Utrecht University

A dual eye tracking study: Uncovering gaze patterns during live conversations with different emotional content

Master Thesis 27,5 ECTS

Author:
Sara Pruthi Kola
7091575
s.pruthikola@students.uu.nl

Supervisor:
-Niilo Valtakari
n.v.valtakari@uu.nl

Second assessor:
-Yentl de Kloe

Master's program: Applied Cognitive Psychology
Faculty of Social and Behavioral Sciences
Utrecht University
Academic year 2020-2021

Abstract

The importance of the gaze has mesmerized writers, psychologists, philosophers and body language experts for ages. Despite the importance of social context, most studies that investigate gaze behavior have employed pictures of faces or social scenes, thereby ignoring the possible dynamics in gaze behavior when social partners are physically present. Hence, the aim of this study is to allow social interaction between participants while recording their gaze. We designed an experiment where the participants were asked to converse about pictures with different emotional content, while their gaze behavior was being recorded by two Tobii eye trackers. The results suggest that there is bias towards looking at specific features of the face, but that this is not caused by the emotional content of the conversations. This outcome might have been influenced by situational and dispositional circumstances.

Keywords: Eye tracking; Emotions; Gaze patterns.

Table of Contents

Introduction

"For eyes can speak and eyes can understand." (Chapman)

"There are often voice and words in a silent look." (Ovid)

The importance of the gaze has mesmerized writers, psychologists, philosophers and body language experts for ages. Since the pioneering work of Yarbus (1967), researchers have tried to find out what information humans gather from other people's faces, by studying eye movements during the face exploration. A consistent finding has been that humans tend to fixate on the eye area of other people's faces. This has been indicated to be because of the extent of information that can be extracted solely from the eyes (Hessels et al., 2017).

Some researchers took the approach of analyzing the occasions during which the participants engage in mutual eye-to-face gaze or gaze aversion, and hence they have been able to pinpoint some of the functions these eye behaviors serve during interactions (Mirenda, Donnellan & Yoder, 1983). They have concluded that gaze behavior serves at least: (a) To communicate the type of the interpersonal relationship (Argyle, 1972; Exline & Winters, 1965; Knapp, 1980). During dyadic interactions gaze behavior can be also used to express inner feelings, especially interpersonal attitudes and the regulation of interpersonal intimacy (Argyle & Dean, 1965; Duncan & Fiske, 1977; Heslin & Patterson, 1982). For instance, nonverbal expressions of intimacy include a wide range of behaviors that reflect both positive affect and engagement, such as gaze, smiling, forward lean, and affirming head nods (Andersen, Guerrero & Jones, 2006). (b) As a display of interest in the other person or the topic of discussion (Exline & Winters, 1965). Fixations that are longer and occur more often are usually taken to be an indication of interest or of a greater level of relevant information, as it seems reasonable that we would spend more time looking at regions that catch our interest (Henderson & Hollingworth, 1999). (c) To get information about the other person's reactions (Argyle, 1972; Knapp, 1980). An anxious state, for example, can be shown by (1) tone of voice, (2) facial expression, (3) posture, (4) gestures and (5) gaze (Argyle, 1972). (d) To show attentiveness (Argyle & Cook, 1976; Exline & Fehr, 1978; Knapp, 1980). Seeing a face with direct gaze engages the observer's attention, perhaps because of the social significance transmitted by eye contact (Baron-Cohen, 1995). (e) To regulate turn-taking during a conversation (Argyle, 1972; Exline & Fehr, 1978; Kendon, 1967; Knapp, 1980). In typical human-human interactions, the speakers are expected to understand when it is good to talk rather than being overtly told when to do so. Eye gaze shows where the speaker's attention is directed, and thus helps the partner to understand what the speaker's intentions are, establishing effortless and effective communication without explicit verbalized expressions (Jokinen, Furukawa, Nishida & Yamamoto, 2013). (f) To indicate violations of norms in communication (Exline & Fehr, 1978). In more recent studies regarding this, the authors claim that changes in eye gaze when participants are being watched respond to demands of social norms (Foulsham et al., 2011; Laidlaw et al., 2011; Gobel et al., 2015). For instance, one subtle type of a social norm is civil inattention (Goffman, 1963), which suggests that the amount of gaze directed to strangers should suffice to recognize their presence, but without implying that they are of particular interest. (g) As a reinforcement - in the operative conditioning of verbal behavior such as looking one's partner in the eye when one says "I love you" or during wedding vows (Argyle, 1972; Argyle & Cook, 1976). (h) To begin or end an interaction - i.e. as long as one can avoid eye gaze in a superficially natural way, it is much easier to dodge interaction (Knapp, 1980). (i) To indicate hardship with encoding or acquiring information - for example, people tend to look away when having difficulty processing information, or deciding what to say (Kendon, 1967; Knapp, 1980). And lastly, (j) To show emotion - for example, positive feelings, such as love and admiration, are shown by increased gazing; there is a tendency to avoid gazing under conditions of guilt, embarrassment, sadness, and anger (Argyle & Cook, 1976; Knapp, 1980). Indeed, gaze can be used to signal relevant emotional states (Baron-Cohen, Wheelwright & Jolliffe, 1997) thus, enabling us to interact empathically.

Humans are very fast in "decoding" emotional expressions, reflecting their significance for successful social interactions (Schyns, Petro, & Smith, 2009; Smith et al., 2005). American psychologist Paul Ekman distinguished certain universal emotions whose visual cues are understood the same way across cultures. For instance, a picture of a grin signals joy to modern urban dwellers and native tribesmen alike. And according to Ekman, anger, disgust, fear, joy, sadness, and surprise are similarly recognizable (Ekman, 2005).

With developments in technology, new techniques for measuring gaze more precisely and objectively have emerged, using devices widely known as eye trackers (Valtakari et al., 2021). The way eye trackers work is by sending out near infrared light that is reflected in the eyes to create corneal reflections. These reflections are recorded by the eye tracker's cameras and it estimates where gaze is directed on the basis of the location of the pupil and the corneal reflections in the camera image.

Several studies have employed eye tracking in order to investigate gaze patterns. Rogers et al. (2018) for instance, used dual eye tracking to uncover personal gaze patterns during social interaction. With dual eye-tracking setups, the temporal and spatial relation between two gaze position signals can be analyzed. This allows for various other measures of gaze behavior, such as those relating to mutual gaze or shared gaze (Valtakari et al., 2021). The results suggested the existence of an eye-mouth gaze continuum. This continuum includes people showing a solid preference for eye gaze or mouth gaze, and others spreading their gaze between the eyes and mouth to varying extents.

Another study done by Schurgin et al. (2014) took the approach of examining gaze behavior while participants differentiated between emotional (joy, anger, fear, sadness, shame, and disgust) and neutral face expressions. The eye movements of the participants mainly fell in five different areas (eyes, upper nose, lower nose, upper lip, nasion). In addition, different fixation patterns appeared for each emotion, such as a focus on the lips for joyful faces and a focus on the eyes for sad faces.

Both gaze behavior and emotion have been found to be associated with the behavioral motivations to approach or avoid eye contact (see Argyle & Cook, 1976; Davidson & Hugdahl, 1995; Harmon-Jones & Segilman, 2001). Positive emotions, anger, and direct gaze, for example, are associated with the motivation to maintain eye contact. On the other hand, negative emotions (besides anger) are associated with the motivation to avoid eye contact.

