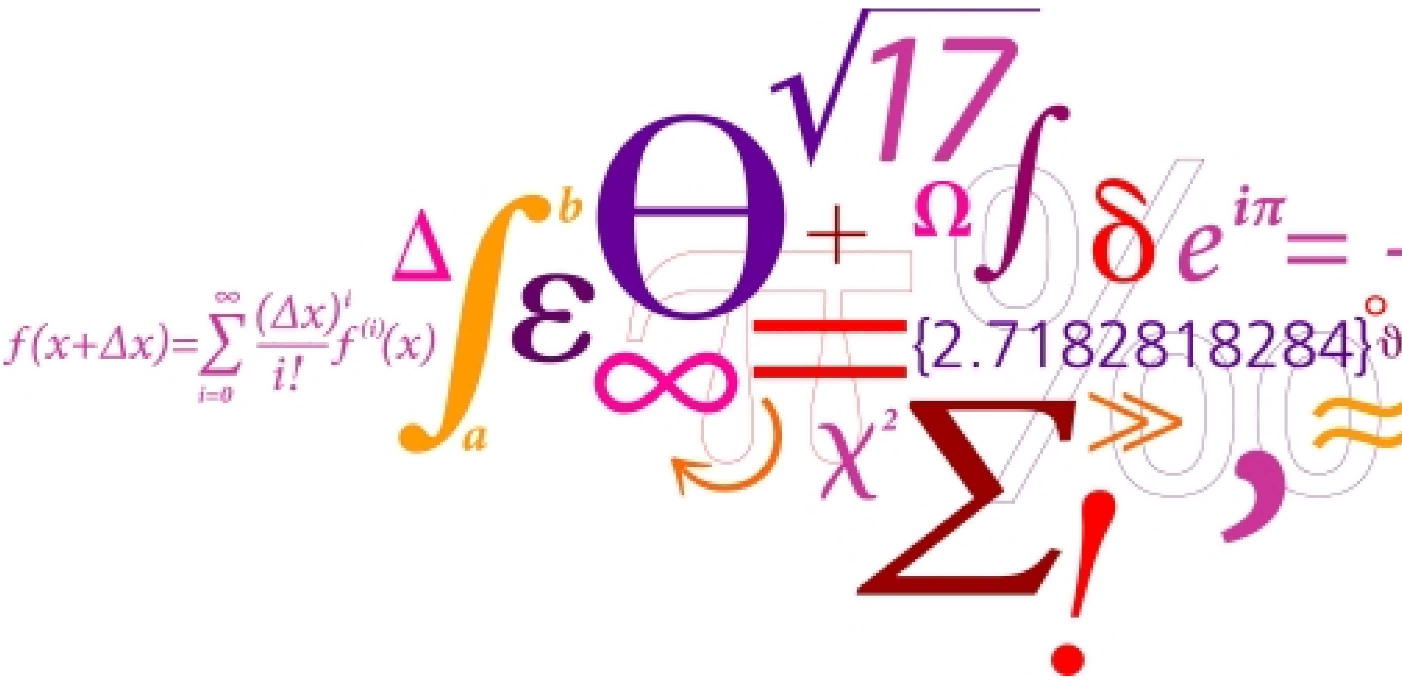
DTU Compute
Department of Applied Mathematics and Computer Science

Expressive Robots and their impact on the performance of a Gaze-cueing task.

02455, Experiment in Cognitive Science

December 2019

Luna Skytte Hansen, Andreas Råskov Madsen and Vimal Suresh Mollyn



Title: Expressive Robots and their impact on the performance of a Gaze-cueing task.

This report was prepared by:

Andreas Råsksov, s183901

Luna Skytte Hansen, s144289
Vimal Suresh Mollyn, s191385

Supervisor:

Invana Konvalinka, Assistant Professor,
Department of Applied Mathematics
and Computer Science, Technical
University of Denmark

Submit Date:
December. 8th, 2019

Contents

Introduction

Background & Motivation

Entertaining and assistive robots are an emerging topic of robotics and Human Robotic Interaction research. A challenging aspect with regards to the compliance of robotic technology is the development of human like appearance and behavior with the end goal of what is called humanoid robotics [2]. Humanoid robots can play a significant role in future research of social and behavioral studies. Robotic technologies provide the prospect of a fully controlled experimental environment with no noise from the variation that is usually introduced when using real humans for stimulation. Robots can also play an important role in the study of animal behavior. Using animals as stimuli is even harder than using humans and the field of animal research can therefore also benefit from using realistic animal robots [3]. The aspect of safety is also a significant advantage that can play a role for the design of an experiment.

The benefits of using robots for both research and daily matters are many. But the challenges of robotic design and interaction is just as apparent. The major challenge lies in getting the same effect and response when using a robot as when using a real human. This is especially true when dealing with research. A study investigated the expected features in a social robot and found that the ability to communicate and express emotions played an important role. Important physical features included the head, mouth and eyes [4].

The spatial cueing paradigm by M. Posner (1980) is a widely used paradigm in cognitive science to asses a measure of attention [5]. The paradigm has been used in different variation for various groups including children [6], neurological patients [7] and autism [8]. The paradigm involves a symbolic or spatial cue followed by a stimuli appearing on either side of the cue to which the participant should respond as fast as possible [5]. The findings of the paradigm claims that the reaction time (RT) is faster for the valid cueing condition compared to the invalid cue condition, even though the cue is not always predictive [9].

One variation of the cueing paradigm is the gaze cueing paradigm in which right or left gaze from a photograph of human is used as cue [10]. Other studies have used different variations of gaze like schematic drawings of gaze [11]. Both studies support the use of gaze to obtain a gaze cueing effect (GCE) and faster reaction time (RT) in the valid cue-target trials.

Gaze cueing plays an important role in social interactions like having a conversations. Gaze cueing acts as a signal in the roles of the conversation; like who is being addressed in the current moment of the conversation. Gaze also signals turns of speaking and it provides information on the speaker's course of topic [12].

A study has also investigated the impact of eye contact during gaze cueing using a humanoid robot [13]. The study performed two experiments with 50% and 25% cue-target validity respectively. Each experiment investigated the two conditions of eye contact vs. no eye contact. The study computed the gaze cueing effect (GCE) as the difference of reaction times between invalid and valid trials for the eye contact and the no eye condition separately. The result of the study showed a positive GCE for experiment 1 for both the eye contact and no eye contact condition, which indicated that the participants responded faster in the valid cue-target trials. They also found a larger GCE for the eye contact condition. A study also suggest that patients diagnosed with autism are more likely to follow the gaze of a robot than a human in a gaze cuing task [8].

Another study investigated the use of happy, angry and fearful emotions in a gaze cuing task in infants in the age of 9-12 month-old [14]. The study measured saccadic latencies in response to a peripheral target and found a positive bias in the gaze-cueing tasks. The results show that orientation to the target cued by a happy face was faster compared to angry and fearful faces and that the gaze cueing effect was present in the happy-face cue conditions and not in the angry and fearful cues. Studies do also suggest that attention and emotions are correlated as amygdala plays a role for both [15].

Another study [16] investigated the impact of a gaze cues expression positive emotions, negative emotions and neutral emotions on the affective evaluation of unfamiliar faces. Main findings however includes no evidence that emotional cues effect the evaluation of other people and no evidence of a main effect of emotion on the reaction time. This is however in dispute with the very intuitive work by Bayliss et. al [17] which showed that objects were rated higher when gazed by a happy expression.

To summarize, the motivation for this study is to investigate the use of robotics in cognitive research. The future will most likely bring more advanced technology and robotics into our daily life. Studies show that gaze cuing plays an important role in social interactions and that emotions impact the compliance and affiliation for humanoid robots. The studies suggest that reaction time can be affected by a robot in a gaze cueing task and that introducing human features like eye contact in a robot can also play a role. This study will therefore explore the use of a simple robot with features like eyes and motion but also emotions in a variation of the well studied gaze cueing paradigm. The study will explore if an emotional robot can cause the same gaze cuing effect as seen in other studies.

Research Question & Hypothesis

The study presented here will aim at investigating the following questions. 1) Can the gaze cueing effect be observed in an experiment using a simple robot as cue. 2) Can a robot expressing emotions compared to a robot with no emotions impact the performance of a subject when used for cueing in a gaze cueing paradigm?

