How do Humans Perceive Gaze Direction?

Janis Hölter

Bachelor Thesis

Faculty of Behavioral, Management and Social Sciences (BMS),

University of Twente

Human Factors and Engineering Psychology

First Supervisor: Dr. Martin Schmettow

Second Supervisor: Dr. Cesco Willemse

24th of June 2025

Abstract

According to the Cooperative Eye Hypothesis, the human eye evolved into its characteristic shape with a contrast between the white sclera and the iris to enable communication via gaze following. Your Eye Tracker (YET) is an eye tracking system that splits the eye into four quadrants and predicts gaze direction from the contrasts in average brightness between the quadrants. This study examines the question whether YET is working with the same mechanism as human gaze perception and what this means for the Cooperative Eye Hypothesis. For this quantitative experimental study, 43 participants were recruited via convenience sampling. Participants had to estimate gaze direction from images of a face in two conditions, the experimental condition, here the face contained the YET-quadrant in place of the eyes, and the control condition in which the image was not manipulated. These stimuli were presented at five exposure times (70ms-1000ms). The resulting gaze perception accuracy was analysed using a Bayesian generalised linear multi-level model. The model revealed that participants were able to detect gaze direction from the manipulated images. However, deviation of their responses was significantly higher in the experimental condition than in the control condition. Gaze perception was consistently higher in the experimental condition from an exposure time of 140ms onwards. Across the random effects of the model, considerable variability was found. From these results it can be deduced that humans are able to perceive gaze direction from the 2x2 quadrants used by the YET. However, accuracy in this was significantly lower compared to gaze perception from unmanipulated face images. Deviation in gaze perception from the 2x2 quadrant eye face was significantly lower from an exposure of 140ms onwards than from an exposure of 70ms. These findings strengthen the Cooperative Eye Hypothesis and indicate that the mechanisms of the YET might resemble the cognitive processes of humans. Practically, this research could be used in fields such as education to teach about gaze perception or in the designing process of intuitive human-robot communication by informing about the characteristics of gaze perception.

How do Humans Perceive Gaze Direction?

Cooperation between humans is an essential part of our evolution as a species. Throughout history, humans had to form partnerships and groups for rearing children, foraging, hunting and mating to survive. Followingly, an evolutionary advantage of collaborating and becoming interdependent developed (Tomasello et al., 2012). For collaborating, it is important to communicate with each other, define the objects of interest and to create a "joint attention". Joint attention, caused by gaze cues such as creating eye contact or by looking at certain objects (Hamilton, 2016), enables individuals to share their thoughts and intentions while further providing a foundation for several basic social cognitive processes including person and object knowledge, rewarding, empathy, and agency (Stephenson et al., 2021). Reflecting the relevance of gaze following and gaze perception, humans start to follow the gaze of others as early as around eleven months of age (Brooks & Meltzoff, 2005). This mechanism becomes entrenched as a person grows older.

To facilitate the social need of gaze following, the human eye might have evolved to enable individuals to follow each other's gaze as proposed by the Cooperative Eye Hypothesis (CEH) (Tomasello et al., 2007). Human eyes are characterised by a white non-pigmented sclera and a pigmented iris. According to Tomasello et al. (2007), the characteristic white sclera evolved to create a contrast to the iris, and therefore, enhances the visibility of the iris and gaze direction. Exploring the role of the white sclera, Yorzinski and Miller (2020) found that the appearance of the white sclera is important in locating a face in an environment. Next to that, humans are faster to evaluate the gaze of eyes with white, depigmented sclera than the gaze of eyes whose sclera matched the pigmentation of the iris (Yorzinski et al., 2021). Yorzinski et al. (2021) argued that the white sclera evolved in a time when human eyes solely had dark irises to increase the contrast in the eye and facilitate gaze perception.

When comparing the human eye with the eyes of our closest relatives, the distinctive evolution proposed by the CEH becomes clearer. Compared to all living primate species,

humans have clearly visible eye features, the largest ratio of exposed sclera, and an outstanding horizontally elongated eye outline (Kano et al., 2021; Kobayashi & Kohshima, 2001). Investigating gaze perception in human related species, it was discovered that great apes do follow gaze to a certain degree as well, however, they rely more on head-movement cues than on gaze (Tomasello et al., 2007). Furthermore, gaze perception in great apes is dependent on ecologically relevant situations (Whitham et al., 2022). In humans, central to gaze cueing and gaze perception are the eyes. In conditions where only the eyes of a person are visible, for example by wearing a mask, people are still able to perceive gaze direction (Dalmaso et al., 2021).

There is a critical debate around the CEH and taking this into account, makes the perspective presented here more nuanced. In their review, Perea-Garcia et al. (2025) raise the critical remarks to the hypothesis that the human eye might be not unique across species as a white sclera developed in other primate species as well, nor it is uniform, as eye appearance across global samples outside the typically studied Western population displays variation. Furthermore, it was argued that that the communicative advantage proposed by proponents of the CEH might be minor and that robust experimental support for the hypothesis is lacking. This more nuanced perspective highlights the need for novel methods to explore the CEH.

Such a novel perspective could be provided by considering the mechanisms of the "Your Eye Tracker" (YET). As not only for human gaze perception, the contrast between sclera and iris is essential. For eye tracking, this contrast can also be the foundation. This is the case for the YET. YET is a low-budget eye tracker that only requires a USB-camera to observe the eye region. In its head mount it relies on commonly available and low-price material such as a ruler, glue and a frame of glasses. The captured region of the eye is then split into four quadrants and for each quadrant, the average brightness is calculated (Bender, 2024). Based on a linear regression model, it is then predicted where the eye is looking (https://github.com/schmettow/YET). On the average brightness, the contrast between sclera and

iris is the deciding factor as the white sclera increases average brightness and the darker iris decreases it. With this simple approach, the YET is comparable to other established eye tracker in accuracy while predicting gaze in the middle of a computer screen. However, towards the edges, accuracy was found to decrease (Bender, 2024).

In their review on main eye tracking methods, Li et al. (2021) provide an overview of various eye tracking techniques. The first attempts of eye tracking reach back into the 19th century when a relationship between electrode potential on the human skin and eye movement was discovered (Du Bois-Reymond, 1849). This led to the development of Electrooculography (EOG). In EOG, the change in the potential difference in the eye is measured with skin surface electrodes and allows to draw conclusions about the movement of the eye (Haslwanter & Clarke, 2010). In the 1960s, soviet psychologist Yarbus studied saccadic movement of the eye (Yarbus, 1967). Here, to accurately record eye position and movement, suction cups mounted to the eyes were developed and allowed for stable recording over extended time periods (Tatler et al., 2010). Nowadays, commonly used and commercially available eye trackers rely mostly on tracking the reflection of the pupil and the cornea with an infrared camera (Li et al., 2021). For this, the eye is illuminated with near-infrared light which is invisible to the human eye and does not distract. The camera captures the reflection from cornea from this reflection, gaze direction can be predicted (Gonzales-Sanchez et al., 2017).

Next to the hardware used for eye tracking, there is also variation in the software component of eye tracking. Machine learning algorithms such as a Random Forest (RF) algorithm or a Convolutional Neural Network (CNN) algorithm are applied to estimate gaze direction. RF algorithms are used to automate the classification of gaze samples and to train the detection of gaze features from which generalisation about unseen images can be taken (Zemblys et al., 2017). A CNN algorithm estimates gaze by treating gaze estimation as a regression task. The CNN technique predicts gaze based on the two inputs, a face component,

in this, gaze characteristics are extracted from the features of the eyes, and a facial landmark component in which the model incorporates characteristics of facial expressions into predicting gaze (Akinyelu & Blignaut, 2022).

Comparing these approaches to eye tracking with the YET, it stands out that they differ from each other. EOG uses electrical signals in the eye to gather information about eye movement, corneal pupil reflection-based methods rely on a light source such as an infrared light to predict gaze direction, RF and CNN algorithm-based methods require complex deep learning for their estimations. Due to the complexity of the mentioned techniques, expertise in the mechanical characteristics of the hardware or in machine learning of the algorithms is required for being able to perform eye tracking. Further, none of the reviewed techniques is directly using the contrast in the eye for gaze perception which is, according to the CEH, the central element in human gaze perception. The YET on the other hand, applies a simple linear-regression model approach based on the contrast in the eye to predict gaze direction and achieves comparable results. Taking the mechanisms of the YET relying on contrast in the eye and the CEH into consideration it could be argued that the YET functionally models the gaze perception process proposed by the CEH and that possible the processes of eye tracking with YET reflect the cognitive processes of human gaze perception.

To further understand the cognitive events of human gaze perception, additional insight can be gained by taking a biological perspective and to focus on the neural processes that are active here. The neural process of gaze perception takes place in fractions of a second. To investigate this, researchers use methods such as Electroencephalography (EEG) to measure the electrical activity in the brain which allows to compare specific neural responses to an event. These Event-Related Potentials (ERP) can be interpreted as markers for the stage of the ongoing cognitive process. For examining gaze perception, two ERPs are relevant in particular: the P1 component and the N170 component. These components seem to be marking brain activity connected to face perception (P1) and gaze direction processing

(N170) (Tautvydaitė et al., 2022). Tautvydaitė and Burra (2024) tested how much time is needed to detect gaze direction irrespective of other facial cues. Therefore, they measured the onset of the P1 component and of the N170 component In their study, they found an onset of the P1 component, indicating a decoding of head orientation after 20ms. In regard to the N170 component, indicating gaze perception, an onset after 140ms was found. This timeline suggests a two-parted gaze perception process in the brain starting with a rapid assessment of head orientation, followed by a more in-depth analysis of the eyes. The measured processing time of 140ms for eye cues might serve as a benchmark for the time humans need to decode the gaze of another person.

This study

To examine the possibility of whether the mechanisms of the YET resemble human gaze detection and its potential implications for the Cooperative Eye Hypothesis, the following research question will be explored in this study: *Is the eye tracker YET working with the same mechanisms as human gaze perception and what does this mean for the Cooperative Eye Hypothesis?*

To explore the research question, it will be tested whether humans are able to detect gaze direction from human faces based on the same principles the YET is applying. Based on the previous review of the gaze perception, the Cooperative Eye Hypothesis and the YET, the following hypotheses arise for testing:

1. Humans are able to detect gaze direction from a face with 2x2 quadrants representing the average brightness of the eye.
2. The accuracy in detecting gaze direction from a face with 2x2 quadrants representing the average brightness of the eye is equal to the accuracy of detecting gaze direction from human eyes.

To further extend this exploration, the amount of exposure time needed for gaze perception from faces based on the same principles YET is tested. For this, 140ms will be

applied as a benchmark. This is based on the onset of the N170 component measured by Electroencephalography (Tautvydaitė & Burra, 2024). Based on these finding it is hypothesized that:

3. Humans are able to perceive gaze direction after being presented for 140ms with a face with 2x2 quadrants representing the average brightness of the eye.

Methods

Design

To research the three hypotheses of this study, a quantitative design is applied to measure the accuracy of participants in recognising the gaze direction of the stimulus and the exposure time needed. The first independent variable is the condition of the presented image which is categorical and can either be experimental, an image of face with 2x2 quadrants of the average brightness placed in the eye region, or control, a nonmanipulated image. The second independent variable is the exposure time with the ordinal values of 70ms, 140ms, 400ms, 600ms, and 1000ms. This was repeatedly measured for each combination of condition and exposure time, thus 5x2 times.

Participants

The sample consisted of 43 participants. Of those participants, 23 identified as male and 20 identified as female. The age of participants ranged from 18 years to 70 years with a mean of 34,2 years. The sample was gathered via non-probability convenience sampling by advertising the participation for the study on the platform Sona Systems of the University of Twente and by inviting personal contacts of the researchers to participate. There was no other exclusion criteria than that participant must be able to see in order to perceive the stimuli as it was assumed that the examined cognitive processes are universal in humans.

Ethics

This study was approved by the ethics committee of the Faculty of Behavioural, Management, and Social Sciences (BMS) of the University of Twente. Participants were informed about the study's purpose, procedure, potential risks and benefits and provided written informed consent. All data was anonymised by assigning a unique identifier to each participant that left no information about the participant's identity.

Material

Images

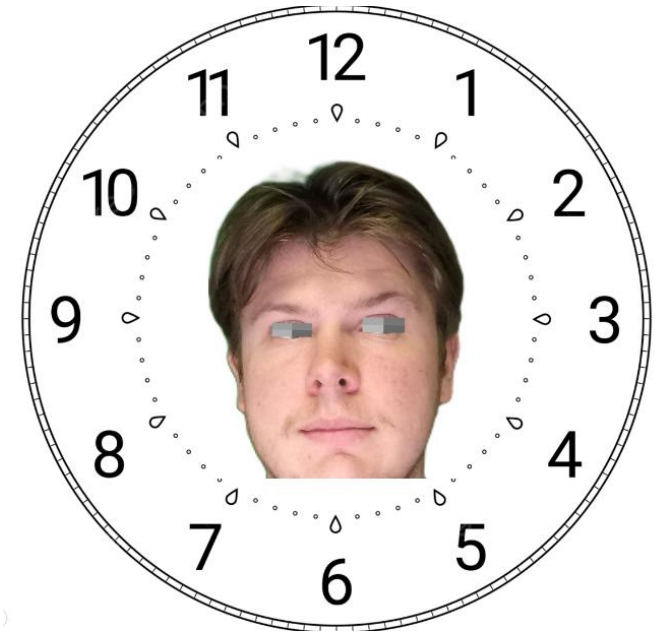
The images used as stimuli in this experiment are portrait photos. These photos used in the experimental condition were edited in such a way that the eyes of the person are turned into four quadrants displaying the average brightness in each of them. In each of the photos, the person is looking towards another direction while keeping their head in the same position. For each, the experimental and the control condition, there were twelve stimuli in which it was gazed at every full hour on the clock. The background of the images was turned white to erase any distractors. Exemplary images for both conditions are displayed in Figure 1 and Figure 2. All 24 stimuli can be found in the Appendix.

Figure 1

Exemplary stimulus for the control condition that is indicating 3 o'clock.

Figure 2

Exemplary stimulus for the experimental condition that is indicating 3 o'clock.



The images were shot with a GoPro HERO5 Black that was placed on a tripod. The camera was placed directly in front of a 55” television screen. The model sat in front of the tv with around 40cm to his face. A clock ranging over the whole screen was displayed to give the model points to focus their gaze on while taking the photos for each gaze around the clock. In order to reduce shadows that could distort the stimulus an LED lamp was placed just below the camera. It was taken care that the eyes of the model were on the same height as the middle of the clock and the camera’s lens was a few centimetres below the eyes. A greenscreen was placed behind the model to replace the background of the photo afterwards. To keep the head as stable as possible a headrest was used. A visual impression of this process can be gained from Figure 3.

Figure 3

Setup of the photo shooting process for the creation of stimuli.

To create the stimuli, an algorithm that used some components of YET's (Your Eye Tracker) algorithm to create the quadrants replacing the eyes was written in Python (Appendix). This standardized the stimulus creation procedure. The algorithm detected the eye region, divided it into four quadrants, calculated the brightness of each quadrant of the eye and replaced the eye with the quadrants such as shown in Figure 1.

Software

The structure of the experiment was designed with PsychoPy (Version 2024.2.4). PsychoPy is an open-source software package that is specialised on creating experiments in psychophysics, experimental psychology and cognitive neuroscience (Peirce et al., 2019). The experiment consisted of ten routines for each combination of the experimental and control condition with the five exposure times of 70ms, 140ms, 400ms, 600ms and 1000ms. The routine was started by clicking the left mouse button on which a three second timer appeared. Then, the stimulus of either experimental or control condition was presented for the respective exposure time. After exposure, participants had to enter the time they perceived as indicated

by the face displayed on the stimulus. For this, a text box appeared to enter the time using the number pad of the computer. This indicated number was verified with the return key which also started the exposure with the next stimulus. The loop type of the stimuli was random, thus, in all ten routines, participants were presented with a randomised stimuli order.

Hardware

The stimuli were presented on two different laptops with 60 Hz displays in the sizes of 15.6 and 17 inches.

Demographic Survey

The demographic data of participants was collected with a survey on the survey tool Qualtrics. In this survey, participants were asked to indicate their assigned participant number from 1 to 43, their age in number and their gender.

Informed Consent

The informed consent was created by adjusting the Informed Consent template of the Faculty of Behavioural, Management and Social Sciences (BMS) of the University of Twente to the context of this study (BMS faculty, 2022). It can be found in the Appendix.

Procedure

The experiment started by inviting the participant to sit down on a table in a quiet room. In front of them, the participant found a sheet with the informed content and a laptop. After filling in the informed consent, participants filled in their participant number, age and gender into the demographic survey. As they finished with this, the researchers explained the procedure of the experiment, the functions of the software and the stimuli. To explain the stimuli, an exemplary image of the experimental stimulus was shown on screen, and the participant was asked to enter which time they perceive to be indicated in the image. After this short tutorial, the starting screen of the experiment was shown and as the participant was ready to start a three second countdown appeared before the presentation of the first stimulus. After exposure with the stimulus, participants entered their perceived time and as they

confirmed their indication the next stimulus was shown. This process was repeated until every stimulus for each full hour on the clock was shown three times in randomised order. This was done for every combination of condition and exposure time. It was started with the experimental condition and an exposure of 70ms. This was followed with the images of the control condition and repeated for each exposure time. Thus, in total 360 observations in ten trials were taken for each participant.

Data Analysis

To prepare the data analysis, a cleaned data set was prepared. The final dataset was reduced to the following variables: the *participant number*, the *condition,* which was either experimental or control, the *presentation time,* which was either 70ms, 140ms, 400ms, 600ms or 1000ms, the *stimulus* for each observation from 1 to 12, the *response* for each observation from 1 to 12, and the dependent variable *deviation* between response and *stimulus*. In total, there were 15,420 observations.

72 observations were excluded as the response was no integer from 1 to 12. To test the data on outlier in the participants, histograms were created to visualize the deviations of participants. In these density plots, no clearly deviating pattern was found and therefore no participant excluded from the dataset.

It had to be taken in mind while calculating the deviation that the responses were given on a circular stimulus. Therefore, it was not possible to simply take the difference between stimulus and response. Instead, first, the absolute value of the difference between stimulus and response was calculated, if the absolute value was larger than six which represents the maximum difference on a circular scale from 1 to 12, this absolute value was subtracted from 12. Finally, the minimum of the two calculations was kept as deviation on clock.

To get an overview of the distribution of the data, box plots with scatters for the observations were created. These box plots displayed the distribution of deviations of the two condition and the five presentation times.

Based upon the variables, this was the formula of the model: Deviation ~ Condition * Exposure + (Condition | Part) + (Condition | Stim).

To test the first and the second hypothesis, the deviation and the effect of the experimental condition on the deviation was reviewed. To test the third hypothesis, the effect of the exposure times on deviation was reviewed.