A study done by Hessels et al. (2017) found that there is a bias for fixating the eye region of a partner that is physically present. Moreover, they reported that the time one partner in a duo spends looking at the eyes, is a good predictor of how long the other partner looks at the eyes.

Despite the importance of social context, most studies that investigate gaze behavior have used pictures of faces or social scenes, thereby overlooking the possible dynamics in gaze behavior when social partners are physically present (Risko et al., 2012). Therefore, an important study done by Laidlaw et al. (2011) showed that the participants looked at a videotaped confederate significantly more often and for an overall longer duration than other participants did toward the same confederate that was physically in the room. Indeed, the willingness of the participants to look at another individual was strongly influenced by whether or not that individual is physically present, and as such, whether the confederate was capable of looking back and engaging in a social interaction with the participant.

The way we detect and percept others' emotions is of crucial importance due to its multidisciplinary application in many areas, such as detecting mental disorders, marketing, human-computer interaction, education, etc. Therefore, the main purpose of this study was to investigate whether and how do emotional states influence gaze behavior. As mentioned above, there have been several studies that have measured emotion recognition using eye tracking through pictures or videos. As they provide with high experimental control, the representativeness of the findings to common conversations is problematic. This was pointed out in a recent study done by Risko, Richardson, & Kingstone (2016), where they found that in live situations (or conversations) gaze behavior might be different to that observed in studies using static stimuli. They explain that there is a dual purpose of gaze, as the eyes focus visual attention to collect information for the individual, while also potentially communicating information to others. Indeed, an awareness of the dual function of gaze can influence an individual to behave differently when aware that their eye movements are observable to another person.

The aim of the present study was to allow social interaction between participants while recording their gaze. In this case, we define social interaction as reciprocity of behavior between two participants, such as, conversing, laughing, performing a similar task, or simply looking at each other. In other words, we intended to investigate gaze patterns to faces during social interaction. Therefore, a two-way video setup was built, capable of recording eye movements from two participants while they looked at each other and had the possibility of engaging in interaction. Participants were placed across each other in the same room. Lastly, this design also allowed us to compare and explore whether the findings from studies with images and videos, apply to a live situation as well. An advantage to our approach is the dual eye-tracking method of our research. Dual eye-tracking setups allow for the simultaneous recording of the gaze of two participants, hence making it possible to measure how the gaze behavior of two individuals relates to one another (Valtakari et al., 2021). This method enables us to report on: a) which facial features people look at when listening to and conversing about topics with different emotional content, and b) the amount of eye contact throughout the conversations with different emotional content.

Considering the literature, the main research question of this study is: (1) Which facial features do people gaze at when in interaction with each other in conversations with different emotional content?, with the first sub-question being: (1) Where do people look at when they are talking about an emotional situation? (i.e. where they look when experiencing joy vs. when experiencing sadness), and the second: (2) How does the gaze behavior of one participant relate to the gaze behavior of the other participant and vice versa?

Based on the studies discussed above, we hypothesize that: 1) People look more at the eyes during conversations with negative emotional content, more at the mouth in conversations with positive emotional content, and we do not expect to find a bias for looking at any specific features of the face (or looking away) during conversations with neutral emotional content. 2)

People look away more when talking about their negative experiences, and at the other person more when talking about their positive experiences. 3) The emotional content of the conversations will have an effect on mutual gaze.

The conclusions drawn from this study could provide with further valuable knowledge on whether different emotional states influence our gaze behavior, as this would be relevant to future uses of emotion recognition technologies, but also areas such as psychology, marketing and education.

Method

Ethics Statement

Ethical approval was received through the Ethics Committee of the Faculty of Social and Behavioral Sciences of Utrecht University, and written informed consent was obtained from each participant prior to the start of the study.

Participants

A total of 26 participants were recruited from different channels, such as an announcement where those interested could sign up to participate to the study, that was posted on the University website.

Participants took part in one dyad each. 13 dyads were created from the 26 participants, and each participant was paired with the person who signed up for the same experimental session. Of the 26 participants, 4 of them were excluded because of technical difficulties. The experiment videos of the two dyads where corrupt, and therefore inutile.

Of the 22 remaining participants, 4 were male and the rest was female. Mean age was 26 years (*SD* = 6,7). 10 participants were Dutch, and their native language was Dutch. The rest of the participants were from other countries. 14 of the participants had a Master's degree.

All participants could choose between receiving monetary reimbursement (7€), or credits (0.75 PPU, if they were bachelor students) for their time. A post-experiment questionnaire was administered to determine participants' experience towards various aspects of the task, and as well as collecting demographic information.

Stimulus

Social scientists have been increasingly interested to investigate emotions under controlled environments such as laboratories. In order to achieve this, a number of emotion elicitation strategies have been used, such as: (a) conversations with trained confederates (e.g. Ax, 1953); (b) hypnosis (e.g. Bower, 1983); (c) repeating phrases (e.g. Velten, 1968); (d) facial muscle movements (e.g. Ekman, Levenson & Friesen, 1983); (e) imagery (e.g. Lang, 1979); (f) music (e.g. Sutherland, Newman, & Rachman, 1982); (g) slides (e.g. Lang, Ohman, & Vaitl, 1988); and (h) films (e.g. Lazarus et al., 1962; McHugo, Smith, & Lanzetta, 1982; Philippot, 1993).

Among these, pictures and film clips are the most widely used and accepted stimuli to elicit emotions. Based on theoretical arguments it was often assumed that the emotional effects of movie clips exceed those of pictures. Differing from this, a study done by Uhrig et al., demonstrates the opposite. More specifically, movie clips were less effective than pictorial stimuli in producing the corresponding emotional states (Uhrig et al., 2016).

Therefore, for this experiment, we decided to use pictures with different emotional content. Benefitting from their versatility, we were able to use them for a twofold purpose. Firstly, to elicit emotions, and secondly, to give the participants a topic to talk about.

These stimulus pictures were obtained from an online open-access stimulus set, called OASIS (Retrieved from https://db.tt/yYTZYCga). This set contains 900 color images that depict a broad spectrum of themes, including humans, animals, objects, and scenes. Images represent both the physical and social worlds. Utilizing a few thousand pixels, they can illustrate an unlimited selection of people, objects, and scenes that can evoke a variety of affective responses, such as happiness, excitement, contentment, sadness, anger, or disgust. What makes this standardized set more reliable is that the data was collected in 2015, and thus OASIS features more current images and reflects more current ratings than other stimulus sets (Kurdi, Shayn & Mahzarin, 2017).

From the 900 pictures, 12 of them were chosen for the experiment. There were 4 pictures for each condition (positive, negative and neutral). We selected 4 in order to give the participants as much freedom possible, as they could pick themselves the picture that they found more relatable and talk about it.

The pictures were chosen after careful consideration. To ensure a fair selection, a representative sample of participants was provided with several sets of pictures prior to the experiment, in order to filter out the pictures that did not yield the expected outcomes based on their response. More specifically, a picture of a smiling dog would normally elicit positive emotions. However, there are people who have lost their dogs and such picture may have an unintended effect.

Figure 2. Examples of the stimulus pictures. The first one elicits negative emotions, the second one elicits positive emotions, and the third one elicits neutral emotions.