It is hypothesised that the gaze cuing effect (GCE) can be observed in a gaze cuing paradigm with 50% valid trials using a simple robot as cue. The null hypothesis to be tested is therefore

$$H_0 : \mu_{valid} - \mu_{invalid} = 0$$

Where μ is the reaction time for valid and invalid cue condition. The hypothesis is tested for both the emotional and unemotional cases.
Further more it is expected that both emotions and validity influences the reaction time in a the gaze cuing task. The reaction time should decrease in the valid cue condition and increase in the invalid cueing condition, when the cue condition is with an emotional robot compared to the control condition of an unemotional robot. This is to be tested with a Generalized Linear Mixed Model.

Methods

Participants

Participants for this study were mainly students from the Technical University of Denmark (DTU). The recruitment was not limited to this demographic, however practicalities lead to this being the case. The recruitment was mainly carried out by personal request to fellow students. The inclusion and exclusion criteria used during recruitment can be seen below.

Inclusion criteria:

- Age 18-50
- Male or female

Exclusion criteria:

- Under 18, over 50
- Unable to understand given information

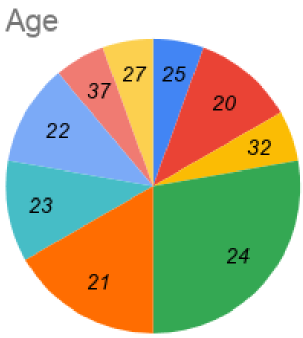


Figure 2.1: Age distribution of participants

In total 18 participants (9 male, 9 female) were recruited and non where excluded. The distribution of age can bee seen in Figure 2.1. The age range was from 20 to 37 and the most frequent age was 24.

Another study on gaze cueing [18] found a difference in RT between valid and invalid cues of $M_{valid} - M_{invalid} = 15.5$ ms and a standard error of mean of $\sqrt{\frac{SEM_{valid}^2 + SEM_{invalid}^2}{2}} = 10.81$ using gaze cueing from a robot initiated by no eye-contact (ie straight-ahead gaze), which can be argued is similar to the design presented in the study of this report. For a desired power of 0.8 and a 0.95 significance level, R-studio can be used to calculate that at least 8.7 participants should be used to obtain the same effect. Other studies have used 24 [18][19], 34 [13], 18 [8] and 44 [20] participants.

Ethical considerations

This study has been approved by the ethics committee at the Department of Applied Mathematics and Computer Science at the Technical University of Denmark. The study was conducted over a 13 week period under the supervision of Assistant Professor Ivana Konvalinka.

Inclusion information and informed consent

The participants were given consent forms (see Appendices A.5-A.6) to fill out before they took part in the experiment. All information regarding data protection and handling was detailed in that document for the participants to go through. The participants were also provided with an information sheet (see Appendices A.3-A.4) pertaining to the experimental procedure, and the tasks that they had to perform. Further instructions regarding analysis and results were given through a debriefing session once data collection was finished.

Data

The information provided is confidential. Only the experimenter and the principal investigator have access to the information provided. The participant's consent forms are kept separate from the observations collected during the study. Data is stored on a secure server in accordance with the University data storage policy. Once the data has been analysed a report of the findings may be submitted for publication. Only broad trends will be reported and it will not be possible to identify any individuals. All the data was anonymized at the source of data collection.

Safety and potential risks

The experiment did not involve unpleasant stimuli or situations or invasive procedures of any kind (including caffeine, drinks, food). It did not elicit pain. It did not involve deprivation of water, food, sleep, etc, drug administration, harmful procedures of any kind, stress inducing procedures, etc. Finally, we did not deceive or mislead the participants. There was no potential risks to this experiment.

Debriefing

The participants were debriefed at the end of data collection, to let them know of further details pertaining to the data analysis and results of the experiment. The participants were also made aware of the purpose of the experiment, and all questions were answered.

Experimental design

The experiment aims at investigating the influence of a simple robot expressing emotions in a gaze cuing task. The experiment is a within-subject 2x2 mixed factorial design, that measures the reaction time to a cue-target stimuli with emotions during the cue or no emotions during the cue. Furthermore the cue can be either congruent(valid) or incongruent(invalid) to the target.

The independent variables are the presence of happy emotions during the cue and validity of the cue. The main dependent variable is the reaction time from the appearance of the target to the response from the participant.

The design of the study enables to investigate 1)If any gaze cueing effect exists for the emotional or unemotional cue conditions, 2)If any main effect of emotions and validity influences the reaction time and 3)If any interaction between emotions and validity exists.

Considering the emotional/unemotional, valid/invalid and right/left a total of 8 trials of different cue-target combination exist for this experiment. The stimuli will be presented to the participant in sessions of 16 trials with two of each cue-target combination. Each session will be repeated five times for each participant. The order of the stimuli will be fully randomized in each session.

Experimental procedure and data collection

After recruitment participants are asked to read the information sheet and sign the consent form. The participant is then instructed in the course of the experiment and any questions are answered. The experiment is carried on a PC with the participant facing the screen, wearing headphones and operating the keyboard left and right keys (see Figure 2.2). The experiment consist of a training session and 5 testing sessions. The purpose of the training session is to introduced the participant to the setup and allow him/her to get used to operating the keyboard. During the training session the participant will see a cross followed by an arrow (pointing left or right) and then an X on the left or right side of the screen. The participants are instructed about this and the fact that the target is not always true. They are also instructed to press either the right or left keyboard key as fast a possible according the position of the X. Figure 2.3 shows the screen presented during the training session. The total duration of a single trial during the training session is five to six seconds, depending on the participants reaction time. A total of 16 trials are performed during training, giving a total of 80 to 96 seconds for the training session.

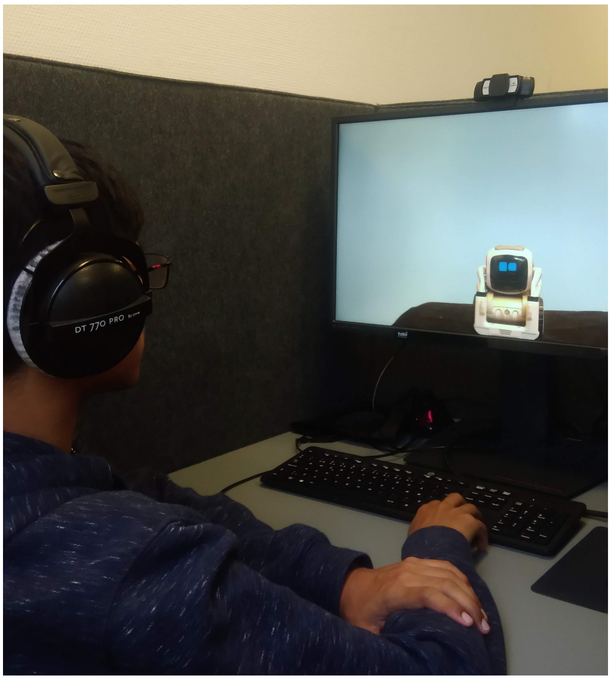
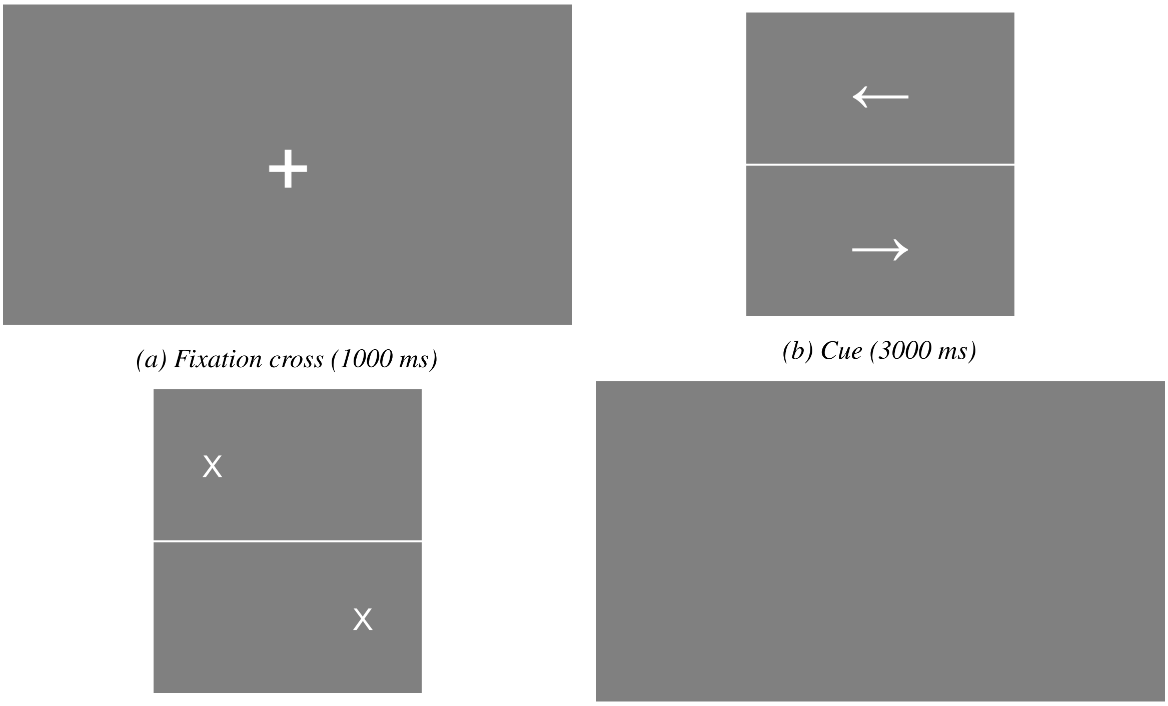


