MacDonald, R and Tatler, B

 Gaze in a real-world social interaction: A dual eye-tracking study

http://researchonline.ljmu.ac.uk/id/eprint/12350/

Article

For more information please contact researchonline@ljmu.ac.uk

Gaze in a real-world social interaction: a dual eye-tracking study

Shortened Title: Eye movements in a social interaction

Macdonald, R. G.[1] & Tatler, B. W.[2]

[1]School of Health Sciences, University of Manchester, Manchester, M13 9PL, UK

[2]School of Psychology, University of Aberdeen, Aberdeen, AB24 2UB, UK

Corresponding author:

Ross Macdonald[1]

Tel: +44 7795 463688

email: ross.macdonald@manchester.ac.uk

Word Count: 7,300

Abstract

People communicate using verbal and non-verbal cues, including gaze cues. Gaze allocation can be influenced by social factors, however most research on gaze cueing has not considered these factors. The presence of social roles was manipulated in a natural, everyday collaborative task whilst eye movements were measured. In pairs, participants worked together to make a cake. Half of the pairs were given roles ("Chef" or "Gatherer") and the other half were not. Across all participants we found, contrary to the results of static image experiments, that participants spent very little time looking at each other, challenging the generalisability of the conclusions from lab-based paradigms. However, participants were more likely than not to look at their partner when receiving an instruction, highlighting the typical coordination of gaze cues and verbal communication in natural interactions. The mean duration of instances in which the partners looked at each other (partner gaze) was longer in the roles condition, and these participants were quicker to align their gaze with their partners (shared gaze). In addition, we found some indication that when hearing spoken instructions, listeners in the roles condition looked at the speaker more than listeners in the no roles condition. We conclude that social context can affect our gaze behaviour during a social interaction.

Keywords: *Social attention; Eye movements; Joint attention; Social interaction; Gaze cues*

Introduction

People have a strong tendency to follow gaze cues (Friesen & Kingstone, 1998; Ricciardelli, Bricolo, Aglioti, & Chelazzi, 2002) and to orient towards the eyes of others when viewing static social scenes (Birmingham, Bischof, & Kingstone, 2009). However, there is also evidence that the language and social context that accompany gaze cues can affect how people orient to and follow these cues (Knoeferle & Kreysa, 2012; Liuzza et al., 2011; Macdonald & Tatler, 2013, 2015). In the real world, gaze cues are rarely used without these accompanying factors, yet most previous research has not considered them together. By observing spontaneous gaze during a natural collaboration using eye-tracking, we aimed to 1) assess whether the findings from studies using static social scenes, namely the tendency to fixate on the eyes of others, are generalisable to real world interactions and 2) to investigate whether our social perceptions of a collaborator affect how we look at them and follow their gaze.

When investigating social attention it is important to make use of real-world interactions, as there is evidence that social presence influences our eye movement behaviour. Despite the evidence that we have a tendency to fixate the faces and eyes of others in static social scenes (Birmingham et al., 2009), we look at faces and eye much less when others are actually present (Laidlaw, Foulsham, Kuhn, & Kingstone, 2011) and although we know people orient attention in the direction of a gaze cue (Ricciardelli et al., 2002), in the real world people have been found to follow the gaze cues of another less when that other can observe them (Gallup, Chong, & Couzin, 2012). In both of these cases, the other person was a stranger who was not already involved in a natural interaction with the participant, therefore it is unclear what, if anything, these findings can tell us about

3

gaze behaviour during a natural social interaction. There is inevitably a lack of control in real world tasks relative to lab-based paradigms. However, given the evidence that conclusions from lab-based social attention experiments do not always generalise to the real-world (Risko, Richardson, & Kingstone, 2016), insights gained from studies using natural settings are not only helpful, but also arguably essential to properly investigate social attention (Kingstone, Smilek, & Eastwood, 2008). Specifically regarding our aims, using a real-world paradigm allowed us to investigate whether the tendency to look at faces in social scenes is generalisable to real social interactions.

The present study is not the first to observe how two interlocutors use their eyes in a natural interaction. Many researchers in the nineteen-sixties and -seventies (e.g., Argyle & Dean, 1965; Efran, 1968; Modigliani, 1971; Watson & Graves, 1966) investigated the factors affecting looking behaviour during interactions. Kendon (1967) stands out as particularly ambitious in his attempt to account for the different functions of gaze throughout an interaction. While much of Kendon's discussion is speculative, he provided behavioural evidence that people look to their interlocutor when they are completing an utterance and look away when beginning an utterance. This was interpreted as a way of signaling turn-taking during a conversation. A more recent study (Ho, Foulsham, & Kingstone, 2015) supported these much earlier findings, but used more precise and reliable measures in the form of modern eye tracking technology. In the current study we further explored how and when two interlocutors looked at each other during an interaction by investigating how these looks were affected by our manipulation of social perceptions.

The manipulation of social perceptions in this real-world study was simple: participant

pairs were either given social roles (a leader and a follower) or they were not. This type of manipulation has been used before in this research area. There is some evidence from early social interactions studies that social perceptions can influence the extent to which we look at another, such as work by Efran (1968) that showed college Freshmen looked more at Seniors than fellow Freshman. More recently, participants carrying out a Posner (1980) task in Italy were shown distracter gaze cueing stimuli made from the faces of Italian political figures, including Silvio Berlusconi (Liuzza et al., 2011). The gaze of Berlusconi was found to cause significantly more interference in the task for in-group than out-group participants. Crosby, Monin, and Richardson (2008) showed that participants were more likely to look at an individual on a monitor if they thought the individual could hear comments that were potentially offensive to that individual. These results show that social factors such as beliefs about another individual can affect how others look at them as well as how others look at external objects whilst communicating with them. Although these results show effects of prior beliefs about others on gaze behaviour, it is still unclear how beliefs about the role or knowledge of another affect the use of gaze cues in natural collaboration. In the present experiment we have made use of dual eye tracking in a real-world collaboration to explore this issue. More specifically, we altered the social role of our participants in order to investigate whether one's social perceptions of a collaborator affects the extent to which they are looked-at and the extent to which gaze following occurs.

In the present real world paradigm the onset of language and gaze cues were not controlled or directed. Instead, natural behaviour was observed and coarse measures were adopted to indicate gaze seeking and following behaviour. In line with a method used previously (Macdonald & Tatler, 2013, 2015), looks to the other person (anywhere on their body)

were used as an indicator of gaze seeking behaviour: the more one looked at their partner the more likely they were seeking information from the gaze of their partner. Additionally, this indicator was measured during instructional phrases, in order to investigate gaze-seeking behaviour when gaze cues were likely to be most helpful in the task. To investigate gaze following, we used shared gaze, which was defined as any instance in which the two participants looked at the same object. The initiator of each instance of shared gaze was also recorded as an indicator of who initiated each shared fixation. Gaze alignment has previously been shown to occur when participants work together on a task (Richardson, Dale, & Kirkham, 2007) and has been shown to increase performance in visual search tasks (Brennan, Hanna, Zelinsky, & Savietta, 2012). If this alignment of gaze can help participants in computer-based tasks, then it is interesting to see if this alignment spontaneously occurs in real-world collaborations and, if so, for how long it occurs. These measures are rough indicators of the phenomena that are under investigation, but are essential first steps for characterising the natural behaviour and for identifying the factors that need to be explored in more controlled experiments in the future.