Results

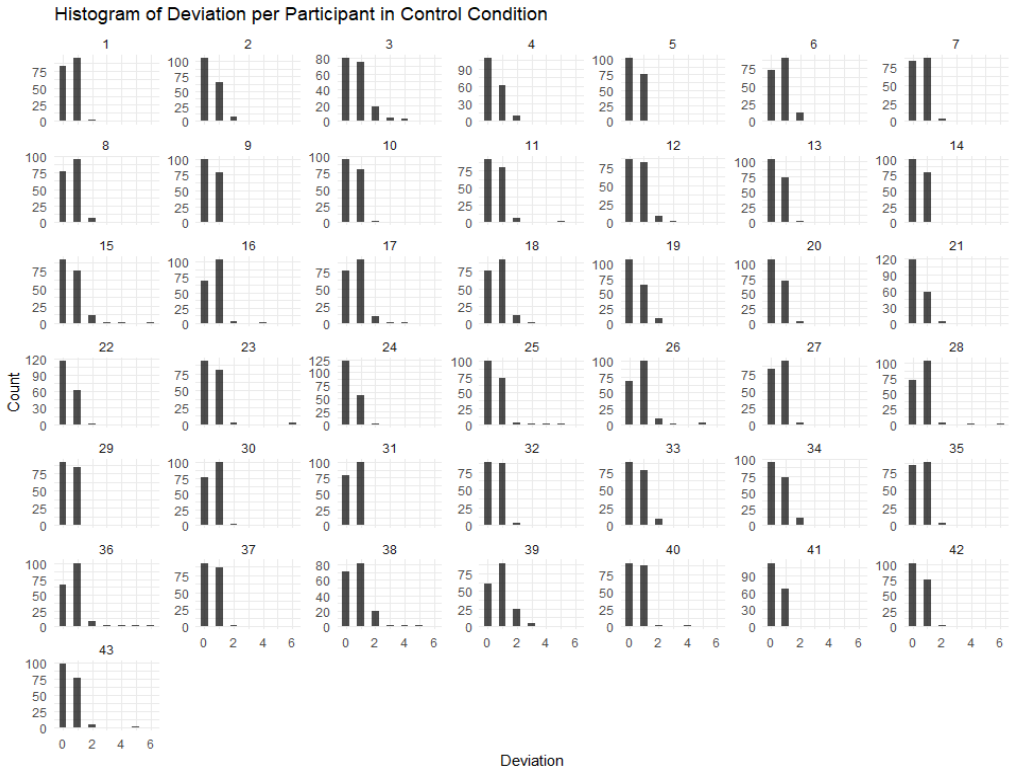
Exploration of Data

Distribution of Deviation per Participant

In the histograms (Figure 4), the distribution of the deviation per participant in the control condition is presented. The distribution shows several similarities for all participants. For each histogram, the peaks are either 0 or 1. Further, all distributions are positively skewed. 27 participants have a peak of 0 deviation in the control condition. The other 16 peak at a deviation of 1. The proportions for deviations of 3,4,5, or 6 are either very low or zero for each participant. The differences between the deviations of 0 and 1 vary, and for participants 21,22 and 24, the bin of 0 as twice as high than the bin for 1. On the other hand, for many others such as participants 3, 7, 12, 29, 32, 35, 37 and 40, the distribution of 0 and 1 is almost equal.

Figure 4

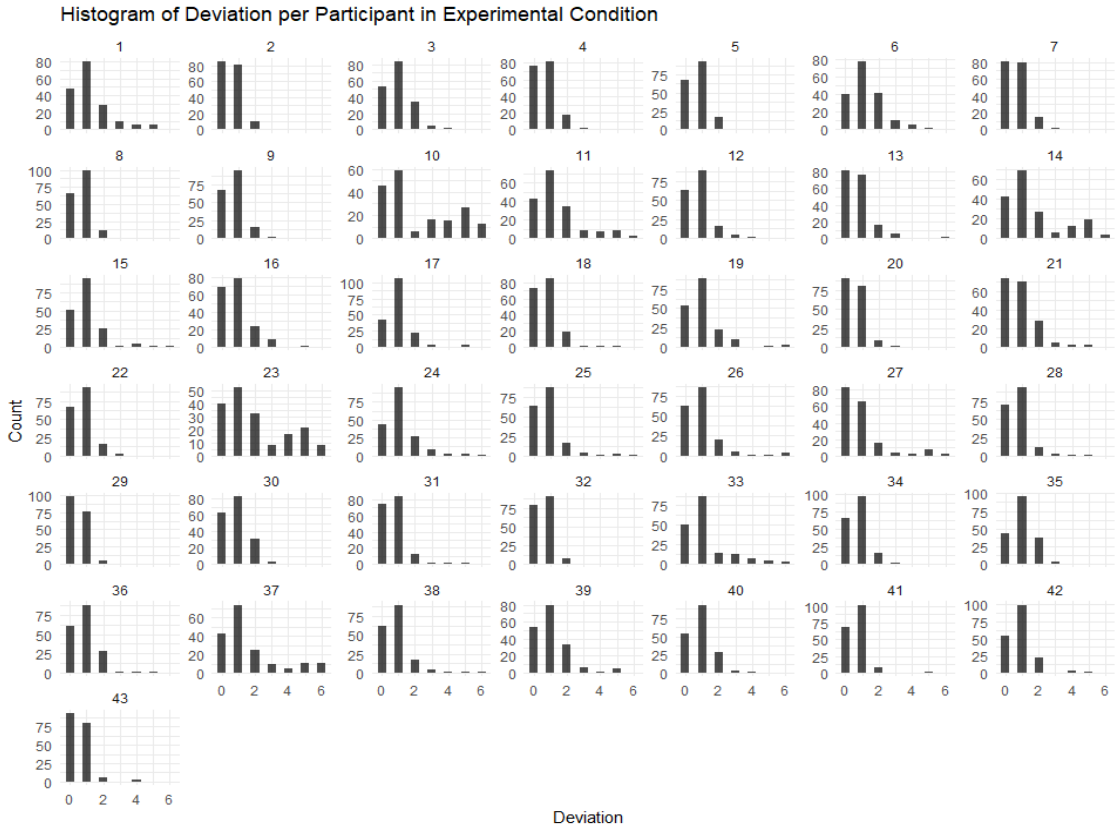
Histograms of Deviation per Participant ID in the Control Condition.



In the distributions of deviation per participants in the experimental condition, there are some similarities and some differences. The histograms of the participants resemble each other by having the highest counts in 0 and 1 deviations (Figure 5). Except for the participants 2, 7, 13, 20, 21, 27, 29, and 43, who have 0 deviations most of time, the other 35 participants have 1 deviations most of the time. There are some participants as well who have proportions of around 10% or more of observations for a deviation of 2 as well. Especially participant 6 stands out as the count of deviation 2 is equal to the count of deviation 0. Examining the counts for deviations 3,4,5, and 6 per participants, it can be noticed that for the majority of participants the proportions are either low or non-existent, however for participant 10, 14, and 23, there is visible rise at a deviation of 5.

Figure 5

Histograms of Deviation per Participant ID in the Control Condition.



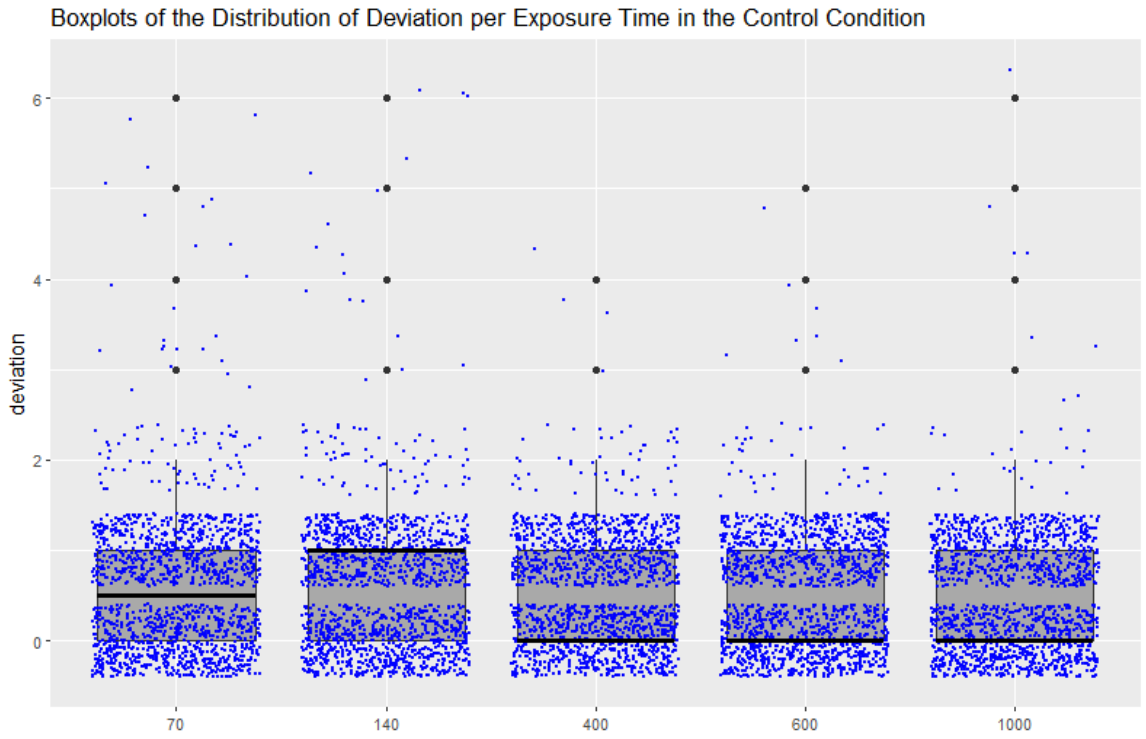
Comparing the distributions of deviation per participant in the control and experimental condition, several similarities can be found. In both conditions, deviations of 0 and 1 display the highest count for each participant. Further, in almost all cases, the distributions are positively skewed from a deviation of 1 onward. In contrast to the control condition, in the experimental condition, the highest count of observations per deviation is mostly at 1 and not at 0. Furthermore, the counts for deviations 0f 2,3,4,5 and 6 are higher in the experimental condition. For a few participants, a rise in counts from deviation 3 to 5 can be seen. These findings indicate that participants were more accurate in gaze perception in the control group and that there is some variation between participants in the experimental group.

Distribution of deviation per condition and presentation time

Throughout all five presentation times in the control condition, the distribution of deviation shows several similarities. This is shown in the boxplots in Figure 6. From 70ms to 1000ms, there are high densities around the low deviations of 0 and 1. The median in each condition lies between 0 and 1. The interquartile range spans for every condition from 0 to 1, and there are no whiskers towards lower deviations, only towards higher deviations. From this, it can be deduced that at least 75% of observations are either 0 or 1. Furthermore, for each exposure time deviations of 2 or higher are outliers. This indicates that the distribution of deviations is positively skewed for each exposure time. The medians differ slightly per exposure time. In 70ms, it is at 0.5 and in 140ms at 1, for the higher exposure times it is at 0. It is also visible that the outliers for the three longer exposure times, the outliers are less than for the shorter exposure times. These findings point out that deviation is slightly less for exposure times of 400ms, 600ms, and 1000ms than for 70ms, and 140ms.

Figure 6

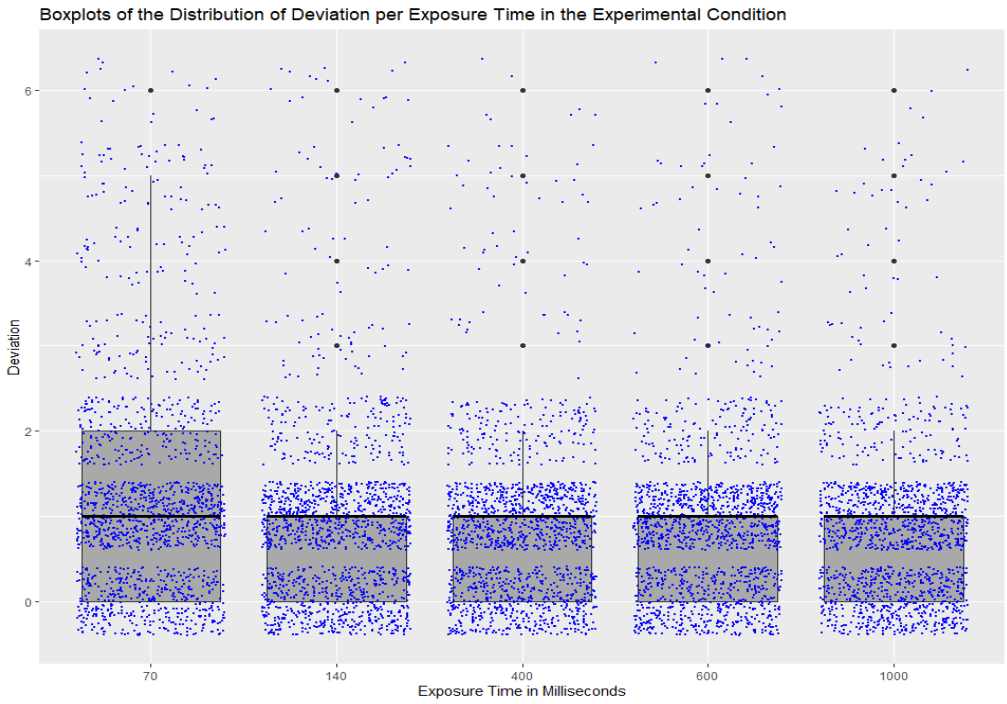
Boxplots of the Distribution of Deviation per Exposure Time in the Control Condition.



In the experimental condition, the distributions per exposure time are most dense at deviations of 0 and 1. This is displayed in the boxplots of Figure 7. The distribution at 70ms is slightly less dense at 0 and 1 in comparison to the other exposures. In all exposure times, the box plots have a median deviation of 1. However, the interquartile range differs between 70ms exposure and the other four. For 70ms, the IQR spans from 0 to 2. In 140ms, 400ms, 600ms and 1000ms, the IQR spans from 0 to 1. In the boxplots, it can be seen that in each exposure time, whiskers only reach towards higher deviations. From this, it might be deduced that at 70ms, at least 75% of observations are a deviation of 0,1 or 2, and that for 140ms, 400ms, 600ms and 1000ms, at least 75% of observations are either a deviation of 0 or 1. For an exposure of 70ms, deviations of 6 are extreme outliers. For the longer exposure times, deviations of 3,4,5, and 6 are extreme outliers. These findings indicate that the distributions of deviations of each exposure time are positively skewed. Furthermore, deviation is decreasing from 70ms to 140ms. From 140ms onwards, no difference in deviations can be observed.

Figure 7

Boxplots of the Distribution of Deviation per Exposure Time in the Experimental Condition.



Comparing the boxplots of both conditions, it stands out that deviation in each plot is centred around deviations of 0 and 1. This centering hints at an ability of participants to perceive gaze direction when presented with faces with YET-eyes and unmanipulated faces. However, the concentration at small deviations was more pronounced in the control condition which points out that participants were more accurate when presented with unmanipulated faces. In the control condition, there were only minor differences between the boxplots per exposure, and thus, no change throughout presentation time can be inferred. On the other hand, in the experimental condition a difference between 70ms and the other four exposure times could be observed as the boxplots indicated a decreasing deviation.

Modeling

Based on the properties of the variables, it was decided to fit a Bayesian generalized linear multi-level model with the formula: Deviation ~ Condition * Exposure + (Condition | Part) + (Condition | Stim). A multi-level model usually consists of fixed and random effects

and is able to account for individual differences in the observations (Schmettow, 2021). From this, predictions can be made at the population level for the whole sample and at the individual level for each singular participant. The fixed effects in this model are the condition, the exposure time and their interaction effect. These population level effects are expected to affect the observations of each participant similarly. The random multi-level effects are each participant and each stimulus per condition. These individual level effects were expected to vary for each participant and for each stimulus per condition. From these effects, the dependent variable deviation was estimated. The dependent variable deviation was treated as non-negative integers from 0 to 6. Respecting this count data with the boundary at 0, a Poisson distribution was applied as family of the model. The link is a logarithmic function. Therefore, the model predicts the log of the expected deviation. The value on the log scale is then exponentiated to obtain the expected value of deviation on the original scale (Schmettow, 2021).

Overdispersion testing

Testing potential overdispersion of the model, it was found that the dispersion ration is 0.815 which pointing towards a slight under dispersion of the model, and thus, less variance in the actual data than predicted by the model. However, the p-value of 1 strongly suggests that there is no evidence of overdispersion found. Followingly, it can be assumed that the model handles variance in the data adequately.

Fixed Effects

For a better overview, the logarithms of the expected intercept and effects were exponentiated (Table 1). Thus, the intercept represents the expected deviation, and the fixed effects can be understood as multiplicative on the intercept. In the control condition with an exposure time of 1000ms represented by the intercept, the expected deviation is 0.44 (CI [0.32, 0.6]). Several deviation increasing factors and no decreasing factors were found. Compared to the intercept, there is a clear effect of the experimental condition which is

expected to increase the deviation approximately 69%. Further, there are clear effects for an exposure of 70ms and 140ms. However, those are not as pronounced and each increases deviation by approximately 17%. There was one clear interaction effect found. If the condition is experimental and the exposure time is 70ms, there is an increasing effect of approximately 31%.

Table 1

Coefficient estimates with 95% credibility limits.

Fixed Effects	*Centre*	*Lower*	*Upper*
Intercept	0.4441424	0.3245297	0.5976229
Condition Experimental	1.6919188	1.2932302	2.2315963
Exposure 600ms	1.0575596	0.9559768	1.1668465
Exposure 400ms	1.0657836	0.9674002	1.1786510
Exposure 140ms	1.1694148	1.0626642	1.2876030
Exposure 70ms	1.1727877	1.0670005	1.2963488
Condition Experimental:Exposure 600ms	1.0407471	0.9204266	1.1839248
Condition Experimental:Exposure 400s	1.0059798	0.8835101	1.1471686
Condition Experimental:Exposure 140ms	1.0115770	0.8968429	1.1493466
Conditionxperimental:Exposure 70ms	1.3145932	1.1599226	1.4844017

In regard to the first hypothesis “humans are able to detect gaze direction from a face with 2x2 quadrants representing the average brightness of the eye”, this means that it can be accepted as the expected deviation for the experimental condition at an exposure of 1000ms is less than 30 degrees, and therefore, indicates that humans were mostly able to detect gaze direction. Furthermore, there is no significant difference in the exposures of 600ms, 400ms and 140ms.

In contrary, the second hypothesis “the accuracy in detecting gaze direction from a face with 2x2 quadrants representing the average brightness of the eye is equal to the accuracy of detecting gaze direction from human eyes” must be rejected because there is a significant

effect of the condition on the accuracy in gaze detection leading to an increase in deviation for the experimental condition.

The third hypothesis “humans are able to perceive gaze direction after being presented for 140ms with a face with 2x2 quadrants representing the average brightness of the eye” can be accepted as there is no significant increase of the expected deviation for an exposure from 140ms onwards. Thus, humans can be assumed to be able to detect gaze direction if they are presented with the stimulus for 140ms or longer.

Random Effects

In Table 2, the standard deviations of logarithms to the population level for the intercept and the experimental condition are displayed. Exponentiated, these standard deviations display factors on the population level. The standard deviations of participants are a factor of 1.22 on the intercept and a factor of 1.42 on the effect of the experimental condition. For the stimuli, the factors are 1.63 on the intercept and a factor of 1.49 on the effect of the experimental condition. The standard deviation of participants on the intercept indicates a rather low variability. On the contrary, there is a moderate variability in the effect of the experimental condition per participant. Focusing on the stimuli it stands out that there is a high variability per stimulus at intercept level and a moderate variability per stimulus in the effect of the experimental condition. These findings it can be deduced that there is considerable variability in the expected deviation that the fixed effects cannot fully account for.

Table 2

Population-level coefficients with random effects standard deviations.