Figure 2.2: Experimental setup (Pilot study)



(a) Fixation cross (1000 ms) *(b) Cue (3000 ms)*

(c) Target (200 ms) *(d) Waiting for response (inf)*

Figure 2.3: Visual stimuli presented to the participants during training

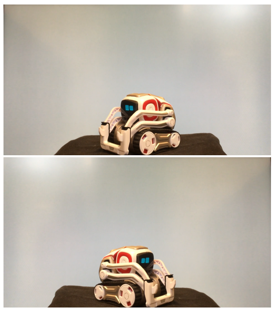
After the training session, the participant is instructed that in the testing sessions the cue will be different from the arrow but still not always true. The participant is told that the testing consist of 5 sessions and if needed the participant can take breaks between these. A wait-screen will appear between the five sessions which allows the participant the pause the experiment. During the testing sessions the participant will first see a fixation cross, followed by a right or left cue and then a right or left target. The cue during the testing is a video of a robot turning either right or left. The robot performs the cue by first expressing a happy emotion with motion and sound or no emotion and then turning left or right and back.

In both the training and the testing sessions the cue can be valid, meaning it cues in the same direction as the target, or invalid, cueing in the opposite direction of the target. The ratio of valid/invalid cues are 50/50.

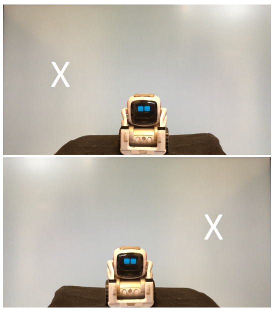
The data collection took place in a lab at the Department of Applied Mathematics and Computer science at the Technical University of Denmark. The lab contains 10 PCs and other studies were also conducted during the data collection of this study. Participants were however told to wear headphones that both enabled the participant to hear the sound of the cueing robot and to block out additional noise.

The experiment is designed using PsychoPy [21] a Python based open-source software for building paradigms mainly for neuroscience and psychology (see the full illustration of the paradigm in Figure A.4 in Appendix A.2). The robot used for the experiment is the Cozmo robot from Anki [22]. The robot can be programmed using the provided app or using Python. For this experiment the robot was programmed using Python to meet the wanted specifications of the cue.

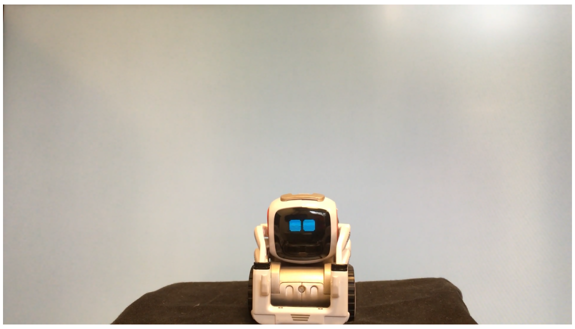
(a) Fixation cross (1000 ms)



(b) Cue (9000 ms)



(c) Target (200 ms)



(d) Waiting for response (inf)

Figure 2.4: Visual stimuli presented to the participants during testing. Note; images of the robot are representing videos.

Pre-processing of data

For each participant a .csv file was generated. After completing the data collection general information from each recording was extracted from the .csv files and merged to a single file. The extracted information includes the following information for each trial:

- *#subject*: The participant's anonymized number
- *Data_code*: The "codename" for the stimuli (cue and target) of the given trial.
- *Stim_dir*: The coordinate (0.5/-0.5) of the target
- *resp_train.rt* and *resp.rt*: The reaction time for the training and testing trails.
- *resp_train.keys* and *resp.keys*: The response key for the training and testing trials.

To ease the analysis of our data, separate data files where further generated as described in the following: The first file (Figure 2.5) contains the data of all trials with the data of the training sessions and outliers removed, see section below. A second file (see Figure 2.6) contains the same information however the trials have been averaged for each condition for each participant to be used in the T-test analysis of the gaze cueing effect.

	A	B	C	D	E	F	G
1		#subject	Data_code	resp.rt	Correctness	Validity	Emotions
2	18	P001	invalid_unemo	0.4348565	1	invalid	unemo
3	19	P001	valid_unemo	0.3035776	1	valid	unemo
4	20	P001	invalid_emo	0.3554292	1	invalid	emo
5	21	P001	invalid_unemo	0.2958308	1	invalid	unemo
6	22	P001	invalid_emo	0.363836	1	invalid	emo
7	23	P001	valid_emo	0.3278637	1	valid	emo
8	24	P001	valid_unemo	0.3255069	1	valid	unemo
9	25	P001	valid_emo	0.3547855	1	valid	emo
10	26	P001	invalid_unemo	0.4269403	1	invalid	unemo

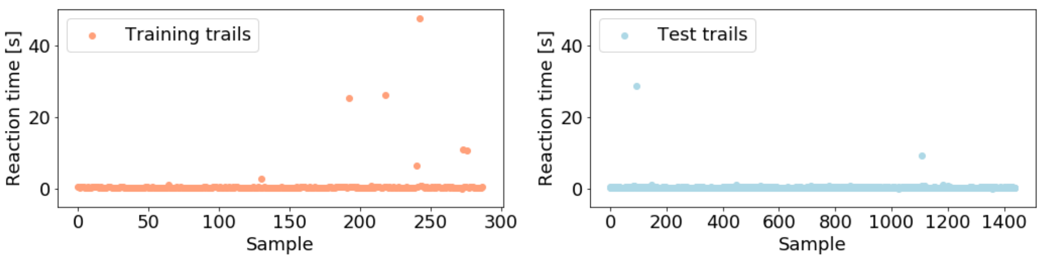
Figure 2.5: Preprocessed dataset

	A	B	C	D	E
1		#subject	Data_code	resp.rt	Correctness
2	0	P001	invalid_emo	0.35575004	1
3	1	P001	invalid_unemo	0.339915195	0.95
4	2	P001	valid_emo	0.337468015	1
5	3	P001	valid_unemo	0.340822805	1
6	4	P002	invalid_emo	0.4165721651	0.95
7	5	P002	invalid_unemo	0.43519385	1

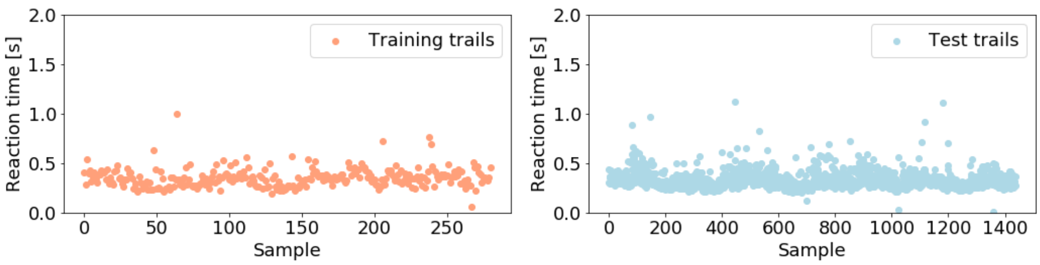
Figure 2.6: Preprocessed dataset, averaged over trials for each condition for each participant.

