The Influence of the Eyes in the Occurrence of the Uncanny Valley Effect

Isabelle M. M. Stikker

Bachelor Thesis

Faculty of Behavioural, Management and Social Sciences,

University of Twente

Human Factors and Engineering Psychology

Supervisor: Dr. Martin Schmettow

Second Supervisor: Dr. Funda Yildirim

June 28, 2023

Abstract

The Uncanny Valley Effect refers to the eerie feeling people experience when encountering highly humanlike entities, such as humanlike robots or digital characters in movies and games. Since the effect has been suggested to have an evolutionary origin, this study will focus on primate faces, to gain further knowledge in the influence of the eyes in the occurrence of the Uncanny Valley effect. The aim is to explore whether incongruence within a face (humanlike eyes combined with a non-human face) is at the root of the phenomenon.

Participants ($N = 85$) filled out a questionnaire and rated primate stimuli on likability and human likeness. The stimuli were grouped into four categories: animals (congruent: animal face with animal eyes), uncanny animals (incongruent: animal face with humanlike eyes), humans (human face with human eyes) and none (inverted faces). Additionally, the upright stimuli were presented with and without the addition of sunglasses covering the primate's eyes. The results showed that the incongruent animal faces were scored lowest on likability. Additionally, covering the eyes with sunglasses increased likability estimates in both animal conditions, but not for humans. Lastly, inverted animal stimuli were rated lower on likability as well as on human likeness. The findings contribute to knowledge in the field of visual face processing mechanisms regarding the Uncanny Valley effect. It is concluded that incongruence within a face is not the sole cause of the Uncanny Valley effect to occur, and it is suggested that people base their likability evaluations on other facial elements in the absence of eyes.

Key words: uncanny valley, configural face processing, eye processing, primates

Contents

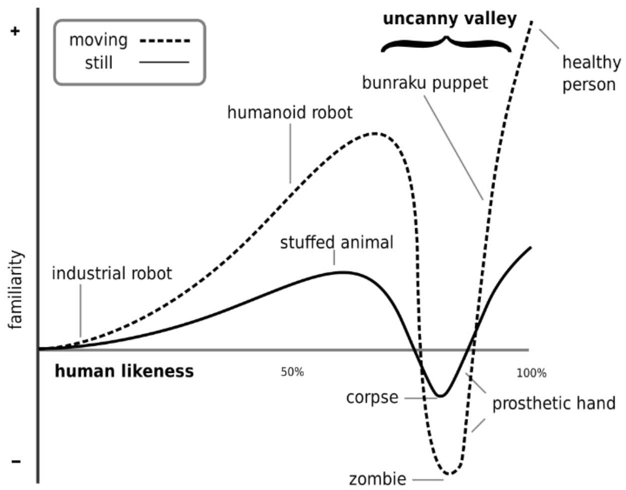
Introduction

It is expected that the presence of robots will increase in the upcoming years, and that they will be able to perform the role of a social companion in environments such as elderly homes. To be able to properly fulfil this function, it is important that humans feel at ease with the social robot. Therefore, research into the design of social robots and their facial features is crucial.

Mori (1970) hypothesised that the more robots resemble humans, the more feelings of likability and trustworthiness we tend to attribute to the robot. But, that this curve would drop at a certain limit when the robot becomes more human-like and that after this point people would show a negative emotional response. When the resemblance is very close to a real human, the curve would go up again. Mori (1970) referred to this phenomenon as the Uncanny Valley effect (UV) and illustrated the effect with the graph depicted in Figure 1.

Figure 1

Curve of the Uncanny Valley Effect as Explained by Mori (1970)



Since Mori's hypothesis several researchers have investigated whether the Uncanny Valley effect actually exists. Mathur and Reichling (2016) along with other researchers (e.g. Mac Dorman et al., 2009; Vaitonyte et al., 2022) have investigated this topic with a range of pictures of humans, androids and robots and have shown that the phenomenon is indeed real. Questions arise what causes this phenomenon and how the robotic design should be adapted to prevent people from experiencing feelings of unease towards the social robot.

The topic is not only relevant in the context of social robots, but also for the design of digital characters for movies and games. In the movie and game industry, the Uncanny Valley effect can make digital characters or animated creatures appear creepy or disturbing to audiences. This can be a significant problem for filmmakers and game developers when this is not intended, as the UV can decrease the enjoyment of watching a movie or playing a game. As the animation and game industries are multi-billion-dollar industries this can have disastrous consequences for their profits. An example of a movie in which the characters were experienced to fall in the uncanny valley is the 2004 animated film, "The Polar Express" (Geller, 2008). The movie used motion capture technology to create realistic human characters, but many viewers found the characters to be creepy and unsettling due to their slightly off facial expressions and movements. Due to the uncanny valley effect people expressed that they did not feel an emotional connection with the characters. On the contrary, knowledge of mechanisms behind the uncanny valley effect can also be used to purposely create eerie characters when designing evil characters that play the role of villains or that should otherwise be reacted to with discomfort.

The aim of this research is to deepen our knowledge into the mechanisms that play a role in causing the uncanny valley effect to occur. This will build to our understanding of visual processing mechanisms and can contribute to the field of robot and character design.

Universal Experience

Koopman and Schmettow (2019) performed a replication study of Mathur and Reichling's (2016) experiment while adding multilevel modelling to investigate whether all participants show the characteristic Uncanny Valley. This would indicate that we are dealing with a universal experience and that the UV is not a result of individual differences. The findings showed a UV for all participants, indicating that the UV is indeed a universal experience.

If the Uncanny Valley phenomenon is indeed a universal experience, this would likely mean that there is an innate cognitive mechanism formed by evolution at play which causes everyone to experience negative feelings when robots become too human-like. This idea prompted Geue and Schmettow (2021) to delve deeper into the evolutionary mechanisms behind the UV by investigating whether the UV would also appear when using primate faces instead of robot-like faces. The results indicated that indeed, a UV was present in almost all participants. Furthermore, they noticed that the stimuli of the primates and robot faces that fell into the uncanny valley area almost all had a clear white sclera (the white in the eyes), indicating that the eyes and especially the presence of a white sclera play a major role in the occurrence of a UV. This poses the idea that the way we process faces is crucial to our understanding of the Uncanny Valley effect.

Facial Processing Theories

Configural Processing

We process faces with a system called configural processing (Kanwisher and Moscovitch, 2000). According to the configural face processing model, face recognition involves the processing of spatial relationships between facial features (configurations), rather than the features (such as eyes, nose and mouth) separately. In other words, facial features are processed concurrently rather than independently.

Maurer et al. (2002) proposed that configural face processing consists of three different processes: “sensitivity to first-order relations”, “sensitivity to second-order relations” and “holistic processing”. Sensitivity to first-order relations refers to the ability to detect and process the typical features that make a face, such as the size and oval/round shape of a face and the characteristic individual facial features: two eyes above a nose above a mouth. Therefore, other researchers (e.g. McKone, 2010; Wang, 2019) speak of *part-based* processing when referring to first-order processing, since it concerns the individual components that make a face. Because there is a system that is alert to features that are prototypical for a face, it often results in false positives. This is called pareidolia and examples can be found everywhere in objects in daily life, for example when perceiving faces in the headlights of a car or in a sink as in Figure 2.

Figure 2

Pareidolia in a Sink

Image credits: adme.ru

Sensitivity to second-order relations involves the processing of the spatial relationships between facial features. This includes analysing the distance, angle, and orientation between different parts of the face, such as the distance between the eyes or the angle between the nose and mouth. Thus, first-order relations enable us to detect faces, while sensitivity to second-order relations is thought to be important in discriminating between different faces, particularly when features are similar. Lastly, holistic processing makes us perceive the face as a whole rather than as a collection of individual parts. The different parts are thought to be 'glued' together into a gestalt. However, McKone (2010) questioned the distinction between holistic processing and the processing of second-order relations as two different processes, and proposed that they should rather be considered as one: holistic/configural processing.

Support for the configural face processing model can be found in the observation that inversion of faces impairs facial recognition (Yin, 1969). In his research Yin showed how inverting images of faces as well as objects impairs recognition of both types of presented stimuli, but that faces were disproportionately affected in this condition. Turning a face upside down affects how we perceive the spatial arrangement of the features of a face. Certainly, the individual features of the face have remained the same as before inversion. In Figure 3 the painting The Gardener (Italian - *L'ortolano*) by Giuseppe Arcimboldo is depicted. Few people would see a face in this bowl of vegetables, however, when the image is viewed upside down (Figure 4), pareidolia would occur, due to the configural face processing mechanism.

Figure 3

The Gardener

Figure 4

The Gardener, Reversed

Further evidence for the configural face processing model can be found in the Thatcher illusion. When in a face only the eyes and mouth are inverted, this elicits strong feelings of discomfort, as the second-order relations are altered. However, when the whole face is inverted, these feelings disappear. This remarkable phenomenon is called the Thatcher illusion (see Figure 5) and shows that when viewing a face upside down our sensitivity to second order relations is affected and feature analysis is used instead.

Figure 5

Thatcher Illusion

There is some debate about the temporal order of first and second order processing. Behavioural and neural experiments have indicated that the mechanisms behind facial recognition are incredibly fast. Research using event-related potentials (ERPs) and fMRI studies have shown that there is a neural reaction (an event-related negative potential) in the fusiform gyrus, the area that is involved in face processing already at 170ms after the presentation of a face stimulus (Maurer et al., 2002). This reaction (called the N170) is associated with first-order processing: recognising a face as a face. Based on their experiment Wang (2019) hypothesises that second-order processing occurs slightly after first-order processing, indicating a temporal order. However, other sources (e.g. Richler et al., 2009) indicate that the two components occur simultaneously and in parallel.