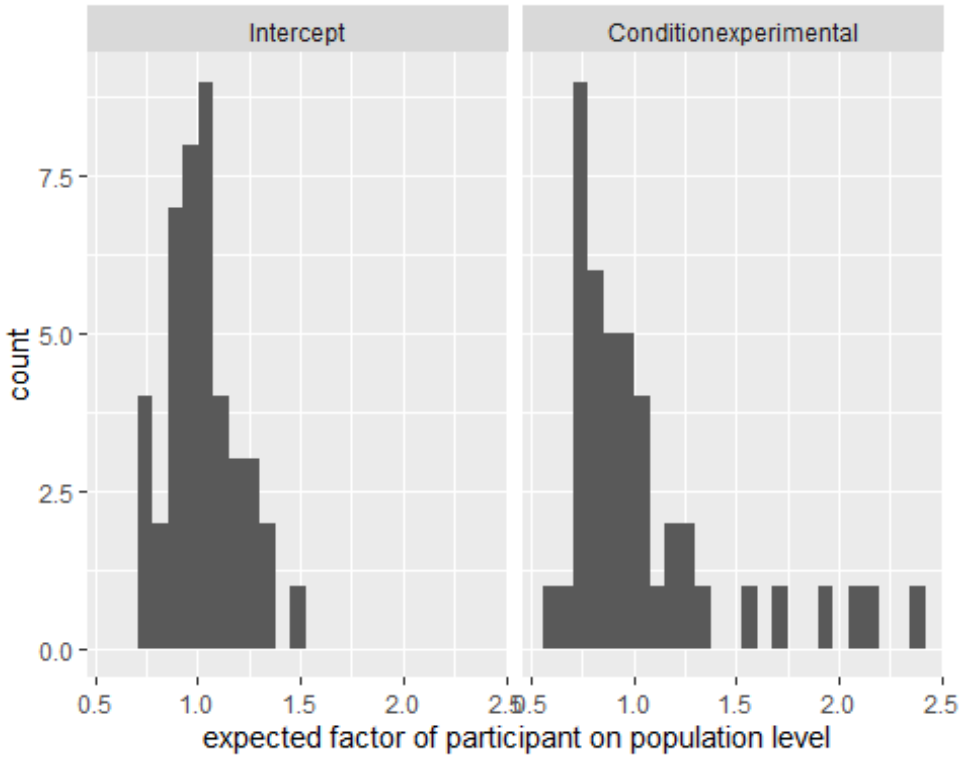
Fixed Effects	*Centre*	*Lower*	*Upper*	*SD Participants*	*SD Stimulus*
Intercept	-0.8116101	-1.1253784	-0.5147953	0.1958375	0.4902023
Condition experimental	0.5258633	0.2571431	0.8027172	0.3532276	0.3974395

Variability in Participants

Taking a closer look on the variability between participants, it appears that the multiplicative effects derived from the log-scale random effects of participants on the intercept is peaking slightly below 1 (Figure 8). The smallest effect is approximately at 0.7 and the highest at 1.5. The distribution is relatively symmetrical which suggests a small variance in the random effects. On the contrary, focusing on the distribution of the multiplicative effects derived from the log-scale random effects on the experimental condition, the distribution is positively skewed and peaking at around 0.75. The smallest effect is approximately 0.6 and the highest 2.4. This strong positive skewness is a characteristic of a log-normal distribution, the expected shape for the exponentiated random effects. These findings indicate that the multiplicative effects derived from the log-scale random effects of participant on the intercept display relatively consistent variance. Thus, confidence in the population-level estimate can be strengthened. On the other hand, the strong skewness in the distribution of the multiplicative participant effects on the fixed effect of the experimental condition indicate heterogeneity. The majority of participants display a decreasing or similar effect on the effect of the experimental condition. This is contrasted by a small number of participants that displays much larger multiplicative effects ranging up to 2.4 times the population-level effect.

Figure 8

Histograms of expected participant factor on population level.



Multiplicative Effects of Stimuli

Examining the multiplicative effects of each stimulus on the intercept that are displayed in Figure 9 and 10, it can be seen that stimuli 1, 5, 8, and 9 increase the intercept and thus have an increasing effect on the dependent variable deviation. The stimuli 3,6, and 10 have an decreasing effect on the dependent variable. No clear effect is observed for stimulus 2, 4, 7, 11, and 12 as their 95% credibility intervals include 1. Focussing on the effects of stimuli on the effect of the experimental condition, the stimuli 3, 6 and 10 display an increase, and the stimuli 7, 8, 9 display a decrease. The 95% credible intervals of stimuli 1, 2, 4, 5, 11, and 12 contain 1 and therefore no clear effect was observed.

Figure 9

Multiplicative Effects of Stimuli on Intercept.

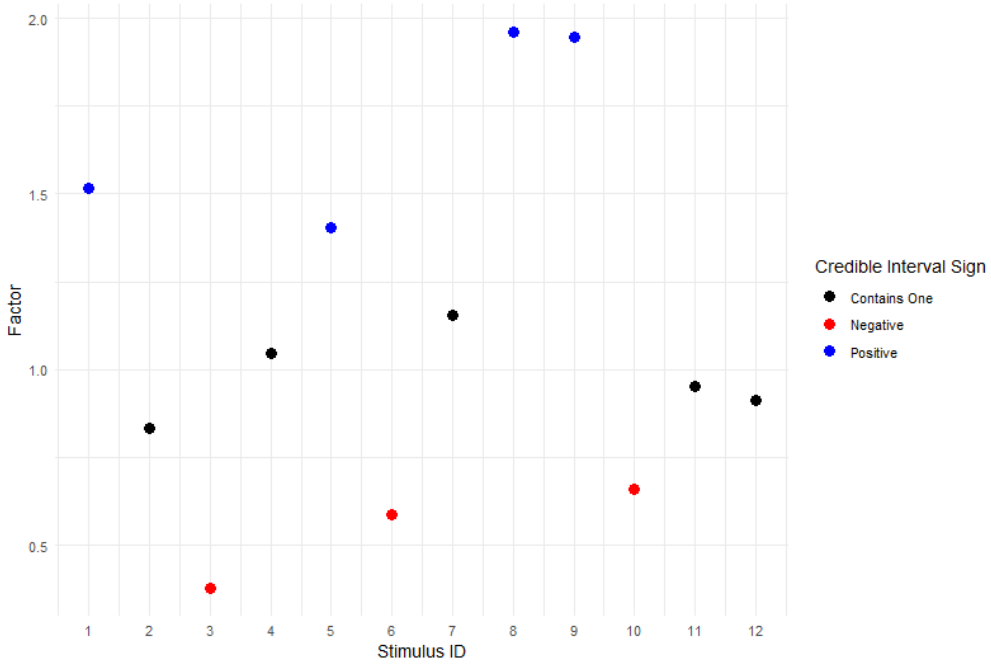
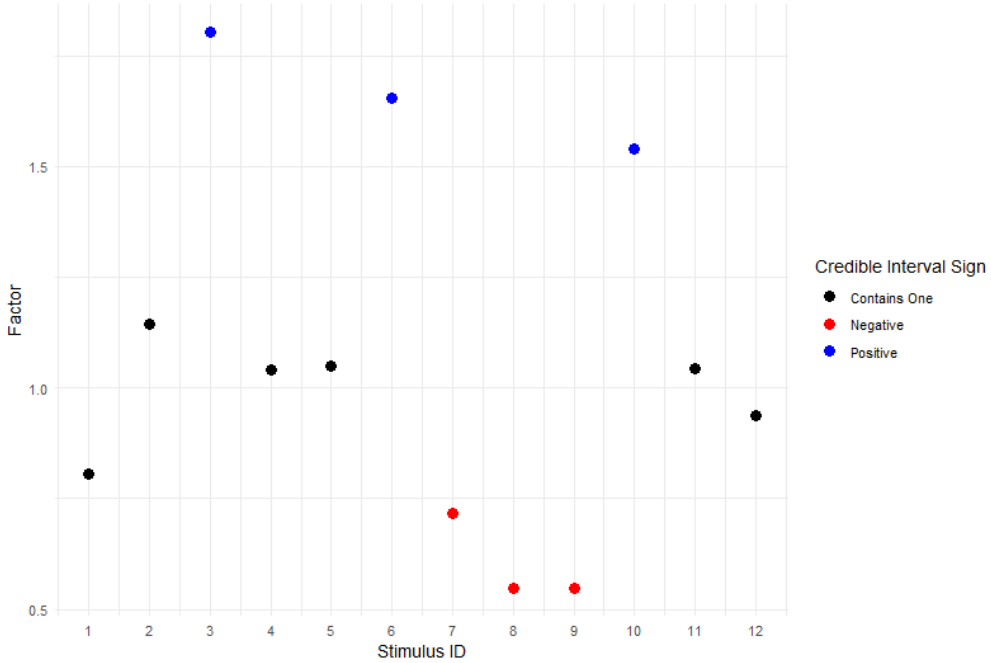


Figure 10

Multiplicative Effect of Stimuli on Experimental Condition Effect.



Discussion

Based on the findings of this study and in regard to the hypotheses, it was found that humans are able to perceive gaze direction from the YET-quadrants. However, in comparison to gaze perception from human eyes, participants were less accurate. The deviation in gaze perception was around 70% higher than in the control condition in which participants were presented with a regular human face. The deviation of telling gaze direction from faces with YET-quadrants instead of eyes decreased strongly as participants were presented with the stimulus for 140ms instead of 70ms. From an exposure with the stimulus for 140ms and longer, the deviation did not differ significantly.

Interpretation of Findings

These findings indicate that the mechanism of the YET might resemble the actual human gaze perception processes as participants were clearly able to detect gaze direction from the YET-quadrant eyes. Furthermore, this can be interpreted as support for the Cooperative Eye Hypothesis because the results show that gaze perception is possible solely by relying on the brightness of four quadrants, and therefore, can be reduced to the contrast in the eye as deciding cue. This is in favour of the hypothesis which states that the eye evolved to enlarge the contrast between sclera and iris making it possible to infer gaze direction from looking into someone else's eyes (Tomasello, 2007). Another point indicating the resemblance between the YET-mechanism and human gaze perception is that there is a large difference in deviation between exposure times of 70ms and 140ms. That deviation is smaller from 140ms onwards and that there is no clear difference between exposure of 140ms, 400ms, 600ms and 1000ms can be understood as a representation of the finding from the study of Tautvydaitė and Burra (2024) that the onset of the event-related potential that is connected to gaze perception takes place at 140ms.

Contrasting to these findings in line with the hypotheses, gaze perception from YET-quadrant eyes was significantly less accurate than gaze perception from a regular face. This is

important to mention as it points out that human gaze perception differs from the one of the eye tracker. While this does not mean that is no connection between the gaze perception of human and eye tracker, it might indicate that there might exist other additional processes other than perceiving the contrast in the eye.

Gaze perception is the product of many factors including social circumstances. For instance, the gaze-cueing effect that humans direct their attention on locations gazed at by another is influenced by the perceived trustworthiness of the cueing face (Driver et al., 1999; Ding et al., 2024). This might be the due to the aspect that, to establish cooperation, it is important to assess whether to other can be trusted or not. Further, humans tend to expect that gaze is directed towards them, thus perceive gaze direction in such a way (Mareschal et al., 2014). Next to these social cues, the spatial perception of another person's head orientation provides significant information about gaze perception as well (Balsdon & Clifford, 2017). Additionally, neurodiversity, and specifically, the autism spectrum disorder (ASD) influence gaze perception. In their review on measurable markers of ASD, Tiede and Walton (2020) mentioned gaze avoidance, and hence an impact on gaze perception and joint attention, as a marker.

From placing the findings of this study into the context of the previous research exploring gaze perception and its factors, it might be deduced that the Cooperative Eye Hypothesis can be strengthened as the contrast enhancement of seeing the YET-quadrants in position of the eye enabled participants to detect the gaze direct. This is in line with the hypothesis which describes that the eye evolved to enhance contrast in the eye. Followingly, it could be argued that the evolution of the eye is reflected in the function of the YET. This claim is further supported by the decrease in deviation for an exposure with the stimulus containing YET-quadrants as eyes for longer than 140ms.

Nevertheless, the difference in deviation for the experimental condition and the control condition is an indication that gaze perception of the YET and of a human being is not

identical. As mentioned above, a number of factors such as social circumstances, head orientation, and individual differences in neurodiversity are influencing human gaze perception, and additionally to the ones referred to, there might be number of further influences too. Hence, this study is focused on the factors of contrast in the eye and exposure time to a face and it is not accounted for other influences that might affect gaze perception as well.

Limitations

Visual Masking

It is important to mention that in the experiment of this study no visual mask was used immediately after the exposure with the stimuli, and thus, the findings regarding the needed exposure time to detect gaze direction must be viewed with caution. In psychophysics, visual masking describes suppression of a visual stimulus by presenting another stimulus which is then the masking stimulus (Colman, 2008). A masking stimulus is either shown immediately before, after, or before and after the actual stimulus to "overwrite" the visual information of the actual stimulus (Bachmann, 2018). These methods are called "forwards masking", "backward masking" or "sandwich masking". Commonly, meaningless patterns or visual noise consisting of geometrical shapes or unrelated images are used as masking materials.

Consequently, as no visual masking was used in this experiment, the presented stimulus was not overwritten which means that the exposure time could be blurred, and the participants had more time to process the information of the stimulus. This is called iconic memory which describes the storage of visual information in the short-term memory after a stimulus was removed (Sperling, 1960). This allows for elongated access to the information. The elongated access to the iconic memory can last for several hundred milliseconds after the exposure depending on the stimulus (Pratte, 2018). In this persistence of iconic memory, Yi et al. (2017) found that colour information can be longer accessed than information about presented numbers. This is an important aspect as in the performed experiment of this study,

participants had to detect gaze direction based on the contrast between the quadrants displaying the respective average brightness, and therefore, had to decide following the colour information.

Further applied to this study, it means that the results focused on the effect of exposure time on deviation might be unreliable and that the earlier statement that the decreasing deviation as exposure time rises is an indication of the resemblance between gaze perception of the YET and humans potentially must be withdrawn. On the other hand, the results show that there was only an increasing effect on deviation for an exposure time of 70ms and that from 140ms onwards deviation did not differ significantly. Thus, an actual correlation between exposure and deviation and followingly a minimal exposure matching the one for the face with regular eyes might be hinted at. On the findings regarding the ability to perceive gaze direction from YET-quadrants and the differences in deviation per eye set, the missing masking should have no influence as the exposure time was not of interest for the hypotheses. In the opposite, the short exposure time still fulfilled the purpose to hinder the participants from memorising the stimuli.

60 Hertz Screens for Presenting Stimuli

The stimuli in the experiment in PsychoPy were presented on screens with 60Hz refreshment rates. This might have produced deviation in the exposure time with the stimuli. A refreshment rate of 60Hz implies that frames refresh every 16.66ms. However, 70ms and 140ms are no multiples of 16.66ms. Thus, the 70ms stimuli might have actually been presented for 83.3ms or five frames and the 140ms stimuli for 150ms or nine frames. For the 400ms, 600ms, and 1000ms stimuli this did not have an effect as they are multiples of 16.66.

Stimuli

Another limitation for this study appears from taking the multiplicative effects of the stimuli into perspective. Here, pronounced differences between the effects of the stimuli were observed. Moderately high standard deviations for the multiplicative effect on intercept and

on the effect of the experimental condition were found. Some stimuli decreased the expected deviation, and others had an increasing effect.

A potential explanation for these different effects could be the additional information in the stimuli that was not accounted for in the model. Exemplary for additional information could be the mimicry of the face in the stimuli. To visualise this, in Figures 11 and 12, the stimulus for the control condition and the experimental condition indicating 6 is shown. Focusing on the eye region, it appears that the eye lids of the face point downwards which is the natural occurrence when a human looks to the bottom. As the visible part of the eye is almost symmetrical in this position, the quadrants in the experimental condition are similar in brightness. Thus, the mimicry potentially had an impact on the deviation per stimulus which was not accounted for in this experiment. In the context of the research question and the hypotheses, this impact of the mimicry information might have affected the gaze perception of the participants as it provided additional information about the gaze direction.

Figure 11

Stimulus indicating 6 o'clock in the Control Condition.

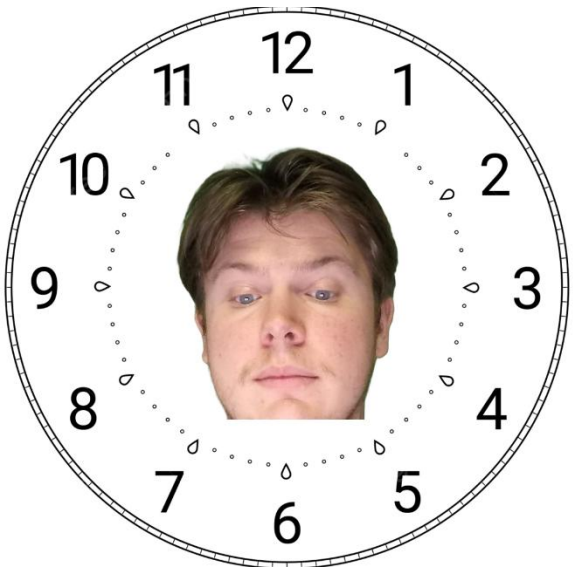
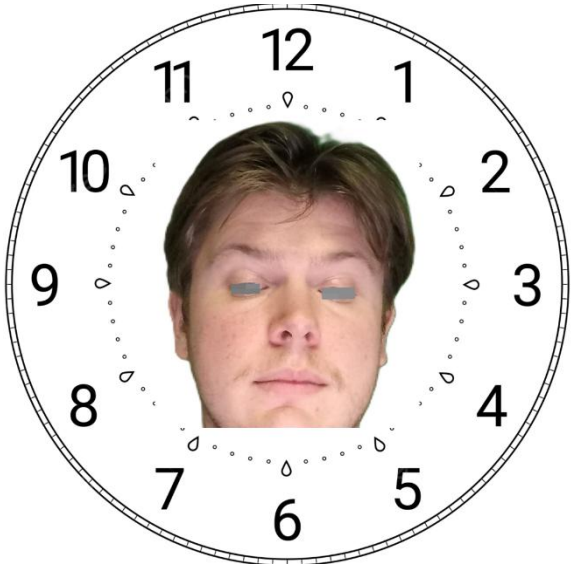


Figure 11

Stimulus indicating 6 o'clock in the Control Condition.



Variability in Participants

There was a high variability in the multiplicative effect of participants on the effect of the experimental condition found. While for most participants the effect of the experimental decreased, and with it the expected deviation, for some participants, an increase was found that ranged up to 2.4 times of the population-level. In this study, it was not accounted for differences between participants as it was assumed that the examined cognitive processes were universal. The found variability does not disprove this assumption, especially as variance in the control condition was found differ little from the population level. However, the variability in participants in the experimental condition might be an indication that there was an effect not accounted for in this study. This potentially affected the found results.

Future research

Addressing limitations

To evaluate the limitations of this study and gain more confidence in the findings of this study, future research is advised to take the following aspects into consideration. First, a study design to explore the minimum of exposure time needed to encode the gaze of a person

must include a visual mask following on the stimuli presentation. This allows to control the time span a participant can access the visual information of the stimuli in their short-term memory. Using the insight from this study, it can be focused on examining the minimum exposure in the context of encoding faces with YET-quadrants as eyes and whether the findings of Tautvydaitė and Burra (2024) that human gaze perception displays an onset after 140ms apply here as well.

Second, the variability between participants effects on the effect of the experimental conditions, the variability between stimuli effects and the respective population-level effects must be further explored. To further find out about the variability in stimuli, an option could be to virtually model a human face with eyes and YET-eyes gazing into different directions. This way, it might be possible to control for the varying mimicry and to keep it identical for each stimulus. Another advise for further exploration is to focus on the role of mimicry in human gaze perception and in eye tracking itself. As described by Kano et al. (2021), the human eye is characterised by a visible eye outline, visible iris, and a horizontally elongated form. In the context of the variability in stimulus it might be interesting to see whether the characteristic horizontally elongated form of the eye influences the accuracy of perceiving the gaze direction. Potential research could explore whether there are differences in accuracy per direction of the gaze. Focusing on eye tracking, the relation of mimicry and gaze direction in relation to the accuracy of gaze detection might be explored. This potential research could focus on finding possibilities to integrate information from mimicry into the eye tracking process.