Outliers

A scatter plot of the reaction time in all trials is seen in Figure 2.7a. The normal reaction time for visual stimuli is under 1 second [23]. It was therefore found safe to exclude any sample of reaction > 2 seconds from the analysis, see Figure 2.7b. A total of 2 outliers in the test trials and 7 in the training trials was found and excluded.



(a) With outliers



(b) With no outliers

Figure 2.7: Scatter plot of all trials

Data analysis and statistical tests

As described earlier, our experiment is a within-subject 2x2 mixed factorial design as shown in figure 2.8. The first part of our hypothesis (the gaze cueing effect) includes 1 factor (Cue Validity) with 2 levels in a within subjects design. From the Q-Q plots of our data (fig 3.3b) we noticed that our data was roughly normally distributed and hence, we chose to test the hypothesis with the paired-samples t-test [24]. The second part of our hypothesis includes 2 factors (Cue Validity and Emotion) with 2 levels each in a within subjects design. We decided to use the General linear mixed model to test this part of the hypothesis, as we found noticed that this model would be able to take care of both fixed and random effects, as well as the interaction effects between the 2 factors [24]. All statistical test will be performed using `RStudio`.

		Factor A: Emotion	
		Emotional	Unemotional
Factor B: Cue Validity	Valid	Valid Cue and Emotional Robot	Valid Cue and Unemotional Robot
	Invalid	Invalid Cue and Emotional Robot	Invalid Cue and Unemotional Robot

Figure 2.8: 2-factor design

Paired T-test

To test the gaze cueing effect, the mean value of the four conditions; emotional valid cue, emotional invalid cue, unemotional valid cue and unemotional invalid cue was calculated for each participant be averaging over trials, see Figure 3.1.

A t-test will be performed to compare the valid and invalid cues in the emotional condition and the unemotional condition. The t-test will be paired because of the within subject design, and two sided because the hypothesis only specifies that valid and invalid cues should be different. The null-hypothesis to test for both the emotional and unemotional condition is therefore that the difference in RT between the valid and invalid condition is zero.

$$H0 : \mu_{valid} - \mu_{invalid} = 0$$

Generalized Linear Mixed Model

A generalized linear mixed model is used to test if the emotions and the validity play a role in the performance of the participant, measured using the reaction time. Generalized Linear Mixed Models (GLMMs) are non-parametric models that account for fixed, random and interaction effects between various factors in an experiment. The fixed effects for our experiment were the emotion of the robot and the validity of the cue. We expect a random effect to exist for the subjects as well as the trial number, and hence we added those to our model as well. Our final model was as follows:

$$Reaction\ Time \sim Validity \cdot Emotions + (1|subject) + (1|trial) \tag{2.1}$$

Results

Figure 3.1 show the mean reaction time for each participant in each of the four paradigm conditions. See Appendix A.1 Figure A.1 for the full dataset.

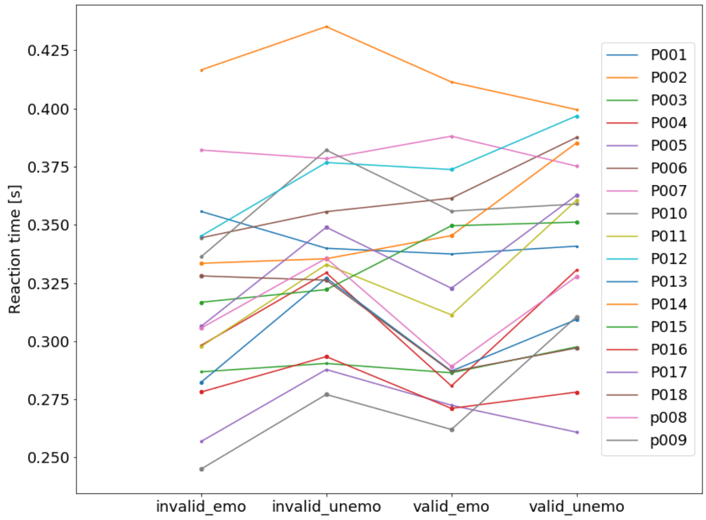
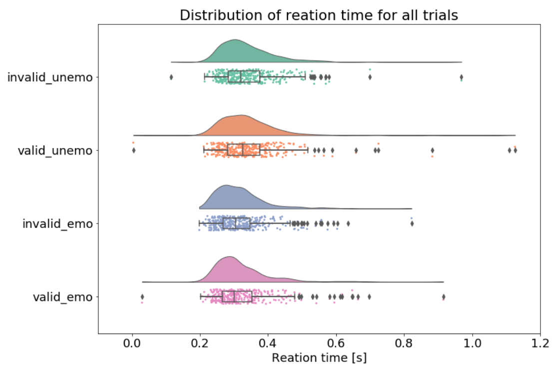


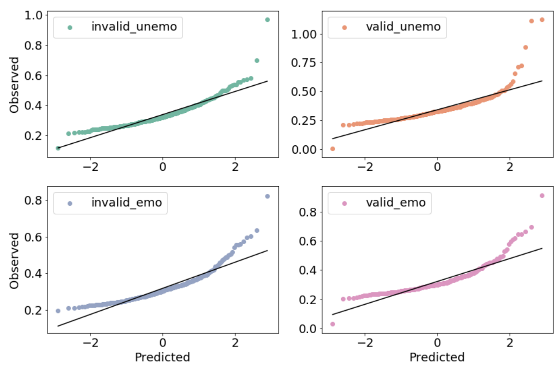
Figure 3.1: Mean reaction time for each participant in each paradigm condition.

Data visualization and distribution

The following figures show the distribution of the data. Figures 3.2 show a rain cloud plot and a QQ-plot of the full dataset. The colors of the rain cloud correspond to the colors used in QQ-plot. The QQ-plot shows that it is difficult to approximate the data by a normal distribution.



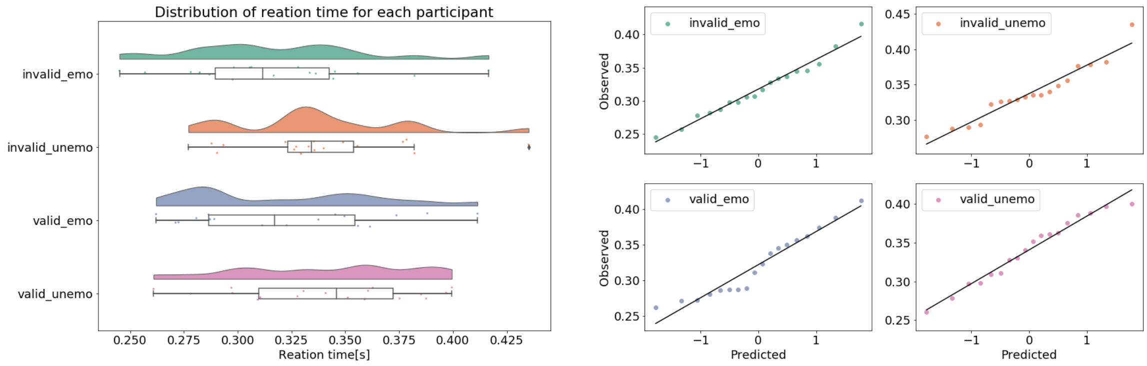
(a) Raindcloud of all trials each condition



(b) QQ-plot of all trials each condition

Figure 3.2: RT distribution of all trials

Figure 3.3 show the distribution of the data averaged over trials for each participant. This data is better approximated by a normal distributed as shown in the QQ-plot.



(a) Raindcloud of each condition averaged over trials *(b) QQ-plot of of each condition averaged over trials*

Figure 3.3: RT distribution of averaged trials

Using a robot to obtain the gaze cuing effect

Figure 3.4 shows the boxplot of the RT averaged over trials for each participant. The left most part of the figure shows the emotional cue condition and the right most part shows the unemotional cue condition. Paired-samples t-tests were performed to test the presence of the gaze-cueing effect in the emotional and non-emotional conditions.