The present study aims to observe collaborations between two people in a natural environment in order to 1) assess whether there is a tendency to fixate on the eyes (or indeed any part) of others during real world interactions and 2) to investigate whether our social perceptions of a collaborator affect how we look at them and follow their gaze. Participants, in pairs, were given a recipe to follow in order to make the batter for a cake. During this collaboration their eye movements were recorded using portable eye-trackers. This paradigm allowed us to observe natural eye movement behaviour during a natural collaboration featuring spoken language and social influences. When coding the data, the time participants spent looking at each other (partner gaze) or at the same object

6

simultaneously (shared gaze) were of particular interest as well as the time taken for participants to look at an item after the onset of the verbal reference to that item. Although the collaboration was devised to be as natural as possible, social context was manipulated in one way: half of the pairs were given roles (chef or gatherer) to fulfil and the other half were not. By manipulating this, interactions between gaze and social context could be investigated. Our hypotheses were that 1) participants would look less at each other than static-image studies would suggest and that, in line with Efram's (1968) findings, 2) the gatherers would look at their partner longer than other participants and would be less likely to initiate shared gaze, indicating a higher level of gaze seeking and following behaviour.

Method

Participants

Twenty-four undergraduate and post-graduate students from the University of Dundee participated in this study (16 women). They were split into twelve pairs to carry out the task. Six pairs were allocated to the roles condition and six were allocated to the no roles condition (see design). In the roles condition, three of the pairs were both women, one pair both men. Of the two mixed pairs, one had a female chef and the other a female gatherer. In the roles condition there were two all female pairs and four mixed pairs. Each individual was recruited separately, using the School of Psychology's participant recruitment system, and all participants were asked if they knew their partner. All but

one pair[1] (a mixed pair in the no roles condition) had never met their partner before. We were limited to 12 pairs due to the practical considerations. Each participant's video data was synced with their partners and coded manually frame-by-frame (see *Video coding and dependent variables* section for more details), which was a highly time-consuming process. Sample sizes are frequently limited in real world eye tracking studies. For a discussion of the reasons for and implications of this for the field please see Scrafton, Stainer, and Tatler (2017). Undergraduate students received course credits for participating. The study was approved by the local ethics committee and was in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

Materials

The experiment took place in a kitchen area on the University of Dundee campus. The kitchen was fully equipped with standard kitchen appliances, but only the oven and microwave were used. All items and foodstuffs that could be removed were removed before testing and the experimental materials were arranged carefully around the kitchen. This included the items and foodstuffs that were to be used for the procedure (a knife, a spoon, a set of kitchen scales, a large bowl, a small bowl, butter, sugar, whisk, eggs, sieve, flour, milk, and a square tin) as well as a selection of distractor items (a jar of instant coffee, a bottle of table sauce, a jar of cooking sauce, a bottle of olive oil, a

---

[1]These participants were separately recruited and we learned that they knew each other once they arrived for testing. Given the social setting of this study, it is possible that their knowing each other may have influenced their behaviour in the task. We found that their results fell within 1.5 standard deviations of the mean for the no roles condition in each of our dependent variables. We therefore included these data in our analyses

bottle of wine, a salt shaker, a box of tea bags, an insulated coffee cup, a box of muesli and an empty shopping bag). All of these items were placed in the same location for each pair of participants. A Recipe Procedure sheet was provided for each pair. This sheet explained, step-by-step, how to make the batter for a Victoria Sponge. There was also a Chef Guidelines sheet and a Gatherer Guidelines sheet for those in the roles condition. These sheets explained the responsibilities and duties for participants in the chef and gatherer roles. Eight toy building blocks (Megabloks) were used to calibrate the eye-trackers.

Design

This experiment used a between-subjects design. The two independent variables for the analysis of shared gaze and the time participants spent looking at each other (partner gaze) were the use of roles (roles or no roles) and the allocation of roles within the roles condition (chef or gatherer). For the analysis of the instructions, the independent variables were the use of roles (roles or no roles) and whether the participant was the speaker or listener of the instruction. It was important to use a between-subjects design for our roles condition, so that any influence of a participant's social perceptions of their partner in one condition did not effect their perceptions and behaviour in the following condition.

Procedure

This experiment required two participants to work together. The experimenter began by fitting a portable eye tracker to the first participant outside the kitchen. At this point

in the roles condition the first participant was given the Chef Guidelines and the second participant was given the Gatherer Guidelines. They were both instructed to read over their sheet and make sure they understood their roles. The Chef Guidelines informed the chef that they were in charge of preparing the recipe and that the gatherer was there to assist them. The sheet explained that the chef was expected to mix and prepare ingredients, following a recipe that they could not show to the gatherer. The chef would not be expected to collect any items or foodstuffs, but to delegate those duties to the gatherer. The chef would also be able to ask the gatherer to assist them with any aspect of the preparation they wished. The Gatherer Guidelines explained that the gatherer would not be expected to make any decisions concerning the preparation, but should instead do as instructed by the chef. Once the participants declared they understood their roles the gatherer was asked to remain outside while the experimenter and the chef entered the kitchen. The experimenter then gave the chef the Recipe Procedure sheet and told the chef where all of the necessary items and foodstuffs were located. The chef was then told they would have approximately three minutes to familiarise themselves with the kitchen and the locations of the items. This was in order to ensure that the chef knew where everything was and so could provide informative and directed instructions to the gatherer, and to ensure that the gatherer expected and received informative instructions and gaze cues. During these three minutes the experimenter fitted another portable eye-tracker to the gatherer. In the no roles condition the second eye-tracker was fitted straight after the first. At this point, in both conditions, the participants were brought into the kitchen and the eye-trackers were switched on.

[Figure 1 about here]

The participants were instructed to stand facing each other with their eyes closed. A flash of light was then emitted from a camera flash in between the two participants. This flash was picked up by all four cameras across the two eye-trackers and was later used as a reference point to synchronise all of the videos. The calibration procedure then took place (see *Eye movements and sound recording* section for more details). Once calibration was complete those in the no roles condition were directed to the Recipe Procedure sheet and informed that all of the items they would require were located around the kitchen. All participants were informed that the experimenter would be standing outside the kitchen, out of sight and that the participants must make no attempt to interact with him. The experimenter then told the participants that they should begin as soon as he was out of the room. The experimenter left and the procedure began. Once the participants had put the batter mixture in the oven (the final step) the experimenter entered and the earlier calibration procedure was repeated followed by another camera flash. The eye-trackers were then switched off and removed and the eye-movement and audio data were later analysed.

Eye movement and sound recording

Participants' eye movements were tracked using Positive Science LLC mobile eye trackers, which allowed free head movement. Each eye tracker has two cameras mounted on the frame of a pair of spectacles: one records the scene from the participant's point of view and the other records the right eye. Data from these cameras were captured on digital camcorders (Figure 1). For one of the eye-trackers these camcorders were stored, alongside a power supply for the eye-tracker, in a lumbar pack worn by the participant.