Apparatus

The experiment was conducted in a lab room of Utrecht University. The setup (Figure 3) for the experiment included: two Tobii Pro Nano remote eye trackers (60 Hz), a measurement computer and two cameras connected to the computer. The computer had its own monitor. The looking direction of the participants was measured using the eye trackers. We also recorded a video image of the participants using the cameras placed behind them. These cameras serve as “scene cameras”, as they give us a view of what is happening in the scene in front of the participants. By transforming the eye tracker coordinates to pixel coordinates on the video frames, we were able to estimate where the participant looked at in the scene that was in front of them.

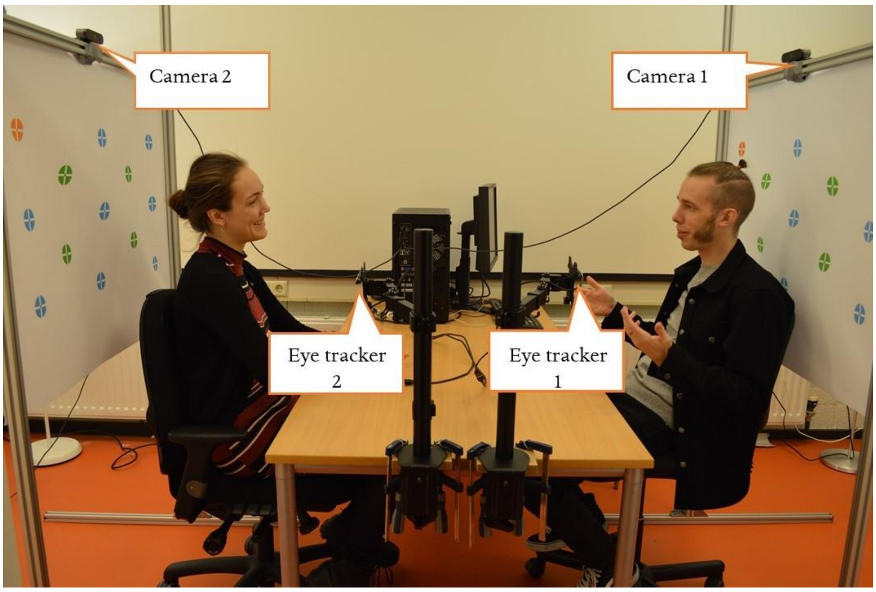


Figure 3. A picture of the set up (before Covid-19).

Procedure

Recruited participants were invited to the experiment room in couples. After they provided informed consent, they were asked to sit across each other, in front of the eye-trackers (respecting the COVID-19 regulations). The eye trackers needed to be correctly positioned so that they could record the gaze of the participant in front of them. Since the experimenter could not approach the participants (due to the COVID-19 regulations), she instructed them to adjust the angle and position of the eye trackers, so that the eye-trackers have a clear view of the participants head and eyes.

Before starting the experiment, calibration procedures took place. Calibration is done so that we can relate the gaze coordinates provided by the eye trackers to the scene the participants are viewing (the pixel coordinates on the video frames). To do so, the experimenter instructed the participants to look at certain points on the white poster behind the participant that is in front of them (see Figure 3). For this experiment, we used a 4-point calibration procedure.

After this was done, the researcher announced the start of the experiment. The separation of the three different conditions was done by giving the participants the pictures that elicit positive, negative and neutral emotions, one (set) at a time. The three sets of four images were placed in front of the participants in the same order. The participants were then instructed to look at the images of the first set and to choose an image they related to the most. After they made their choice they were asked to converse about the image as naturally as they would normally do, for 3-4 minutes. This was repeated for the two other sets as well. Such freedom of the task was allowed in order to make sure that the participants feel as comfortable as possible.

The order of the conditions was counterbalanced for every experiment. What this means is that the all the pairs were presented to different orders of the conditions. This was done to eliminate a possible carryover effect, that may transfer from one experimental condition to another. For instance, participants may perform differently in later conditions than the ones in the beginning of the experiment.

The amount of time that they spend interacting was equal for every topic (condition) and this whole period of interaction lasted approximately twelve minutes. We aimed for 4 minutes of conversation for every condition, thus, the experimenter timed the participants and told them when they could move on to the next set of pictures.

After finishing the task, the participants were asked to fill in a questionnaire regarding their experience during the experiment and another one regarding their demographic characteristics. Subsequently, the participants received their compensation as they were thanked for participation, and they were also asked if they have any questions or concerns regarding the experiment.

Results

1.1 Data quality

The eye tracking setups used to study human interaction vary mainly in relation to the nature of the eye tracking signal (head- or world-centered), and the extent to which they allow freedom of movement for the participants. Therefore, different types of eye tracking setups (head-free, head boxed or head-restricted setups) place restrictions on the research questions that can be answered about human interaction (see Valtakari et al., 2021). For our experiment, we chose to use a setup with an unrestrained head, commonly known as a head-boxed setup (see Figure 3). Head boxed setups allow for movement of the head, inside the area of the virtual head box (i.e., the area of allowed head movement). Having participants converse with each other face-to-face in an experiment, is maybe often more similar to a typical conversation that a person might have on a usual day rather than, e.g. putting screens between them. During this experiment the participants did not need to be in motion, but nevertheless we wanted to allow for a normal amount of head movement during seated conversations. We also wanted to have sufficient data quality to be able to distinguish between features of the face. Therefore, we decided that using a head-free setup was the optimal choice for the research questions. However, as this setup allows for more freedom of motion for the participants, this goes on the expense of data quality.

The (spatial) quality of the tracked gaze can be measured in terms of accuracy and precision. If higher accuracy is more important, using data from one eye might be better. But, if higher precision is more important, averaging the data from both eyes might be better (see Hooge et al., 2019). In our case, since we are dealing with large areas of interest (face features & other parts of the scene) we opted for better accuracy, hence we only used data from the left eye (as seen in figure 4). The reason we only used the left eye is because when we ran the transformation script for all participants, the calibration accuracy for the left eye was usually better than the right eye based on what we saw in the result figures. In addition, for one participant we only had data for the left eye.

A common problem with head-boxed setups without screens is the parallax error. In order for the eye tracker to correctly report where someone is gazing, that person will have to remain inside the eye tracker's calibrated plane. If people communicating with each other do not stay within the calibrated plane (e.g. by moving backward or forward), the recorded gaze position will deviate from the true gaze position, and this can result in a parallax error (Valtakari et al., 2021). As the plane we calibrated the data with is slightly behind the participant, (the white poster with dots in Figure 3) we knew that such error would exist. In order to address this problem of the offset (the difference between the true gaze position and the recorded gaze position), we manually inspected the gaze distribution for each participant and shifted the distribution to be centered at the face of the participant based on the assumption that the participants will look at the face of the other (see Figure 4).

Figure 4. An example: The before (first picture) & after (second picture) result photos of the manual adjustments of the gaze distributions. The orange dots represent the gaze data from the participants left eye.

1.2 Data reduction

What we determine as the start of a condition is when the participants start looking at the sets of pictures, since the time they spend trying to choose one, is very inconsistent. We extracted 3.5 minutes of data for each condition, to make it equal for each participant.

To make sense of the eye tracking data, we needed to first relate them to the video recordings. In other words, we wanted to find out where each participant looked at for each frame on their respective scene camera recording. To do this, we first used an optimization algorithm to linearly transform the eye tracking coordinates to pixel coordinates on the video frames based on the calibration procedure done preceding the experiments. Next, we calculated where participants looked at on each frame of the scene camera by determining the 1-2 samples (averaging if needed) of eye tracking data closest in time to each video frame.