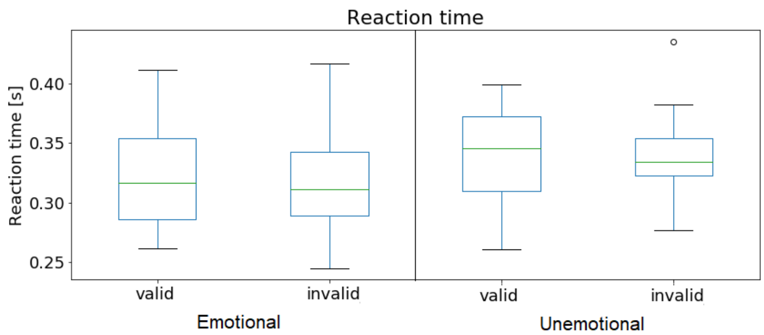


Figure 3.4: Box plot of mean reaction times for each condition.

Gaze cueing	t-value	df	Mean of difference	95% Confidence interval	p-value	Significance
Emotional	0.96167	17	0.0043	[-0.0051,0.01372]	0.3497	None
Unemotional	0.512194	17	0.0031	[-0.0094,0.01556]	0.6084	None

Table 3.1: T-test of gaze cuing effect, valid vs invalid. See figure 3.4

The result of a paired t-test between emotion valid and invalid cue condition seen in the left most part of figure 3.4 gave $p = 0.3497$ (see Table 3.1). The result of a paired t-test between unemotional valid and invalid cue condition seen in the right most part of figure 3.4 gave $p = 0.6084$ (see table 3.1). See complete output from `RStudio` analysis in Appendix Figure A.2.

Effect of happy emotions on reaction time

Figure 3.5 shows the observed trend for the emotional and the unemotional cue condition for the valid and invalid cues. The figure suggests a trend that the RT of the unemotional cue condition, both valid and invalid, is higher than the emotional cue condition.

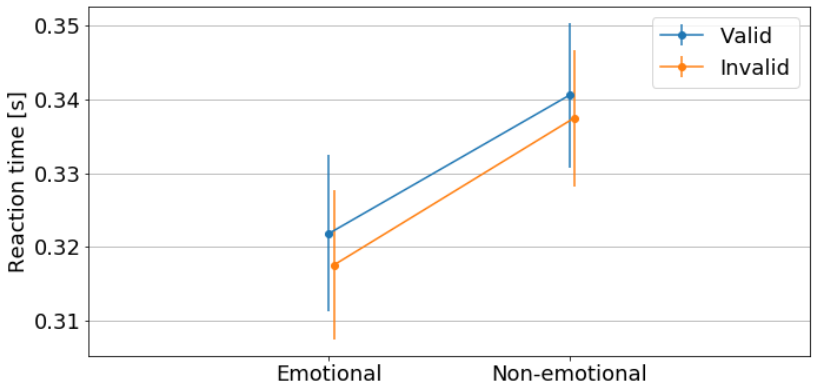


Figure 3.5: Trends for the RT in each condition averaged over trials. Standard error of mean is used for the errorbars.

As stated earlier, a Generalized Linear Mixed Model was fit on our data, with the robot's emotion and cue validity as fixed within-subjects factors and trials and subjects as random factors 2.1. The tables 3.2 and 3.3 show the results of the random effects and the fixed effect in the GLMM. See complete output from `RStudio` analysis in Appendix A.1 Figure A.3.

Groups	**Name**	**Variance**	**Std. Dev.**
Trial	(Intercept)	0.003662	0.06052
Subject	(Intercept)	0.003869	0.06220
Residual		0.005816	0.07626

Table 3.2: Random Effects of the Generalized Linear Mixed Model

There was a main effect of the robot's emotion on the reaction time ($p = 0.000221 < 0.01$). The reaction time did not vary significantly with the other fixed factors or interactions ($p > 0.05$).

	Estimate	**Std. Error**	**t-value**	**Pr(>|t|)**	**Significance**
(Intercept)	-1.20109	0.04750	-25.289	<2e-16	High
Validity (Valid)	0.01738	0.01742	0.998	0.318405	None
Emotion (Unemotional)	0.06603	0.01722	3.836	0.000125	High
Interaction (Valid:Unemotional)	-0.01709	0.02412	-0.708	0.478691	None

Table 3.3: Fixed Effects of the Generalized Linear Mixed Model

Discussion & Conclusion

The work of the report was to explore the use of a simple robot expressing emotions in a gaze cueing task. We did not find any evidence to support the hypothesis that a robot, neither emotional or non emotional, will have a gaze cueing effect. However our GLMM suggests that an emotional robot will improve the overall reaction time by 21.27 milliseconds when compared to an unemotional robot. The validity was not found to impact the reaction time. Further studies should be made to investigate this effect.

Can a robot trigger a gaze cueing effect?

The result of the paired t-test did not show any gaze cueing effect in either the emotional or unemotional cue condition. The test compared the valid cue and the invalid cue and found a p-value of 0.35 for the emotional cues and a p-value of 0.6 for the unemotional cues. This indicates that there is no significant difference between the valid and invalid cues for either the emotional or unemotional cues.

The fact that we are not seeing a gaze cueing effect in our study at all does not support what other studies have observed with regards to gaze cueing task using human gaze [9], [10],[8]. However another study also didn't find any GCE with the use of a robot with no eye-contact [13], which can be argued is similar to the conditions presented in this study.

A happy robot impacts the reaction time in a gaze cueing task

Our results showed a significant main effect of the Emotion factor on the reaction time of the participant. From the estimates of the fixed effects (table 3.3), we can see that an unemotional robot increases the reaction time by about 21.3 milliseconds when compared to an emotional robot, which is significant as seen from the p-value ($p<0.01$). In other words, emotional robots induce faster reaction times, thus verifying the second part of our hypothesis.

A valid cue increases the reaction time by about 4.5 milliseconds when compared to an invalid cue, which is deemed to be insignificant from the p-value ($p>0.05$). The interaction effect between the fixed effects of validity and emotion decreases the reaction time by around 3.5 milliseconds and hence is also deemed to be insignificant from the p-value ($p>0.05$).

Furthermore, the random effects that we included in our model (2.1) do not seem to describe a significant part of the variance (<1%) in the reaction times of the participants, as can be seen from table 3.2.

Limitations

A number of limitations and post study considerations have been made clear during the work of this study. The original design of the study included the robot in a live version to emphasize the human-robot interaction. However a video of the robot was used instead of a real life robot because of certain problems faced to make the robot perform reliably in test settings. Hence, a video of the robot was used. This may have affected our results, and our original hypothesis was stated based on the assumption that emotions would make the robot more relatable. Since we used a video this may have affected this effect. However this should affect both our emotional and unemotional state. This may explained why no gaze cueing effect was detected for both states. It could also be that the cue was more distracting than it was acting as a cue.

The design of the study included a fixed duration between the cue and the target appearing. This could influence the results of the study as it makes it easier for the participant to stop paying attention to the task and just learn when to respond by pressing any key. Other studies do included a variation to this factor of the design in order to prevent this from happening.

Instruction on how to perform the experiment were given to the participant before the beginning the test. The participant was told that a cue would appear on either the left or right side of the screen, followed by a target. The participant was told to respond as quickly as possible by pressing either the left of right keyboard key according to the position of the target. No instruction on what finger to use for responding was given and different strategies for pressing the keys was observed during data collection. Some participants used two fingers from the same hand, one from each hand and some used just one finger. Clearer instructions on this should have been given to eliminate any effect caused by this.

The data included samples from 18 University students. This sample is not representative for the general population and should therefore not be used for any statements about the general population. The number of participants is a bit lower than what other studies have used, but higher than the calculated sample size. It is expected that a higher number of participant would decrease the variance and approximate a normal distribution better than what is currently observed.

Advancement of science

The findings of this study did not suggest any gaze cueing effect using a simple robot as cue, however the expression of emotions does seem to have an effect on the performance in the task. This confirms that implementing human features in a robot like movements and emotions can play an important role when developing humanoid robotics. As the gaze cueing effect does not seem to be apparent, different studies should be conducted to explore the social interaction between humans and robots.