As there was considerable variability found in the participants regarding their deviation in the experimental condition, future research might explore the underlying reasons for this deviation. To do this, it might be focused on finding characteristics of participants that potentially impact gaze perception from faces with YET-eyes.

Unanswered questions

From the finding of this study that the mechanisms of the YET might resemble the cognitive processes of detecting gaze direction from the contrast in the eye, the question arises whether the mechanism of the YET resembles the cognitive processes in gaze perception closer than other types of eye trackers. To examine this question, the different types of existing eye tracking techniques might be compared and a similar study to this could be performed for the other eye trackers. It could be tested whether humans are able to perceive gaze direction from a visualisation of their mechanisms.

Another question arising from the findings of this study is if iconic memory works better when presented with a stimulus displaying an unmanipulated human face than when presented with a face with YET-eyes. In this study, there was a stronger decreasing effect for the shortest exposure time of 70ms found in the experimental condition in comparison to the control condition. Further, it is the case that several types of information such as colour and numeric information impact iconic memory differently (Yi et al., 2017). This might raise the assumption that the stimuli in the control condition and the experimental condition differ in their effect on iconic memory. Future research might explore this question to deepen the knowledge about the characteristics of the used stimuli.

Implications

Theoretical Implications

The findings of this study imply that the Cooperative Eye Hypothesis can be strengthened. Participants were able to extract information about gaze direction from eye that were reduced to 2x2 quadrants that display the average brightness in their part of the eye. This minimalizes the eye in its function of indicating gaze direction into a basic contrast between the for directions: up, down, left and right. The ability of participant to extract the necessary information from the YET-eyes to perceive gaze direction supports the claim of the

Cooperative Eye Hypothesis that the eye evolved by increasing the contrast in the eye to enable gaze following.

Practical Implications

From a practical perspective, these findings could be used in areas that focus on communication and non-verbal communication cues. The YET and the visualisation of the 2x2 quadrants it is using to predict gaze direction can be used a helpful tool to give an impression of the mechanisms of human gaze perception. A potential application of this knowledge is in educational environments to teach about human gaze perception. Another potential field for using the information from this study could be the design of the eyes and faces of humanoid robots. Here, the findings can be possibly applied to tailor humans-robot communication to the cognitive mechanism of the human user. In this process, an important focus point, emerging from this study, is to create eyes with clear contrast to enable communication between robots and humans based on gaze cues.

References

Akinyelu, A. A. & Blignaut, P. (2022). Convolutional Neural Network-Based Technique for Gaze Estimation on Mobile Devices. *Frontiers in Artificial Intelligence*, *4*. https://doi.org/10.3389/frai.2021.796825

Balsdon, T. & Clifford, C. W. G. (2017). A bias-minimising measure of the influence of head orientation on perceived gaze direction. Scientific Reports, 7(1). https://doi.org/10.1038/srep41685

Bachmann, T. (2018). Visual masking: Contributions from and comments on Bruce Bridgeman. *Consciousness And Cognition*, *64*, 13–18. https://doi.org/10.1016/j.concog.2018.04.013

Behavioural, Management and Social Sciences Faculty of the University of Twente. *Informed consent form template for research with human participants*. https://www.utwente.nl/en/bms/research/ethics-domainHSS/explanation-webapplication/informed-consent-procedure/

Bender, F.R. (2024) Evaluation of the low budget do it yourself eye tracking system: Your Eye Tracker (YET).

Brooks, R. & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Colman, A. M. (2008). A Dictionary of Psychology. In *Oxford University Press eBooks*. https://doi.org/10.1093/acref/9780199534067.001.0001

Dalmaso, M., Zhang, X., Galfano, G., & Castelli, L. (2021). Face Masks Do Not Alter Gaze Cueing of Attention: Evidence From the COVID-19 Pandemic. *I-Perception*, *12*(6). https://doi-org.ezproxy2.utwente.nl/10.1177/20416695211058480

Ding, X., Gan, J., Xu, L., Zhou, X., Gao, D. & Sun, Y. (2024). Not to follow because of distrust: perceived trust modulates the gaze cueing effect. Psychological Research, 88(8), 2195–2210. https://doi.org/10.1007/s00426-024-02000-4

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E. & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. Visual Cognition, 6(5), 509–540. https://doi.org/10.1080/135062899394920

Du Bois-Reymond, E. (1849). *Untersuchungen über thierische Elektricität* [Investigations on animal electricity]. G. Reimer.

Gonzalez-Sanchez, J., Baydogan, M., Chavez-Echeagaray, M. E., Atkinson, R. K. & Burleson, W. (2017). Affect Measurement: A Roadmap Through Approaches, Technologies, and Data Analysis. In Elsevier eBooks (S. 255–288). https://doi.org/10.1016/b978-0-12-801851-4.00011-2

Hamilton, A. F. C. (2016). Gazing at me: the importance of social meaning in understanding direct-gaze cues. *Philosophical Transactions Of The Royal Society B Biological Sciences*, *371*(1686), 20150080. https://doi.org/10.1098/rstb.2015.0080

Haslwanter, T., & Clarke, A. H. (2010). Eye movement measurement: electro-oculography and video-oculography. Handbook of Clinical Neurophysiology, 9, 61-79. https://doi.org/10.1016/S1567-4231(10)09005-2

Kano, F., Furuichi, T., Hashimoto, C., Krupenye, C., Leinwand, J. G., Hopper, L. M., Martin, C. F., Otsuka, R. & Tajima, T. (2021). What is unique about the human eye? Comparative image analysis on the external eye morphology of human and nonhuman great apes. Evolution And Human Behavior, 43(3), 169–180. https://doi.org/10.1016/j.evolhumbehav.2021.12.004

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal Of Human Evolution*, *40*(5), 419–435. https://doi.org/10.1006/jhev.2001.0468

Li, Z., Guo, P. & Song, C. (2021). A Review of Main Eye Movement Tracking Methods. *Journal Of Physics Conference Series*, *1802*(4), 042066. https://doi.org/10.1088/1742-6596/1802/4/042066

McKay, K. T., Grainger, S. A., Coundouris, S. P., Skorich, D. P., Phillips, L. H., & Henry, J. D. (2021). Visual attentional orienting by eye gaze: A meta-analytic review of the gaze-cueing effect. *Psychological Bulletin, 147*(12), 1269–1289. https://doi.org/10.1037/bul0000353

Mareschal, I., Otsuka, Y. & Clifford, C. W. G. (2014). A generalized tendency toward direct gaze with uncertainty. Journal Of Vision, 14(12), 27. https://doi.org/10.1167/14.12.27

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E. & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Perea-García, J. O., Teuben, A. & Caspar, K. R. (2025). Look past the cooperative eye hypothesis: reconsidering the evolution of human eye appearance. *Biological Reviews/Biological Reviews Of The Cambridge Philosophical Society*. https://doi.org/10.1111/brv.70033

Pratte, M. S. (2018). Iconic Memories Die a Sudden Death. *Psychological Science*, *29*(6), 877–887. https://doi.org/10.1177/0956797617747118

Schmettow, M. (2021). Generalized Linear Models. In *New statistics for design researchers: A Bayesian Workflow in Tidy R*. Human–Computer Interaction Series. Springer, Cham. https://doi.org/10.1007/978-3-030-46380-9_6

Schmettow, M. (2021). Multilevel models. In *New statistics for design researchers: A Bayesian Workflow in Tidy R*. Human–Computer Interaction Series. Springer, Cham. https://doi.org/10.1007/978-3-030-46380-9_6

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74*(11), 1–29. https://doi.org/10.1037/h0093759

Stephenson, L. J., Edwards, S. G., & Bayliss, A. P. (2021). From Gaze Perception to Social Cognition: The Shared-Attention System. *Perspectives On Psychological Science*, *16*(3), 553–576. https://doi.org/10.1177/1745691620953773

Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., & Velichkovsky, B. M. (2010). Yarbus, eye movements, and vision. *i-Perception*, 1(1), 7–27. https://doi.org/10.1068/i0382

Tautvydaitė, D. & Burra, N. (2024). The Timing of Gaze Direction Perception: ERP Decoding and Task Modulation. NeuroImage, 295, 120659. https://doi.org/10.1016/j.neuroimage.2024.120659

Tautvydaitė, D., Mares, I., Rahman, M. S., Burra, N. & Senju, A. (2022). Effect of perceived eye gaze on the N170 component – A systematic review. Neuroscience & Biobehavioral Reviews, 143, 104913. https://doi.org/10.1016/j.neubiorev.2022.104913

Tiede, G. M. & Walton, K. M. (2020). Social endophenotypes in autism spectrum disorder: A scoping review. Development And Psychopathology, 33(4), 1381–1409. https://doi.org/10.1017/s0954579420000577

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of human evolution*, *52*(3), 314–320. https://doi.org/10.1016/j.jhevol.2006.10.001

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two Key Steps in the Evolution of Human Cooperation. *Current Anthropology*, *53*(6), 673–692. https://doi.org/10.1086/668207

Whitham, W., Schapiro, S. J., Troscianko, J., & Yorzinski, J. L. (2022). Chimpanzee (Pan troglodytes) gaze is conspicuous at ecologically-relevant distances. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-13273-3

Wolf, W., Thielhelm, J. & Tomasello, M. (2023). Five-year-old children show cooperative preferences for faces with white sclera. *Journal Of Experimental Child Psychology*, *225*, 105532. https://doi.org/10.1016/j.jecp.2022.105532

Yarbus, A. L. (1967). *Eye Movements and Vision*. Plenum Press

Yi, W., Kang, M. & Lee, K. (2017). Visual attribute modulates the time course of iconic memory decay. *Visual Cognition*, *26*(4), 223–230. https://doi.org/10.1080/13506285.2017.1416007

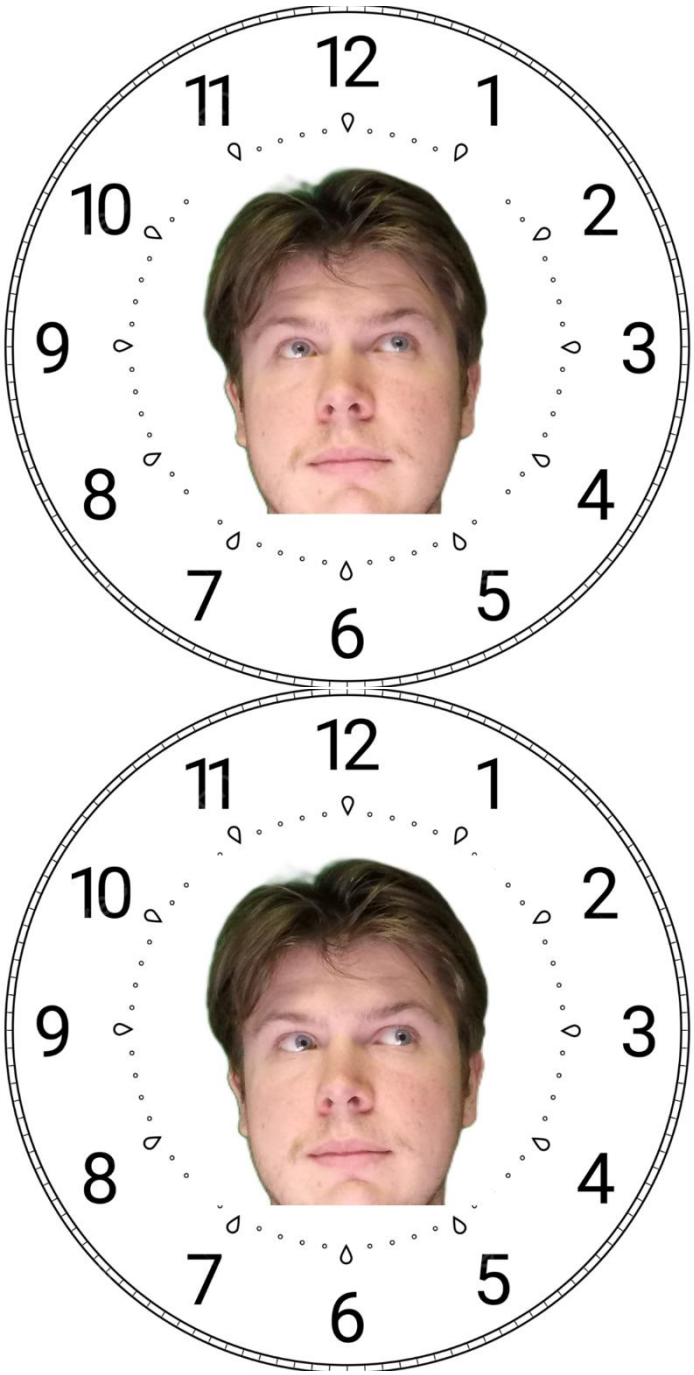
Yorzinski, J. L. & Miller, J. (2020). Sclera color enhances gaze perception in humans. *PLoS ONE*, *15*(2), e0228275. https://doi.org/10.1371/journal.pone.0228275

Yorzinski, J. L., Thorstenson, C. A. & Nguyen, T. P. (2021). Sclera and Iris Color Interact to Influence Gaze Perception. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.632616

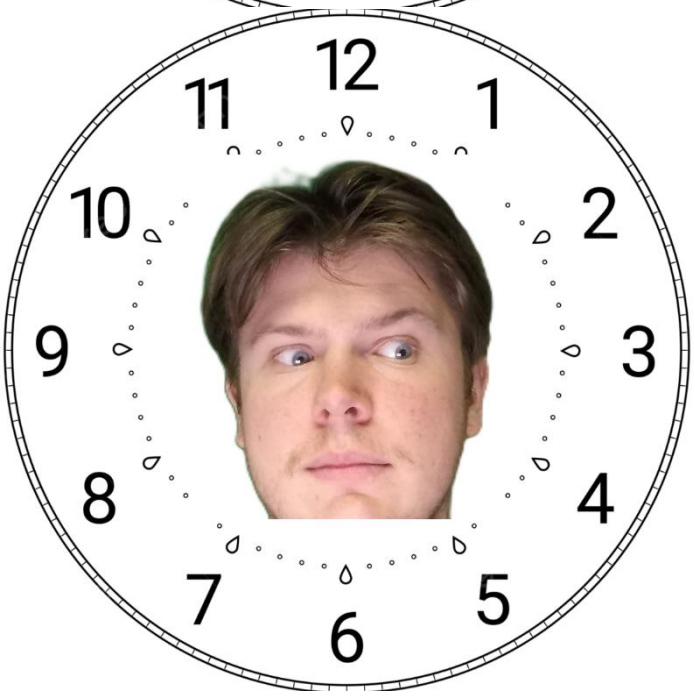
Zemblys, R., Niehorster, D. C., Komogortsev, O. & Holmqvist, K. (2017). Using machine learning to detect events in eye-tracking data. *Behavior Research Methods*, *50*(1), 160–181. https://doi.org/10.3758/s13428-017-0860-3

Appendix

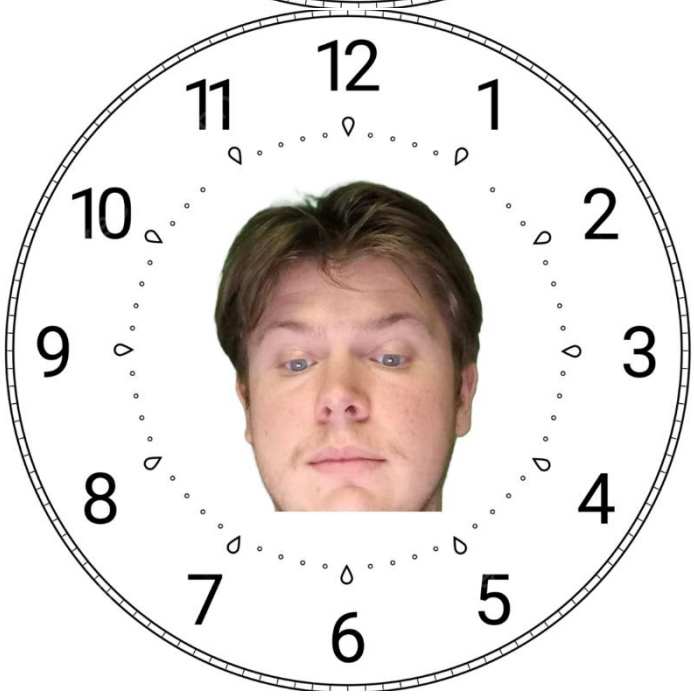
Stimuli Control Condition from 1 to 12 o'clock