We wanted to know how much participants looked at the eyes and mouth of the person they were interacting with, and whether there are any differences across the conditions. Therefore, we initially needed to construct the areas of interest (AOIs). In our case, AOIs can be defined as the link between eye movement measures with the parts of the scene (i.e., the time spent looking at specific parts of the other participants face).

To construct the AOI's, we first examined them with the facial behavior analysis toolkit OpenFace to automatically determine the location of the eyes and mouth in the video frames (Baltrušaitis, Robinson, & Morency, 2013; Baltrušaitis, Robinson, & Morency, 2016). For each of these three locations (both eyes and the mouth), we defined three AOI's by calculating circles using the locations as the center point and the distance between the center of the eyes and the mouth as the radius. We used this radius as it allowed us to define relatively large AOI's that have been shown to be more robust to noise (see Hessels et al., 2019). A non-AOI (Other AOI) was used for all frames in which gaze was available, but not on any of the two other AOIs; this includes the background and the upper body of the participant. Looking outside the scene was counted as "other" as well. Three circles were created, one for each AOI (two for the eyes and one for the mouth). The left and right eye AOI's were merged to one general "eyes" AOI. Gaze was defined as being on an AOI if the gaze point was within a circle. If the gaze point was within multiple circles, then it was attributed to the circle which had the closest center.

Subsequent video frames for which gaze was at the same AOI were merged into dwells. Dwells were defined as a minimum of seven consecutive video frames (120 ms) in which the gaze position was on the same AOI (similar to Hessels et al., 2019). Total dwell time to an AOI during a condition was calculated as a relative measure by dividing the sum duration of all dwells to that AOI during the condition by the length of the condition (3.5 minutes). The terminology and definitions for the total dwell time were obtained from Holmqvist et al. (2011). Finally, we calculated how much each pair looked at each other simultaneously throughout the conditions. If one participant looks at the face of the other while the other also looks back at their face, this is defined as mutual gaze. An episode of mutual gaze was also defined as a minimum of seven consecutive frames of the participants looking at each other's face.

1.3 Statistical analysis

The first hypothesis was that the participants will show bias towards fixating specific areas of the scene, depending on the emotional content of the conversations. To determine whether such bias for fixating the eyes, the mouth, or fixating irregularly exists, a repeated-measures ANOVA (multiple measures of the same variable taken on the subjects under different conditions) on total dwell times, with the factor AOI (Eye AOI, Mouth AOI and Other AOI) and the condition (Positive, Negative or Neutral), was used. The effect of emotional content of the conversations on gaze behavior, was non-significant $F(2,38)= 0.36$, $p=0.69$. There was also a non-significant interaction between the different emotional conditions and gaze behavior, $F(2,38)= 0.17$, $p=0.83$. However, the effect of the AOI was significant $F(1,19)=18.952$, $p=0.0003$. What this means is that the participants have shown bias towards looking at specific AOIs, but that was not related to the emotional content of the conversations. Relative dwell times (%) per AOI & per condition are given in Figure 5.

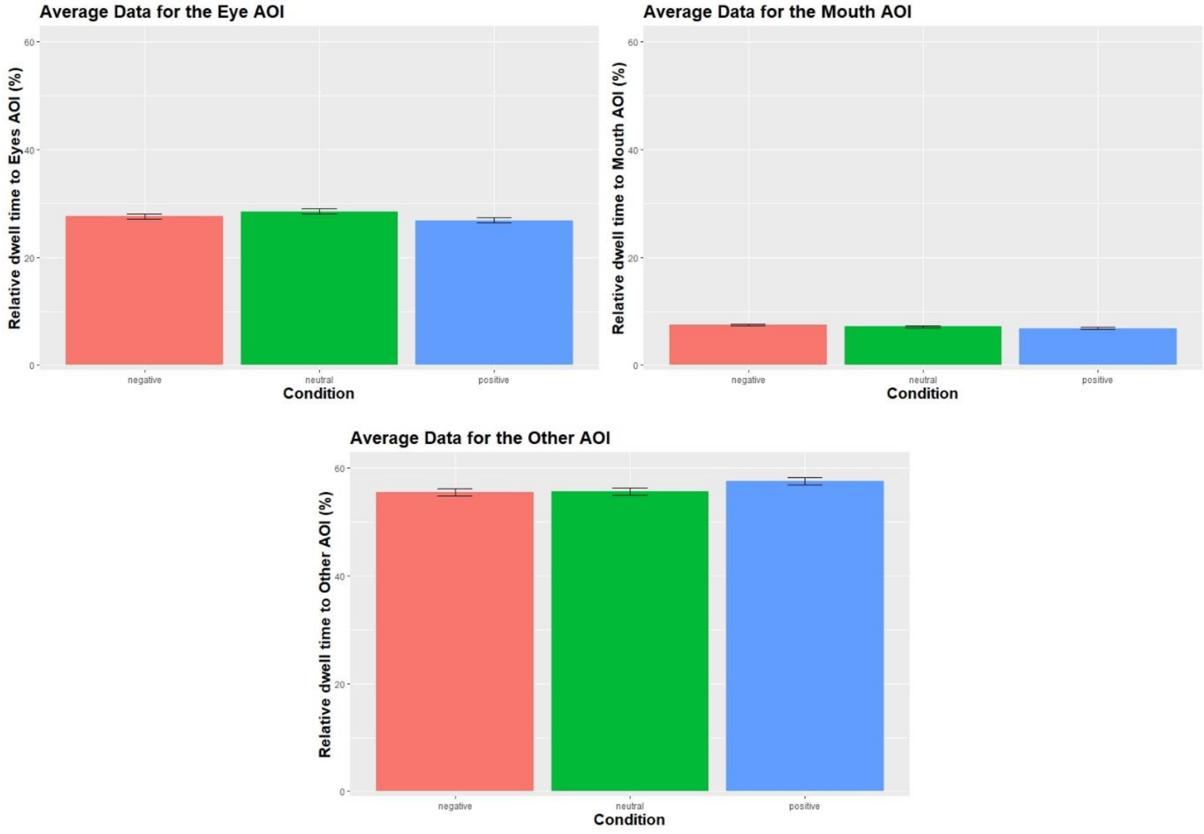
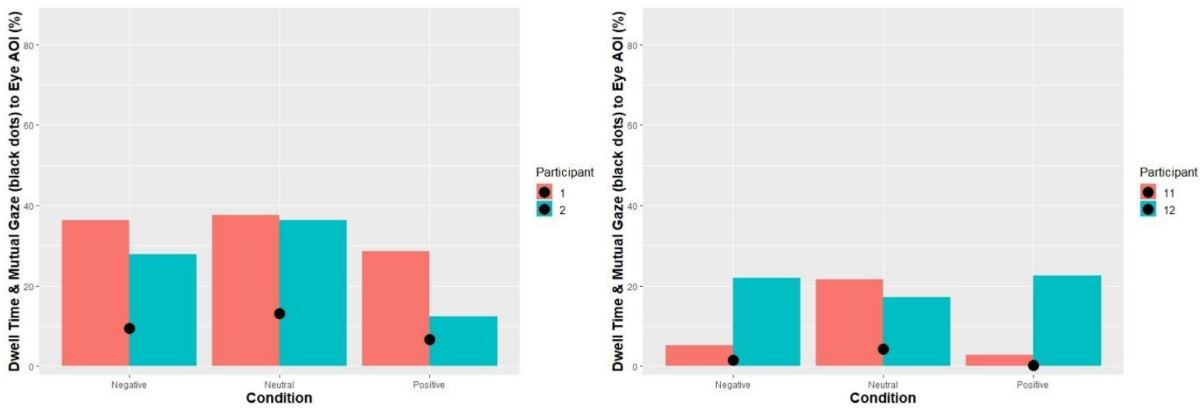