Appendices

The appendices include the following:

- Appendix A.1: Additional figures
- Appendix A.2: Illustrations of the paradigm
- Appendix A.3 and A.4: Information sheet for the participant in Danish and English.
- Appendix A.5 and A.6: Consent form in Danish and English.

Figures

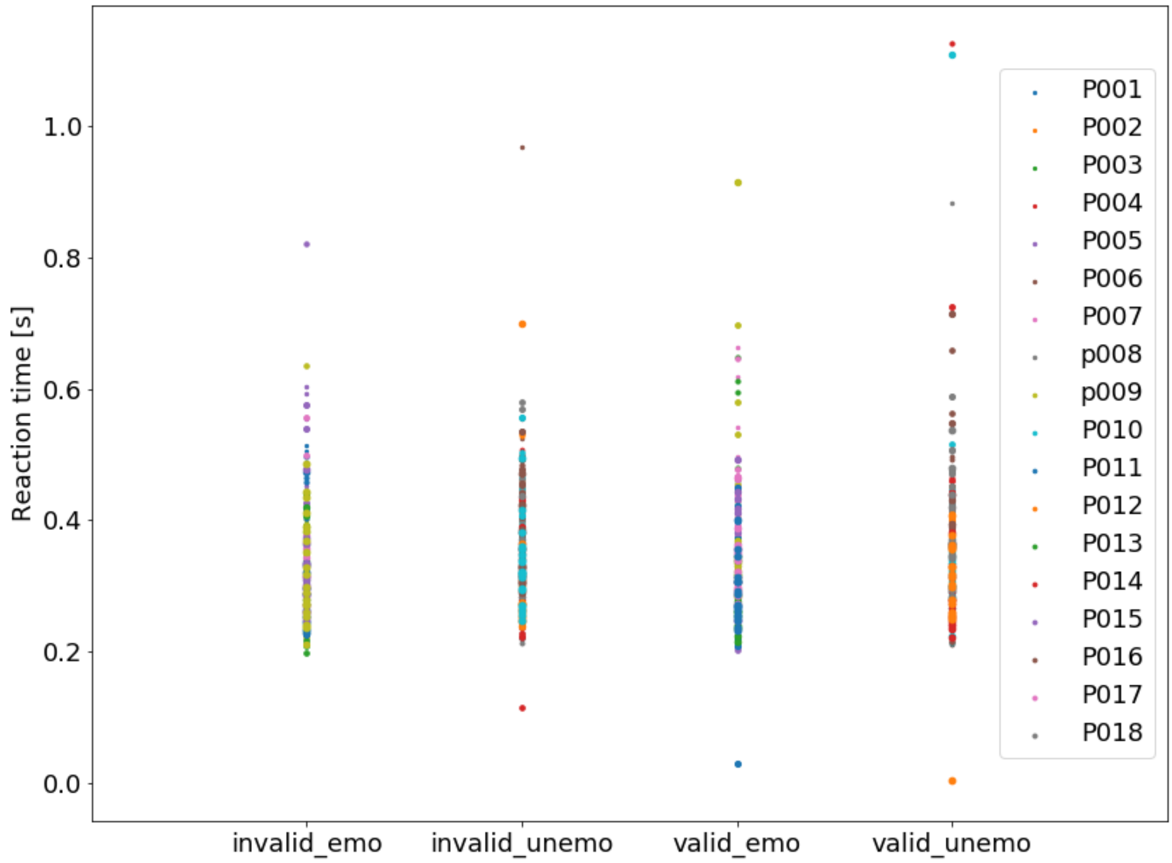


Figure A.1: Reaction time in all trials

```
> t.test(valid_emo$resp.rt,invalid_emo$resp.rt, paired = TRUE, alternative = "two.sided")

        Paired t-test

data:  valid_emo$resp.rt and invalid_emo$resp.rt
t = 0.96167, df = 17, p-value = 0.3497
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.005129692  0.013722748
sample estimates:
mean of the differences
            0.004296528

> t.test(valid_unemo$resp.rt,invalid_unemo$resp.rt, paired = TRUE, alternative = "two.sided")

        Paired t-test

data:  valid_unemo$resp.rt and invalid_unemo$resp.rt
t = 0.52194, df = 17, p-value = 0.6084
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.009388899  0.015561176
sample estimates:
mean of the differences
            0.003086139
```

Figure A.2: RStudio output of paired T-tests

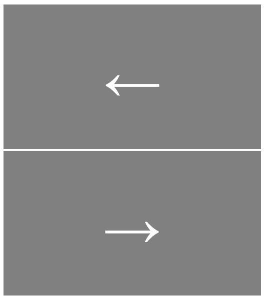
```
> summary(data.model)
Linear mixed model fit by REML ['lmerMod']
Formula: resp.rt ~ Validity * Emotions + (1 | X.subject) + (1 | X)
   Data: data

REML criterion at convergence: -3205

Scaled residuals:
    Min      1Q  Median      3Q     Max
-5.4463 -0.5097 -0.1818  0.3024  9.9557

Random effects:
 Groups    Name        Variance  Std.Dev.
 X         (Intercept) 0.0004437 0.02107
 X.subject (Intercept) 0.0015417 0.03926
 Residual              0.0056548 0.07520
Number of obs: 1438, groups:  X, 99; X.subject, 18

Fixed effects:
                            Estimate Std. Error t value
(Intercept)                 0.316987   0.010360  30.598
Validityvalid               0.004528   0.005722   0.791
Emotionsunemo               0.021268   0.005744   3.703
Validityvalid:Emotionsunemo -0.003353   0.008105  -0.414

Correlation of Fixed Effects:
            (Intr) Vldtyv Emtnsn
Validityvld -0.278
Emotionsunm -0.278  0.505
Vldtyvld:Em  0.197 -0.708 -0.709
```

Figure A.3: RStudio output of GLMM

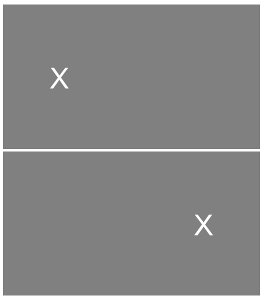
Paradigm

(a) Start

(b) Fixation cross (1000 ms)



(c) Cue (3000 ms)



(d) Target (200 ms)

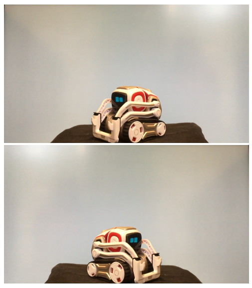
(e) Waiting for response (inf)

(f) Training trial (repeated 16 times)

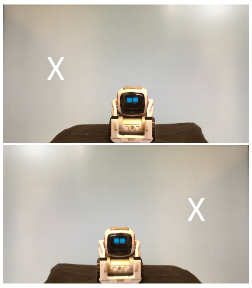
(g) Wait screen

(h) Start testing session

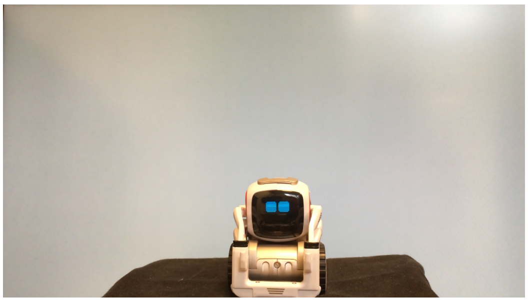
(i) Fixation cross (1000 ms)



(j) Cue (9000 ms)



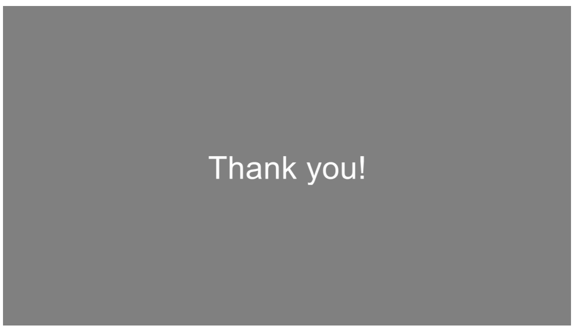
(k) Target (200 ms)



(l) Waiting for response (inf)

(m) Testing trial (repeated 16 times)

(n) Testing session (repeated 5 times)



(o) End screen

Figure A.4: Visual stimuli presented to the participants during the paradigm

Information Sheet (DK) [1]

***Experiment's tittle:* Expressive Robots and their impact on the performance of a Gaze-cueing task**

Du er blevet inviteret til at deltage i en forskningsundersøgelse. Før du beslutter dig for, om du vil deltage eller ej, er det vigtigt at du forstår, hvorfor forskningen udføres, og hvad den vil involvere. Tag dig tid til at læse følgende information omhyggeligt før du beslutter dig for at deltage i denne undersøgelse. Du er velkommen til at stille spørgsmål, hvis der er noget, der ikke er klart, eller hvis du ønsker mere information.
Du vil blive bedt om at udføre en simpel computer-opgave, der kræver, at du reagerer på et objekt, der dukker op på computerskærmen ved hjælp af det tastatur, der bliver leveret til dig. Du vil interagere med en robot der udviser følelser (uden at være i kontakt med den). Fuld instruktioner om, hvordan du udfører opgaven, vil blive givet til dig inden. Hver opgave vil vare cirka 5 minutter, og hele eksperimentet forventes at vare mindre end 60 minutter.