The camcorders connected to the second tracker were again stored alongside a power supply, but were stored in a light backpack worn by the participant. This eye tracker also has a small microphone attached to the frame. This microphone recorded sound throughout the experiment and was able to pick-up the voices of both participants. Gaze direction was estimated off-line using Yarbus software provided by Positive Science, LLC, which tracks the pupil and corneal reflection. This offline calibration method has the advantage of allowing for multiple attempts at gaze estimation, but the obvious disadvantage of not being able to repeat the calibration procedure. Gaze estimation was successful for all participants. Calibration was carried out in two stages, one looking down at a counter and the other looking across the room. These two stages were used because by tracking one eye, the vergence of the eyes that occurs as participants focus on objects at different distances cannot be directly measured, and may result in poor calibration over different depths. Instead the model was fitted to fixations on both proximal and distal points to minimise this issue. The first stage involved the participant fixating on Megabloks in a circular-array as the experimenter pointed to them. The second stage involved the participants, at a distance of approximately three metres, fixating points on the experimenter's body as he called them out. These included the tip of the experimenter's left index finger as it pointed up and to the left and the opposite for his right index finger, the experimenter's nose, the tips of the experimenter's left and right shoes and finally the top of the experimenter's thumb whilst it was positioned centrally over his abdomen. If the tracker estimates in the scene video fell on the correct blocks and correct body positions the calibration was deemed adequate. Eye movement data were recorded at 30Hz with a spatial accuracy of about 1 degree. Once videos for both participants were rendered with the eye movement information, Quicktime Pro was used

to synchronise both videos in to one movie file, ready for analysis.

Video coding and dependent variables

Eye tracking data were coded manually offline using Quicktime Media Player and audio information was extracted using Audacity sound editing software. The first three dependent variables investigated when participants looked at the same objects (shared gaze). These variables were (1) the proportion of task time in which shared gaze occurred, (2) the proportion of shared gaze events initiated by person A and (3) the time difference between the start of a shared gaze event and time the initiator first fixated the object. For these analyses, in each pair, one participant was labelled person A and the other was labelled person B. In the roles condition person A was the chef and B was the gatherer. Since there were not any defined roles in the no roles condition, participants in this condition were arbitrarily allocated as person A or B. The coding of these data was split between the lead experimenter and three undergraduate volunteers from the School of Psychology. All coders went through the videos frame-by-frame and marked down frame references for the beginning and end of each shared gaze event (any frame in which both participants fixated the same object). Additionally the initiator of each shared gaze event was noted, along with the frame references for the time the initiator began fixating on the shared attention location.

The video coders also noted the times the participants looked at each other (partner gaze) throughout the task. These data were used for the first two partner gaze measures: (4) the proportion of task time spent on partner gaze and (5) the mean duration of partner gaze events. The lead experimenter provided thorough written guidelines for coding for

the volunteers and provided them with the opportunity to ask questions at any time. To begin, all four coders coded the same movie file and these were all compared by the lead experimenter to ensure a consistent and high quality of coding. Here we found that all total times varied by less than 5%. We identified specific areas of disagreement between coders and reiterated and developed the coding process. When eye-tracking, there are very brief and frequent instances of tracking-loss, for example when a participants blinks, or simply when the tracker fails to identify the pupil. Fortunately the video outputs retain a synchronised video of the eyeball, meaning that we can ascertain whether the loss was due to a tracker error, blink or a sudden saccade away from the point of fixation. If, during a shared or partner gaze event, a blink occurred or if there was tracking loss with no change in eyeball position, we counted this time in our measure.

The remaining two partner gaze measures involved additional coding. Instances where one of the participants issued verbal instructions to the other were identified and analysed. These instructions were coded using Audacity sound editing software and the Quicktime movie files. For each pair, each instruction statement was numbered and transcribed, noting the speaker. The start time and end time (to the nearest frame reference) of each instruction was coded, along with the time that the speaker first looked (if at all) at the listener and the time the listener first looked at the speaker. In the roles condition, the speaker was always the chef and the listener always the gatherer. This was not directly instructed by the experimenter, but was a natural consequence of the roles that participants were given; the chef would ask the gatherer to collect items and the gatherer would listen to these requests. In the no roles condition the participant who gave the instruction was considered to be the speaker. Therefore the identity of the speaker and listener would switch throughout each movie in the no roles condition. Due

to the combination of the complexity of this process and the lower work-load relative to earlier coding (coding involved isolating instructions statements and investigating them, rather than comprehensive frame-by-frame notation), the instruction statements were coded entirely by the lead author. From coding these data, two dependent variables for the analysis of partner gaze were used: (6) the proportion of instructions in which one or more partner gaze event occurred and when these events occurred and (7) the proportion of partner gaze events that involved the listener looking at the speaker. Due to the rarity of partner gaze events throughout the trials we did not carry out any combined partner gaze/shared gaze analysis, such as the number of partner gaze events followed by shared gaze events across trials. In line with our previously used methods (Macdonald & Tatler, 2013, 2015) we used partner gaze and shared gaze as indirect indicators of gaze seeking and following respectively.

Statistical analysis

For some of the time measures, the distribution of data was skewed to lower values. In these cases the data were $\log_{10}$-transformed. An independent-samples t-test was used to compare the proportion of task time in which shared gaze occurred between the roles and no roles condition and a 2 (roles condition) x 2 (participant) ANOVA was carried out on the proportion of task time spent engaged in partner gaze. For all other analyses the lme4 package (Bates, Mächler, Bolker, & Walker, 2014) in the R statistical programming environment (R-Core-Team, 2015) was used to run linear mixed effects models (LMMs) or general linear mixed effects models (GLMMs). The details of the approach to GLMMS and LMMs used in this study are outlined in this section.

Random factors Typically, GLMMs and LMMs use subject and item as random factors, providing a benefit over AVOVAs in which subject and item analysis must be carried out separately. In the present study, there was no item equivalent and the task was carried out in pairs, so subject pair was the sole random factor. Following the guidance of Barr, Levy, Scheepers, and Tily (2013), random slope models were used, allowing the models to consider the different effects that the fixed factors have on different subject pairs. Models that are maximal were used where possible.

Simplifying models The maximal (most complex) models can often fail to converge if there are not enough observations in the data. In these cases, simpler models were attempted. Models were simplified, when necessary, by first removing correlations between random-slopes. After this, the slope of any interaction from the random factor was removed. If the model still failed to converge, fixed factor slopes were removed one-by-one. The most complex model that successfully converged was always used (Table 1).

[Table 1 about here]

Calculating p-values To calculate the p-value for an effect of any given fixed factor, a second model was produced, which was identical to the original LMM except for the omission of the fixed factor under investigation. This model was then compared to the original model using the *anova()* function in the car library (Fox & Weisberg, 2010). The p-value from these model comparisons was used as the p-value for the effect of the fixed factor.

Results

Task Duration

The experiment began once the experimenter left the kitchen and ended when the oven door closed, therefore duration varied depending on how quickly the pair could complete the task. Overall, mean task duration was 9 minutes 28 seconds. Those in the no roles condition took longer to finish the task (10 minutes 24 seconds, SD $=$ 1 minute 41 seconds) than those in the roles condition (8 minutes 31 seconds, SD $=$ 1 minute 32 seconds), but this difference did not reach significance, $t(10) = 2.070$, $p = .066$. Due to the variation in task duration, we have used proportion of task time throughout our analyses, rather than using absolute duration.