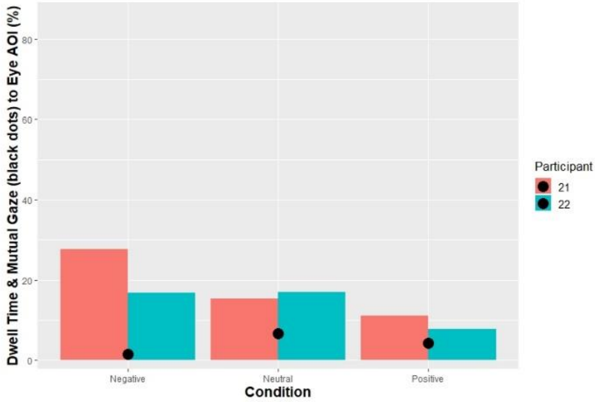
Figure 5. Relative dwell times per AOI (%) & per condition

The second hypothesis was that the participants will show a bias towards looking away during the conversations with negative emotional content, and a bias towards looking at the face of the other participant during the conversations with positive emotional content. To answer this, we first added the total dwell times of the Eye AOI and the Mouth AOI into a new variable called the Face AOI. Hence, we compared the total dwell times of the Other-AOI with those of the Face AOIs. To determine whether there is a bias for looking away, or looking at the face of the respondent, a repeated-measures ANOVA on total dwell times, with the factor AOI (Face AOI and Other AOI) and condition (Positive, Negative, and Neutral), was used. Once again, the effect of the emotional content of the conversations on gaze behavior, was non-significant $F(1,19)= 0.35$, $p=0.55$. There was also a non-significant interaction between the different emotional conditions and gaze behavior, $F(1,19)= 0.35$, $p=0.55$. However, the effect of the AOI was significant $F(1,19)=8.42$, $p=0.009$.

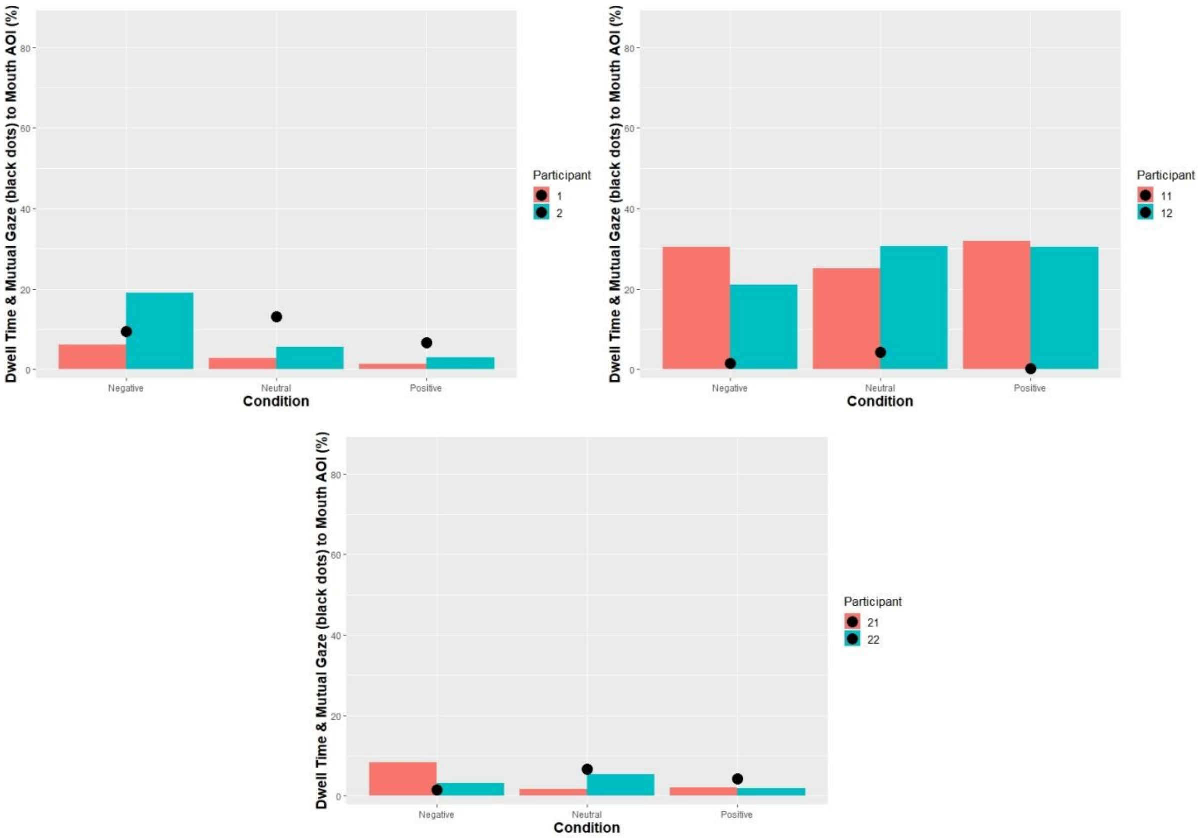
Lastly, regarding our last hypothesis that the different conditions (of the conversations) will have an effect on mutual gaze, we compared how much each pair looked at each other across the positive and negative conditions. To determine if there are significant differences between pairs, we conducted a repeated-measures ANOVA on total mutual dwell times with the factor condition (positive, negative, and neutral). The different conditions did not show a significant effect on mutual gaze $F(2, 18)=1.01$, $p=0.38$. However, when we plot the total dwell times and mutual gaze with the different condition for each pair separately, we see the pattern of mutual gaze being slightly higher during the neutral condition (this is the case for the majority of the pairs). Mutual gazes & total dwell times per AOI for the pairs where such pattern is more evident, are given in Figure 6. Average amount of mutual gaze per condition is given in Figure 7.