Behøver jeg at deltage

Deltagelse i denne undersøgelse er helt frivillig. Du er ikke på nogen måde forpligtet til at deltage. Vi værdsætter meget de data, som du bidrager med. Hvis du beslutter at deltage får du dette informationsark til at opbevare og vil blive bedt om at underskrive en samtykkeformular. Du har ret til at trække dig ud af eksperimentet til enhver tid og uden at give en grund.

Hvad sker der med de information du bidrager med.

De oplysninger, du giver, er fortrolige. Ingen bortset fra forsøgslederne og vores vejleder har adgang til de oplysninger du giver. Din samtykkeformular holdes adskilt fra observationer indsamlet i løbet af studiet. Data gemmes i i overensstemmelse med Universitets data-lagringspolitik. Når dataene er blevet analyseret, kan en rapport om fundne forelægges for offentliggørelse. Kun generelle tendenser vil rapporteres, og det vil ikke være muligt at identificere nogen individer. En oversigt over resultaterne vil være tilgængelig fra forsøgslederne på anmodning, når undersøgelsen er afsluttet.

Hvis du har spørgsmål eller har brug for yderligere oplysninger, bedes du kontakte forsøgslederne eller vores vejleder.

Forsøglederens navne: Andreas, Luna and Vimal
E-mail: s183901@student.dtu.dk, s144289@student.dtu.dk, s191385@student.dtu.dk

Tak for at du tog dig tid til at læse denne informationsformular og overvejer at deltage i undersøgelsen. Dette informations ark må du beholde. Hvis du ønsker at deltage i undersøgelsen, skal du underskrive samtykkeformularen.

Information Sheet (EN) [1]

Title of Experiment: **Expressive Robots and their impact on the performance of a Gaze-cueing task**

You are being invited to take part in a research study. Before you decide whether or not to take part, it is important for you to understand why the research is being conducted and what it will involve. Please take the time to read the following information carefully and decide if you want to take part in this study. Please feel free to ask questions if there is anything that is not clear or if you would like more information.
You will be asked to perform a simple computerised task that requires you to respond to a trigger on the computer screen using the keyboard provided to you. You will be interacting with a social robot that expresses emotions (in a non-contact fashion). Full instructions on how to complete the task will be given to you. Each task will last roughly 5 minutes and the entire study will last less than 60 minutes.

Do I have to take part

Participation in this study is totally voluntary, you are under no obligation to take part in this study. The data that you provide will be very useful for our study. If you decide to take part you will be given this information sheet to keep and will be asked to sign a consent form. You have the right to withdraw from the study at any time and without giving a reason.

What happens to the information I provide?

The information you provide will be confidential. No one apart from the experimenter's and expermental suppervicer will have access to the information you provide. Your consent form will be kept separate from the observations collected during the course of the study. Data will be stored in accordance with the University data storage policy. Once the data is analysed a report of the findings may be submitted for publication. Only broad trends will be reported and it will not be possible to identify any individuals. A summary of the results will be available from the experimenter on request once the study is complete.

If you have any questions or require any further information, please contact the experimenter or research supervisor.

Names of the experimenters: Andreas, Luna and Vimal
E-mail: s183901@student.dtu.dk, s144289@student.dtu.dk, s191385@student.dtu.dk

Thank you for taking the time to read this Participant Information Form and considering taking part in the study. This Participant Information Form is for you to keep. If you do wish to take part in the study, please sign the consent form.

Informed consent (DK) [1]

Samtykkeformular

Tak for din interesse i dette projekt. Vi vil minde dig om at den data du leverer i løbet af dette projekt, blive behandlet i fortrolighed og kun vil blive brugt til forskningsformål. Desuden vil du som deltager i denne forskning aldrig kunne blive identificeret i nogen udgivet resultater (f.eks. Rapporter, forskningsartikler), der stammer fra dette projekt, og dine data vil aldrig kunne identificeres eller ses af nogen anden part uden for forskerteamet.

Samtykkeformular
(Deltageren skal selv udfylde denne formular)

Experimentes title: **Expressive Robots and their impact on the performance of a Gaze-cueing task**

- ☐ Jeg bekræfter at jeg har læst og forstået informations arket for dette experiment.
- ☐ Jeg har haft mulighed for at stille spørgsmål og alle mine spørgsmål er blevet besvaret fuldestgørende.
- ☐ Jeg er forstået med at min deltagelse i forsøget er frivilig og at jeg til enhver tid kan trække min deltagelse tilbage uden at give nogen grund.
- ☐ Jeg har modtaget nok informationer om forsøget.
- ☐ Jeg giver mit samtygge til at deltage i forsøget.

"Dette experiment er blevet forklaret til mig tilfredstillende. og jeg ønsker at deltage. Jeg forstår at jeg til enhver tid kan trække min deltagelse tilbage."

____________________ __________________ ____________________

Deltagers navn Dato Signatur

"Jeg har forklaret experimentet til deltageren og han/hun har givet samtygge til at deltage."

____________________ __________________ ____________________

Forsøglederens navn Dato Signatur

Informed consent (EN) [1]

Consent Form

Thank you for your interest in this project. Just to remind you, the data you provide in the course of this project will be treated in the strictest confidence and will be used for research purposes only. Furthermore, as a participant in this research you will never be identified in any outputs (e.g., reports, research articles) that arise from this project and your data will never be identifiable or viewed by any other party outside the research team.

CONSENT FORM
(The participant should complete the whole of this sheet himself/herself)

Title of Experiment: **Expressive Robots and their impact on the performance of a Gaze-cueing task**

- ☐ I confirm that I have read and understand the information sheet for the above experiment.
- ☐ I have had opportunities to ask questions and my questions have fully been answered.
- ☐ I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason.
- ☐ I have received enough information about the experiment.
- ☐ I agree to take part in the above experiment.

"This experiment has been explained to me to my satisfaction, and I agree to take part. I understand that I am free to withdraw at any time."

__________________ _______________ __________________

Name of the Participant Date Signature

"I have explained the experiment to the above participant and he/she has agreed to take part."

__________________ _______________ __________________

Name of the Experimenter Date Signature

Bibliography

[1] "Consent forms." https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjBlbud0-HkAhXqkYsKHcdoDfcQFjAAegQIABAC&url=https%3A%2F%2Fosf.io%2F5a7vu%2F%3Faction%3Ddownload%26version%3D2&usg=AOvVaw0xkolSPUkuCQ3N4wdL_r_b.

[2] "Humanoid Robotics - IEEE Robotics and Automation Society - IEEE Robotics and Automation Society."

[3] A. Frohnwieser, J. C. Murray, T. W. Pike, and A. Wilkinson, "Using robots to understand animal cognition," *Journal of the Experimental Analysis of Behavior*, vol. 105, pp. 14–22, 1 2016.

[4] M. C. Martini, R. Murtza, and E. Wiese, "Minimal Physical Features Required for Social Robots," 2015.

[5] M. I. Posner, "Orienting of attention.," *The Quarterly journal of experimental psychology*, vol. 32, no. 1, pp. 3–25, 1980.

[6] X. Zhang, W. Fu, L. Xue, J. Zhao, and Z. Wang, "Children with Mathematical Learning Difficulties are Sluggish in Disengaging Attention," *Frontiers in Psychology*, vol. 10, no. APR, 2019.