Shared gaze analysis

Shared gaze was defined in this experiment as any instance in which both participants were fixating the same *object* at the same time, regardless of the part of the object that was fixated. Shared gaze was regarded as an indicator of shared attention in the task. Figure 2a shows the mean proportion of task time in which shared gaze occurred total duration of shared gaze/ the total task duration) for subject pairs in the roles and no roles conditions. Subject pairs in the roles condition spent a higher mean proportion of time on shared attention (.26, 95% CI $=$ .07) than pairs in the no roles condition (.21, 95% CI $=$ .06), but this difference was not found to be significant, $t(10) = 1.286$, $p =$ .228.

[Figure 2 about here]

We found no significant difference in the proportion of shared gaze events that were initiated by person A (Figure 2b) between the no roles condition (.51, 95% CI = .04), where letter assignment was arbitrary, and the roles condition (.57, 95% CI = .04), where person A was always the chef, $\beta = .211$, $SE = .39$, $z = .557$, $p = .591$ (Model 1).

To measure the mean time taken for an instance of shared gaze to occur, the difference between the time the initiator looked at any object and the time that shared gaze on that object began was calculated. Figure 3 shows the mean values of this time difference for both person A and person B in the roles and no roles condition.

[Figure 3 about here]

Figure 3 shows that the mean time gap was smaller in the roles condition for both person A (833 ms, 95% CI = 127) and person B (713 ms, 95% CI = 120) than in the no roles condition (person A = 1,033 ms, 95% CI = 204, person B = 1,227 ms, 95% CI = 278). An LMM of the $\log_{10}$ time difference showed that the time differences were significantly smaller in the roles condition compared to the no roles condition, $\beta = -.097$, $SE = .049$, $t = -1.980$, $p = .048$ (Model 2). Figure 3 also shows there was a difference in mean time difference between person A and person B in the roles but not the no roles condition. However, the LMM showed that there was no significant effect of initiator, $\beta = .036$, $SE = .045$, $t = .800$, $p = .365$, nor any interaction, $\beta = .083$, $SE = .091$, $t = .920$, $p = .328$.

Partner gaze analysis

Any instance in which one participant looked at the other in the task is referred to as partner gaze. The proportion of task time in which partner gaze occurred is shown in Figure 4.

[Figure 4 about here]

The proportion of task time spent on partner gaze (total duration of partner gaze/ the total task duration) in the roles condition was numerically higher for person B (.037, 95% CI = .025) than person A (.018, 95% CI = .020), but in the no roles condition the proportion for person A (.023, 95% CI = .022) and person B (.021, 95% CI = .020) were similar. A 2 x 2 ANOVA found no significant main effect of roles condition, $F(1,20)$ = .402, p = .533, or person looking, $F(1,20)$ = .976, p = .335, nor any significant interaction, $F(1,20)$ = 1.451, p = .242.

We focused specifically on the duration of individual instances of partner gaze. The means of these durations across our conditions are shown in Figure 5.

[Figure 5 about here]

The mean duration of partner gaze was found to be higher in the roles condition for both person A (660 ms, 95% CI = 145) and person B (740 ms, 95% CI = 100) than in the no roles condition (person A = 560 ms, 95% CI = 101; person B = 540 ms, 95% CI = 94). An LMM of the $\log_{10}$ partner gaze time showed a significant effect of roles, β = .119, SE = .053, t = 2.240, p = .036 (Model 3). There was no significant effects of person, β = .026, SE = .075, t = .350, p = .696, nor was there a significant interaction, β = .032, SE = .150, t = .210, p = .863. These results show that not only were participants

19

in the roles condition (regardless of role) quicker to orient to an item fixated by their partner (Figure 3), but that their looks to their partner tended to last longer. Combined these findings suggest that participants in the roles condition may have been attending more to their partners than those in the no roles condition.

[Figure 6 about here]

Given our necessarily small sample size, we also present the mean partner gaze durations of person B for each of the 12 pairs (Figure 6). Although no significant interaction was found, we focused on person B because of our hypothesis that gatherers would look for longer at their partner than non-gatherers. We can see from Figure 6 that, with the exception of pair 1, those in the no roles condition had consistently shorter looks to their partner than the gatherers in the roles condition.

Partner gaze was also analysed specifically for times in which one participant was giving the other participant an instruction. Because these instances of partner gaze are coupled with verbal instructions, these instances are likely to be related to communication. Figure 7a shows the proportion of instructions in which at least one instance of partner gaze occurred for the roles and no roles condition. A GLMM showed that the proportion of instructions with partner gaze was not significantly different between the no roles (.32, 95% CI $=$ .17) and roles (.51, 95% CI $=$ 11) conditions, $\beta = $ -.808, $SE = $ .568, $z = $ 1.423, $p = $ .155 (Model 4).

[Figure 7 about here]

Figure 7b shows that the proportion of instructions with partner gaze in which the listener looked at the speaker was higher in the roles condition (.97, 95% CI $=$ .05) than the no

roles condition (.64, 95% CI $=$ .34). A GLMM showed the difference between roles to be significant, $\beta = 3.099$, $SE = 1.299$, $z = 2.387$, $p = .017$ (Model 5). This indicates that listeners in the roles condition (gatherer) looked more at their speaking partner (chef) compared to listeners in the no roles condition, suggesting that the subordinate role of gatherer increased looking-while-listening behaviour.

[Figure 8 about here]

As with the partner gaze duration data, we have included the data by pair for the proportion of instructions in which the listener looked at the speaker (Figure 8). Although we found a significant effect of role, Figure 8 indicates that this difference is likely driven by just two pairs in the roles condition, thus we must be cautious when interpreting these results.

Discussion

The present study 1) investigated whether the tendency to look at others in social scenes is generalisable to the real world and 2) explored the effects of manipulating the presence of social roles on gaze in a natural collaborative task. Overall, partner gaze was rare in this study, and the presence of roles in the collaboration influenced how participants engaged in shared gaze as well as partner gaze. The results of this study provide evidence that the presence of social roles affects the extent to which participants visually engage with a collaborator and their surroundings. The naturalistic setting we used was necessary for these findings, which can provide the foundation for later real-world experiments on social attention.

There was no significant difference between the proportion of time that partner gaze occurred across roles conditions. Participants spent far less time (between 2-4%) looking at each other than they spent on shared gaze (between 20-25%). These results appear to be at odds with the results of some previous lab-based studies. People have been shown to have a preference for looking at eyes when viewing pictures of people (Yarbus, 1967) or social scenes (Birmingham et al., 2009; Zwickel & Võ, 2010), however in this task participants spent very little time looking at their partners. Given the potential informativeness of the eyes (Tomasello, Hare, Lehmann, & Call, 2007) and the ease with which people can interpret gaze direction (Anderson, Risko, & Kingstone, 2011) this finding may seem surprising. However, studies using real people as stimuli may offer an explanation. Laidlaw et al. (2011) showed that people were less likely to look at a present confederate than the same confederate on a video monitor and Gallup et al. (2012) found that people were less likely to follow the gaze of strangers who could see them than strangers who could not. The authors of these studies concluded that this was due to there being potential consequences (social interaction) to looking at the present confederate or the on-coming stranger. The present study differed from these other real world studies in one important aspect: the participants were already involved in an interaction. The low proportion of time spent on partner gaze suggests that the social influences that dissuade people from looking at strangers also have an effect during an interaction. This tendency to avoid long looks to a collaborator can therefore not simply be explained by the wish to avoid a social interaction, as the interaction was already initiated. To speculate, it may be that participants avoided spending long looking at each other in order to follow social norms. Alternatively, it may simply be that in this task, there were more useful places to look than the collaborator. The specific

collaborative task used in a real world study will of course influence the eye-movement behaviour of participants, as different tasks place different demands on them. In the present task there were a number of objects that required attention and manipulation, which therefore required participants to avert attention from their partners. In a less demanding task it is possible that participants would spend longer looking at each other. Wu, Bischof, and Kingstone (2013) have shown that in an interaction in which looking at another is congruent with social norms (dining with another), looking time is increased relative to an interaction in which looking is incongruent with social norms (dining near, but not with, another). However, while we are not able to generalise about the specific amount of partner gaze during any given real-world interaction, our results do show that our participants successfully completed a social collaboration while very rarely looking at each other. This challenges the generalisability of findings from lab-based studies that appear to show a preference for fixating on the faces of others in social scenes (Birmingham et al., 2009; Zwickel & Võ, 2010).