Eye Processing

In any case, eye tracking studies have shown that people quickly fixate on the eyes when they encounter a face (Bagepally, 2015). The eyes play an important role in communication and face processing (Kobayashi and Kohshima, 1997; Seyama and Nagayama, 2007; MacDorman et al., 2009; Lewkowicz and Ghazanfar, 2011). Human eyes are considered to be special since the colourful iris largely contrasts with the white surrounding it. This is thought to have the functions of being able to follow someone's gaze direction (important for communication) as well as increasing human's visual field by enabling larger movement of the eyes (Kobayashi and Kohshima, 1997).

The sensitivity to the eyes can already be shown in infants of 12 months of age (Lewkowicz and Ghazanfar, 2011). In their research into the developmental origin of the Uncanny Valley effect, Lewkowicz and Ghazanfar measured visual preference for different faces in different age groups of infants (6, 8, 10 and 12 months old). In their experimental setup, they used human faces and avatar faces, and for a third category scaled up the size of the avatars' eyes to 150% and called this category unrealistic avatars. Previously, Seyama and Nagayama (2007) and MacDorman et al. (2009) had shown that this manipulation would lead to the Uncanny Valley effect in adults. Lewkowicz and Ghazanfar found that the Uncanny Valley effect appeared for the unrealistic avatar faces (with increased eyes) at the age of 12 months old, but not yet at 6- 8- and 10-month-olds. This suggests that humans have innate neural structures within the brain that are responsible for face processing. However, this innate structure needs to develop during a sensitive period in which infants are exposed to a range of faces through which they acquire a face prototype and notice that something is off when encountering a face with increased eyes.

Human Likeness and Configural Processing

The Uncanny Valley effect has been found to depend on the extent to which the object resembles a human, expressed as its human likeness (e.g., Kim et al., 2020; MacDorman et al., 2009). Often in research into the UV, this factor is evaluated by observers and expressed as a number on the human-mechano (huMech) scale. Generally, a human likeness score of roughly 70%-80% elicits the strongest feelings of eeriness (Mathur and Reichling, 2016; Koopman and Schmettow, 2019).

According to Green et al. (2008), human likeness of a face increases people's sensitivity to facial proportions. If, upon closer inspection an encountered face does not appear to be human, this could elicit feelings of eeriness. Ho et al. (2008) noted that a robot is most unsettling when human features raise the expectation that we are dealing with a human, while nonhuman features cannot confirm these expectations. This might explain the research findings of Geue and Schmettow (2021) since almost all stimuli that were found in the uncanny valley had visible white sclera. The reasoning might be that upon perceiving a primate that is close in ancestral closeness and thus has more similarities with a human, first-order processing information indicates that a highly human entity is present. However, other information concerning second-order facial relations might establish that the face is in fact non-human. When we then fixate on the eyes the presence of a clear white sclera could result in a mismatch of what is expected, since having visible white sclera has long been considered to be uniquely human (Clark et al., 2023). This is known as the perceptual mismatch theory (Diel and MacDorman, 2021).

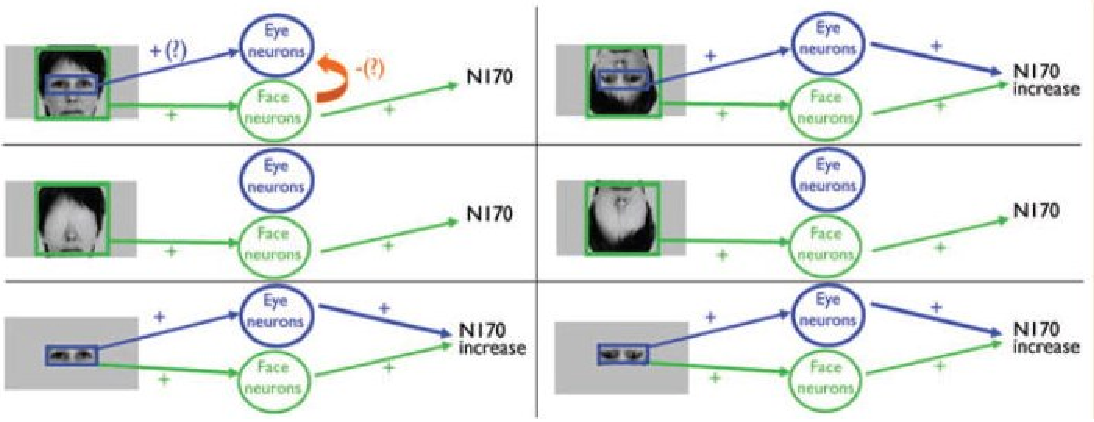
Specific Category Confusion theory

Research on the role of the eyes in face processing with event related potentials might be able to support the beforementioned theory. The aforementioned N170 response which elicits in the brain upon perceiving a face had been measured in participants through different conditions. First of all, as multiple researchers have established, viewing an upright face elicits the expected N170 response (Nemrodov & Itier, 2011). However, when the eyes are removed from the picture of a face, the N170 still appears. Furthermore, presenting solely the eyes, in an isolated condition, results in an increased amplitude for the N170 response. Similar findings (increased N170 response) are found when an inverted face is presented, or when the eyes are inverted while presented in isolated condition. Lastly, inverting the eyeless face results in a normal N170 response.

Itier and Batty (2009) developed a neural model to explain why these differences might occur. According to their hypothesis, there are face-sensitive neurons and eye-sensitive neurons present in the brain. Face-sensitive neurons respond to faces, as well as isolated eyes, while eye-sensitive neurons respond only to eyes. When both types fire in reaction to perceiving a face with visible eyes, this would lead to a higher amplitude of the N170, as compared to when only face-sensitive neurons fire, or only eye-sensitive neurons. They further hypothesize that the occurrence of a N170 response to a normal upright face (and not an increased N170 amplitude) possibly means that there is an inhibition mechanism at play. This mechanism would make face-sensitive neurons inhibit the eye-sensitive neurons leading to the N170 response. The neural model is visualized in Figure 6.

Figure 6

Neural Model of Face Processing Adopted from Itier and Batty (2009)



These research findings could support the beforementioned findings. First of all, when encountering an image of a primate face, face neurons would respond while eye neurons are inhibited. Afterwards, when this information as well as the information from second-order relations are processed, a conclusion could be that one is facing an animal face. Directly after, when fixation on the eyes occurs, eye neurons might react as if one would be facing human eyes, due to the highly human-like characteristic of eyes with visible white sclera. This incongruence between information might be at the root of the experienced feelings of eeriness. As far as the researcher is aware, this theory is not yet investigated in the context of the Uncanny Valley effect and the proposed theory will therefore be referred to as the Specific Category Confusion theory.

Secondly, when inverting a primate face both eye-sensitive neurons and face-sensitive neurons would respond according to the neural model. Additionally, since inverted faces are shown to be further processed by feature analysis, this could explain why no feelings of eeriness appear for inverted faces, as there is no case of Specific Category Confusion.

Lastly, when there are no eyes present in a face, or when the eyes cannot be observed by an observant (for example when someone is wearing sunglasses) this would lead to the normal N170 response and would not lead to feelings of eeriness, since there is again no expectation of Specific Category Confusion.

The Current Study

As research by Geue and Schmettow (2021) had shown, the stimuli of primates that were found to be in the UV almost all had visible white sclera, but a non-human skull. The literature has shown that the eyes play a significant role in the UV to occur, when the eyes are manipulated to look abnormal. However, in the research done by Geue and Schmettow (2021) the eyes were not manipulated and the UV still occurred in primate faces as well as robotic faces. The question arises whether the eyes were the determining factor in these findings. Therefore, this research will explore whether the UV will also occur when the eyes cannot be evaluated by the observer. Additionally, this research will investigate whether inversion affects people's perception of human-likeness and likability towards primate faces. This research will add to our understanding of what facial features are at the root of the UV, to deepen our knowledge in the field of face processing as well as the uncanny valley effect.

There are three independent variables and two dependent variables. The type of stimulus (trigger category), the presence of the eyes and inversion will be the independent variables, while the evaluated likability and human-likeness will be the dependent variables. With type of stimulus is meant if the stimulus is a human, a non-human primate with non-human eyes, or a non-human primate with humanlike eyes (visible white sclera).

The hypotheses that will be examined are the following:

H1: The eeriness response (low likability) is expected to be most extreme when human-like eyes (visible white sclera) are combined with a non-human face (monkey/ape face), due to the Specific Category Confusion theory.

H2: As sunglasses will prevent the mismatch between the detected facial properties and eyes, the reported likability will be higher when the primate eyes are not visible (i.e., covered by sunglasses) than when they are visible for the incongruent faces (ape face with human eyes).

H3: For the inverted *congruent* images, the likability is expected to be scored lower than for the upright congruent faces, since inversion impairs configural processing. Conversely, likability for the inverted *incongruent* faces is expected to be scored higher than for the upright incongruent faces, since the uncomfortable eeriness response will not be experienced when no mismatch occurs for inverted stimuli.

H4: Human-likeness is expected to be scored lower for the inverted images, as compared to the stimuli in both other conditions, since configural processing will be impaired.

The four different factors can be described and ordered on predicted likability as follows:

1. (highest likability) Human face (Human): described as a stimulus with a human face which is not inverted and has human eyes or wears sunglasses.
2. Animal face (Animal): described as a stimulus with a non-human face, which is not inverted and does not have human eyes, but can have sunglasses.
3. Inverted face (none): described as an inverted stimulus.
4. (lowest likability) Incongruent face (Uncanny animal): described as a stimulus with a non-human face and human eyes, which is not inverted.