A)

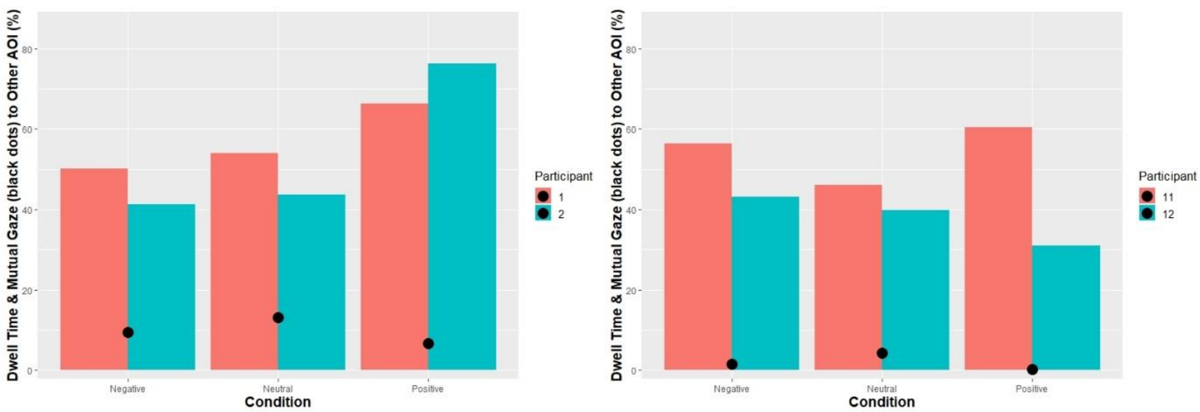



Dwell Time & Mutual Gaze (black dots) to Eye AOI (%)
0
20
40
60
80
Negative
Neutral
Positive
Condition
Participant
21
22

B)


Dwell Time & Mutual Gaze (black dots) to Mouth AOI (%)
0
20
40
60
80
Negative
Neutral
Positive
Condition
Participant
1
2
Dwell Time & Mutual Gaze (black dots) to Mouth AOI (%)
0
20
40
60
80
Negative
Neutral
Positive
Condition
Participant
11
12
Dwell Time & Mutual Gaze (black dots) to Mouth AOI (%)
0
20
40
60
80
Negative
Neutral
Positive
Condition
Participant
21
22

C)


Dwell Time & Mutual Gaze (black dots) to Other AOI (%)
0
20
40
60
80
Negative
Neutral
Positive
Condition
Participant
1
2
Dwell Time & Mutual Gaze (black dots) to Other AOI (%)
0
20
40
60
80
Negative
Neutral
Positive
Condition
Participant
11
12

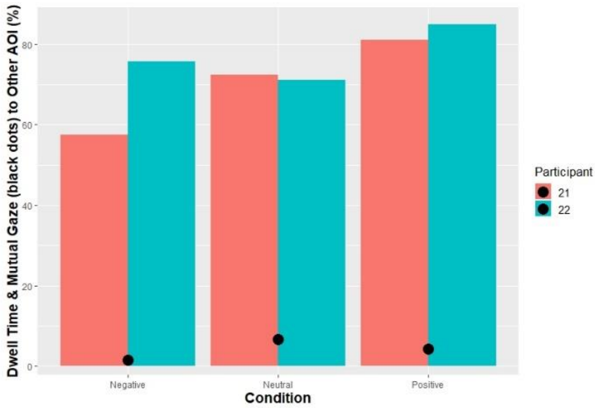


Figure 6. A) Dwell time & Mutual Gazes per Eye AOI, B) Dwell Time & Mutual Gazes per Mouth AOI, C) Dwell Time & Mutual Gazes per Other AOI (the black dots represent mutual gaze)

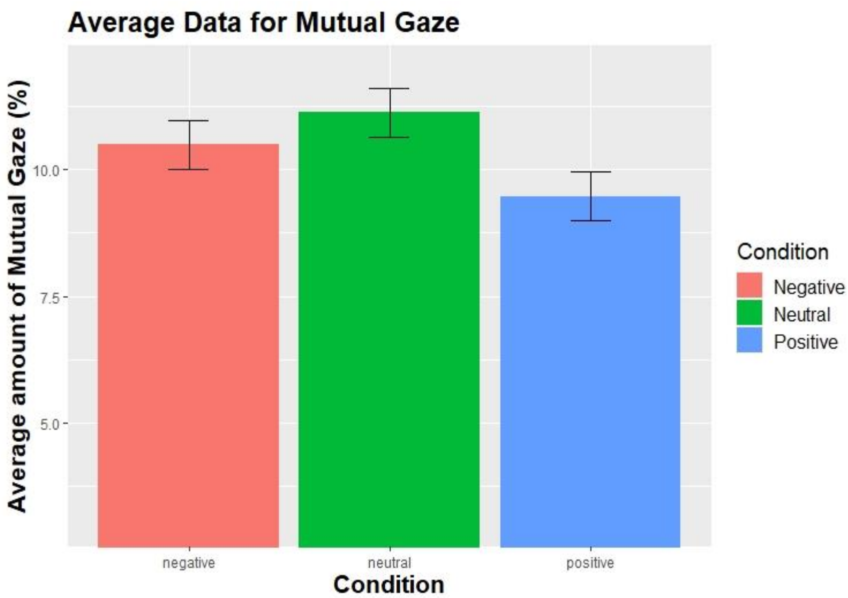


Figure 7. Average amount of mutual gaze per condition (%)

The data gathered from the questionnaires reported that the participants felt the emotions we aimed at. Deriving from a Likert scale, the participants reported to feel the positive emotions “very much” (61%) for the positive condition, the neutral emotions “quite a bit” (57%) for the neutral condition, and the negative emotions “quite a bit” & “moderately” (35%) for the negative condition (see Appendix). Three participants were excluded from this analysis, since their questionnaires were not completed.

Discussion

As discussed in the introduction, we have seen that human gaze is multifunctional. But the most important functions, can be narrowed down to two - we can both perceive information from others and use our gaze to signal to others (Argyle & Cook, 1976; Gobel et al., 2015; Risko et al., 2016). "The eye cannot take unless at the same time it gives," wrote Simmel. This emphasizes the role of the eyes, as a powerful tool for social interactions, with a "uniquely sociological function" (Simmel, 1921, p. 356-361).

As our primary goal was to uncover gaze patterns in conversations with emotional content we gave the participants pictures with different emotional content, and asked them to converse freely while their eye movements where being recorded using a dual eye tracking setup. Such approach was needed, since even though there are studies that investigate the link between gaze behavior and emotions, this is done with static stimuli. In these experimental environments, signals are sent only one-way (from the picture/video, to the participant) and the dual function of gaze is completely lost. Therefore, the interaction is lost as well. However, recent studies have shown consistency in eye movement patterns when looking at pictures or videos of faces that express different emotions. More specifically, Rogers et al. (2018) found that there is a bias for an eye-mouth gaze continuum, Schurgin et al. (2014) found that there is a bias for fixating the eyes when looking at sad faces and a bias for fixating the lips when looking at joyful faces. On another note, Hessels et al. (2014) found that there is a bias for fixating the eyes when the partner is physically present. This present study extends upon this work by finding that consistent eye gaze patterns, such as partners looking more at each other's eyes than mouth (see Figure 5) also occur during face-to-face conversation, but these patterns are not necessarily dependent on emotions.

As seen above in Figure 7, a pattern of gazing at the other participants eyes emerges slightly more during the neutral condition. This is in contrast with our expectations of not finding biases for fixating at specific AOIs during the neutral condition. On the other hand, this pattern makes sense, specifically because the participants did not know each other before the experiment. Talking about topics that have little to no emotional content might have made the interaction easier, thus allowing for more eye contact (Figure 7). This might also come by as a demand of a social norm, an extension of civil inattention. As mentioned before, civil inattention is "the process whereby strangers who are in close proximity demonstrate that they are aware of one another, without imposing on each other" (Finkelstein, 2007, p. 109). Thus, the participants might want to show that they are aware of the other but also give the latter a sense of privacy when he/she is experiencing a certain emotion (positive or negative be it). Eventually, it also seems expected that the pairs show more mutual gaze for conditions when there is more dwell time to the eyes of the other participant.