While we argue that the social presence of the collaborator caused the low partner-gaze time, it is of course possible that it was the observation of the experimenter (via the eye trackers) that provided a type of social presence. In other words, the participants may have not wanted to be seen looking at their partner. There is evidence that being observed affects viewing behaviour; Risko and Kingstone (2011) found that participants were less likely to look at a provocative swimsuit calendar when they believed their eye movements were being monitored. We argue, however, that experimental observation is less likely to have an effect on the present study. Firstly, being observed looking at a person with whom one is interacting is intuitively much less undesirable than being observed looking at a lewd calendar, particularly when the task-at-hand involves collaborating with this person.

Secondly, Laidlaw et al. (2011) found that participants wearing eye-trackers fixated on a person presented on a monitor much more than a baseline object, suggesting that being observed (via an eye tracker) did not deter participants from looking at this person. Therefore, we consider it more likely that it was the presence of the collaborator that led to the low partner gaze time rather than the inferred observation of the experimenters.

The low partner gaze times present an obvious question; if people rarely look at each other in an interaction, can they still utilise gaze cues? Although the results cannot provide a definitive answer, there are three main arguments for the ability to utilise gaze cues in these circumstances. Firstly, it may be that any gaze cues were followed covertly. A previous study has shown that gaze cues can be followed and affect language comprehension, even when they are not directly fixated (Knoeferle & Kreysa, 2012). More recently, using a real-world paradigm Laidlaw, Rothwell, and Kingstone (2016) found that naive participants were more likely to look up when a confederate raised his hand to wave (social action) than when he raised his hand to put a phone to his ear (non-social action). This finding suggests that in the real-world people, to some extent, can use convert attention to detect social signals from others. This may well have occurred in the present experiment.

The second possibility is that successfully identifying a gaze cue may not need a large overall proportion of time spent on partner gaze, but may require effective individual partner gaze events. Therefore instead of spending a long time looking at the eyes of another, an effective way to follow gaze cues may be to look to the eyes briefly and infrequently, but at appropriate times. One way to investigate this possibility is to consider the features of a partner gaze event that may make the event more effective at

perceiving nonverbal signals, such as gaze cues. One feature that could have reasonably influenced effectiveness in this way is the length of individual partner gaze events. In this study, the mean length of partner gaze events was found to be significantly and consistently longer in the roles condition than the no roles condition. This may indicate that participants in the roles condition were more effectively attending to the non-verbal cues of their partners, resulting in the shorter gaps found between initiator fixation time and onset of shared attention in the roles condition.

Third, it may be the case that eyes are generally not sought out during a task, but are used when required, for example, during instructions. To consider this possibility, partner gaze was measured when the gaze allocation of a partner may have been particularly informative: during spoken instructions. Previous research has shown that looking behaviour varies during speech, with more looks directed towards an interlocutor when they begin speaking (Kendon, 1967; Ho et al., 2015). This may well be the result of seeking gaze information and may also result in an increase of gaze cue utilisation. Participants in the roles condition were overall no more likely to look at their partner than those in the no roles condition, but when there were instances of partner gaze during an instruction, listeners looked at the speaker significantly more in the roles condition compared to the no roles condition. This finding may indicate that our preference for looking at others can be affected by social context. However, although we found a significant effect in our model, analysis of the individual pairs showed that this difference between the roles conditions was not consistent.

As with partner gaze, the overall time participant pairs fixated the same object (shared gaze) was not found to vary significantly across roles condition. It was expected that more

shared gaze events would be initiated by person A in the roles condition because person A was the leader (chef) of the pair and initiating a shared gaze event was an indicator of providing a gaze cue that was followed. However this difference was not found to be significant. The mean time difference between the initiator of a shared gaze event fixating the gazed-at item and the initiation of the shared gaze event was significantly shorter for pairs in the roles condition compared to those in the no roles condition. The cause of this difference is unclear, but to speculate, one explanation is that the presence of distinct social roles increased the extent to which each participant attended to the other. Early research on non-verbal communication in real social interactions found that dominant interactors looked more at partners when talking and subordinate interactors looked more at their partners when listening, perhaps because the dominant interactor wanted to ensure they were being listened to and the subordinate wanted to signal that they were listening (Ellyson, Dovidio, & Fehr, 1981). It may be the case that those in the roles condition attended to their partners more and this led to them responding more quickly to the gaze location of their partners.

When interpreting the present data it is of course important to consider the statistical power of our analyses; we used 12 pairs of participants, which is a small sample relative to those typically used in laboratory-based eye-tracking experiments. These smaller sample sizes are standard in real-world eye-tracking studies, largely due to the time-consuming nature of frame-by-frame coding and annotation (Scrafton et al., 2017). Nevertheless, analyses on these smaller samples obviously lack the power of those on larger samples and limit the extent to which the findings are generalisable. We have gone some way to mitigate this problem by using GLMMs and making use of each observation, rather than comparing condition means comprised of varying numbers of observations, however

26

the issue of limited statistical power remains. As discussed in the Introduction, real world eye-tracking studies like the present experiment are invaluable to the investigation of social attention, as lab-based studies can strip away some of the myriad factors that influence our looking behaviour. While lacking generalisablity, the findings from our study can be used to inform the design of more controlled, less-naturalistic studies on social attention.

This study 1) investigated whether the tendency to look at others is generalisable to the real world and 2) explored the effect of social roles on gaze behaviour in a natural collaboration using dual portable eye-trackers. In this real social collaborative setting, people spent very little time looking at each other, challenging the generalisability of the conclusions from lab-based paradigms (Birmingham, Bischof, & Kingstone, 2007; Birmingham et al., 2009; Zwickel & Võ, 2010). Furthermore, this experiment provides evidence that those in the roles condition were more closely attending to where their partners were looking. Listeners were found to look more at a speaker providing verbal instructions if the speaker was playing the role of a chef, but this effect was not consistent across all participants. These results show that social context can affect gaze behaviour and provide a strong case for further investigation of gaze cueing behaviour in highly naturalistic environments in order to properly understand how social attention works in natural social contexts.

Acknowledgements

video coding and we also wish to thank two anonymous reviewers for their contribution to the paper.