Methods

Procedure

Participants were able to access the online questionnaire by clicking on a weblink that lead them to Qualtrics. Hereafter, they were welcomed with an information screen that informed the participants on the content and purpose of the study (see Appendix A). They were told the goal of the questionnaire was to gain insight in emotional reactions towards different faces. The decision to not directly fully inform the participants that the study researches the uncanny valley effect was made to prevent possible bias. After the information screen the participants were asked to give informed consent to participate in the study. This study was approved by the Ethics Committee of the Faculty of Behavioural and Management and Social Sciences at the University of Twente. If the participant had given informed consent, they were asked about their gender and age followed by a short instruction indicating that they were going to evaluate 100 faces by finishing the sentence "To me, this face seems…" and "How human-like does this face appear to you?". The instructions can be found in Appendix B.

They were then presented with a randomised total of 100 stimuli of biological faces of which some were presented normally, others upside down and others were presented with the addition of sunglasses that covered the eyes of the face. There was a total of 132 stimuli, but the decision to not show all stimuli to the participants was made to prevent people from not finishing the survey when they found it was taking too long. In addition, since this research is partly a replication study of the study performed by Geue and Schmettow (2021) who used 100 stimuli, the decision was made to scale the original 132 stimuli down to a total of 100.

Each facial stimulus was followed by two questions regarding the perceived likability of the stimulus and the perceived human likeness. The questions could be answered using a slider. The survey ended with a debriefing screen in which the full purpose and aim of the study was explained and where the participant was thanked for their participation (see Appendix C). They were also asked to give their final consent for their data to be used for the analysis and were given the option to opt out of the study.

Measures

Each picture was followed by two questions presented in random order regarding the likability and human-likeness of the stimuli. The questions could be answered on a visual analogue scale that enabled the participants to use a slider to indicate how they felt about the images. For likability the task was to finish the sentence "To me, this face seems…" with on the one hand (-100) the description *less friendly, more unpleasant, creepy* and on the other hand (+100) the description *more friendly, more pleasant, less creepy* (derived from the one-item likeability scale by Mathur and Reichling (2016)). Using the slider, participants were able to indicate on a scale of -100 to +100 how they felt about the presented picture. Regarding the human likeness the question "How human-like does this face appear to you?" was asked, with 0 indicating the face did not resemble a human at all, and 100 meaning that the presented face was a human.

Design

This study used a within subject design with independent variables *trigger category, visibility of the eyes* and *inversion* and dependent variables *likability* and *human likeness.* All subjects were presented to all three types of stimuli: upright images with visible eyes, upright images with eyes that are covered by sunglasses and inverted images with visible eyes. These consisted of all three trigger categories (human, animal, uncanny animal).

Data Analysis

For the data analysis the program Rstudio (version 4.3.0) was used. The R script used for analysis can be found in Appendix D.

To prepare for the data analysis, all stimuli were categorised based on their face/skull type, which could be either human or non-human. The same was done for the eyes of the stimuli which could also be categorised as human (when visible white sclera was present) or non-human (when there was no white in the eyes). Then, the stimuli were categorised on the variable visibility of the eyes by indicating if the face is wearing sunglasses (True or False) and if they were presented inverted (True or False). Based on these variables, four trigger categories were created: Human (human face + human eyes), Animal (non-human face + non-human eyes), Uncanny Animal (non-human face + human eyes) and none (inverted stimuli).

Then, a factorial model with these four levels was employed, using a Bayesian generalized linear model with dependent variable likability to test if the four groups were rated differently on likability. The same was done for the dependent variable human likeness.

Afterwards, to test the second hypothesis another Bayesian generalized linear model was constructed. The dependent variable likability was regressed on three factors: human eyes (true or false), human skull (true or false) and sunglasses (true or false) and consisted of the interaction term of the three factors.

The third hypothesis was tested by creating a similar model, but with the variable inverted instead of sunglasses. Then, the fourth hypothesis was tested by creating another Bayesian generalized linear model. The dependent variable human likeness was regressed on three factors: human eyes (true or false), human skull (true or false) and inversion (true or false) and consisted of the interaction term of the three factors.

Materials

Using the online tool Qualtrics a questionnaire was made. The questionnaire consisted of a welcome screen with information about the study, an informed consent page, questions regarding gender and age, stimuli followed by questions concerning likability and human likeness, and ended with a debriefing.

Stimuli

The participants were presented with pictures of biological (primate) faces that were derived from the data set of Geue and Schmettow (2021). Their dataset originally consisted of 89 primate faces, but was scaled down to 44 pictures to prevent the experiment from becoming too long. The original dataset was ordered on human likeness (based on expert rating scores) and then every second stimulus was chosen for the new data set. By doing this, the new data set still consisted of the total range of human likeness. These 44 pictures were manipulated in PowerPoint to cover the eyes of the primates with pictures of sunglasses (derived from *Vecteezy.com)* to create the variable visibility of the eyes. Additionally, the original 44 pictures were inverted to create the variable inversion (see Figure 7 for three example stimuli). In total, the dataset consisted of 132 pictures, of which a randomised 100 stimuli were shown in the experiment.

Figure 7

Example Stimuli of an Upright Face, Face with Sunglasses and Inverted Face

Participants

148 people initiated the questionnaire, of these 85 people completed the full questionnaire and could be used for further analysis. 28 (32.9%) people were male and 57 (67.1%) were female. The mean age was 23.1 years (*SD* = 7.4). The participants were recruited using convenience and snowball sampling through the researcher's social network as well as through the SONA test subject recruiting system of the University of Twente's faculty of Behavioural, Management and Social Science (BMS). Students who took part through SONA received credit points they need to finish their studies. To be eligible for the study participants needed to be at least 16 years old and have sufficient understanding of English.

Since Koopman and Schmettow (2019) had shown how the uncanny valley effect is a universal phenomenon and the research question did not require any demographical data, except gender and age no further demographical data were collected.

Results

Descriptive Statistics

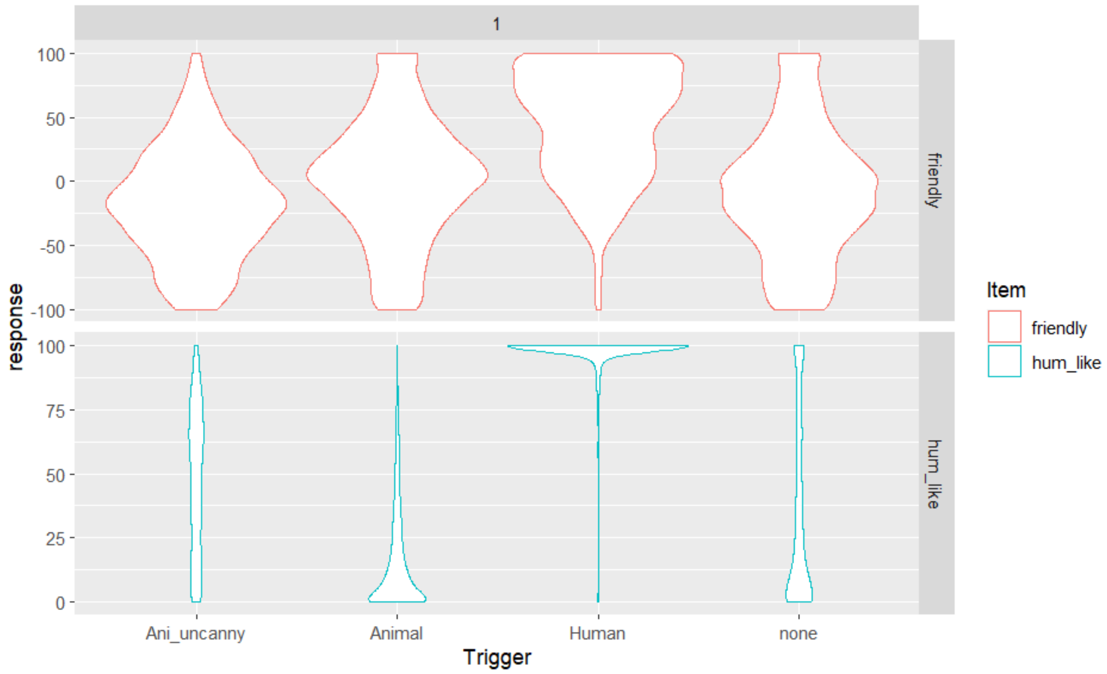
We investigated whether the eyes play a significant role in causing the uncanny valley phenomenon to occur in primate stimuli. Therefore, we tested whether there was a significant difference between the likability rating that participants gave facial stimuli of primates that could be divided into four categories: congruent faces (human face with human eyes, or non-human face with non-human eyes), inverted faces and incongruent faces (non-human face with human eyes). We used a Bayesian generalized linear model (GLM) to compare the likability scores of the four groups. Additionally, to test whether the uncanny valley effect disappears in the presence of sunglasses (when the eyes cannot be evaluated), we used another Bayesian GLM. Lastly, another Bayesian GLM was created to investigate the effect of inversion on human likeness scores and on likability.

The following violin graph (Figure 8) illustrates the associations on the raw data. The four graphs represent the four categories: incongruent faces (Ani_uncanny), animal faces (Animal), human faces (Human) and inverted faces (none). On the y-axis the responses are presented, with in red the likability or friendliness scores and in blue the human likeness scores.

This first look at the data shows us that there are some clear differences in human likeness between the groups, with the animal and inverted group scoring relatively low, humans (not surprisingly) very high, and the uncanny group showing a more spread distribution. It is also visible that for uncanny animals most responses seem to be in the typical 70-80% of human likeness region. The friendliness scores also show differences, especially with the human group scoring higher than the other groups. Moreover, the uncanny group seems to have a lower modus and mean, as compared to the other groups.

