmettow/Designometrics\_1">https://www.rpubs.com/schmettow/Designometrics\_1</a> [Accessed 22 May 2021].

Steckenfinger, S. A., & Ghazanfar, A. A. (2009). Monkey visual behavior falls into the uncanny valley. *Proceedings of the National Academy of Sciences*, *106*(43), 18362-18366.

Strait, M. K., Floerke, V. A., Ju, W., Maddox, K., Remedios, J. D., Jung, M. F., & Urry, H. L. (2017). Understanding the uncanny: both atypical features and category ambiguity provoke aversion toward human-like robots. *Frontiers in psychology*, *8*, 1366.

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.

Tung, F. W., & Chang, T. Y. (2013, July). Exploring children's attitudes towards static and moving humanoid robots. In *International Conference on Human-Computer Interaction* (pp. 237-245). Springer, Berlin, Heidelberg.

Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2013). Disgust: evolved function and structure. *Psychological review*, *120*(1), 65.

Voland, E., & Grammer, K. (2003). Evolutionary aesthetics. Springer Science & Business Media.

Vernot, B., & Akey, J. M. (2014). Resurrecting surviving Neandertal lineages from modern human genomes. *Science*, *343*(6174), 1017-1021.

Zhang, J., Li, S., Zhang, J. Y., Du, F., Qi, Y., & Liu, X. (2020, July). A Literature Review of the Research on the Uncanny Valley. In *International Conference on Human-Computer Interaction* (pp. 255-268). Springer, Cham.

### **Appendix A: Consent Form**

Information for participation in the research study on humans' emotional response towards different faces

You are being invited to participate in a research study conducted as part of a bachelor thesis within the Faculty of Behavioural, Management and Social Sciences at the University of Twente. The aim of this research is to assess humans' emotional response towards different faces. Completing the study will take approximately 30 minutes and involves viewing images and answering survey questions.

### Risks

To our understanding, there are no risks associated with this research study, however, some images might be sensitive or provoke personal discomfort. Given your participation in this study is entirely voluntary, you are free to withdraw at any time without providing any reasons or experiencing any disadvantages.

### Handling of data and confidentiality

To the best of our ability, your answers in this study will remain confidential. We will minimise any risks by assuring that your participation in this study is anonymous, as no information will be collected that would allow personal identification. All data will be stored securely and will not be shared with anyone outside the research team. Any comments you make may be quoted in an anonymised form in the research papers resulting from this study.

### Contact information

If you have any questions or concerns about this survey or wish for your data to be deleted you can contact the researchers via email:

Lara Geue (l.geue@student.utwente.nl), Milan Bischoff (m.bischoff@student.utwente.nl), Marcel Pertenbreiter (m.pertenbreiter@student.utwente.nl), Jana Westermann (j.m.l.westermann@student.utwente.nl).

If you have questions about your rights as a research participant or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher, please contact the secretary of the Ethics Committee of the Faculty of Behavioural, Management and Social Sciences of the University of Twente (ethicscomittee-bms@utwente.nl).

### Consent

I confirm that I have read and understood the information provided above and consent voluntarily to be a participant in this study. I further confirm that I possess a sufficient level of proficiency in English and that I am at least 18 years old.

### I consent.

■ I do not consent to take part in this study.

# **Appendix B: Scale Descriptions and Translations**

Translations of the eeriness scales									
English	Dutch	German							
Uninspiring - Spine-tingling	Oninteressant - Opwindend	Uninteressant - Aufregend							
Boring - Shocking	Saai - Schokkend	Langweilig - Schockierend							
Predictable - Thrilling	Vorspelbaar - Spannend	Vorhersehbar - Spannend							
Bland - Uncanny	Flauw - Verontrustend	Fade - Beunruhigend							
Unemotional - Hair-raising	Emotieloos - Doodeng	Emotionslos -Haarsträubend							

Likability scale translations		
English	Dutch	German
Less friendly, more unpleasant, creepy	Minder vriendelijk, onaangenaamer, griezelig	Weniger freundlich, unangenehmer, gruselig
More friendly and pleasant, less creepy	Vriendelijker, aangenaamer, minder griezelig	Freundlicher, angenehmer, weniger gruselig

**Appendix C: Data analysis protocol** 

### Experiment

Stimuli

```
Stimuli <-
    readxl::read_excel("Data/BA21/Stimuli.xlsx") %>%
    mutate(Stimulus = str_c("S", as.character(Stimulus))) %>%
    mutate(Set = factor(Set, 1:4, labels = c("Primate", "Human", "Tri23", "Robo
t"))) %>%
    mutate(humLike = rowMeans(select(., starts_with("H_like")), na.rm = T)) %>%
    mutate(valence = rowMeans(select(., starts_with("E_valence")), na.rm = T))
```

Stimuli %>%

sample\_n(20)