Table captions

Table 1. Structure of the GLMM models used in the experiment

Figure captions

Figure 1: Two still images from the rendered eye tracker movies. The top still shows an example of *partner gaze*, with person B (left) looking at person A. The bottom still is from the same pair 19 frames (640ms) later and shows an example of *shared gaze*, with both participants looking at a plastic spoon

Figure 2: The (a) Proportion of task time in which shared gaze occurred and (b) the proportion of shared gaze events initiated by person A in the no roles and roles conditions. Standard error is shown in both graphs. Mean task time was 9 minutes 28 seconds.

Figure 3: The mean time gap between the beginning of the initiator's fixation to shared gaze location and the initiation of shared gaze. Results are shown (with standard error) for shared gaze events initiated by person A and person B in the no roles and roles conditions.

Figure 4: The proportion of task time in which partner gaze occurred for person A and person B in the no roles and roles conditions. Error bars show standard error. Mean task time was 9 minutes 28 seconds.

Figure 5: The mean duration of partner gaze events. Results are shown (with standard error) for partner gaze events in which person A looked at person B and person B looked at person A in the no roles and roles conditions.

Figure 6: The mean duration of partner gaze events for each of the 12 pairs of participants.

Figure 7: The a) Proportion of instructions in which partner gaze occurred and b) the

proportion of instructions with partner gaze in which the listener looked at speaker in the no roles and roles conditions. Standard error is shown in both graphs.

Figure 8: The proportion of instructions with partner gaze in which the listener looked at speaker for each of the 12 pairs.

Table 1. Structures of the GLMM and LMM models used in the experiment

Models
1) Shared Gaze Initiator ~ Roles + (1 | Subject Pair)
*2) Mean Gap Duration ~ Roles * Initiator + (1 + Initiator || Subject Pair)*
*3) Mean Gaze Duration ~ Roles * Person + (1 + Person || Subject Pair)*
4) Look to Partner ~ Roles + (1 | Subject Pair)
5) Listener Look ~ Roles + (1 | Subject Pair)

Figure 1: *Two still images from the rendered eye tracker movies. The top still shows an example of partner gaze, with person B (left) looking at person A. The bottom still is from the same pair 19 frames (640ms) later and shows an example of shared gaze, with both participants looking at a plastic spoon*

Figure 2: *The (a) Proportion of task time in which shared gaze occurred and (b) the proportion of shared gaze events initiated by person A in the no roles and roles conditions. Standard error is shown in both graphs. Mean task time was 9 minutes 28 seconds.*