Figure 8

Violin Graph Representing the Friendliness (Likability) Scores and Human Likeness Scores for the Four Groups



Next, we calculated the bivariate relationship between the dependent variables human likeness and likability. A weak positive correlation was found between human likeness and likability ($r = .26$). This weak positive correlation indicates that participants that found a stimulus to be human like more often also found the stimulus to be likable. However, the weak correlation does show that the two variables clearly measure different constructs.

Inferential statistics

Next, we used Bayesian generalized linear regression to create a coefficient table in which the coefficient estimates for the four groups are presented, along with the credibility limits (see Table 1). The intercept represents the predicted likability score for the Uncanny animal group, while the center values for the other groups represent the expected change in the likability score relative to the reference category (Uncanny animal).

Table 1

Coefficient Estimates for Likability within the Four Groups with 95% Credibility Limits

	Center	Lower	Upper
Intercept (Uncanny animal)	-17.1	-19.4	-14.9
Animal	17.9	15.1	20.8
Human	62.0	58.1	66.1
none	9.01	6.18	11.8

In Table 2 the predicted values for the four categories are presented, including their lower and upper limit. As can be seen, the predicted value for the incongruent group (Uncanny animal) is clearly negative, as well as the inverted group (none), while the likability score for human stimuli is clearly positive. For the animal group the predicted score is 0.747, but the credibility limits show that the true value lies between -0.908 and 2.36 within the credibility interval of 95%. The table shows that the likability scores for the incongruent group are lowest, and the lack of overlap (see Figure 9) of the credibility limits with the other groups shows that there is a clear difference between the groups.

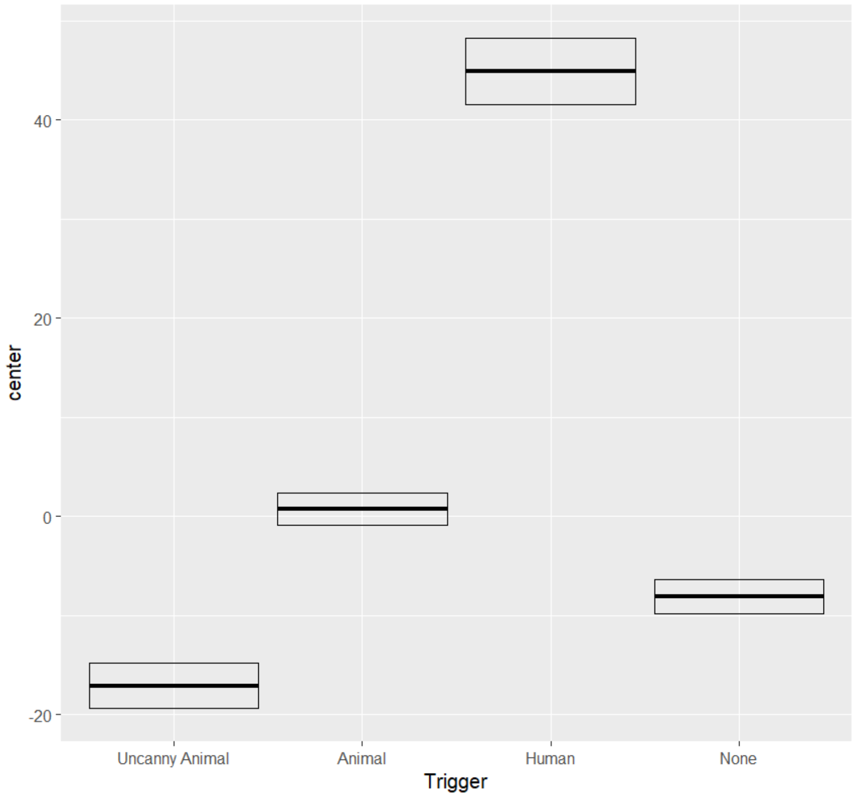
Table 2

Likability Estimates Within the Four Groups with 95% Credibility Limits

	Center	Lower	Upper
Uncanny animal	-17.1	-19.4	-14.8
Animal	0.747	-0.908	2.36
Human	44.9	41.6	48.3
none	-8.15	-9.88	-6.41

Figure 9

Error Bar Plot Visualizing the Likability Estimates Within the Four Trigger Groups with 95% Credibility Limits



We also calculated the predicted human likeness scores which are presented in Table 3 along with their lower and upper credibility limit. In order, the human stimuli show highest human likeness, followed by the incongruent category, the inverted category and lastly the animal category. The data show that there is no overlap between the values of the different categories (see Figure 10), indicating that they can indeed be evaluated as four different groups.

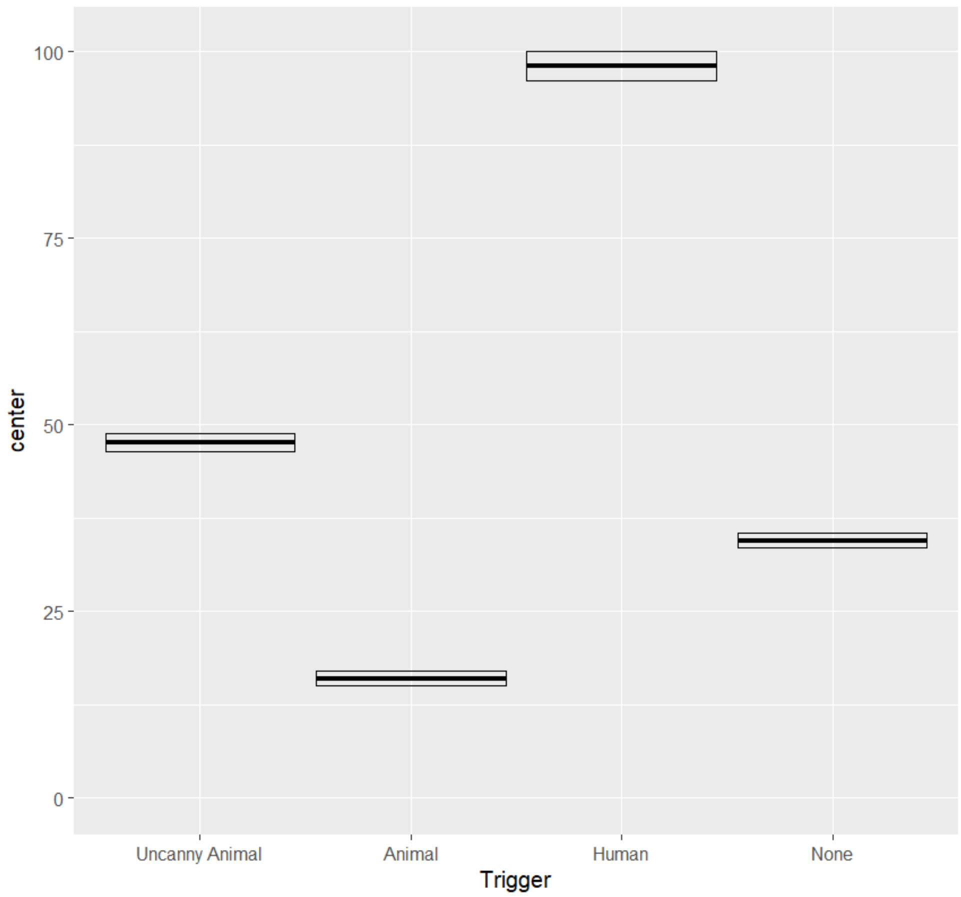
Table 3

Human Likeness Estimates Within the Four Groups With 95% Credibility Limits

	Center	Lower	Upper
Uncanny animal	47.6	46.3	48.9
Animal	16.0	15.0	16.9
Human	98.1	96.1	100
None	34.5	33.4	35.5

Figure 10

Error Bar Plot Visualizing the Human Likeness Estimates Within the Four Trigger Groups with 95% Credibility Limits



Then, the effect of sunglasses on the likability evaluation of incongruent faces was evaluated using a Bayesian generalized linear model. The results are presented in Table 4. As can be seen, likability estimates for both animal groups are higher with sunglasses than without. While the opposite is true for humans. Figure 11 shows again a lack of overlap between the credibility limits of the different groups.