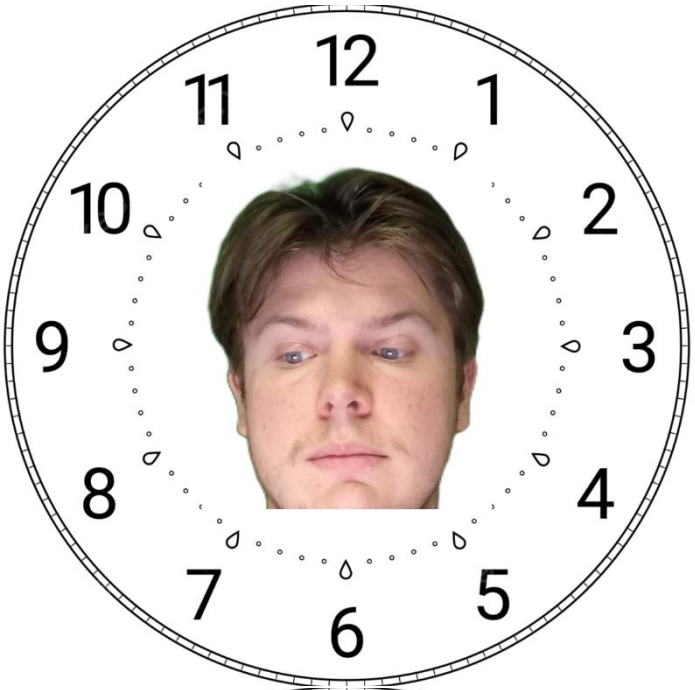


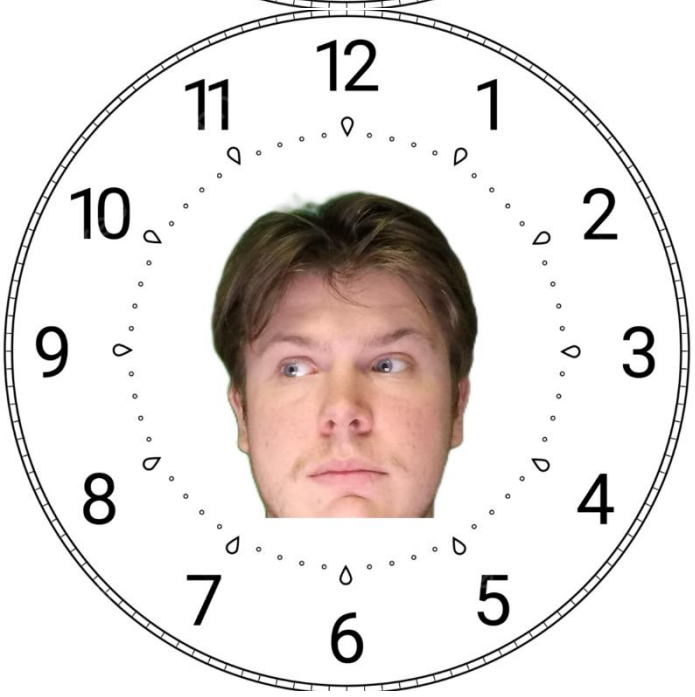

12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11

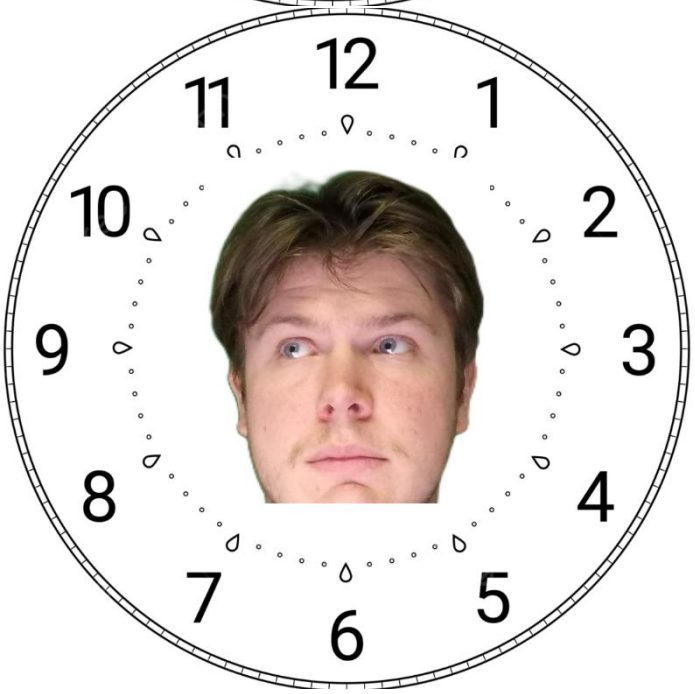

12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11

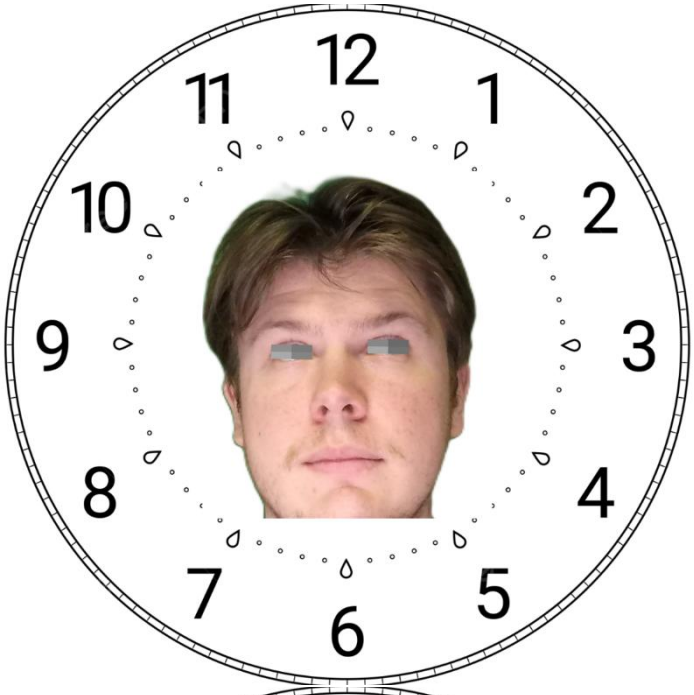

12
1
2
3
4
5
6
7
8
9
10
11

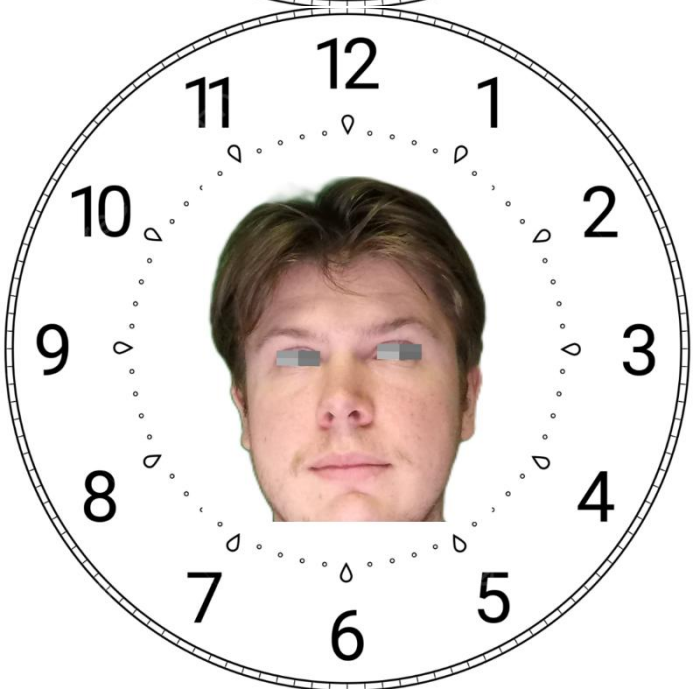

12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11

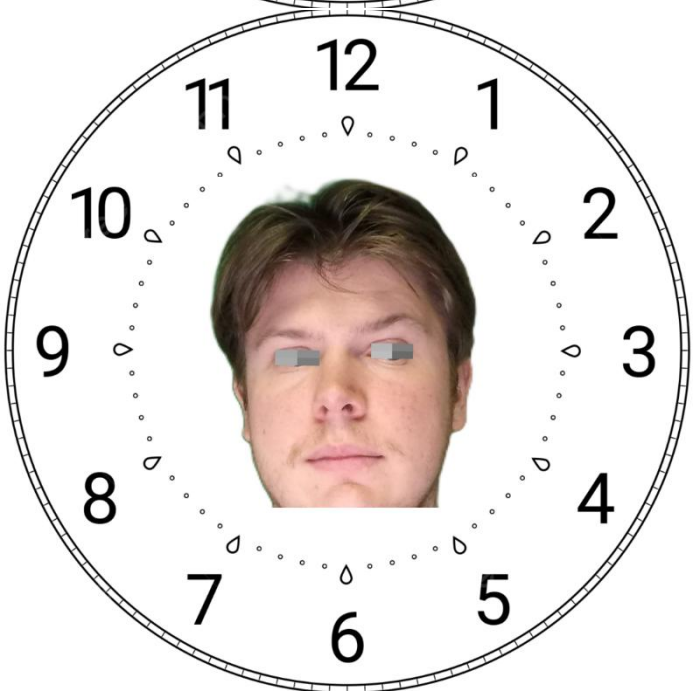

12
1
2
3
4
5
6
7
8
9
10
11

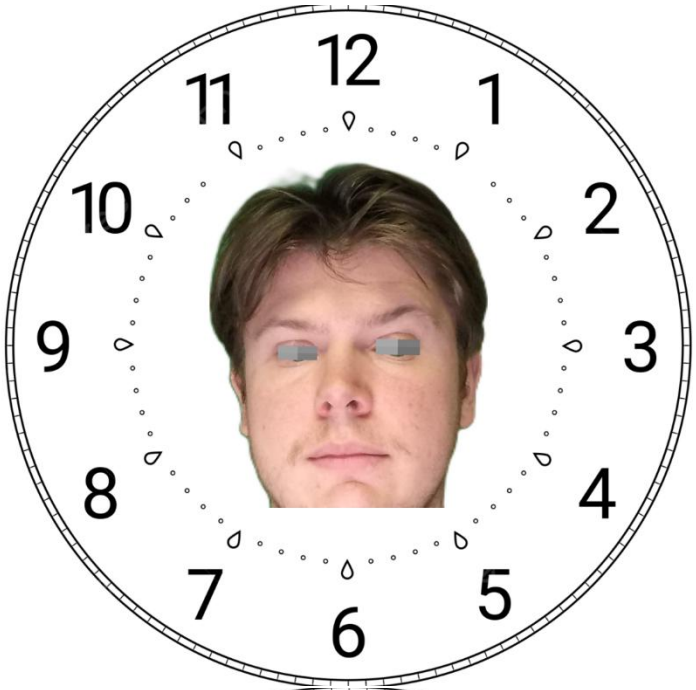
Stimuli Experimental Condition from 1 to 12 o'clock

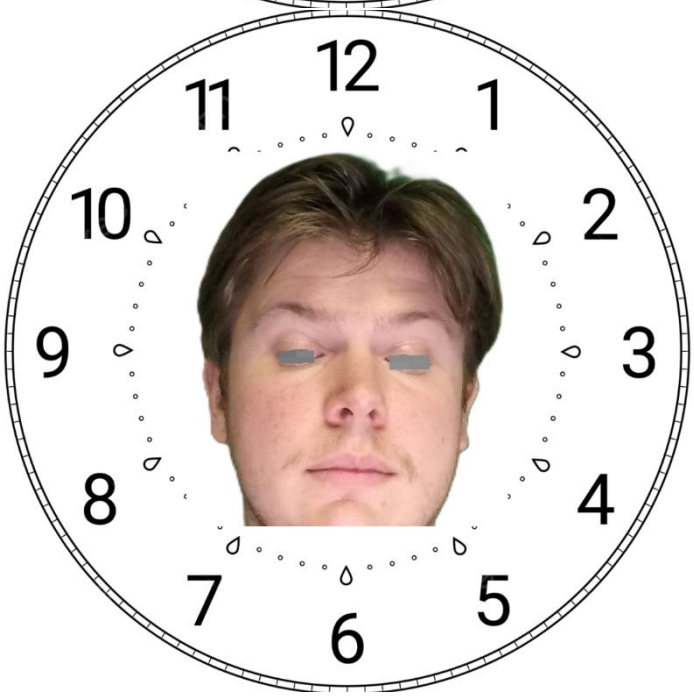


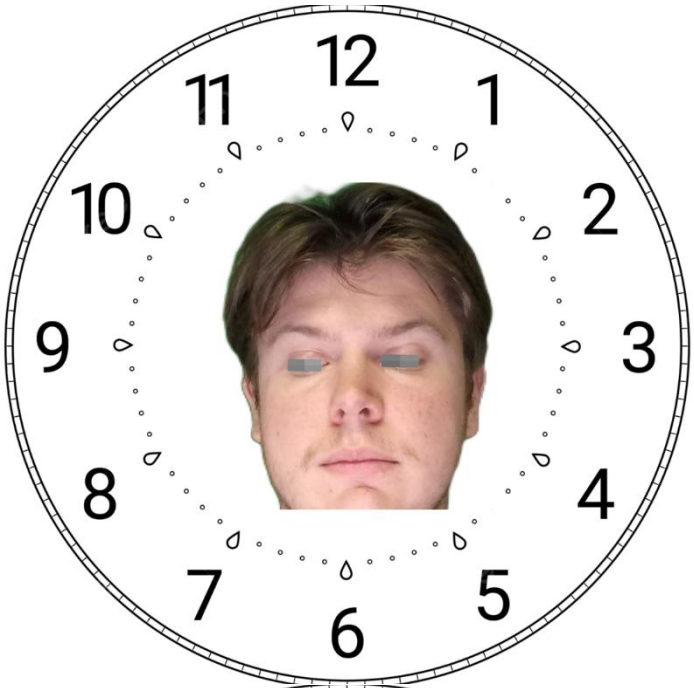


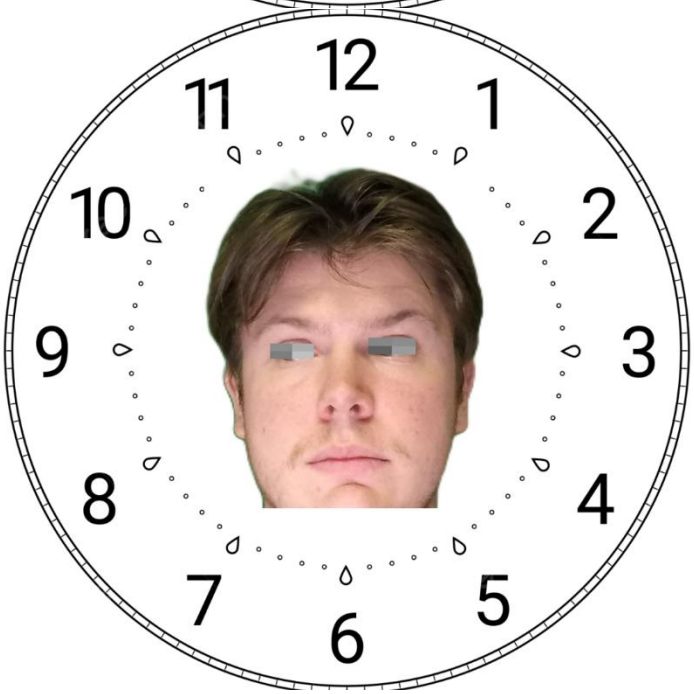

12
1
2
3
4
5
6
7
8
9
10
11

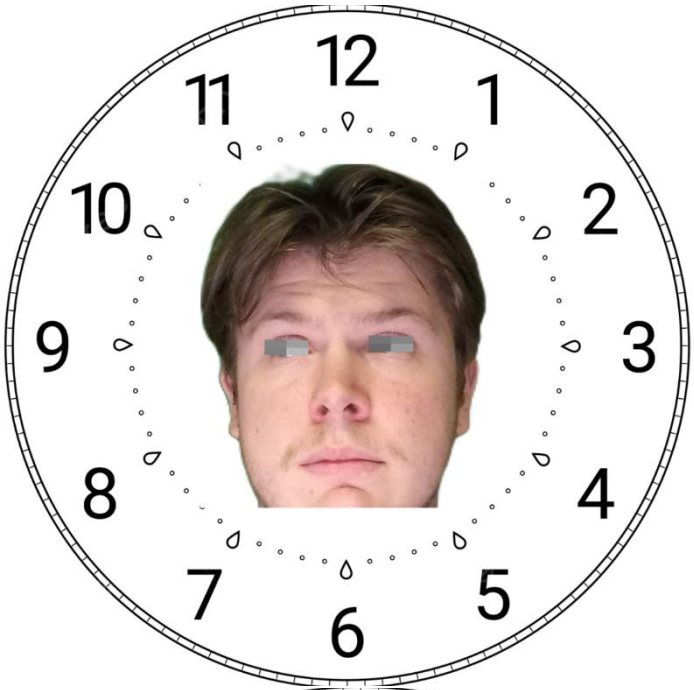

12
1
2
3
4
5
6
7
8
9
10
11

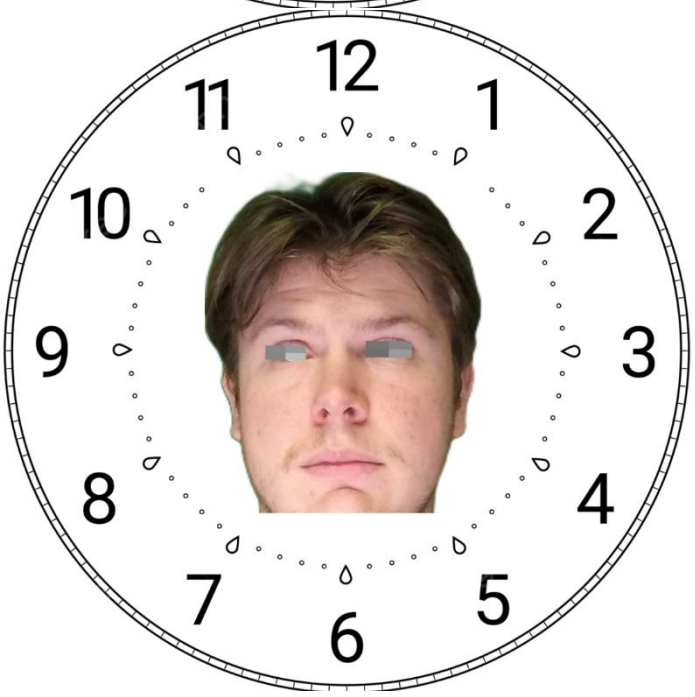

12
1
2
3
4
5
6
7
8
9
10
11

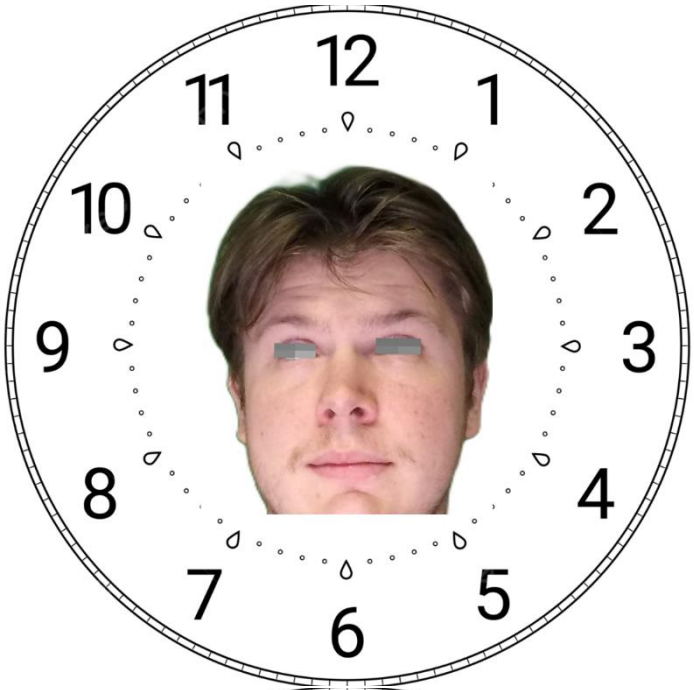

12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11


12
1
2
3
4
5
6
7
8
9
10
11

Python code for stimuli creation

```
import cv2 as cv

import numpy as np

import os


# Define input and output directories

input_dir =
r"C:\Users\minko\Documents\Uni\M12\Yet\Experiments\Rawmaterial\60_Seconds"

output_dir = r"C:\Users\minko\Documents\Uni\M12\Yet\Experiments\Edited_material"

os.makedirs(output_dir, exist_ok=True)  # Ensure the output directory exists


# Get list of image files in input directory (only JPG and PNG)

image_files = [f for f in os.listdir(input_dir) if f.lower().endswith(('.jpg', '.png'))]

image_files.sort()  # Sort to maintain order


#YEC takes and jpg or png image and detect the eyes with the image.

#For each eye the average brightness is calculated and croped onto the eyes, split into four
regions.

#To make this programm functional the file directionary in line 29 needs to be changed, the
image is stored in the current directionary.


def quad_bright(frame):
    """Splits an image into four quadrants and calculates their average brightness."""
    h, w = frame.shape  # Get dimensions (height first in OpenCV)


    NW = np.mean(frame[0:int(h / 2), 0:int(w / 2)])

    NE = np.mean(frame[0:int(h / 2), int(w / 2):w])

    SW = np.mean(frame[int(h / 2):h, 0:int(w / 2)])

    SE = np.mean(frame[int(h / 2):h, int(w / 2):w])


    return (NW, NE, SW, SE)
```

```
def generate_brightness_image(brightness, size):
    """Creates a grayscale image filled with a given brightness value."""
    img = np.full(size, int(brightness), dtype=np.uint8)
    return img

# Load Haar cascades
"""Uses eye_cascade as this is more precicse, but this does not work with glasses."""
face_cascade = cv.CascadeClassifier(cv.data.haarcascades +
"haarcascade_frontalface_default.xml")
eye_cascade = cv.CascadeClassifier(cv.data.haarcascades + "haarcascade_eye.xml")

# Process each image in the folder
for i, filename in enumerate(image_files, start=1):
    image_path = os.path.join(input_dir, filename)
    image = cv.imread(image_path)

    if image is None:
        print(f"Error: Could not read {filename}. Skipping...")
        continue

    # Convert to grayscale
    gray = cv.cvtColor(image, cv.COLOR_BGR2GRAY)

    # Detect faces in the image
    faces = face_cascade.detectMultiScale(gray, scaleFactor=1.3, minNeighbors=4)
    print(f"Processing {filename} - Detected faces: {len(faces)}")

    # Loop over detected faces
    for (x, y, w, h) in faces:
        roi_gray = gray[y:y + h, x:x + w]  # Face region in grayscale
        roi_color = image[y:y + h, x:x + w]  # Face region in color
```

```
    # Detect eyes in the face region
    """scaleFactor checks for false positives minNeighbors for false negatives minSize can be used to exclude false detection such as noses"""
    eyes = eye_cascade.detectMultiScale(roi_gray, scaleFactor=1.3, minNeighbors=10, minSize=(40, 40))

    if len(eyes) == 0:
        print("No eyes detected in this face.")

    # Loop over detected eyes
    """uses an enumerate to keep track of item and index"""
    """rename ex, ey, ew and eh to allign better with inital NW, NE, SW, SE"""
    for (ex, ey, ew, eh) in eyes:
        # Crop the eye region (excluding unnecessary parts)
        crop_start_y = max(0, int(ey + 0.36 * eh))
        crop_end_y = min(h, int(ey + 0.64 * eh))
        crop_start_x = max(0, int(ex + 0.15 * ew))
        crop_end_x = min(w, int(ex + 0.95 * ew))
        eye_region = roi_gray[crop_start_y:crop_end_y, crop_start_x:crop_end_x]

        # Ensure cropped region is valid
        if eye_region.size == 0:
            print("Warning: Eye region is empty after cropping. Skipping...")
            continue

        # Compute brightness
        """check if we should use percived brigthness?"""
        quadrants = quad_bright(eye_region)

        # Generate quadrant brightness images
        h_half, w_half = eye_region.shape[0] // 2, eye_region.shape[1] // 2
```

```
        quad_images = [
            generate_brightness_image(quadrants[0], (h_half, w_half)),  # NW
            generate_brightness_image(quadrants[1], (h_half, w_half)),  # NE
            generate_brightness_image(quadrants[2], (h_half, w_half)),  # SW
            generate_brightness_image(quadrants[3], (h_half, w_half))   # SE
        ]

        # Stack quadrants to create final brightness image
        top_row = np.hstack((quad_images[0], quad_images[1]))  # NW | NE
        bottom_row = np.hstack((quad_images[2], quad_images[3]))  # SW | SE
        quadrant_img = np.vstack((top_row, bottom_row))  # Full quadrant image

        # Resize quadrant image to match detected eye size
        quadrant_img = cv.resize(quadrant_img, (crop_end_x - crop_start_x, crop_end_y - crop_start_y))

        # Insert quadrant image back into original eye position
        roi_color[crop_start_y:crop_end_y, crop_start_x:crop_end_x] = cv.cvtColor(quadrant_img, cv.COLOR_GRAY2BGR)

    # Save the modified image with a sequential number
    output_image_path = os.path.join(output_dir, f"{i}.jpg")
    success = cv.imwrite(output_image_path, image)
    if success:
        print(f"Saved modified image as {output_image_path}")
    else:
        print(f"Error: Could not save {output_image_path}")

print("Processing complete.")
```