[7] M. Dalmaso, L. Castelli, K. Priftis, M. Buccheri, D. Primon, S. Tronco, and G. Galfano, "Space-based and object-centered gaze cuing of attention in right hemisphere-damaged patients," *Frontiers in Psychology*, vol. 6, no. AUG, pp. 1–12, 2015.

[8] E. Wiese, H. J. Müller, and A. Wykowska, "Using a gaze-cueing paradigm to examine social cognitive mechanisms of individuals with autism observing robot and human faces," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8755, pp. 370–379, 2014.

[9] A. Frischen, "Gaze Cueing of Attention," *Psychol Bull*, vol. 133, no. 4, pp. 694–724, 2007.

[10] J. Driver, G. Davis, P. Ricciardelli, P. Kidd, E. Maxwell, and S. Baron-Cohen, *Gaze perception triggers reflexive visuospatial orienting*, vol. 6. 1999.

[11] C. K. Friesen and A. Kingstone, "The eyes have it! Reflexive orienting is triggered by nonpredictive gaze," *Psychonomic Bulletin and Review*, vol. 5, no. 3, pp. 490–495, 1998.

[12] B. Mutlu, T. Kanda, J. Forlizzi, J. Hodgins, and H. Ishiguro, "Conversational gaze mechanisms for humanlike robots," *ACM Transactions on Interactive Intelligent Systems*, vol. 1, no. 2, 2012.

[13] K. Kompatsiari, F. Ciardo, V. Tikhanoff, G. Metta, and A. Wykowska, "On the role of eye contact in gaze cueing," *Scientific Reports*, vol. 8, no. 1, pp. 1–10, 2018.

[14] A. Niedźwiecka and P. Tomalski, "Gaze-cueing effect depends on facial expression of emotion in 9- to 12-month-old infants," *Frontiers in Psychology*, vol. 6, no. FEB, 2015.

[15] S. Yamaguchi and K. Onoda, "Interaction between emotion and attention systems," 2012.

[16] T. L. Landes, Y. Kashima, and P. D. Howe, "Investigating the effect of gaze cues and emotional expressions on the affective evaluations of unfamiliar faces," *PLoS ONE*, vol. 11, 9 2016.

[17] A. P. Bayliss, A. Frischen, M. J. Fenske, and S. P. Tipper, "Affective evaluations of objects are influenced by observed gaze direction and emotional expression," *Cognition*, vol. 104, pp. 644–653, 9 2007.

[18] K. Kompatsiari, J. Perez-Osorio, D. De Tommaso, G. Metta, and A. Wykowska, "Neuroscientifically-Grounded Research for Improved Human-Robot Interaction," *IEEE International Conference on Intelligent Robots and Systems*, pp. 3403–3408, 2018.

[19] A. Marotta, J. Lupiáñez, D. Martella, and M. Casagrande, "Eye gaze versus arrows as spatial cues: Two qualitatively different modes of attentional selection," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 38, no. 2, pp. 326–335, 2012.

[20] D. A. Hayward and J. Ristic, "Measuring attention using the Posner cuing paradigm: the role of across and within trial target probabilities," *Frontiers in Human Neuroscience*, vol. 7, 2013.

[21] "Home — PsychoPy v3.0."

[22] Anki, "Cozmo | Meet Cozmo." https://anki.com/en-us/cozmo.html, 2018.

[23] A. Jain, R. Bansal, A. Kumar, and K. Singh, "A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students," *International Journal of Applied and Basic Medical Research*, vol. 5, no. 2, p. 124, 2015.

[24] I. Konvalinka, "Week10 - Data Analysis considerations and Linear Effect Mixed Models," tech. rep., Techincal University of Denmark. Deparment of Applied Mathematics and Data Science., 2019.

Ekstra Appendix

Extra appendix to korect mistakes made in the report:

- Appendix B correct version of sample size:
- Appendix B: correct GLLM, replace table 3.2 and 3.3

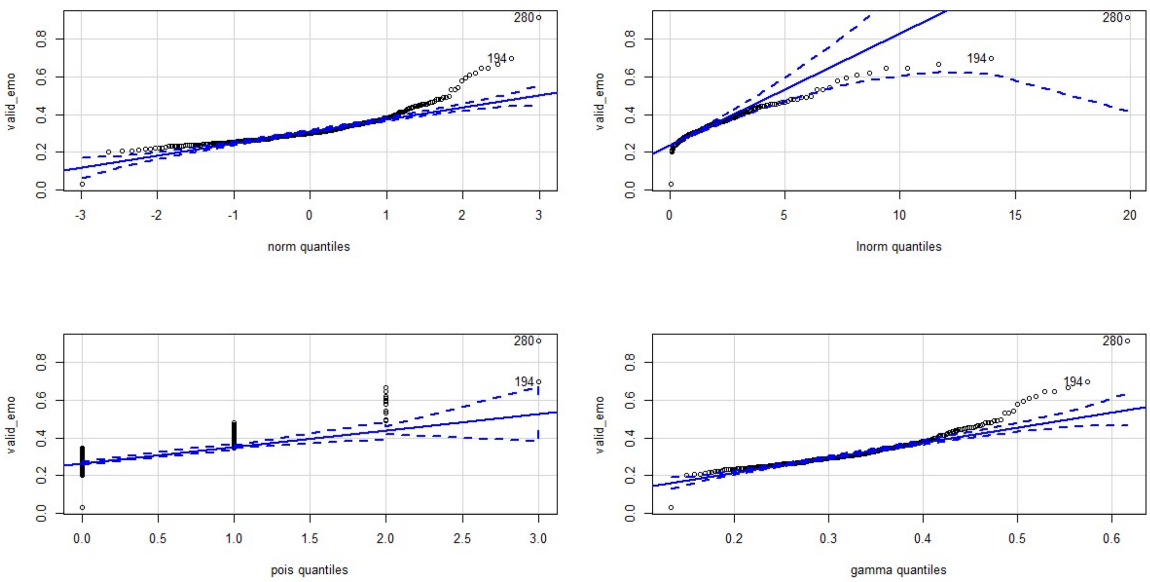
```
#input from Neuroscientifically-grounded research for improved human-robot interaction
M_valid=435
M_invalid=450.5

Delta_M=M_valid-M_invalid

SEM_invalid=11.21
SEM_valid=10.41

Partispants=24

sd_invalid=SEM_invalid*sqrt(Partispants)
sd_valid=SEM_valid*sqrt(Partispants)

sd_poled=sqrt((sd_invalid^2+sd_valid)/2)

power.t.test(delta = Delta_M,sd= sd_poled,sig.level = 0.05, power = 0.8)
```

(a) Code for calculating sample size

```
     Two-sample t test power calculation

              n = 101.1651
          delta = 15.5
             sd = 39.15952
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

(b) code output

Figure B.1: Correct samplesize



(a) The logarithmic data in qqplot

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [
glmerMod]
 Family: gaussian  ( log )
Formula: resp.rt ~ Validity * Emotions + (1 | X.subject) + (1 | X)
   Data: data

     AIC      BIC   logLik deviance df.resid
 -3314.5  -3277.6   1664.3  -3328.5     1431

Scaled residuals:
    Min      1Q  Median      3Q     Max
-5.4211 -0.5007 -0.1764  0.3120  9.8348

Random effects:
 Groups    Name        Variance Std.Dev.
 X         (Intercept) 0.003662 0.06052
 X.subject (Intercept) 0.003869 0.06220
 Residual              0.005816 0.07626
Number of obs: 1438, groups:  X, 99; X.subject, 18

Fixed effects:
                            Estimate Std. Error t value Pr(>|z|)
(Intercept)                 -1.20109    0.04750 -25.289  < 2e-16 ***
Validityvalid                0.01738    0.01742   0.998 0.318405
Emotionsunemo                0.06603    0.01722   3.836 0.000125 ***
Validityvalid:Emotionsunemo -0.01709    0.02412  -0.708 0.478691
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) Vldtyv Emtnsn
Validityvld -0.190
Emotionsunm -0.193  0.525
Vldtyvld:Em  0.139 -0.727 -0.714
```

(b) RStudio output of GLMM

Figure B.2: Corret GLMM