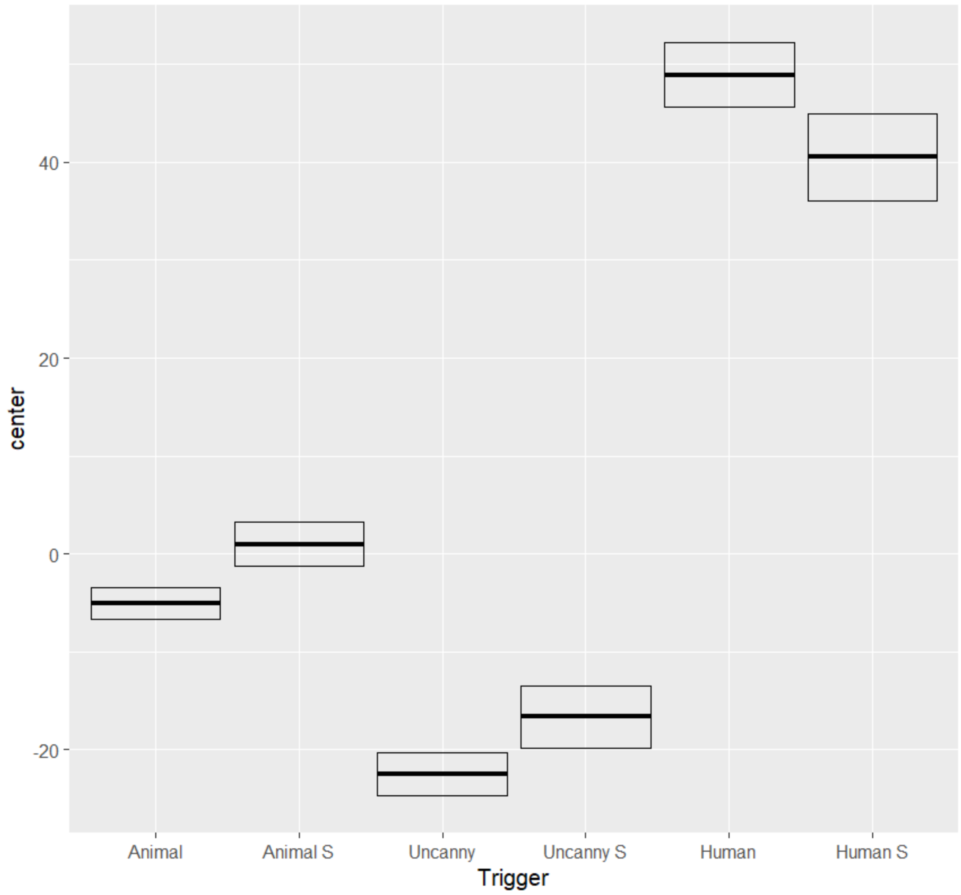
Table 4

Likability Estimates With and Without Sunglasses With 95% Credibility Limits

	Center	Lower	Upper
Uncanny animal without sunglasses	-22.4	-24.5	-20.5
Uncanny animal with sunglasses	-16.6	-19.7	-13.4
Animal without sunglasses	-5.01	-6.65	-3.41
Animal with sunglasses	0.996	-1.21	3.15
Human without sunglasses	48.9	45.6	52.1
Human with sunglasses	40.5	36.2	45.0

Figure 11

Error Bar Plot Visualizing the Likability Estimates for the Four Trigger Groups With (S) and Without Sunglasses With 95% Credibility Limits



Next, the effect of inversion on the evaluation of incongruent faces was evaluated, as well as the effect of inversion on the evaluation of congruent human and animal faces. The results are presented in Table 5. The results show that both animal groups are evaluated lower on likability when presented inverted. For humans the credibility intervals overlap (see Figure 12), indicating no clear difference between inverted or upright ratings for likability.

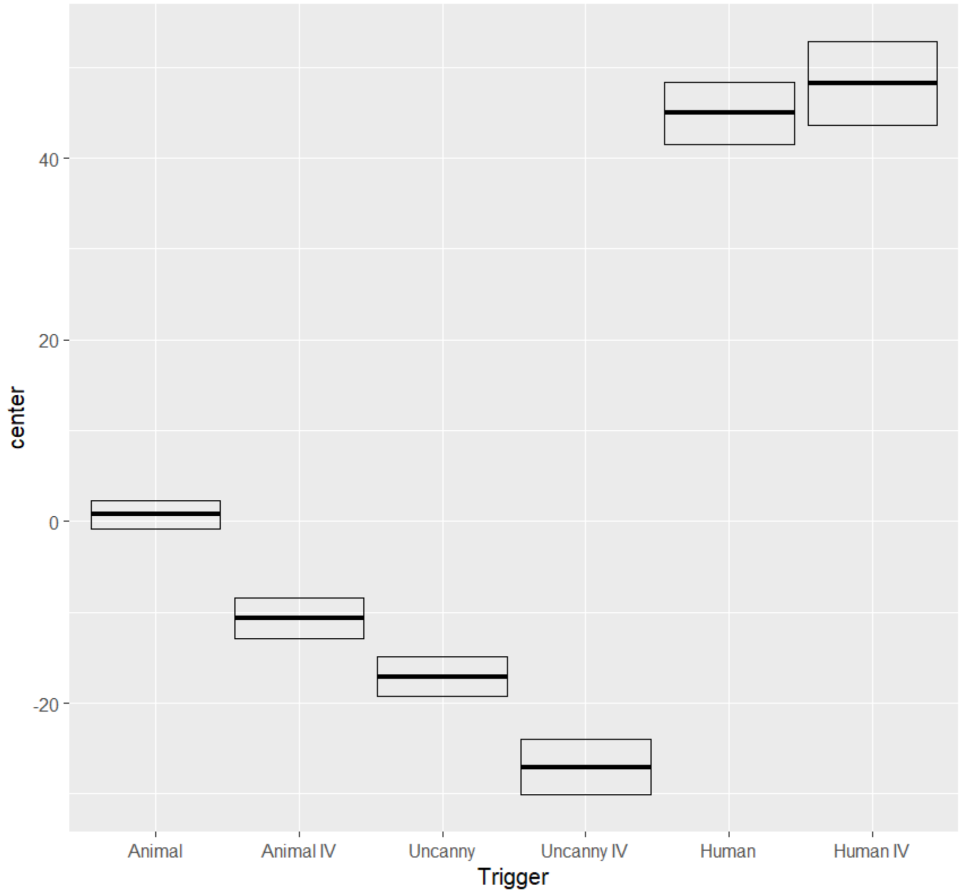
Table 5

Likability Estimates for Upright and Inverted Stimuli With 95% Credibility Limits

	Center	Lower	Upper
Uncanny animal upright	-17.1	-19.3	-15.0
Uncanny animal inverted	-27.0	-30.1	-24.0
Animal upright	0.760	-0.859	2.27
Animal inverted	-10.6	-12.9	-8.37
Human upright	44.9	41.5	48.3
Human inverted	48.3	43.6	52.8

Figure 12

Error Bar Plot Visualizing the Likability Estimates for Upright and Inverted (IV) Stimuli with 95% Credibility Limits



Then, we investigated whether inverted stimuli are evaluated as less human-like compared to upright stimuli (see Table 6). Here again, both animal groups are evaluated lower on human likeness when presented inverted, while no clear effect is found for human stimuli (see Figure 13).

Table 6

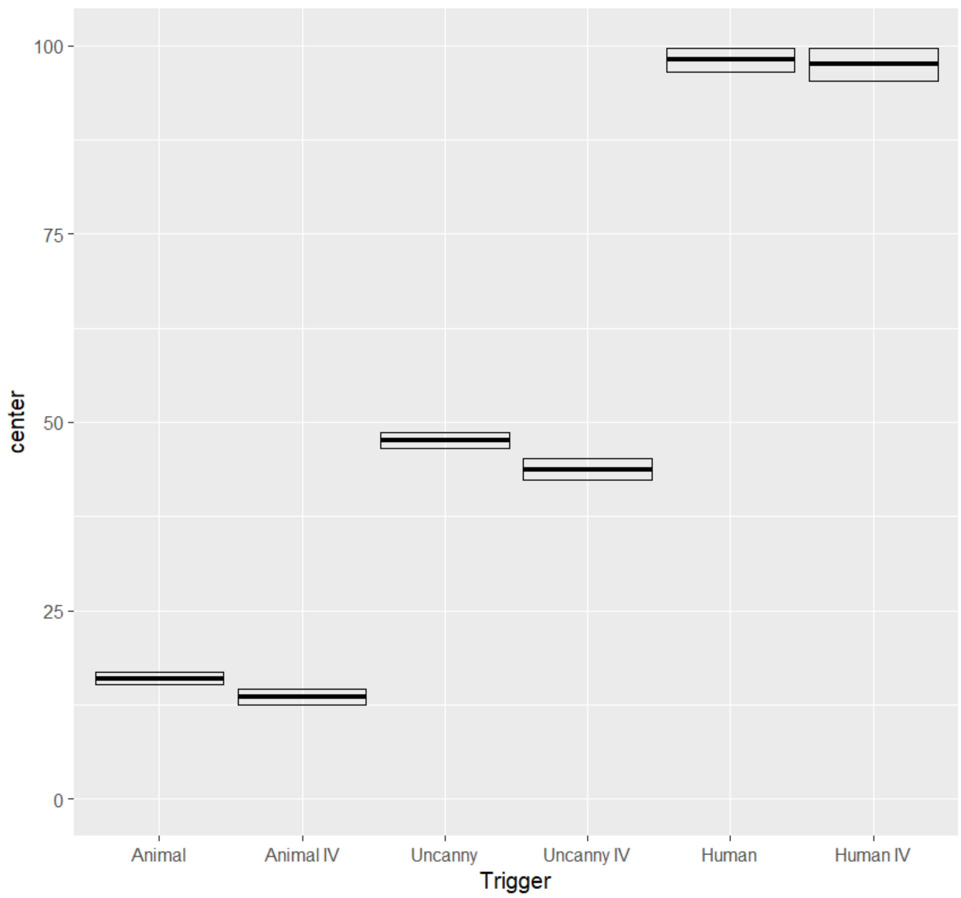
Human Likeness Estimates for Upright and Inverted Stimuli With 95% Credibility Limits

	Center	Lower	Upper
Uncanny animal upright	47.6	46.6	48.6
Uncanny animal inverted	43.8	42.3	45.1
Animal upright	16.0	15.2	16.8

Animal inverted	13.5	12.4	14.6
Human upright	97.5	95.3	99.7
Human inverted	98.1	96.5	99.7

Figure 13

Error Bar Plot Visualizing the Human Likeness Estimates for Upright and Inverted (IV) Stimuli with 95% Credibility Limits



Discussion

The aim of this research was to investigate the effect of the eyes on the occurrence of the uncanny valley effect in primate stimuli, and to explore the effect of inversion on likability and human likeness. To test the Specific Category Confusion theory, we categorised the stimuli into four groups: animals (congruent: animal face with animal eyes), humans (congruent: human face with human eyes), uncanny animals (incongruent: animal face with human eyes) and none (inverted stimuli) and hypothesised that the incongruent uncanny animal group would score lowest on likability. The research findings support this hypothesis (see Table 1 and 2), since the likability estimates for this group were clearly lowest. This supports Geue and Schmettow's (2021) explorative hypothesis that stimuli with visible white sclera would be evaluated lower on likability.