Figure 3: *The mean time gap between the beginning of the initiator's fixation to shared gaze location and the initiation of shared gaze. Results are shown (with standard error) for shared gaze events initiated by person A and person B in the no roles and roles conditions.*
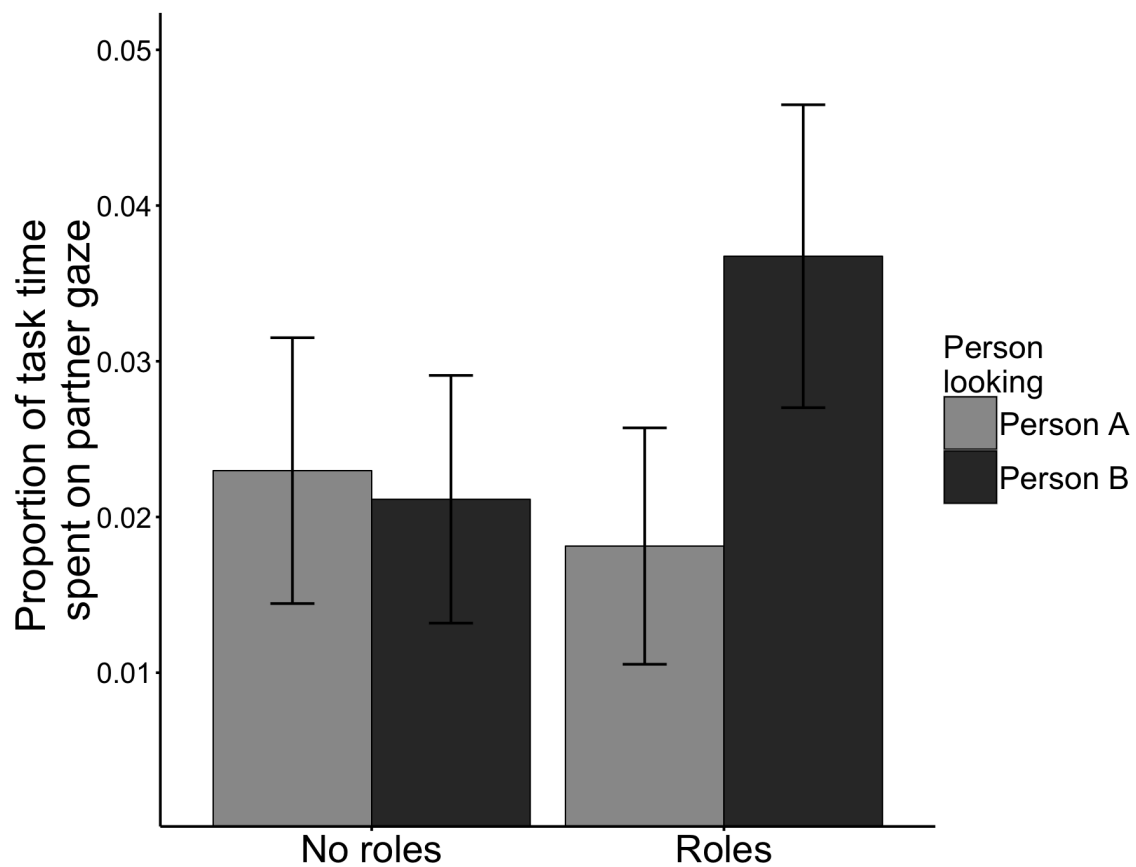
Figure 4: *The proportion of task time in which partner gaze occurred for person A and person B in the no roles and roles conditions. Error bars show standard error. Mean task time was 9 minutes 28 seconds.*
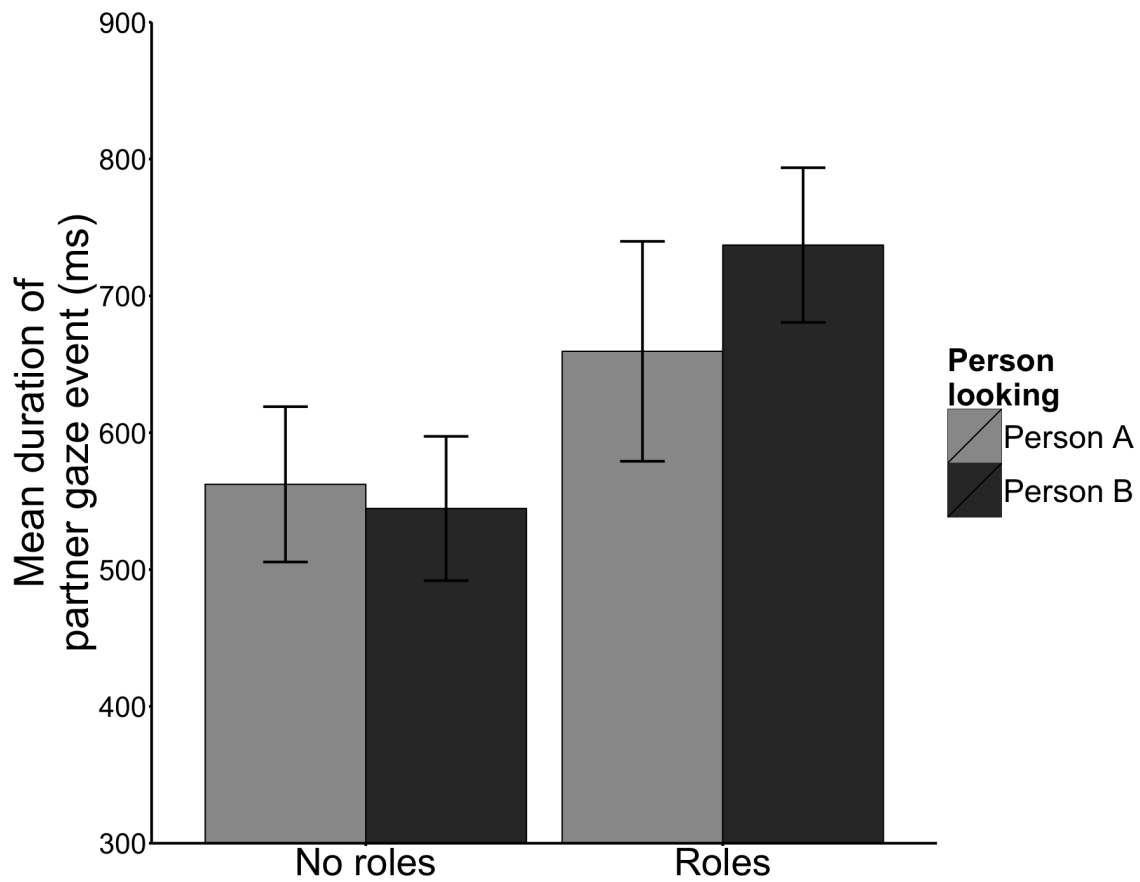
Figure 5: *The mean duration of partner gaze events. Results are shown (with standard error) for partner gaze events in which person A looked at person B and person B looked at person A in the no roles and roles conditions.*
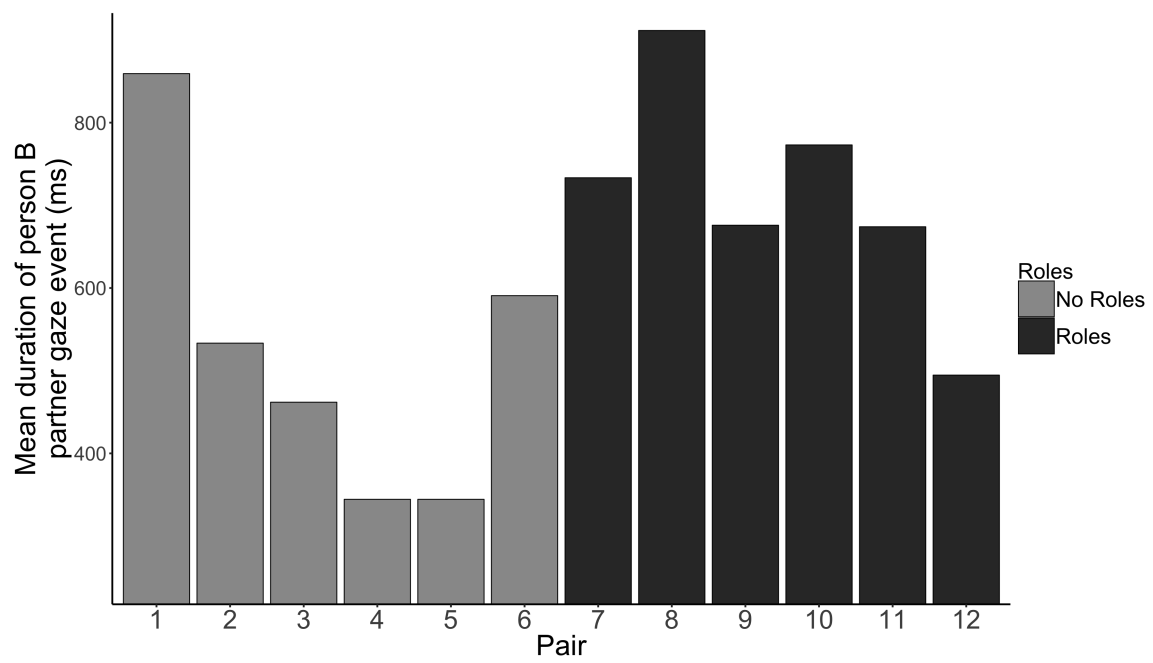
Figure 6: *The mean duration of partner gaze events for each of the 12 pairs of partici-pants.*
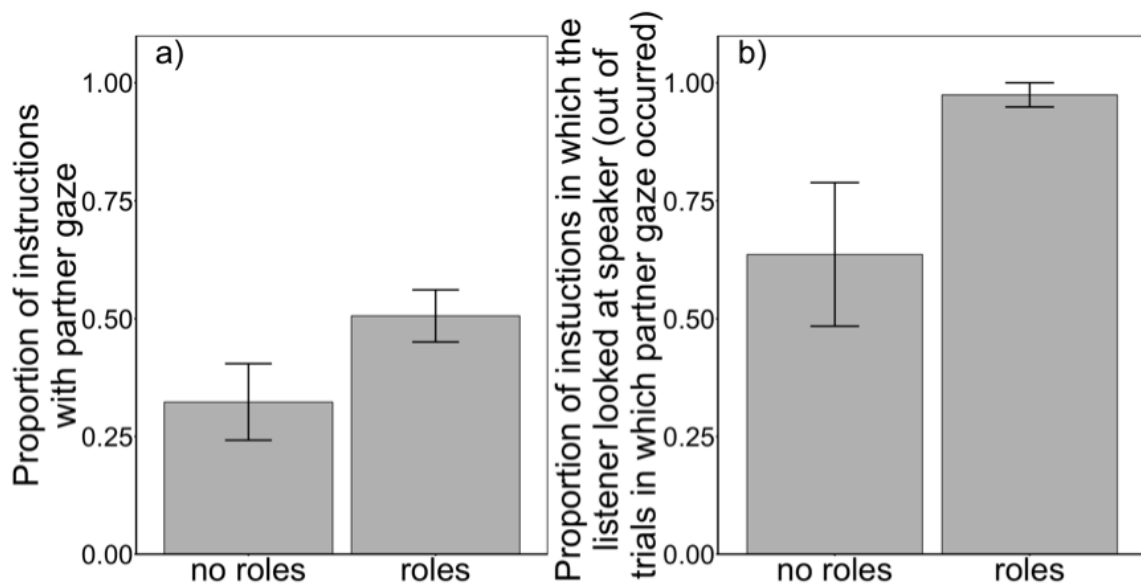
Figure 7: *The a) Proportion of instructions in which partner gaze occurred and b) the proportion of instructions with partner gaze in which the listener looked at speaker in the no roles and roles conditions. Standard error is shown in both graphs.*
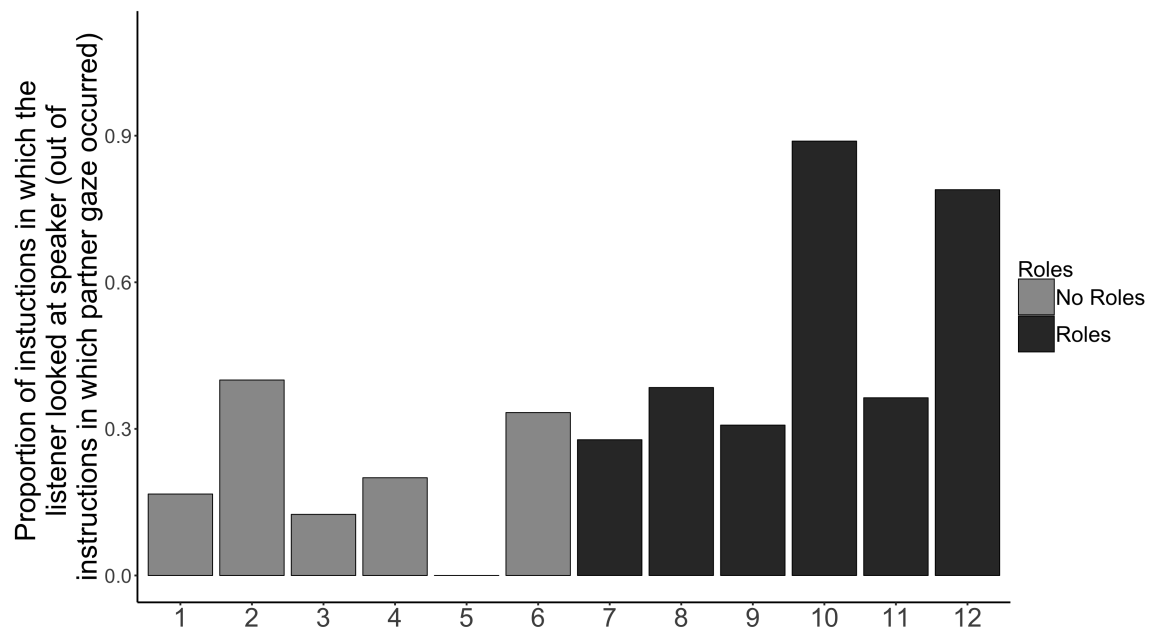
Figure 8: *The proportion of instructions with partner gaze in which the listener looked at speaker for each of the 12 pairs.*