R code of data cleaning

```
#libraries

library(readr)

library(tidyverse)

library(dplyr)

library(tidyr)



#reducing unneccessary columns

dataset_reducedcols <- subset(combined_trials_data, select = -c(key_resp.duration_raw,
key_resp.keys_raw, key_resp.rt_raw, thisRepN_raw, thisTrialN_raw, key_resp_2.keys_raw,
key_resp_2.rt_raw, key_resp_3.duration_raw, key_resp_3.keys_raw, key_resp_3.rt_raw,
key_resp_4.duration_raw, key_resp_4.keys_raw, key_resp_4.rt_raw,
key_resp_5.duration_raw, key_resp_5.keys_raw, key_resp_5.rt_raw,
key_resp_6.duration_raw, key_resp_6.keys_raw, key_resp_6.rt_raw,
key_resp_7.duration_raw, key_resp_2.duration_raw, key_resp_7.keys_raw,
key_resp_7.rt_raw, key_resp_8.duration_raw, key_resp_8.keys_raw, key_resp_8.rt_raw,
key_resp_9.duration_raw, key_resp_9.keys_raw, key_resp_9.rt_raw,
key_resp_10.duration_raw, key_resp_10.keys_raw, key_resp_10.rt_raw))


dataset_reducedcols <- dataset_reducedcols[!grepl("expEnd", dataset_reducedcols$imageFile,
ignore.case = TRUE), ]


#create new column with correct responses


# Create a mapping vector

response_mapping <- c(

  "twelve.png" = 12, "twelve" = 12,

  "one.png" = 1, "one" = 1,

  "two.png" = 2, "two" = 2,

  "three.png" = 3, "three" = 3,

  "four.png" = 4, "four" = 4,

  "five.png" = 5, "five" = 5,

  "six.png" = 6, "six" = 6,

  "seven.png" = 7, "seven" = 7,
```

```
  "eight.png" = 8, "eight" = 8,
  "fine.png" = 9, "nine" = 9,
  "ten.png" = 10, "ten" = 10,
  "eleven.png" = 11, "eleven" = 11)

# Clean the `imageFile_raw` column by removing backslashes (if present)
dataset_reducedcols$imageFile_raw <- gsub("\\\\", "", dataset_reducedcols$imageFile_raw)

# Create the correct_response column using the mapping
dataset_reducedcols$correct_response <-
response_mapping[dataset_reducedcols$imageFile_raw]



#rename response columns to name containing presentation time and condition
dataset_reducedcols <- dataset_reducedcols %>%
  rename(response_exp_70ms = Response.text_raw, response_con_70ms = textbox.text_raw,
response_exp_140ms = textbox_2.text_raw, response_con_140ms = textbox_3.text_raw,
response_exp_400ms = textbox_4.text_raw, response_con_400ms = textbox_5.text_raw,
response_exp_600ms = textbox_6.text_raw, response_con_600ms = textbox_7.text_raw,
response_exp_1000ms = textbox_8.text_raw, response_con_1000ms = textbox_9.text_raw)



#assign participant numbers

# Calculate the number of rows per participant
rows_per_participant <- 450

# Calculate the total number of participants
total_participants <- 43

# Create a vector of participant numbers in the correct order
```

```
participant_numbers <- c(
  10:19, 1, 20:29, 2, 30:39, 3, 40:43, 4, 5:9
)

# Repeat each participant number for the appropriate number of rows
participant_column <- rep(participant_numbers, each = rows_per_participant)

# Truncate the vector if the total number of rows is not a multiple of rows_per_participant
if (length(participant_column) > nrow(dataset_reducedcols)) {
  participant_column <- participant_column[1:nrow(dataset_reducedcols)]
}

# Add the participant column to the data frame
dataset_reducedcols$participant_number <- participant_column



#reorder columns
dataset_reducedcols <- dataset_reducedcols %>%
  select(participant_number, correct_response, response_exp_70ms, response_con_70ms, response_exp_140ms, response_con_140ms, response_exp_400ms, response_con_400ms, response_exp_600ms, response_con_600ms, response_exp_1000ms, response_con_1000ms, thisN_raw, imageFile, imageFile_raw)

#create new column with responses

# 1. Create combined_data and combined_temp
dataset_reducedcols <- dataset_reducedcols %>%
  mutate(
    combined_data = unite(
      .,
      col = "combined_temp",
```

```
    response_exp_70ms, response_con_70ms, response_exp_140ms, response_con_140ms,
    response_exp_400ms, response_con_400ms, response_exp_600ms, response_con_600ms,
    response_exp_1000ms, response_con_1000ms,
    sep = " - ",
    na.rm = TRUE,
    remove = FALSE
  )$combined_temp
)

# 2. THEN, remove combined_temp
dataset_reducedcols <- dataset_reducedcols %>%
  select(-combined_temp)



#add condition column

# Define the pattern of conditions
condition_pattern <- c(rep("experimental", 45), rep("control", 45))

# Calculate how many times to repeat the pattern
num_repeats <- ceiling(nrow(dataset_reducedcols) / length(condition_pattern))

# Create the condition vector by repeating the pattern
condition_vector <- rep(condition_pattern, times = num_repeats)

# Truncate the vector to match the number of rows in the dataframe
condition_vector <- condition_vector[1:nrow(dataset_reducedcols)]

# Add the condition column to the dataframe
dataset_reducedcols$condition <- condition_vector
```

```
#add presentation time column

# Define the pattern of presentation times
time_pattern <- rep(c(70, 140, 400, 600, 1000), each = 90)

# Calculate how many times to repeat the pattern
num_repeats <- ceiling(nrow(dataset_reducedcols) / length(time_pattern))

# Create the presentation time vector by repeating the pattern
time_vector <- rep(time_pattern, times = num_repeats)

# Truncate the vector to match the number of rows in the dataframe
time_vector <- time_vector[1:nrow(dataset_reducedcols)]

# Add the presentation time column to the dataframe
dataset_reducedcols$presentation_time_ms <- time_vector



#remove unimportant rows

# Filter the dataframe to keep only rows where thisN_raw is between 0 and 35
dataset_reducedcols <- dataset_reducedcols %>%
  filter(thisN_raw >= 0 & thisN_raw <= 35)



#only keep neccessary columns
```

```
dataset_reducedcols_total <- dataset_reducedcols %>%
  select(participant_number, correct_response, combined_data, condition,
presentation_time_ms)



#rename combined_data to actual_response
dataset_reducedcols_total <- dataset_reducedcols_total %>%
  rename(actual_response = combined_data)



#delete invalid responses
# Define the numbers you want to match exactly (as strings)
exact_matches <- as.character(1:12)

# Filter the rows where the 'actual_response' column exactly matches any of the specified
numbers
dataset_reducedcols_total <- dataset_reducedcols_total %>%
  filter(actual_response %in% exact_matches)



#rename correct response to simulus and actual response to response
dataset_reducedcols_total <- dataset_reducedcols_total %>%
  rename(stimulus = correct_response, response = actual_response)

str(dataset_reducedcols_total)

#Transfer numbers responses into degrees of deviation
#stim =a, resp=b
#a-b = c -> abs(c) = d, 12-d = e -> d oder e kleiner? -> in tabelle eintragen in spalte Deviation
# Assuming your data frame is dataset_reducedcols_total
```

```
# Convert the 'response' column to numeric
dataset_reducedcols_total <- dataset_reducedcols_total %>%
  mutate(
    response = as.numeric(response)
  )

# Now, run your transformation algorithm
dataset_reducedcols_total <- dataset_reducedcols_total %>%
  mutate(
    c = stimulus - response,      # Subtract 'response' (b) from 'stimulus' (a)
    d = abs(c),                   # Calculate the absolute value of 'c'
    e = 12 - d,                   # Subtract 'd' from 12
    deviation = ifelse(d <= e, d, e) # Save the smaller value of 'd' and 'e' in 'deviation'
    # If d and e are equal, d is saved.
  )

#remove c,d,e columns
dataset_reducedcols_total <- dataset_reducedcols_total %>%
  select(participant_number, stimulus, response, condition, presentation_time_ms, deviation)

# Create a new dataset without the 'deviation' column
dataset_cleaned_43 <- dataset_cleaned[, !(colnames(dataset_cleaned) == "deviation")]

columns_to_keep <- c("participant_number", "stimulus", "response", "condition",
"presentation_time_ms") # Replace with your actual column names
dataset_cleaned_43 <- dataset_cleaned[, columns_to_keep]
write_csv(dataset_cleaned_43,"/Users/hoelt/OneDrive/Desktop/Data
Analysis/cleaning/dataset_cleaned_43.csv")
```

R markdown of results section

Results

Janis Hölter

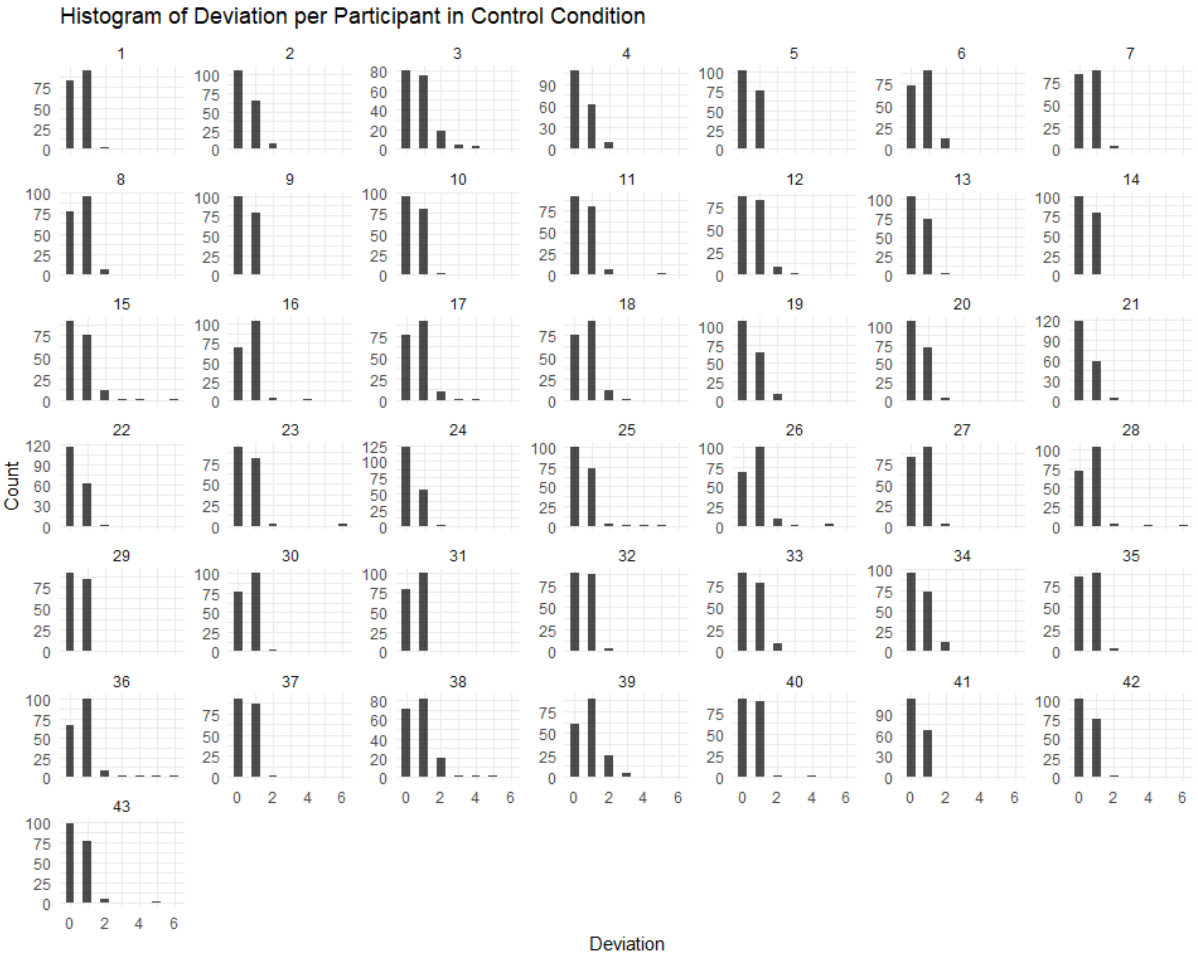
2025-05-28

Exploration of Data

Distribution of Deviation per Participant

In the histograms, the distribution of the deviation per participant in the control condition is presented. It can be seen that the distribution shows several similarities for all participants. For each histogram, the peaks are either 0 or 1. Further, all distributions are positively skewed. 27 participants have a peak of 0 deviation in the control condition. The other 16 peak at a deviation of 1. The proportions for deviations of 3,4,5 ,or 6 are either very low or zero for each participant. The differences between the deviations of 0 and 1 vary, and for participants 21,22 and 24, the bin of 0 as twice as high than the bin for 1. On the other hand, for many others such as participants 3, 7, 12, 29, 32, 35, 37 and 40, the distribution of 0 and 1 is almost equal.

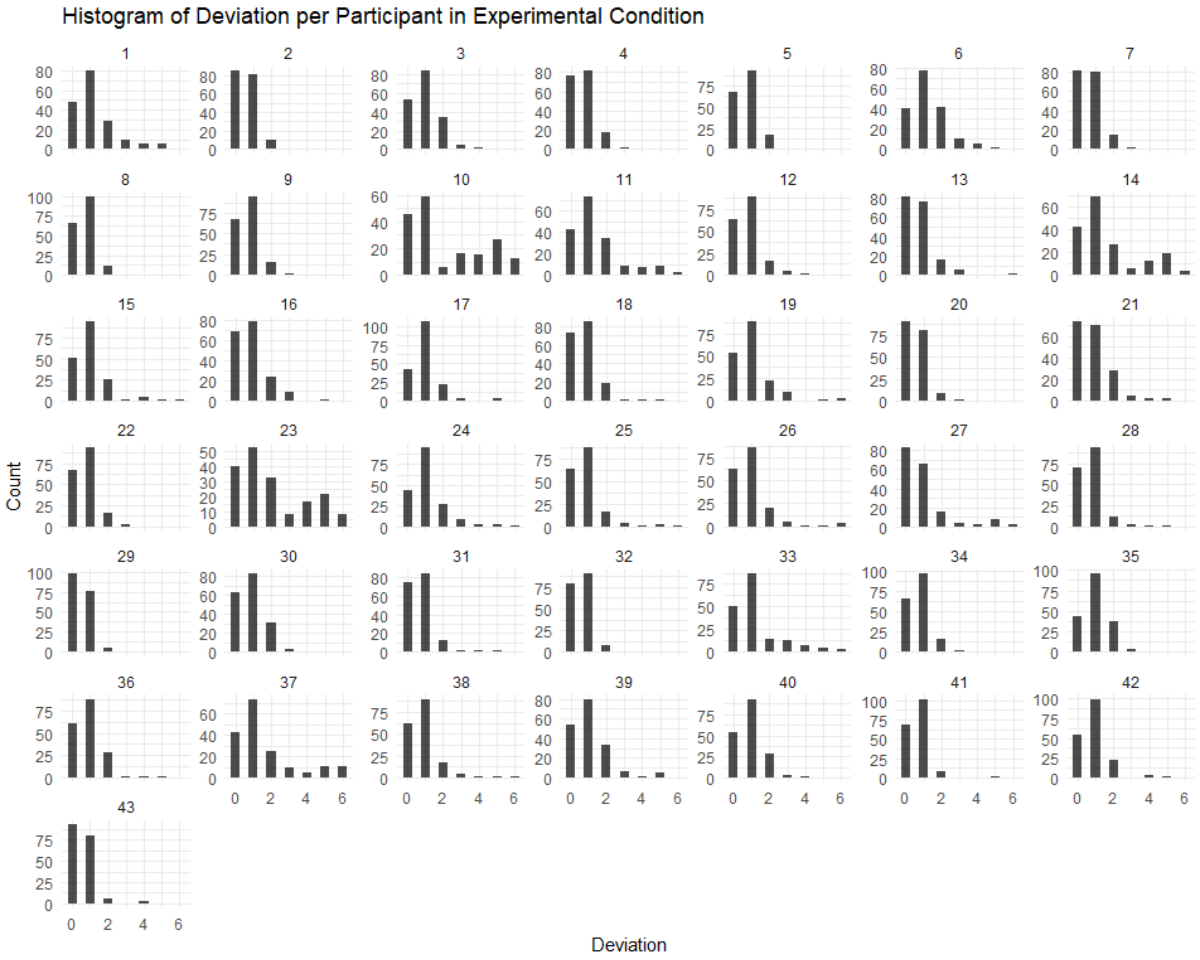
```
library(ggplot2)
library(dplyr)

##
## Attache Paket: 'dplyr'

## Die folgenden Objekte sind maskiert von 'package:stats':
##
##     filter, lag

## Die folgenden Objekte sind maskiert von 'package:base':
##
##     intersect, setdiff, setequal, union

dataset_cleaned <- read.csv("dataset_cleaned.csv")
df_control <- dataset_cleaned %>%
  filter(condition == "control")


ggplot(df_control, aes(x = deviation)) +
  geom_histogram(fill = "black", alpha = 0.7, binwidth = 0.5) + # Using binwidth from previous adv
ice
  facet_wrap(~ participant_number, scales = "free_y") +
  labs(title = "Histogram of Deviation per Participant in Control Condition",
       x = "Deviation",
       y = "Count") +
  theme_minimal()
```