The Effect of Sunglasses on Likability

The additional question was if this effect would then disappear when the eyes in incongruent faces would be covered by sunglasses, preventing the perceptual mismatch between a non-human face and humanlike eyes to occur. According to the findings the effect does not completely disappear, since the effect estimate for Uncanny animals with sunglasses is still negative within the credibility interval, and is still quite different from the expected effects for originally congruent animals with sunglasses.

However, the results do show that the uncanny animal group with sunglasses have higher values than without sunglasses, indicating a higher estimate for likability. This would show slight support that the eyes play a significant role in causing the uncanny feeling to occur. On the other hand, the same effect is present for the animals with sunglasses, as they also have higher predicted values as compared to animals without sunglasses. Therefore, it cannot be concluded that the reason why the uncanny animal group would be evaluated higher on likability in the presence of sunglasses (as compared to without) is because the eyes are covered, which would remove the uncanny feeling as hypothesised by the Specific Category Confusion theory.

There are several possible explanations for these findings. First of all, it could be possible that both animal groups are scored higher on likability when wearing sunglasses because wearing sunglasses makes them look funny and therefore more friendly and less creepy. Future research could therefore include a question regarding funniness of the stimuli.

Secondly, another possibility is that the eyes are simply not the determining factor in the uncanny feeling to occur. If the effect is visible in upright stimuli for which the eyes are visible, but also when the eyes are not visible, it seems like other factors within the face of incongruent stimuli might still be evaluated as uncanny, such as the general high human likeness, or facial expression of the stimuli. Other ways to test the effect of the eyes could be to digitally remove the eyes from the faces, although an eyeless face will most likely automatically be evaluated as creepy. Furthermore, the eyes could be made invisible by adding a black bar covering the eyes, however, people might have the association of criminals with this manipulation. Lastly, the eyes could also be digitally manipulated to be either black (animal like) or white (human like) to investigate the Specific Category Confusion theory.

Dollee and Schmettow (2022) researched the effect of this manipulation on likability scores. They found that both white eyes in an animal face and black eyes in a human face were evaluated as eerie, but that black eyes in a human face were found to be more eerie. This does show that the eyes play an important role in likability perceptions and provides support for the Specific Category Confusion theory. It seems that upon adding sunglasses people base their judgment on other factors, and still evaluate more human like ape faces as more eerie than stimuli that were lower in human likeness. The results showed that the originally incongruent faces were still evaluated as less likable than the animals and humans. Therefore, it could be possible that in the absence of eyes, we focus on other parts of the face to determine the likability and human likeness of a face. To investigate this hypothesis, an eye tracking study could be performed to explore where people look when eyes are covered by sunglasses.

The dataset consisted of stimuli that differed in their human likeness, as evaluated before the research. Stimuli that had a higher resemblance to humans have facial features that are more similar to human’s facial features, such as skull nose and mouth. Additionally, the stimuli that were previously determined as high in human likeness almost all had visible white sclera. One of the main research questions was if the mismatch between the visible white sclera in ape faces was at the root of the uncanny valley effect. However, the dataset did not contain any monkey faces (low in human likeness) with white sclera. Neither did the dataset contain stimuli high in human likeness with dark eyes, or humans with dark eyes. Therefore, a future dataset would have to be more inclusive throughout the whole spectrum by including stimuli of monkeys for which the eyes are manipulated so that they also have white sclera, and stimuli of highly human-like apes with dark eyes, as well as humans with dark eyes. This way, it can be determined if white sclera is indeed a main factor for the uncanny valley effect to occur, or if this is only the case within a face that is high in human likeness. In addition, to rule out the effect of facial expressions it would be more useful to include only neutral looking faces.

Another explanation for the findings that uncanny animals with sunglasses were not rated as friendly as congruent animals with sunglasses is that people recognise the stimuli from the condition without sunglasses. Research has shown that people show higher recall rates for faces that are rated low in likability (Geiger & Balas, 2021). Therefore, the uncanny faces might have been remembered better and (unconsciously) still elicited an eerie feeling even in the absence of the eyes. This would only possibly explain the data in case participants first viewed the face without sunglasses and later with sunglasses. But, since the stimuli were presented randomly and not all 132 stimuli were shown to all participants this could only partly explain the findings. Future research could prevent this from happening by making sure the sunglass condition is presented before the other conditions, or by including only one of the three conditions of the same stimulus.

Interestingly, for humans the data showed a decrease in likability when adding sunglasses to the face within the credibility interval of 95%. As the credibility intervals do not overlap, it seems like humans with sunglasses are indeed evaluated as less likable as compared to humans without sunglasses. Meanwhile, for both the congruent animal group and the uncanny animal group the addition of sunglasses seemed to increase likability estimates. These findings imply that when evaluating human faces sunglasses negatively affect our perceptions of likability towards a person. This can be explained in the context of communication and risk perception theories, that explain the importance of eye contact in human communication as well as to determine possible threats. We can extract a lot of information from someone's eyes, such as gaze direction and emotions. A negotiation study found that people rate their negotiation partner as less trustworthy, less sincere and more manipulative when this person was wearing sunglasses in interpersonal communication (Pezzuti et al., 2011).

The model sunglasses might have also strengthened these findings, since they are dark and quite serious. Other models might have also made humans seem friendlier, as they have also made animals friendlier. Further research could include stimuli with sunglasses that differ in their degree of transparency. The sunglasses lowest in transparency should make it impossible to view the eyes of the stimulus, similar to the sunglasses from this experiment. When viewing a stimulus wearing sunglasses highest in transparency, the eyes of the stimulus should be easily visible. This way, it can be investigated whether animals are evaluated as more likable when wearing sunglasses just because they wear sunglasses, or if the possibility to view the eyes plays a role.

The Effect of Inversion on Likability and Human Likeness

Lastly, concerning the third and fourth hypothesis, it was expected that the inverted images would be scored lower on human likeness than their upright counterpart, since configural face processing has been shown to be impaired upon inverting faces. This effect was found for the uncanny animal group, as well as the animal group. Since configural face processing would be impaired in inverted faces, it was also expected that the uncanny animal group would be evaluated as more likable upon inversion, since the hypothesised Specific Category Confusion would not occur when configural face processing is not possible. The findings do not support the hypothesis regarding likability, since the effect estimations for the inverted uncanny group are lower than for the upright group, which indicates lower predictions for likability. The same effect is visible for the congruent animal group, while there is no clear difference for the humans, since the credibility intervals for human upright and inverted stimuli show large overlap. Apparently, animals of both categories were rated as less friendly, more unpleasant or creepy when inverted.

One limitation concerning the setup of the study for investigating the effect of inversion was that most participants filled out the questionnaire on their phone. Some participants expressed that they were confused about the inverted stimuli and turned their phone to be able to look at the stimuli normally and then answered the questions. This could have affected the effect size of inversion on human likeness, which might have been greater if participants would not have been able to turn their phone. An option to prevent this from happening could be to ask participants to fill out the survey on a laptop, but nowadays everyone carries their phone around and asking participants to use a laptop could greatly decrease the number of participants. Therefore, future research could either indicate beforehand that some stimuli will be presented upside down and ask participants not to turn their phones, or the stimuli can be presented only for a few seconds so participants do not have enough time to turn their phone and are prevented from evaluating the stimulus from its original angle.

Limitations

Apart from the limitations already mentioned before, another limitation was the duration of the survey. Since the survey took longer than intended there was a high dropout rate. A total of 148 people initiated the survey, while only 85 people finished the complete survey. Lowering the total number of stimuli could have helped to shorten the time needed for completion and increase the number of participants. Additionally, adding a progress bar to indicate how many questions are still left to answer could promote participants to continue with the questionnaire. The difference between the initial number of participants and those who completed the survey could have affected the representativeness of the sample and limit the generalizability of the findings through what is known as self-selection. However, since the credibility limits of the findings are small and research has shown that the uncanny valley effect is a universal phenomenon (Koopman and Schmettow, 2019), this risk is low.

Lastly, it might be possible that due to the length participants got less serious about answering the questions if they felt it was taking too much time. This is based on the finding that a few participants indicated 0 for human likeness of human faces. Although this could also have a different explanation such as a lack of understanding of the task, technical problems or a general disinterest in the study.

Furthermore, as explained before, the dataset did not include the full set of possible combinations of conditions. Additionally, there were no normality checks performed, however, the data presented in Figure 8 already clearly showed that the data is not normally distributed. One of the main aims was to find whether the results that ape faces with visible white sclera would show lowest likability ratings as found by Geue and Schmettow (2021) could be replicated. Within this aim and the scope of this research deeper analyses did not fit. However, future analyses could investigate to what extent the findings could also be found on participant level by employing a multi-level model.

Future Research

To summarise the main suggested recommendations for future research, it would be useful when creating a future dataset to include stimuli from all different conditions to be able to investigate the specific effect of the eyes on likability ratings. Furthermore, it would be interesting to create stimuli with sunglasses that differ in their transparency, so that it can be tested whether the effect depends on the absence of the eyes, or whether the findings were found because monkeys with sunglasses are perceived to be more funny and thus more likable. Additionally, eye tracking devices could provide insight in where people look upon viewing a primate face with sunglasses. It might be useful to exclude stimuli that have too extreme facial expressions, since that highly influences how people perceive likability. To overcome the limitations regarding duration and inversion, adding a time limit to the stimuli could shorten the survey and prevent people from turning their phone when encountering an inverted stimulus. In addition, it forces participants to base their verdict on what knowledge they have required using configural face processing, since they do not have the time to separately inspect all individual features within the face. Lastly, an insightful future experiment could make use of EEG equipment measuring the N170 response to investigate the Specific Category Confusion theory using primate faces, instead of only human faces.