Sti		H_li	H_li	H_li	H_li					E_v	E_v	E_v	E_v	Ancest	Et	hu	va
m		ken	ken	ken	ken	Е	Е	Е	Е	ale	ale	ale	ale	oralCl	hn	m	le
ul	Se	ess_	ess_	ess_	ess_	_	_	_	_	nce	nce	nce	nce	osene	ici	Li	nc
us	t	1	2	3	4	1	2	3	4	_1	_2	_3	_4	SS	ty	ke	e
S7 8	Pr im at e	55	65	65	55	1	4	4	1	65	80	40	70	4	0	60 .0 0	6 3. 7 5
S2 5	Pr im at e	97	96	93	95	1	1	1	1	37	60	20	25	1	0	95 .2 5	3 5. 5 0
S7 2	Pr im at e	82	75	65	70	0	0	0	0	-10	0	0	0	3	0	73 .0 0	- 2. 5 0
S4 0	Pr im at e	18	50	45	50	4	4	4	4	10	-20	-20	5	5	0	40 .7 5	- 6. 2 5
S5 7	Pr im at e	40	50	55	50	0	0	0	0	0	0	0	5	4	0	48 .7 5	1. 2 5
S8 4	Pr im at e	97	90	94	90	5	5	5	4	80	- 100	-95	-60	2	0	92 .7 5	- 4 3. 7 5

S3 8	R ob ot	37	NA	NA	NA	4	4	4	4	30	0	-20	-15	10	0	37 .0 0	- 1. 2 5
S5 4	Pr im at e	2	10	15	5	5	0	0	0	-26	0	0	0	9	0	8. 00	- 6. 5 0
S9 8	H u m an	100	100	100	100	4	4	2	4	85	0	-95	-70	0	1	10 0. 00	- 2 0. 0 0
S4 2	Pr im at e	94	90	90	85	1	1	1	4	33	50	30	5	2	0	89 .7 5	2 9. 5 0
S8 5	Pr im at e	91	90	95	95	4	4	2	2	-25	-30	-90	-50	2	0	92 .7 5	- 4 8. 7 5
S4 9	Tr i2 3	100	100	98	97	1	1	1	1	60	70	50	10	0	0	98 .7 5	4 7. 5 0
S5 3	Pr im at e	83	90	89	80	0	1	1	1	13	60	40	5	2	0	85 .5 0	2 9. 5 0
S7 6	Pr im at e	80	80	82	80	0	5	0	0	0	-50	0	0	3	0	80 .5 0	- 1 2. 5 0
S6 1	Pr im at e	1	20	15	3	2	2	0	0	-20	-10	0	0	8	0	9. 75	- 7. 5 0
S9 3	H u m an	100	100	100	100	5	6	6	6	-85	-60	-90	-70	0	1	10 0. 00	- 7 6. 2 5
S1 2	Pr im	24	45	25	25	4	4	2	4	70	-20	-80	-70	7	0	29 .7 5	- 2 5.

	at e																0 0
S6 9	Pr im at e	28	45	35	30	4	4	4	4	55	20	10	25	7	0	34 .5 0	2 7. 5 0
S4 3	Pr im at e	45	50	60	40	0	3	0	0	-10	-50	0	0	4	0	48 .7 5	- 1 5. 0 0
S3 7	Pr im at e	70	85	80	70	6	6	1	1	-30	-20	90	20	3	0	76 .2 5	1 5. 0 0
Stim se co	uli % lect( rrr::	\$>% start corre	:s_wit elate(	:h("H_ )	like	"),	Ar	nce	sto	ralCl	osene	ss) %	>%				

```
##
```

```
## Correlation method: 'pearson'
```

## Missing treated using: 'pairwise.complete.obs'

	H_likeness_	H_likeness_	H_likeness_	H_likeness_	AncestoralClosene				
term	1	2	3	4	SS				
H_likeness_1	NA	0.9554019	0.9436330	0.9328772	-0.7707424				
H_likeness_2	0.9554019	NA	0.9303407	0.9426051	-0.9422645				
H_likeness_3	0.9436330	0.9303407	NA	0.9438880	-0.9379207				
H_likeness_4	0.9328772	0.9426051	0.9438880	NA	-0.9268578				
AncestoralClosene ss	-0.7707424	-0.9422645	-0.9379207	-0.9268578	NA				
<pre>ss Stimuli %&gt;% select(starts_with("H_like"), AncestoralCloseness) %&gt;% corrr::correlate() %&gt;% corrr::network_plot()</pre>									

##

## Correlation method: 'pearson'

## Missing treated using: 'pairwise.complete.obs'



### Reading data

```
response = mascutils::rescale unit(as.numeric(response)),
         response = mascutils::rescale centered(response, scale = .999),
         humLike = mascutils::rescale_unit(humLike),
         #Stimulus = str_extract(Trial, "^S\\d{2,3}_\\d"),
                 = str_extract(Trial, ),
         #Item
         #humLike = as.numeric(str_extract(Stimulus, "\\d{1,3}")),
         humLike 2 = humLike^2,
         humLike_3 = humLike^3) %>%
  bayr::as_tbl_obs()
BA21
BA21 prim <- BA21 %>%
  filter(Set == "Primate")
BA21 %>%
  group_by(Part) %>%
  summarize(N = n())
BA21 %>%
  group_by(Stimulus) %>%
  summarize(N = n())
BA21_agg <-
  BA21 prim %>%
  filter(Scale == "Eeriness") %>%
  group_by(Stimulus, Set, valence, humLike) %>%
  summarize(mean resp = mean(response, na.rm = T)) %>%
  ungroup() %>%
  as_tbl_obs()
```

```
BA21_agg
```