In the distributions of deviation per participants in the experimental condition, there are some similarities and some differences. The histograms of the participants resemble each other by having the highest counts in 0 and 1 deviations. Except for the participants 2, 7, 13, 20, 21, 27, 29, and 43, who have 0 deviations most of time, the other 35 participants have 1 deviations most of the time. There are some participants as well who have proportions of around 10% or more of observations for a deviation of 2 as well. Especially participant 6 stands out as the count of deviation 2 is equal to the count of deviation 0. Examining the counts for deviations 3,4,5, and 6 per participants, it can be noticed that for the majority of participants the proportions are either low or non-existent, however for participant 10, 14, and 23, there is visible rise at a deviation of 5.

```
library(ggplot2)
library(dplyr)

df_experimental <- dataset_cleaned %>%
  filter(condition == "experimental")

ggplot(df_experimental, aes(x = deviation)) +
  geom_histogram(fill = "black", alpha = 0.7, binwidth = 0.5) + # Using binwidth from previous adv
ice
  facet_wrap(~ participant_number, scales = "free_y") +
  labs(title = "Histogram of Deviation per Participant in Experimental Condition",
       x = "Deviation",
```

```
    y = "Count") +
theme_minimal()
```



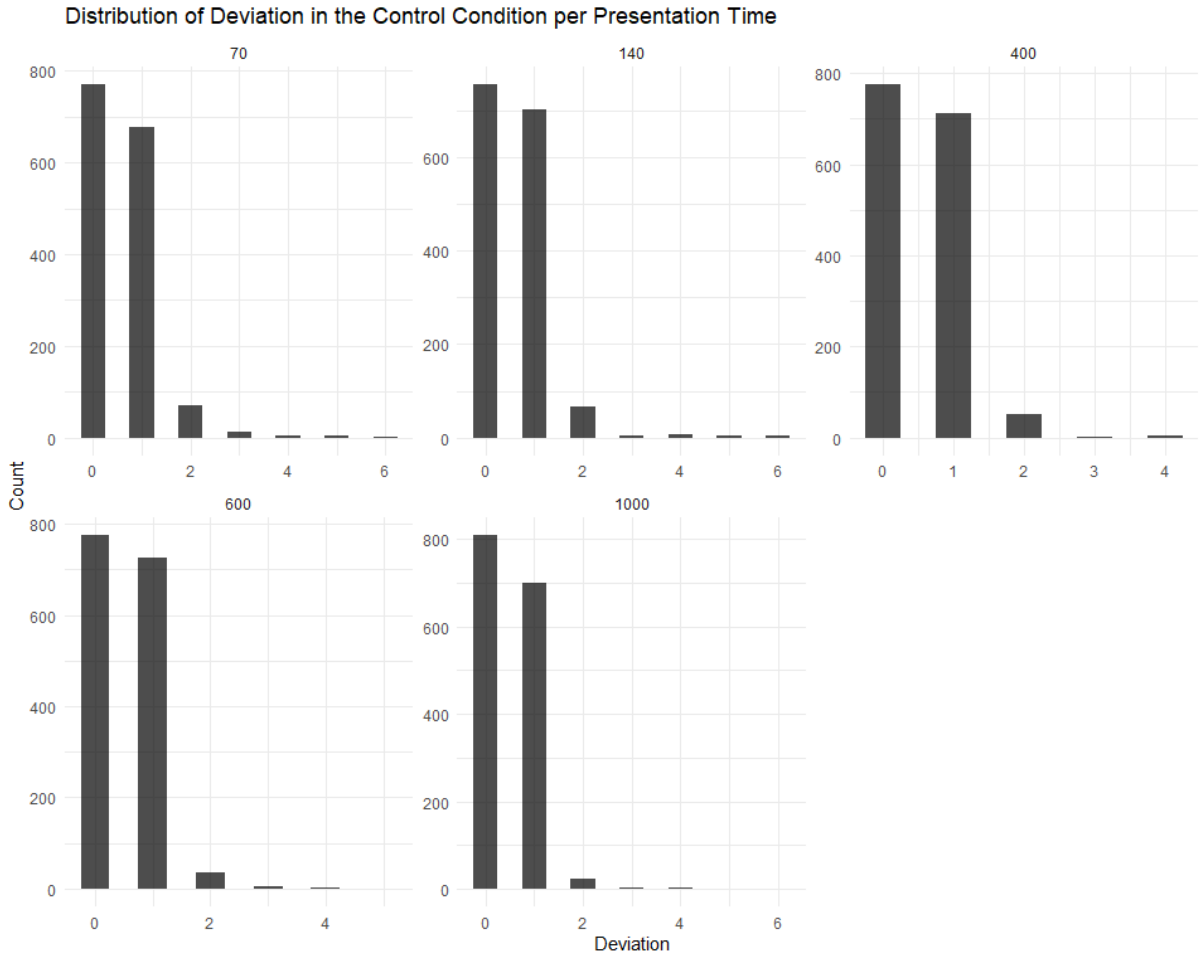
Comparing the distributions of deviation per participant in the control and experimental condition, several similarities can be found. In both conditions, deviations of 0 and 1 display the highest count for each participant. Further, in almost all cases, the distributions are positively skewed from a deviation of 1 onward. In contrast to the control condition, in the experimental condition, the highest count of observations per deviation is mostly at 1 and not at 0. Furthermore, the counts for deviations 0f 2,3,4,5 and 6 are higher in the experimental condition. For a few participants, a rise in counts from deviation 3 to 5 can be seen. These findings indicate that participants were more accurate in gaze perception in the control group and that there is some variation between participants in the experimental group.

Distribution of deviation per condition and presentation time

Throughout all five presentation times in the control condition, the distribution of deviation shows several similarities. From 70ms to 1000ms, there are high densities around the low deviations of 0 and 1. The median in each condition lies between 0 and 1. The interquartile range spans for every condition from 0 to 1, and there are no whiskers towards lower deviations, only towards higher deviations. From this, it can be deduced that at least 75% of observations are either 0 or 1. Furthermore, for each exposure time it can be seen that deviations of 2 or higher are outliers. This indicates that the distribution of deviations is positively skewed for each exposure time. The medians differ slightly per exposure time. In 70ms, it is at 0.5 and in 140ms at 1, for the higher exposure times it is at 0. It is also visible that the outliers for the three longer exposure times, the outliers are less than for the shorter exposure times. These findings point

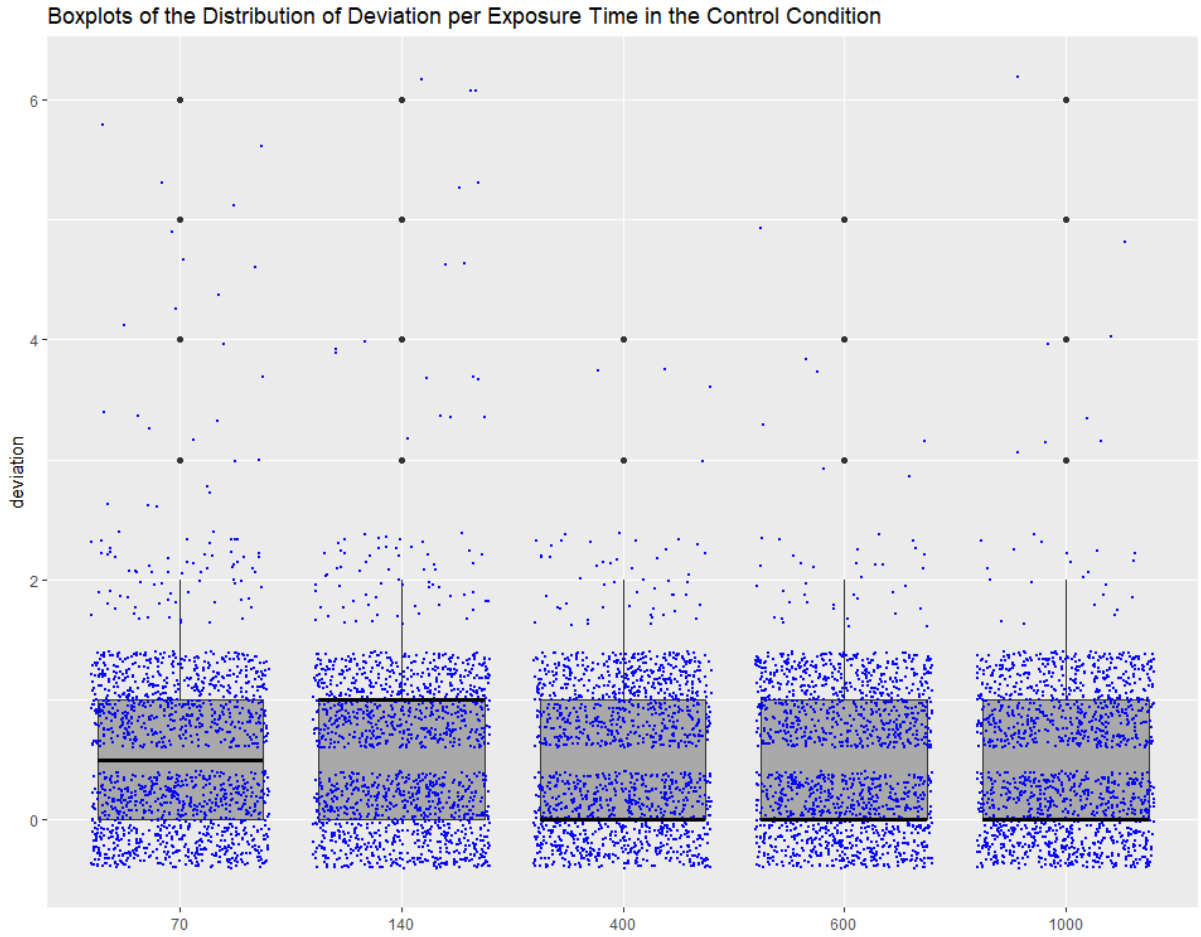
out that deviation is slightly less for exposure times of 400ms, 600ms, and 1000ms than for 70ms, and 140ms.

```
library(ggplot2)
library(dplyr)
dataset_cleaned <- read.csv("dataset_cleaned.csv")
df_control <- dataset_cleaned %>%
 filter(condition == "control")
presentation_times <- c("70", "140", "400", "600", "1000")
df_filtered_control <- df_control %>%
 filter(presentation_time_ms %in% presentation_times)
df_filtered_control$presentation_time_ms <- factor(df_filtered_control$presentation_time_ms,
                                                   levels = presentation_times)
ggplot(df_filtered_control, aes(x = deviation)) +
 geom_histogram(fill = "black", alpha = 0.7, binwidth = 0.5) +
 facet_wrap(~ presentation_time_ms, scales = "free") +
 labs(title = "Distribution of Deviation in the Control Condition per Presentation Time",
    x = "Deviation",
    y = "Count") +
 theme_minimal()
```




```
library(hrbrthemes)
library(viridis)

## Lade nötiges Paket: viridisLite
```

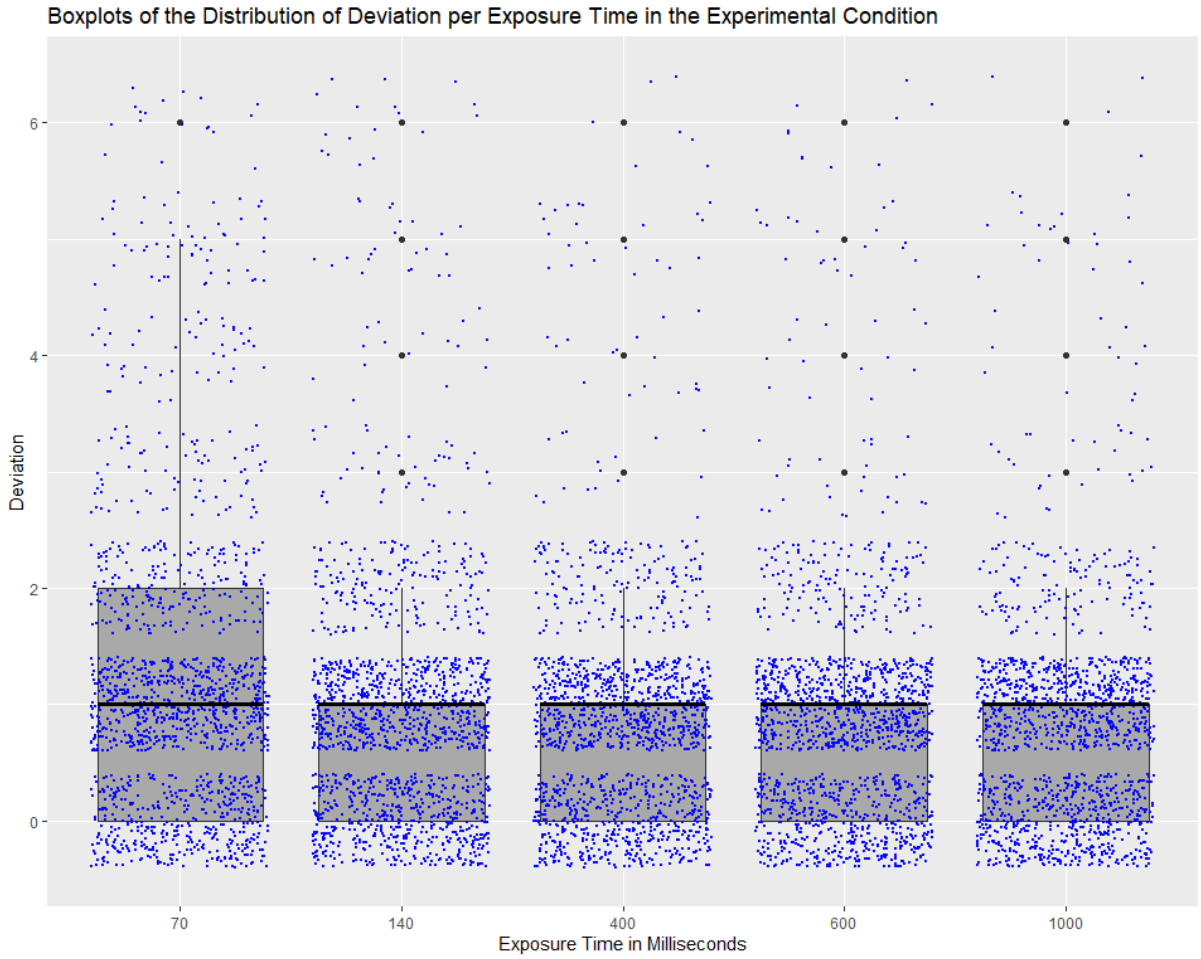
```r
df_filtered_control %>%
 ggplot(aes(x = presentation_time_ms, y = deviation)) +
 geom_boxplot(fill = "darkgrey") +
 geom_jitter(color = "blue", size = 0.3, alpha = 0.9) + # Changed size to linewidth here too for con
sistency
  stat_summary(fun.data = function(x) data.frame(y = median(x), ymin = median(x), ymax = m
edian(x)),
          geom = "crossbar",
          width = 0.75, # Controls the width of the median line
          color = "black") + # Changed size to linewidth
 ggtitle("Boxplots of the Distribution of Deviation per Exposure Time in the Control Condition") +
 xlab("")
```



In the experimental condition, the distributions per exposure time are most dense at deviations of 0 and 1. The distribution at 70ms is slightly less dense at 0 and 1 in comparison to the other exposures. In all exposure times, the box plots have a median deviation of 1. However, the interquartile range differs between 70ms exposure and the other four. For 70ms, the IQR spans from 0 to 2. In 140ms, 400ms, 600ms and 1000ms, the IQR spans from 0 to 1. In the boxplots, it can be seen that in each exposure time, whiskers only reach towards higher deviations. From this, it might be deduced that at 70ms, at least 75% of observations are a deviation of 0,1 or 2, and that for 140ms, 400ms, 600ms and 1000ms, at least 75% of observations are either a deviation of 0 or 1. For an exposure of 70ms, deviations of 6 are extreme outliers. For the longer exposure times, deviations of 3,4,5, and 6 are extreme outliers. These findings indicate that the distributions of deviations of each exposure time are positively skewed. Furthermore, deviation

is decreasing from 70ms to 140ms. From 140ms onwards, no difference in deviations can be observed.

```
library(ggplot2)
library(dplyr)
library(hrbrthemes)
library(viridis)

dataset_cleaned <- read.csv("dataset_cleaned.csv")
df_experimental <- dataset_cleaned %>%
 filter(condition == "experimental")
presentation_times <- c("70", "140", "400", "600", "1000")
df_filtered_experimental <- df_experimental %>%
 filter(presentation_time_ms %in% presentation_times)
df_filtered_experimental$presentation_time_ms <- factor(df_filtered_experimental$presentatio
n_time_ms,
                                levels = presentation_times)

df_filtered_experimental %>%
 ggplot(aes(x = presentation_time_ms, y = deviation)) +
 geom_boxplot(fill = "darkgrey") +
 geom_jitter(color = "blue", size = 0.3, alpha = 0.9) +
  stat_summary(fun.data = function(x) data.frame(y = median(x), ymin = median(x), ymax = m
edian(x)),
          geom = "crossbar",
          width = 0.75,
          color = "black") +
 ggtitle("Boxplots of the Distribution of Deviation per Exposure Time in the Experimental Conditi
on") +
 xlab("Exposure Time in Milliseconds") +
 ylab("Deviation")
```



Comparing the boxplots of both conditions, it stands out that deviation in each plot is centered around deviations of 0 and 1. This hints at an ability of participants to perceive gaze direction when presented with faces with YET-eyes and unmanipulated faces. However, the concentration at small deviations was more pronounced in the control condition which points out that participants were more accurate when presented with unmanipulaeted faces. In the control condition, there were only minor differences between the boxplots per exposure, and thus, no change throughout presentation time can be inferred. On the other hand, in the experimental condition a difference between 70ms and the other four exposure times could be observed as the boxplots indicated a decreasing deviation.