Conclusion

This study investigated the effect of incongruence between eyes and face of primate stimuli on likability estimates, as well as the effect of inverting or adding sunglasses to the stimuli on likability estimates. While incongruence within the face resulted in lowest likability estimates, this effect did not fully disappear in the condition with sunglasses. This indicates that the hypothesised Specific Category Confusion Theory on its own cannot predict the Uncanny Valley effect to occur. The results add to our understanding of how faces are processed regarding likability as they show that the eyes on their own are not the sole cause of the UV to appear. Further research is necessary to explore under which specific conditions the UV occurs.

References

Bagepally, B. S. (2015). Gaze pattern on spontaneous human face perception: An eye tracker study. *Journal of the Indian Academy of Applied Psychology, 41*(Special Issue 3), 127-130.

Clark, I. R., Lee, K. C., Poux, T., Langergraber, K. E., Mitani, J. C., Watts, D., Reed, J., & Sandel, A. A. (2023). White sclera is present in chimpanzees and other mammals. *Journal of Human Evolution, 176, 103322.* https://doi.org/10.1016/j.jhevol.2022.103322

Diel, A., & MacDorman, K. F. (2021). Creepy cats and strange high houses: Support for configural processing in testing predictions of nine uncanny valley theories. *Journal of Vision, 21(4), 1.* https://doi.org/10.1167/jov.21.4.1

Dollee, I. & Schmettow, M. (2022). *The Role of the Eyes in the Uncanny Valley Effect: Does Incongruence between Eyes and Face Influence Uncanniness?* [Bachelor’s thesis, University of Twente]. http://essay.utwente.nl/91623

Geiger, A. R., & Balas, B. (2021). Robot face memorability is affected by uncanny appearance. *Computers in human behavior reports*, 100153. https://doi.org/10.1016/j.chbr.2021.100153

Geller, T. (2008). Overcoming the Uncanny Valley. *IEEE Computer Graphics and Applications*, *28*(4), 11–17. https://doi.org/10.1109/mcg.2008.79

Geue, L. & Schmettow, M. (2021). *From robots to primates: tracing the uncanny valley effect to its evolutionary origin* [Bachelor’s thesis, University of Twente]. http://essay.utwente.nl/8756

Green, R. C., MacDorman, K. F., Ho, C., & Vasudevan, S. K. (2008). Sensitivity to the proportions of faces that vary in human likeness. *Computers in Human Behavior*, *24*(5), 2456–2474. https://doi.org/10.1016/j.chb.2008.02.019

Itier RJ, Batty M. Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience and Biobehavioral Reviews.* 2009;33:843–863. doi: 10.1016/j.neubiorev.2009.02.004.

Kanwisher, N., & Moscovitch, M. (2000). The cognitive neuroscience of face processing: an introduction. *Cognitive Neuropsychology*, *17*(1–3), 1–11. https://doi.org/10.1080/026432900380454

Kim, B., Bruce, M., Brown, L., De Visser, E. J., & Phillips, E. J. (2020). A Comprehensive Approach to Validating the Uncanny Valley using the Anthropomorphic RoBOT (ABOT) Database. *Systems and Information Engineering Design Symposium*. https://doi.org/10.1109/sieds49339.2020.9106675

Kobayashi, H., Kohshima, S. Unique morphology of the human eye. *Nature* 387, 767–768 (1997). https://doi.org/10.1038/42842

Koopman, R., & Schmettow, M. (2019). *The Uncanny Valley as a universal experience: A Replication study using multilevel modelling* [Bachelor’s thesis, University of Twente]. https://essay.utwente.nl/77172/

Lewkowicz, D. J., & Ghazanfar, A. A. (2011). The development of the uncanny valley in infants. *Developmental Psychobiology*, *54*(2), 124–132. https://doi.org/10.1002/dev.20583

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

MacDorman, K. F., Green, R. C., Ho, C., & Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior, 25(3), 695–710. https://doi.org/10.1016/j.chb.2008.12.026*

Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260. https://doi.org/10.1016/s1364-6613(02)01903-4

McKone E. (2010). “Face and object recognition: how do they differ?” in *Tutorials in Visual Cognition ed. Coltheart V. London: Psychology Press, 261–303.*

Mori, M. (1970). The uncanny valley, *7*(4), 33–35.

https://doi.org/10.1109/MRA.2012.2192811

Pezzuti, T., Pechmann, C., Borges, A., & Pirouz, D. M. (2011). Sunglasses, Hierarchy, and Negotiations: Gender Differences in Eye Gaze During Interpersonal Communication. *ACR North American Advances*. http://www.acrwebsite.org/volumes/v38/acr_v38_16124.pdf

Richler J. J., Mack M. L., Gauthier I., Palmeri T. J. (2009). Holistic processing of faces happens at a glance. *Vision Res.* 49 2856–2861. 10.1016/j.visres.2009.08.025

Schmettow, M. (2021). *New statistics for design researchers*.

Seyama, J., & Nagayama, R. S. (2007). The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces. *Presence: Teleoperators & Virtual Environments*, *16*(4), 337–351. https://doi.org/10.1162/pres.16.4.337

Vaitonytė, J., Alimardani, M., & Louwerse, M. M. (2022). Scoping review of the neural evidence on the uncanny valley. *Computers in human behavior reports*, *9*, 100263. https://doi.org/10.1016/j.chbr.2022.100263

Wang C. C. (2019). Investigating the Time Course of Part-Based and Holistic Processing in Face Perception. *Frontiers in psychology, 9, 2630. https://doi.org/10.3389/fpsyg.2018.02630*

Yin, R.K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.

Appendix A

Information Sheet and Informed Consent

Dear participant,

Thank you for your participation in this research study. Please read the following information carefully before deciding whether to participate. This information sheet is intended to provide you with an understanding of the research project, its benefits and risks, and the procedures involved in participating. If you have any questions or concerns, please contact the researcher before agreeing to participate.

Purpose of the Research: The purpose of this research project is to gain insight in emotional reactions towards different faces. The study aims to deepen our knowledge on face perception and will take you approximately 20 minutes to complete.

Benefits and Risks of Participating: By participating in this study, you will contribute to advancing knowledge in the field of face perception. If you participate through SONA, you will be rewarded with research credits. Otherwise, there are no direct benefits to you for participating in this study, but your participation will help us to better understand facial processing mechanisms. There are no known risks associated with participating in this study, but, as with any online related activity the risk of a breach is always possible. If, however, you experience any discomfort or distress during your participation, you may withdraw from the study at any time without any penalty or loss of benefits.

Ethics Approval: This research project has been reviewed and approved by the Ethics Committee of the Faculty of Behavioural and Management and Social Sciences at the University of Twente. This approval ensures that the research is conducted in accordance with ethical principles and guidelines.

Procedures for Withdrawal: Participation in this study is voluntary. If you decide to participate, you may withdraw at any time without any penalty or loss of benefits. If you choose to withdraw from the study, your data will be removed from the analysis.

Collection and Processing of Personal Information: Upon completing the questionnaire, your data will be anonymised, and you won’t be able to be identified by the research team.

Therefore, once the study is over, the research team will no longer be able to identify you and delete your data. Once the survey is completed, you will be offered a final opportunity to withdraw from the study in case you wish to do so and your data will then be deleted.

Usage of Data and Confidentiality: The data collected during this study will be used solely for research purposes. The data will be stored anonymised to protect your privacy. The data will be stored securely, and access to it will be restricted to authorized researchers. Data may be archived for future research use. If the research data is published, no personally identifiable information will be included.

Retention Period: The research data will be retained for at least 10 years after the completion of the study in accordance with the data retention policy of our organisation. We do stress that the data will be stored anonymously, and no personal identifiable information will be kept.

Contact Information: If you have any questions or concerns about this study, please contact the researcher:
Isabelle Stikker; i.m.m.stikker@student.utwente.nl
or the researcher's supervisor:
dr. Martin Schmettow: m.schmettow@utwente.nl

If you have any complaints about the study, please contact the Ethics Committee of the Faculty of Behavioural and Management and Social Sciences at the University of Twente: ethicscommittee-bms@utwente.nl

Informed consent

- I have read and understood the information provided.
- I consent voluntarily to be a participant in this study and understand that I can withdraw from the study at any time, without having to give a reason and without any consequences.
- I am aware that I can contact the researcher in case I have any concerns or questions about the study.
- I agree that my answers will be saved and used for the purpose of the study and research.

- I understand that my responses will be anonymous, and only anonymous versions of the data will be presented, stored or shared.
- I confirm that I possess a sufficient level of proficiency in English and that I am at least 16 years old.
- I give my consent to taking part in the study which involves evaluating faces on the emotional impact they have on me by filling out survey questions. I understand this survey will take approximately 20 minutes.

Do you agree to all the above-mentioned statements?

Appendix B

Study Description

Please read the following information carefully.

On the next page the study will start. You will be presented with biological (primate) faces. Afterwards, you will be asked to indicate your response towards the presented face by answering two questions. For one, you will need to complete the sentence “To me, this face seems…” by moving a slider. Herein, -100 means "less friendly, more unpleasant or creepy" and +100 means "more friendly, more pleasant or less creepy".