References

Anderson, N. C., Risko, E. F., & Kingstone, A. (2011). Exploiting human sensitivity to gaze for tracking the eyes. *Behavior research methods*, *43*(3), 843–852.

Argyle, M., & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 289–304.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Birmingham, E., Bischof, W., & Kingstone, A. (2007). Why do we look at people?s eyes. *Journal of Eye Movement Research*, *1*(1), 1–6.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Get real! resolving the debate about equivalent social stimuli. *Visual Cognition*, *17*(6-7), 904–924.

Brennan, S. E., Hanna, J., Zelinsky, G., & Savietta, K. J. (2012). Eye gaze cues for coordination in collaborative tasks. In *Duet 2012 workshop: Dual eye tracking in csce. 2012 acm conference on computer supported cooperative work* (Vol. 9).

Crosby, J. R., Monin, B., & Richardson, D. (2008). Where do we look during potentially offensive behavior? *Psychological Science*, *19*(3), 226–228.

Efran, J. S. (1968). Looking for approval: Effects on visual behavior of approbation from persons differing in importance. *Journal of Personality and Social Psychology*, *10*(1), 21.

Ellyson, S. L., Dovidio, J. F., & Fehr, B. (1981). Visual behavior and dominance in

women and men. In *Gender and nonverbal behavior* (pp. 63–79). Springer.

Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition*, *117*(3), 319–331.

Fox, J., & Weisberg, S. (2010). *An r companion to applied regression*. Sage.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! reflexive orienting is triggered by nonpredictive gaze. *Psychonomic bulletin & review*, *5*(3), 490–495.

Gallup, A. C., Chong, A., & Couzin, I. D. (2012). The directional flow of visual information transfer between pedestrians. *Biology letters*, *8*(4), 520–522.

Ho, S., Foulsham, T., & Kingstone, A. (2015). Speaking and listening with the eyes: gaze signaling during dyadic interactions. *PloS one*, *10*(8), e0136905.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta psychologica*, *26*, 22–63.

Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive ethology: A new approach for studying human cognition. *British Journal of Psychology*, *99*(3), 317–340.

Knoeferle, H. K. P., & Kreysa, H. (2012). Can speaker gaze modulate syntactic structuring and thematic role assignment during spoken sentence comprehension? *Frontiers in Psychology*, *3*, 538.

Laidlaw, K. E., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. *Proceedings of the National Academy of Sciences*, *108*(14), 5548–5553.

Laidlaw, K. E., Rothwell, A., & Kingstone, A. (2016). Camouflaged attention: covert attention is critical to social communication in natural settings. *Evolution and*

Human Behavior.

Liuzza, M. T., Cazzato, V., Vecchione, M., Crostella, F., Caprara, G. V., & Aglioti, S. M. (2011). Follow my eyes: the gaze of politicians reflexively captures the gaze of ingroup voters. *PLoS One*, *6*(9), e25117.

Macdonald, R. G., & Tatler, B. W. (2013). Do as eye say: Gaze cueing and language in a real-world social interaction. *Journal of vision*, *13*(4), 6–6.

Macdonald, R. G., & Tatler, B. W. (2015). Referent expressions and gaze: Reference type influences real-world gaze cue utilization. *Journal of experimental psychology: human perception and performance*, *41*(2), 565.

Modigliani, A. (1971). Embarrassment, facework, and eye contact: Testing a theory of embarrassment. *Journal of Personality and social Psychology*, *17*(1), 15.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, *32*(1), 3–25.

R-Core-Team. (2015). R: A language and environment for statistical computing (r foundation for statistical computing, vienna, 2012). *URL: http://www.R-project.org*.

Ricciardelli, P., Bricolo, E., Aglioti, S. M., & Chelazzi, L. (2002). My eyes want to look where your eyes are looking: Exploring the tendency to imitate another individual's gaze. *Neuroreport*, *13*(17), 2259–2264.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination common ground and the coupling of eye movements during dialogue. *Psychological science*, *18*(5), 407–413.

Risko, E. F., & Kingstone, A. (2011). Eyes wide shut: implied social presence, eye tracking and attention. *Attention, Perception, & Psychophysics*, *73*(2), 291–296.

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.

Scrafton, S., Stainer, M. J., & Tatler, B. W. (2017). Coordinating vision and action in natural behaviour: Differences in spatiotemporal coupling in everyday tasks. *Canadian Journal of Experimental Psychology*.

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution*, *52*(3), 314–320.

Watson, O. M., & Graves, T. D. (1966). Quantitative research in proxemic behavior. *American Anthropologist*, *68*(4), 971–985.

Wu, D. W.-L., Bischof, W. F., & Kingstone, A. (2013). Looking while eating: The importance of social context to social attention. *Scientific reports*, *3*.

Yarbus, A. L. (1967). *Eye movements during perception of complex objects*. Springer.

Zwickel, J., & Võ, M. L.-H. (2010). How the presence of persons biases eye movements. *Psychonomic Bulletin & Review*, *17*(2), 257–262.