Modeling

Based on the properties of the variables, it was decided to fit a Bayesian generalized linear multi-level model. A multi-level model usually consists of fixed and random effects, and is able to account for individual differences in the observations (Schmettow, 2021). From this, predictions can be made at the population level for the whole sample and at the individual level for each singular participant. The fixed effects in this model are the condition, the exposure time and their interaction effect. These population level effects are expected to affect the observations of each participant similarly. The random multi-level effects are each participant and each stimulus per condition. These individual level effects were expected to vary for each participant and for each stimulus per condition. From these effects, the dependent variable deviation was estimated. The dependent variable deviation was treated as non-negative integers from 0 to 6. Respecting this count data with the boundary at 0, a Poisson distribution was applied as family of the model. The link is a logarithmic function. Therefore, the model predicts the log of the

expected deviation. The value on the log scale is then exponentiated to obtain the expected value of deviation on the original scale (Schmettow, 2021).

```
model_1 <- readRDS("model_fit.rds")
formula(model_1)

## deviation ~ Condition * Exposure + (Condition | Part) + (Condition |
##     Stim)
```

Testing potential overdispersion of the model, it was found that the dispersion ration is 0.815 which pointing towards a slight underdispersion of the model, and thus, less variance in the actual data than predicted by the model. However, the p-value of 1 strongly suggests that there is no evidence of overdispersion found. Followingly, it can be assumed that the model handles variance in the data adequately.

```
## # Overdispersion test
##
##       dispersion ratio =    0.815
##  Pearson's Chi-Squared = 12547.428
##              p-value =       1

## No overdispersion detected.
```

Fixed Effects

In the control condition with an exposure time of 1000ms represented by the intercept, the expected deviation is 0.44 (CI [0.32, 0.6]). Several deviation increasing factors and no decreasing factors were found. Compared to the intercept, there is a clear effect of the experimental condition which is expected to increase the deviation approximately 69%. Further, there are clear effects for an exposure of 70ms and 140ms. However, those are not as pronounced and each increases deviation by approximately 17%. If the condition is experimental, there is an increasing effect of 70ms exposure time of approximately 31%.

```
library(rstanarm)

## Lade nötiges Paket: Rcpp

## This is rstanarm version 2.32.1

## - See https://mc-stan.org/rstanarm/articles/priors for changes to default priors!

## - Default priors may change, so it's safest to specify priors, even if equivalent to the defaults.

## - For execution on a local, multicore CPU with excess RAM we recommend calling

##   options(mc.cores = parallel::detectCores())

library(bayr)

## Registered S3 methods overwritten by 'bayr':
##   method        from
##   coef.stanreg   rstanarm
##   predict.stanreg rstanarm

##
## Attache Paket: 'bayr'
```

```
## Die folgenden Objekte sind maskiert von 'package:rstanarm':
##
##    fixef, ranef
```

```
model_1 <- readRDS("model_fit.rds")
fixef(model_1, mean.func = exp)
```

Coefficient estimates with 95% credibility limits

fixef	center	lower	upper
Intercept	0.4441424	0.3245297	0.5976229
Conditionexperimental	1.6919188	1.2932302	2.2315963
Exposure600	1.0575596	0.9559768	1.1668465
Exposure400	1.0657836	0.9674002	1.1786510
Exposure140	1.1694148	1.0626642	1.2876030
Exposure70	1.1727877	1.0670005	1.2963488
Conditionexperimental:Exposure600	1.0407471	0.9204266	1.1839248
Conditionexperimental:Exposure400	1.0059798	0.8835101	1.1471686
Conditionexperimental:Exposure140	1.0115770	0.8968429	1.1493466
Conditionexperimental:Exposure70	1.3145932	1.1599226	1.4844017

In regards to the first hypothesis “humans are able to detect gaze direction from a face with 2x2 quadrants representing the average brightness of the eye”, this means that it can be accepted as the expected deviation for the experimental condition at an exposure of 1000ms is less than 30 degrees, and therefore, indicates that humans were mostly able to detect gaze direction. Furthermore, there is no significant difference in the exposures of 600ms, 400ms and 140ms. In contrary, the second hypothesis “the accuracy in detecting gaze direction from a face with 2x2 quadrants representing the average brightness of the eye is equal to the accuracy of detecting gaze direction from human eyes” must be rejected because there is a significant effect of the condition on the accuracy in gaze detection leading to an increase in deviation for the experimental condition. The third hypothesis “humans are able to perceive gaze direction after being presented for 140ms with a face with 2x2 quadrants representing the average brightness of the eye” can be accepted as there is no significant increase of the expected deviation for an exposure from 140ms onwards. Thus, humans can be assumed to be able to detect gaze direction if they are presented with the stimulus for 140ms or longer.

Random Effects

In Table ..., the standard deviations of logarithms to the population level for the intercept and the experimental condition are displayed. Exponentiated, these standard deviations display factors on the population level. The standard deviations of participants are a factor of 1.22 on the intercept and a factor of 1.42 on the effect of the experimental condition. For the stimuli, the factors are 1.63 on the intercept and a factor of 1.49 on the effect of the experimental condition. The standard deviation of participants on the intercept indicates a rather low variability. On the contrary, there is a moderate variability in the effect of the experimental condition per participant. Focusing on the stimuli it stands out that there is a high variability per stimulus at

intercept level and a moderate variability per stimulus in the effect of the experimental condition. These findings it can be deduced that there is considerable variability in the expected deviation that the fixed effects cannot fully account for.

```
library(bayr)
raef_sd <- fixef_ml(model_1)
raef_sd %>%
 filter(fixef == "Intercept" | fixef == "Conditionexperimental")
```

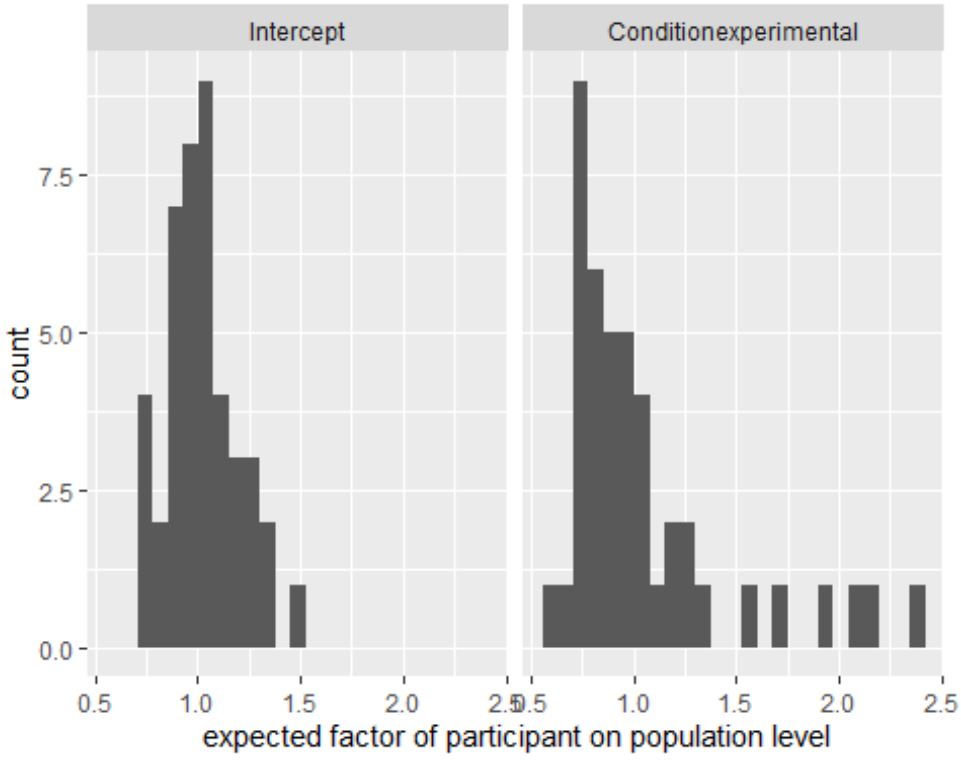
Population-level coefficients with random effects standard deviations

fixef	center	lower	upper	SD_Part	SD_Stim
Intercept	-0.8116101	-1.1253784	-0.5147953	0.1958375	0.4902023
Conditionexperimental	0.5258633	0.2571431	0.8027172	0.3532276	0.3974395

Variability in Participants

Taking a closer look on the variability between participants, it appears that the multiplicative effects derived from the log-scale random effects of participants on the intercept is peaking slightly below 1. The smallest effect is approximately at 0.7 and the highest at 1.5. The distribution is relatively symmetrical which suggests a small variance in the random effects. On the contrary, focusing on the distribution of the multiplicative effects derived from the log-scale random effects on the experimental condition, the distribution is positively skewed and peaking at around 0.75. The smallest effect is approximately 0.6 and the highest 2.4. This strong positive skewness is a characteristic of a log-normal distribution, the expected shape for the exponentiated random effects. These findings indicate that the multiplicative effects derived from the log-scale random effects of participant on the intercept display relatively consistent variance. Thus, confidence in the population-level estimate can be strenghtened. On the other hand, the strong skewness in the distribution of the multiplicative participant effects on the fixed effect of the experimental condition indicate heterogeneity. The majority of participants display a decreasing or similar effect on the effect of the experimental condition. This is contrasted by a small number of participants that displays much larger multiplicative effects ranging up to 2.4 times the population-level effect.

```
library(dplyr)
library(ggplot2)
raef_part <- ranef(model_1, mean.func = exp)
raef_part <- raef_part%>% filter(re_factor == "Part")

raef_part %>% rename(Part = re_entity, `deviation` = center) %>%
 mutate(fixef = factor(fixef, levels = c("Intercept", "Conditionexperimental"))) %>%
 ggplot(aes(x = deviation)) +
 facet_grid(~fixef) +
 geom_histogram(bins = 25) +
 xlab("expected factor of participant on population level")
```

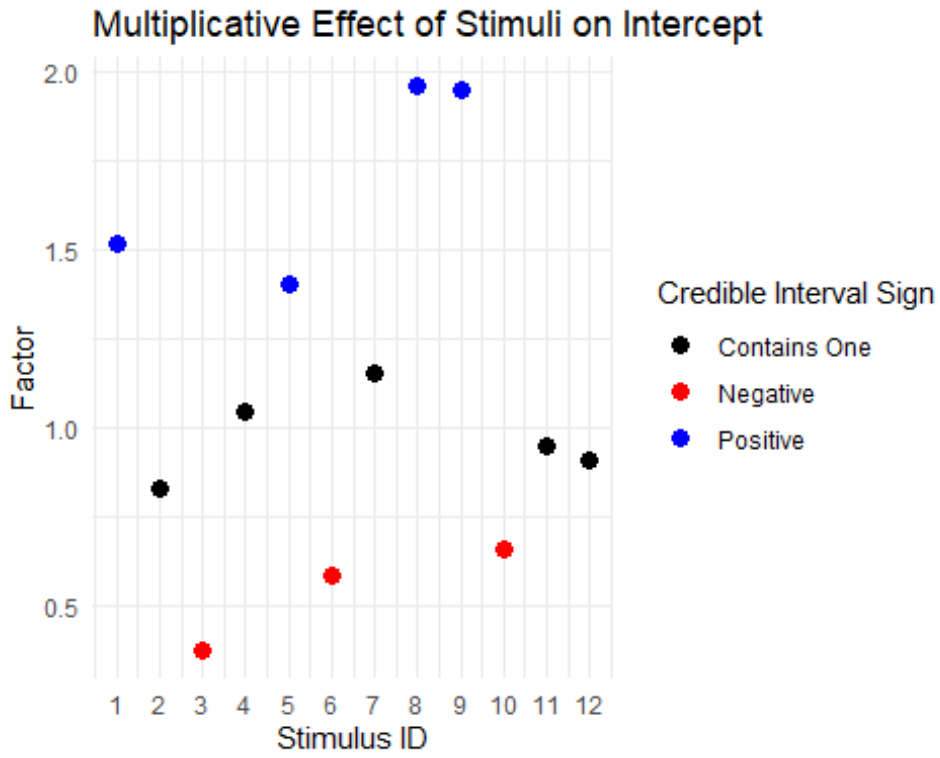


####

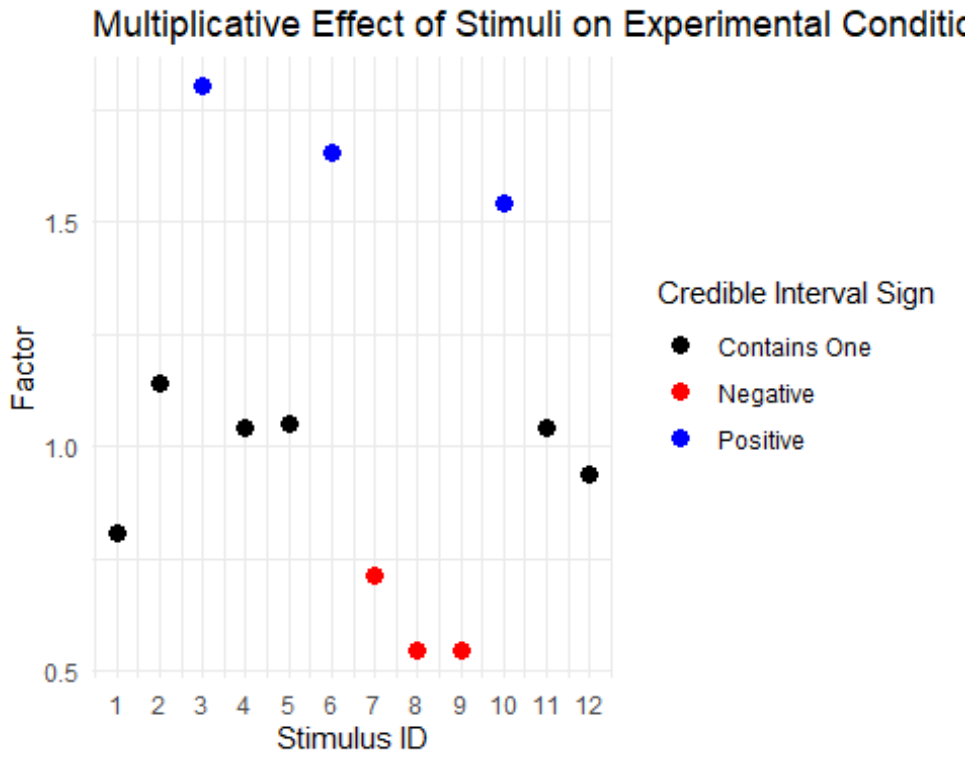
Multiplicative Effects of Stimuli Examining the multiplicative effects of each stimulus on the intercept, it can be seen that stimuli 1, 5, 8, and 9 increase the intercept and thus have an increasing effect on the dependent variable deviation. The stimuli 3,6, and 10 have an decreasing effect on the dependent variable. No clear effect is observed for stimulus 2, 4, 7, 11, and 12 as their 95% credibility intervals include 1. Focussing on the effects of stimuli on the effect of the experimental condition, the stimuli 3, 6 and 10 display an increase, and the stimuli 7, 8, 9 display a decrease. The 95% credible intervals of stimulus 1, 2, 4, 5, 11, and 12 contain 1 and therefore no clear effect was observed.

```
library(ggplot2)
library(dplyr)
library(bayr)

raef_exp <- ranef(model_1, mean.func = exp)
raef_exp_stim <- raef_exp%>% filter(re_factor == "Stim")
raef_exp_stim_int <- raef_exp_stim%>% filter(fixef == "Intercept")
raef_exp_stim_conex <- raef_exp_stim%>% filter(fixef == "Conditionexperimental")

raef_exp_stim_int <- raef_exp_stim_int %>%
 mutate(
  ci_sign = case_when(
   lower > 1 & upper > 1 ~ "Positive",   # Both bounds are positive
   lower < 1 & upper < 1 ~ "Negative",   # Both bounds are negative
   TRUE            ~ "Contains One" # Interval spans across zero (or one bound is zero)
  )
 )
raef_exp_stim_conex <- raef_exp_stim_conex %>%
 mutate(
  ci_sign = case_when(
```

```
    lower > 1 & upper > 1 ~ "Positive",   # Both bounds are positive
    lower < 1 & upper < 1 ~ "Negative",   # Both bounds are negative
    TRUE            ~ "Contains One" # Interval spans across zero (or one bound is zero)
   )
 )
scaplot_re_stim_int <- ggplot(raef_exp_stim_int, aes(x = as.numeric(re_entity), y = center, color
= ci_sign)) +
 geom_point(size = 3) +
 labs(
  title = "Multiplicative Effect of Stimuli on Intercept",
  x = "Stimulus ID",
  y = "Factor"
 ) +
 theme_minimal() +
 scale_color_manual(
  values = c("Positive" = "blue", "Negative" = "red", "Contains One" = "black"),
  name = "Credible Interval Sign"
 ) +
 scale_x_continuous(
  breaks = seq(1, 12, by = 1), limits = c(1, 12)
 )
scaplot_re_stim_conex <- ggplot(raef_exp_stim_conex, aes(x = as.numeric(re_entity), y = cente
r, color = ci_sign)) +
 geom_point(size = 3) +
 labs(
  title = "Multiplicative Effect of Stimuli on Experimental Condition Effect",
  x = "Stimulus ID",
  y = "Factor"
 ) +
 theme_minimal() +
 scale_color_manual(
  values = c("Positive" = "blue", "Negative" = "red", "Contains One" = "black"),
  name = "Credible Interval Sign"
 ) +
 scale_x_continuous(
  breaks = seq(1, 12, by = 1), limits = c(1, 12)
 )

print(scaplot_re_stim_int)
```




```
print(scaplot_re_stim_conex)
```



Informed Consent Form

Consent Form for *How do humans perceive gaze direction*
YOU WILL BE GIVEN A COPY OF THIS INFORMED CONSENT FORM

Please tick the appropriate boxes	Yes	No
Taking part in the study		
I have read and understood the study information dated [/ /], or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	☐	☐
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions, and I can withdraw from the study at any time, without having to give a reason.	☐	☐
I understand that taking part in the study involves collecting my age and gender as well as anonymous responses given during the experiment	☐	☐
Use of the information in the study		
I understand that information I provide will be used for writing a bachelor thesis and building potential future research questions	☐	☐
I understand that personal information collected about me that can identify me, such as [e.g. my name or where I live], will not be shared beyond the study team.	☐	☐
Future use and reuse of the information by others		
I give permission for the data, age and gender, that I provide to be archived in an excel file so it can be used for future research and learning.	☐	☐

Signatures

_____ _____ _____

Name of participant **Signature** **Date**

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

_____ _____ _____

Researcher name [printed] **Signature** **Date**

Study contact details for further information:

Julian Großerichter j.groserichter@student.utwente.nl

Janis Hölter j.holter@student.utwente.nl

Contact Information for Questions about Your Rights as a Research Participant: 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(s), please contact the Secretary of the Ethics Committee/domain Humanities & Social Sciences of the Faculty of Behavioural, Management and Social Sciences at the University of Twente by ethicscommittee-hss@utwente.nl

Responses to Demographic Data Survey

Other, please indicate;

What is your gender?: Other, please indicate; - Text 44

Other, please indicate;

What is your age?

63

23

26

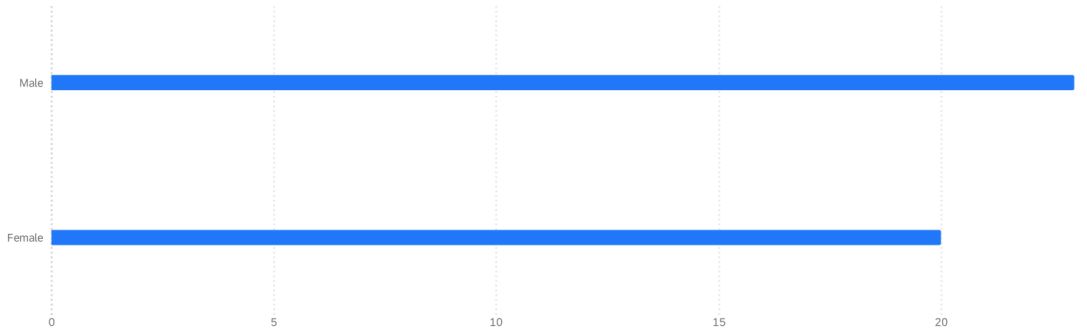
21

23

64

25

What is your gender? 43



What is your gender? 43

Q3 - What is your gender? - Selected Choice	Count	Count
Male	53%	23
Female	47%	20

What is your gender? 43

Q3 - What is your gender? - Selected Choice	Average (Q3 - What is your gender? - Selected Choice)	Minimum (Q3 - What is your gender? - Selected Choice)	Maximum (Q3 - What is your gender? - Selected Choice)	Count
Female	2.00	2.00	2.00	20
Male	1.00	1.00	1.00	23

What is your age?

25

25

23

19

70

67

18

63

53

58

32

35

20

23

60

24

22

22

52

22

25

What is your participant number?

16

37

2

42

What is your age? 44

What is your age?

21

22

60

23

24

62

64

20

23

24

23

37

21

18

25

What is your participant number?

30

38

36

7

39

6

21

5

4

1

32

15

33

11

9

24

31

10

3

27

19

demographic data / Page 1

Responses: 44

What is your participant number? 44

What is your participant number?

8

40

22

35

26

18

43

25

20

34

41

17

13

14

23

28

12

29