### Visualisation

### **Responses averaged over stimuli**



## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'



### **Population-level model**

Four polynomial models are estimated, from cubic to grand mean. Valence of facial expression is included as a control (except M 5).

```
clu(M_poly_3)
```

parameter	fixef	center	lower	upper
Intercept	Intercept	0.5760378	0.5110060	0.6426081
poly(hum_like, 3)1	poly(hum_like, 3)1	0.5122479	0.0501577	0.9618364
poly(hum_like, 3)2	poly(hum_like, 3)2	0.2039400	-0.2525294	0.6671140
poly(hum_like, 3)3	poly(hum_like, 3)3	-0.0764050	-0.4945528	0.3797763





**Model selection** 

Approximate relative predictive accuracy for all four *population-level* models using leave-one-out approximation

```
Loo_poly <-
   list(loo(M_poly_0),
        loo(M_poly_1),
        loo(M_poly_2),
        loo(M_poly_3))
compare_IC(Loo_poly)</pre>
```

Model ranking by predictive accuracy

Model	IC	Estimate	SE	diff_IC
M_poly_1	looic	-4.559410	9.863243	0.000000
M_poly_2	looic	-3.651307	9.411036	0.908103
M_poly_3	looic	-1.364203	9.659988	3.195206
M_poly_0	looic	-1.119816	9.272131	3.439594

On population level, the cubic model is not the preferred one.

Using the MCMC method to measure the probability of an UV curve on population level:

cat("The probability of the population level being a UV curve is: ", mean(P\_w ide\$is\_uncanny))

## The probability of the population level being a UV curve is: 0.72475

Multi-level model

```
PP_poly_4 <- post_pred(M_poly_4, thin = 5)</pre>
```

fixef\_ml(M\_poly\_4)

Population-level coefficients with random effects standard deviations

)5
34
22
)3
<u>.</u>
· · · · · · · · · · · · · · · · · · ·

```
## `geom_smooth()` using method = 'loess' and formula 'y \sim x'
```





**Distributional Beta regression** 

```
fixef_ml(M_poly_5)
```

Population-level coefficients with random effects standard deviations

fixef	center	lower	upper	SD_Part	
Intercept	0.2423457	0.1349185	0.3537239	0.1725790	NA
NA	1.2259843	1.0541065	1.4012498	NA	0.5545595
valence	0.0035227	0.0026571	0.0044081	0.0012479	NA
humLike	2.1916328	1.4888111	2.8686954	0.1213779	NA
humLike_2	-5.0889238	-6.6040510	-3.5435657	0.1514233	NA
humLike_3	3.3125477	2.3226758	4.2920712	0.1318868	NA
<pre>PP_poly_5 %&gt;%   predict() %&gt;%   left_join(BA21, by = "Obs") %&gt;%   ggplot(aes(x = humLike)) +   geom_smooth(aes(y = center, color = "cubic model",</pre>					
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'					





write\_csv(T\_predicted, file = "T\_predicted.csv")

## Universality

```
P univ uncanny <-
  posterior(M_poly_5) %>%
  #as_tibble() %>%
  dplyr::filter(is.na(nonlin), fixef != "valence") %>%
  re scores() %>%
  select(iter, Part = re_entity, fixef, value) %>%
  tidyr::spread(key = "fixef", value = "value") %>%
  select(iter, Part,
         humLike_0 = Intercept,
         humLike 1 = humLike, humLike 2, humLike 3) %>%
  mutate(
    trough = trough(select(., humLike_0:humLike_3)),
    shoulder = shoulder(select(., humLike_0:humLike_3)),
    has_trough = !is.na(trough),
    has_shoulder = !is.na(shoulder),
    shoulder left = trough > shoulder,
    is uncanny = has trough & has shoulder & shoulder left
  )
P_univ_uncanny %>%
  select(Part:shoulder) %>%
  pivot longer(humLike 0:shoulder,
               names_to = "parameter",
               values to = "value") %>%
  group_by(Part, parameter) %>%
  summarize(center = median(value, na.rm = T)) %>%
  pivot_wider(names_from = "parameter",
              values from = "center") %>%
  write_csv(file = "Participant_level.csv")
```

Individual positions of shoulder and trough

```
P_univ_uncanny %>%
ggplot(aes(x = Part)) +
geom_violin(aes(y = trough, color = "Trough")) +
geom_violin(aes(y = shoulder, color = "Shoulder")) +
# theme(axis.text.x = element_text(angle = 90)) +
coord_flip()
```







save(BA21, BA21\_agg, PP\_poly\_4, M\_poly\_4, P\_univ\_uncanny, M\_poly\_3, PP\_poly\_3, M\_poly\_2, M\_poly\_1, M\_poly\_0, Loo\_poly, M\_poly\_5, PP\_poly\_5, P\_univ\_uncanny, M\_6, M\_7, file = "BA21.Rda")