Regarding the other two conditions (positive and negative), we see that the participants did gaze to specific parts of the face throughout the conditions, just not in a matching amount compared between the pairs (see Figure 6). Indeed, such pattern arose also by examining where the participants looked at in the scene in front of them. Apparently, the participants seem to gaze at the face of the other participant when the latter is gazing away. This contradicts the idea that positive emotions are associated with the motivation to maintain eye contact and that negative emotions are associated with the motivation to avoid eye contact. However, we have to take into account that only one pair of the participants knew each other beforehand. The rest of them are just getting acquainted and asked to converse about emotional topics with someone who they don’t know.

An intriguing observation based on examining the video data, is also that the participants tended to smile even when talking about a traumatic event from their life or the loss of a loved one. Considering that these experiments took place after almost one year of restricted social contact due to Covid, it appears that the participants where very open when it came to discussing such matters, but yet they did keep a part of the shield up - their body language.

Although the participants showed they felt the negative emotions on the post-experimental questionnaire, the rate they reported them is smaller than the positive and neutral ones. This needs to be considered when examining the results, while keeping in mind that we do not know if the participants actually felt the negative emotions less, or they reported them less in order to “perform well” on the questionnaire. Our hypotheses were based on studies that employed static stimuli to evoke emotions. It is therefore expected that the results might differ in live conversations as the pictures and videos used expose the participants to an exaggerated display of emotions (i.e. sad faces or happy faces). This might not be the case in dynamic two-way interactions and it could rather be that when in interaction with someone, the expression of positive or negative emotions is more internalized when compared to e.g., expressions used in studies with static pictures. Thus, exploring different ways to make the participants feel the aimed emotions while allowing social interaction, can help in further understanding their gaze behavior.

The consequences that the individual differences in gaze patterns and in expressing emotions might have for feeling emotion and gaze behavior during conversation, create a large area for future inquiry. The way different situational factors such as mood, relationship status between individuals, and gender of people conversing influence deviations away from one's default pattern is an interesting area for future studies. For example, a research done by Coutrot et al., (2016) has reported gender differences with women distributing gaze between the eyes and mouth to a greater extent compared with men. Contextual influences are also important. More specifically, based on prior research there are expected differences in what represents the most typical gaze patterns across cultures (Akechi et al., 2013 & Senju et al., 2012). In our case half of our participants were Dutch. As they had to speak English during the experiment, they had to translate what they wanted to say. Hence, "when people are engaged in thought they tend to disengage from the external world by fixating on an empty part of the visual field, such as a blank wall, or out the window at the sky" (Salvi & Bowden, 2016, para. 1). Future studies examining the effect of emotions on gaze patterns during live conversations will require a bigger sample size than the one we had, to effectively explore the influence of situational and contextual influences.

References

Andersen, P. A., Guerrero, L. K., & Jones, S. M. (2006). *Nonverbal Behavior in Intimate Interactions and Intimate Relationships.* In V. Manusov & M. L. Patterson (Eds.), *The Sage handbook of nonverbal communication* (p. 259–277). Sage Publications, Inc. https://doi.org/10.4135/9781412976152.n14

Akechi H., Senju A., Uibo H., Kikuchi Y., Hasegawa T., Hietanen J.K. (2013). Attention to eye contact in the West and East: autonomic responses and evaluative ratings. *PloS One* 8, e59312 DOI: 10.1371/journal.pone.0059312

Argyle, M. (1972). Non-verbal communication in human social interaction. In R. Hinde (Ed.), Non-verbal communication. New York: Cambridge University Press. https://doi.org/10.1017/S0080443600011079

Argyle, M., & Cook, M. (1976). Gaze and mutual gaze. New York: Cambridge University Press. https://doi.org/10.1017/S0007125000073980

Argyle, M., & Dean, J. (1965). Eye-contact, distance, and affiliation. Sociometry, 28, 289-304. https://doi.org/10.2307/2786027

Ax A. F. (1953). The physiological differentiation between fear and anger in humans. Psychosom Med. 15, 433–442. 10.1097/00006842-195309000-00007 DOI: 10.1097/00006842-195309000-00007

Baltrušaitis, T., Robinson, P., & Morency, L. -P. (2013). Constrained local neural fields for robust facial landmark detection in the wild. In *2013 IEEE International Conference on Computer Vision Workshops (ICCVW)* (pp. 354–361). IEEE.

Baltrušaitis, T., Robinson, P., & Morency, L. -P. (2016). OpenFace: An open source facial behavior analysis toolkit. In *IEEE winter conference on applications of computer vision* (pp. 1–10).

Baron-Cohen S. (1995) The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In: Moore C, Dunham PJ, editors. *Joint attention: Its origins and role in development.* Hillsdale, NJ: Erlbaum; pp. 41–59

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults and adults with autism or Asperger syndrome. Visual Cognition, 4, 311–331. https://doi.org/10.1080/713756761

Bower G. (1983). Affect and cognition. *Philos. Trans. R. Soc. Lond.* 302B, 387–402. 10.1098/rstb.1983.0062. https://doi.org/10.1098/rstb.1983.0062

Coutrot, A., Binetti, N., Harrison, C., Mareschal, I. & Johnston, A. Face exploration dynamics differentiate men and women. *Journal of Vision* 16, 1–19 (2016).

Davidson, R.J., & Hugdahl, K. (Eds.). (1995). Brain asymmetry. Cambridge, MA: MIT Press

Duncan, S., Jr., & Fiske, D. (1977). Face-to-face interaction. Hillsdale, N.J.: Erlbaum.

Ekman P., Levenson R. W., Friesen W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science* 221, 1208–1210. 10.1126/science.6612338. DOI: 10.1126/science.6612338 https://doi.org/10.1126/science.6612338

Ekman, P., & Rosenberg, E. L., (Eds.). (2005). Series in affective science. What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS) *(2nd ed.)*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195179644.001.0001

Exline, R. V., & Fehr, B. J. (1978). Applications of semiosis to the study of visual interaction. In A. W. Siegman & S. Feldstein (Eds.), Nonverbal behavior and communication. New York: Wiley.

Exline, R. V., & Winters, L. C. (1965). Affective relations and mutual glances in dyads. In S. S. Tomkins & C. E. Izard (Eds.), Affect, cognition, and personality. New York: Springer.

Finkelstein J. (2007). *The Art of Self-Invention.* p. 109

Foulsham, T., Walker, E., and Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vis. Res.* 51, 1920–1931. DOI: 10.1016/j.visres.2011.07.002

Gobel, M. S., Kim, H. S., and Richardson, D. C. (2015). The dual function of social gaze. *Cognition* 136, 359–364. https://doi.org/10.1016/j.cognition.2014.11.040

Goffman E. (1963). *Behavior in Public Places.* New York, NY: Simon and Schuster.

Harmon-Jones, E., & Sigelman, J. (2001). State anger and prefrontal brain activity: Evidence that insult-related relative left-prefrontal activation is associated with experienced anger and aggression. Journal of Personality and Social Psychology, 80, 797– 803 DOI: 10.1037//0022-3514.80.5.797

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. Annual Review of Psychology, 50(1), 243–271. https://doi.org/10.1146/annurev.psych.50.1.243

Heslin R., Patterson M.L. (1982) The Role of Nonverbal Cues in Social Influence. In: Nonverbal Behavior and Social Psychology. Perspectives in Social Psychology (A Series of Texts and Monographs). Springer, Boston, MA. https://doi.org/10.1007/978-1-4684-4181-9_3

Hessels, R. S., Benjamins, J. S., Cornelissen, T. H. W., & Hooge, I. T. C. (2018). A validation of automatically-generated areas-of-interest in videos of a face for eye-tracking research. *Frontiers in Psychology, 9*(1367), 1–8.