For the other question, you will indicate to what extent the face resembles a human, according to you, by answering the question "How human-like does this face appear to you?”. Again, you will use a slider for this. Herein, 0 means "not human" and would mean that the face **does not have any resemblance** to a human. 100 means "human" and indicates that according to you, the face presented is a **real human**. There are no right or wrong answers, it is about your perception. You will evaluate a total of 100 faces.

Thank you for your participation, you may click the arrow when you are ready to start the study.

Appendix C

Debriefing

Debriefing & Study Objective

Thank you for your participation!

With your participation you contribute to research into the Uncanny Valley effect. More specifically, this research investigates the influence of the eyes in the occurrence of the uncanny valley effect, and to what extent inversion affects the phenomenon to occur.
In case you have any questions regarding the study or would like to contact the researcher for different reasons, you can do this via the following email address:
Isabelle Stikker: i.m.m.stikker@student.utwente.nl

This is also the last moment to opt out of the study in case you no longer consent to your data being used for further analysis.

Can we use the data you provided for further analysis?

0 Yes, I consent to my data being used for further analysis

0 No, I want my data to be deleted

Appendix D
R Script for Data Analysis

````
---

title: "BT Isabelle Stikker Data Analysis"

author: "M. Schmettow"

format: html

editor: visual

---


```{r}

library(tidyverse)

library(readr)

library(readxl)


UPD_DATA = F

```


## Preparation


### Reading Items


```{r eval = UPD_DATA}

read_excel("data/Qualtrics_data_Isabelle_2023.xlsx", sheet = "data", n_max = 1) ->

  D_raw_items
````
````

D_raw_items %>%
  select(`1_1...22`:`44_3...284`) %>%
  pivot_longer(`1_1...22`:`44_3...284`,
               names_to = "Trial",
               values_to = "ItemText") %>%
  separate(Trial, into = c("Face", "Manipulation"), remove = F) %>%
  mutate(Stimulus = str_c(Face, Manipulation, sep = ".")) %>%
  mutate(Item = if_else(str_detect(ItemText, "less friendly"),
                        "friendly",
                        "hum_like")) ->
  D_Items

D_Items %>%
  print()

```

### Reading Stimulus level data

... from Geue

```{r eval = UPD_DATA}
read_excel("data/Isabelle_Stikker_stimuli_table_new.xlsx",
           sheet = "Prep Stimuli qualtrics 2023") %>%

````
  select(Face = New_stimulus_name,
         OrigFace = Stimulus,
         Set,
         hum_like = `H likeness total`,
         anc_close = `AncestoralCloseness`) %>%
  mutate(Face = as.character(Face)) ->
  D_Geue
```
````

and yours

````
```{r eval = UPD_DATA}
read_excel("data/Isabelle_Stikker_stimuli_table_new.xlsx",
           sheet = "Stimuli qualtrics 2023") %>%
  rename(Stimulus = New_stimulus_name,
         human_skull = human_face,
         Face_type = Face,
         Eyes_type = Eyes) %>%
  mutate(Stimulus = as.character(Stimulus)) %>%
  mutate(across(human_skull:inverted, ~ .x == "true")) ->
  D_Stim
```
````

Defining trigger categories
````

Here goes the theory.

````
```{r eval = UPD_DATA}
D_Stim %>%
  mutate(Trigger = if_else(inverted,
                           "none",
                           if_else(human_skull,
                                   if_else(!human_eyes,
                                           "Hum_uncanny",
                                           "Human"),
                                   if_else(human_eyes,
                                           "Ani_uncanny",
                                           "Animal")))) %>%
  select(Stimulus, Trigger) ->
  D_Theory


D_Theory %>%
  group_by(Trigger) %>%
  summarize(n())
```
````

Reading responses

```
```{r eval = UPD_DATA}
read_excel("data/Qualtrics_data_Isabelle_2023.xlsx", sheet = "data") ->
  D_raw


D_raw %>%
  slice(-1) %>%
  select(Duration = `Duration (in seconds)`,
         Finished,
         Consent = `Informed consent`,
         `1_1...22`:`44_3...284`) %>%
  mutate(Duration = as.numeric(Duration)) %>%
    filter(Consent == "Yes",
         Finished == "True",
         as.numeric(Duration) > 200)->
  D_0



nrow(D_0)
```


```{r eval = UPD_DATA}
D_0 %>%
  mutate(Part = 1:n()) %>%
  select(Part, `1_1...22`:`44_3...284`) ->
```
```

~~~~
  D_1


head(D_1)
```


### Joining it all


```{r eval = UPD_DATA}
D_1 %>%
  pivot_longer(-Part,
               names_to = "Trial",
               values_to = "response") %>%
  mutate(response = as.numeric(response)) %>%
  left_join(D_Items, by = "Trial") %>%
  left_join(D_Geue, by = "Face") %>%
  left_join(D_Stim, by = "Stimulus") %>%
  left_join(D_Theory, by = "Stimulus") %>%
  select(Part, Item, Stimulus, Face, Manipulation,
         Face_type, Eyes_type, human_skull:inverted,
         hum_like, anc_close, Trigger, response) ->
  IS_1


sample_n(IS_1, 12)
~~~~

```

### Separating the two scales

- original hum_like is renamed

```{r eval = UPD_DATA}
IS_1 %>%
  rename(hum_like_orig = hum_like) %>%
  pivot_wider(names_from = Item,
              values_from = response) ->
  IS_2

sample_n(IS_2, 5)
```

Saving Data

```{r eval = UPD_DATA}
save(IS_1, IS_2, file = "IS_23.Rda")
write_csv(IS_1, file = "IS_1.csv")
write_csv(IS_2, file = "IS_2.csv")
```
```

Exploration

```{r}
load("data/IS_23.Rda")
```

```{r}
IS_1 %>%
  ggplot(aes(x = Trigger,
             y = response,
             col = Item)) +
  facet_grid(Item~1, scales = "free_y") +
  geom_violin()
```

```{r}
IS_1 %>%
  group_by(Part, Trigger, Item) %>%
  summarize(avg_response = mean(response, na.rm = T)) %>%
  ggplot(aes(x = Trigger,
             y = avg_response)) +
  facet_grid(Item ~ 1) +
  geom_line(aes(group = Part))
```

```{r}
IS_2 %>%
  ggplot(aes(x = hum_like,
             y = friendly)) +
  facet_grid(Trigger ~ 1) +
  geom_smooth()
```

Modelling

```{r}
library(rstanarm)
library(brms)
# devtools::install_github("schmettow/bayr")
library(bayr)
options(mc.cores = 4)
```

```{r}
M_0 <- stan_glm(friendly ~ Trigger, data = IS_2)
```

```{r}
```

```
fixef(M_0)
```

Results section

```{r}
M_1 <- stan_glm(friendly ~ 0 + Trigger, data = IS_2)
fixef(M_1)
```

##Error bar plots

```{r}
F_9 <- stan_glm(friendly ~ 0 + Trigger, data = IS_2)
fixef(F_9)
coef(F_9) %>%
  rename(Trigger = fixef) %>%


  ggplot(aes(
    x = Trigger,
    y = center, ymin = lower, ymax = upper
  )) +
  scale_x_discrete(labels =
                    c('Uncanny Animal','Animal', 'Human', 'None') )
  +
  geom_crossbar()


F_10 <- stan_glm(hum_like ~ 0 + Trigger, data = IS_2)
```

```
fixef(F_10)

coef(F_10) %>%
  rename(Trigger = fixef) %>%

  ggplot(aes(
    x = Trigger,
    y = center, ymin = lower, ymax = upper
  )) +
  scale_x_discrete(labels =
                    c('Uncanny Animal','Animal', 'Human', 'None') )+
  ylim(0, 101)+
  geom_crossbar()

F_11 <- stan_glm(friendly ~ 0 + human_eyes : human_skull : sunglasses, data = IS_2)
fixef(F_11)

coef(F_11) %>%
  rename(Trigger = fixef) %>%

  ggplot(aes(
    x = Trigger,
    y = center, ymin = lower, ymax = upper
  )) +
```

```
  scale_x_discrete(labels =
                   c('Animal','Animal S', 'Uncanny', 'Uncanny S',
                     'Human', 'Human S')) +
  geom_crossbar()

F_12 <- stan_glm(friendly ~ 0 + human_eyes : human_skull : inverted, data = IS_2)
fixef(F_12)
coef(F_12) %>%
  rename(Trigger = fixef) %>%

  ggplot(aes(
    x = Trigger,
    y = center, ymin = lower, ymax = upper
  )) +
  scale_x_discrete(labels =
                   c('Animal','Animal IV', 'Uncanny', 'Uncanny IV',
                     'Human', 'Human IV')) +
  geom_crossbar()

F_13 <- stan_glm(hum_like ~ 0 + human_eyes : human_skull : inverted, data = IS_2)
fixef(F_13)
coef(F_13) %>%
  rename(Trigger = fixef) %>%
```

```
  ggplot(aes(
    x = Trigger,
    y = center, ymin = lower, ymax = upper
  )) +
  scale_x_discrete(labels =
                     c('Animal','Animal IV', 'Uncanny', 'Uncanny IV',
                       'Human', 'Human IV')) +
  ylim(0, 100)+
  geom_crossbar()

```

##Hypotheses 2, 3 and 4

```{r}
H_2 <- stan_glm(friendly ~ 0 + human_eyes : human_skull : sunglasses, data = IS_2)
fixef(H_2)


H_3 <- stan_glm(friendly ~  0 + human_eyes:human_skull:inverted, data = IS_2)
fixef(H_3)


H_4 <- stan_glm(hum_like ~ 0 + human_eyes:human_skull:inverted, data = IS_2)
fixef(H_4)
```