Hessels, R. S., Cornelissen, T. H. W., Hooge, I. T. C., & Kemner, C. (2017). Gaze Behavior to Faces During Dyadic Interaction. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale. Advance online publication. http://dx.doi.org/10.1037/cep0000113

Hessels, R. S., Holleman, G. A., Kingstone, A., Hooge, I. T. C., & Kemner, C. (2019) Gaze allocation in face-to-face communication. https://doi.org/10.17605/OSF.IO/K32VZ

Hooge, I.T.C., Holleman, G.A., Haukes, N.C. *et al.* Gaze tracking accuracy in humans: One eye is sometimes better than two. *Behav Res* 51, 2712–2721 (2019). https://doi.org/10.3758/s13428-018-1135-3

Holmqvist, K., Nystrom, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* Oxford University Press.

Jokinen, K., Furukawa, H., Nishida, M., and Yamamoto, S. 2013. Gaze and turn-taking behavior in casual conversational interactions. ACM Trans. Interact. Intell. Syst. 3, 2, Article 12 (July 2013). http://dx.doi.org/10.1145/2499474.2499481

Kendon, A. (1967). Some functions of gaze-direction in social interaction. Acta Psychologica, 26, 22-63. https://doi.org/10.1016/0001-6918(67)90005-4

Knapp, M. L. (1980). Essentials of nonverbal communication. New York: Holt, Rinehard and Winston.

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods, 49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

Laidlaw, K. E. W., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 108*(14), 5548–5553. https://doi.org/10.1073/pnas.1017022108

Laidlaw, K., Foulsham, T., Kuhn, G., and Kingstone, A. (2011). Potential social interactions are important to social attention. *Proc. Natl. Acad. Sci.* 108, 5548–5553. https://doi.org/10.1073/pnas.1017022108

Lang P. J., Bradley M. M., Cuthbert B. N. (2005). International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual. Technical Report A-6. Gainesville, FL: University of Florida.

Lang, P., Ohman, A., & Vaitl, D. (1988). The international affective picture system Gainesville, FL: University of Florida, Centre for Research in Psychophysiology.

Lazarus, R. S., Speisman, J. C., Mordkoff, A. M., & Davison, L. A. A. (1962) Laboratory study of psychological stress produced by a motion picture film. Psychol. Monogr., 76(34, Whole No. 553).

Lim, J.Z.; Mountstephens, J.; Teo, J. (2020) Emotion Recognition Using Eye-Tracking: Taxonomy, Review and Current Challenges. Sensors, 20, 2384. DOI: 10.3390/s20082384

McHugo, G. J., Smith, C. A., & Lanzetta, J. T. (1982). The structure of self-reports of emotional responses to film segments. *Motivation and Emotion, 6*(4), 365–385. https://doi.org/10.1007/BF00998191

Mirenda, P.L., Donnellan, A.M. & Yoder, D.E. (1983) Gaze behavior: A new look at an old problem. *J Autism Dev Disord* 13**,** 397–409. https://doi.org/10.1007/BF01531588

Philippot, P. (1993). Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition and Emotion, 7*(2), 171–193. https://doi.org/10.1080/02699939308409183

Risko, E. F., Laidlaw, K. E. W., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: Toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience, 6*, Article 143. https://doi.org/10.3389/fnhum.2012.00143

Risko, E. F., Richardson, D. C., & Kingstone, A. (2016). Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. *Current Directions in Psychological Science, 25*(1), 70-74. https://doi.org/10.1177/0963721415617806

Rogers, S. L., Speelman, C. P., Guidetti, O., Longmuir, M. (2018). Using dual eye tracking to uncover personal gaze patterns during social interaction. Scientific Reports, 8, 4271. DOI: 10.1038/s41598-018-22726-7

Salvi C., Bowden E.M. (2016) Looking for Creativity: Where Do We Look When We Look for New Ideas? Front Psychol. 15;7:161. DOI: 10.3389/fpsyg.2016.00161

Schurgin M.W., Nelson J., Iida S., Ohira H., Chiao J.Y., Franconeri S.L. (2014). Eye movements during emotion recognition in faces. J Vis.;14(13):14. doi: 10.1167/14.13.14 PMID: 25406159. DOI: 10.1167/14.13.14

Schyns, P. G., Petro, L. S., & Smith, M. L. (2009). Transmission of facial expressions of emotion coevolved with their efficient decoding in the brain: Behavioral and brain evidence. PLoS One, 4(5), e5625. https://doi.org/10.1371/journal.pone.0005625

Senju A., Vernetti A., Kikuchi Y., Akechi H., Hasegawa T., Johnson M.H. (2012) Cultural background modulates how we look at other persons' gaze. International Journal of Behavioral Development 37, 131–136. DOI: 10.1177/0165025412465360

Simmel, G. (1921). “Sociology of the senses: visual interaction,” in *Introduction to the Science of Sociology*, eds R. E. Park and E. W. Burgess (Chicago, IL: University of Chicago Press), 356–361.

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. Psychological Science, 16(3), 184–189. https://doi.org/10.1111/j.0956-7976.2005.00801.x

Sutherland G., Newman B., Rachman S. (1982). Experimental investigations of the relationship between mood and intrusive unwanted cognitions. *Br. J. Med. Psychol.* https://doi.org/10.1111/j.2044-8341.1982.tb01491.x

Uhrig, M. K., Trautmann, N., Baumgärtner, U., Treede, R.-D., Henrich, F., Hiller, W., & Marschall, S. (2016). Emotion elicitation: A comparison of pictures and films. *Frontiers in Psychology, 7,* Article 180. https://doi.org/10.3389/fpsyg.2016.00180

Valtakari, N.V., Hooge, I.T.C., Viktorsson, C. *et al.* Eye tracking in human interaction: Possibilities and limitations. *Behav Res* (2021). https://doi.org/10.3758/s13428-020-01517-x

Velten E. (1968). A laboratory task for the induction of mood states. *Behav. Res. Ther.* 6, 473–482. 10.1016/0005-7967(68)90028-4. https://doi.org/10.1016/0005-7967(68)90028-4

Appendix

The table below illustrates the results of the post experimental questionnaire.

Condition	Emotion	Not at all (%)	Slightly (%)	Somewhat (%)	Moderately (%)	Quite a bit (%)	Very much (%)	An extreme amount (%)	Total
Positive	Positive	0%	0%	4%	0%	13%	61%	22%	100%
	Neutral	13%	26%	26%	30%	0%	4%	0%	100%
	Negative	87%	9%	0%	0%	0%	4%	0%	100%
Neutral	Positive	0%	17%	39%	35%	4%	4%	0%	100%
	Neutral	0%	0%	4%	9%	57%	30%	0%	100%
	Negative	74%	26%	0%	0%	0%	0%	0%	100%
Negative	Positive	22%	43%	22%	13%	0%	0%	0%	100%
	Neutral	9%	17%	26%	30%	13%	4%	0%	100%
	Negative	0%	0%	9%	35%	35%	22%	0%	100%