How the Eyes Tell Lies: Social Gaze During a Preference Task

Tom Foulsham, Maria Lock

Department of Psychology, University of Essex

Received 6 October 2013; received in revised form 23 June 2014; accepted 15 September 2014

Abstract

Social attention is thought to require detecting the eyes of others and following their gaze. To be effective, observers must also be able to infer the person’s thoughts and feelings about what he or she is looking at, but this has only rarely been investigated in laboratory studies. In this study, participants’ eye movements were recorded while they chose which of four patterns they preferred. New observers were subsequently able to reliably guess the preference response by watching a replay of the fixations. Moreover, when asked to mislead the person guessing, participants changed their looking behavior and guessing success was reduced. In a second experiment, naïve participants could also guess the preference of the original observers but were unable to identify trials which were lies. These results confirm that people can spontaneously use the gaze of others to infer their judgments, but also that these inferences are open to deception.

Keywords: Social attention; Eye movements; Deception; Theory of mind

1. Introduction

It is widely acknowledged that responding to the gaze of others is an important, and perhaps critical, component of social cognition (Emery, 2000). This view holds that humans are adept at detecting the eyes and following their direction, and that this behavior may underlie higher order social inferences and theory of mind (Baron-Cohen, 1995). In experimental studies of cognition, gaze following has largely been measured in individuals performing an attentional cueing paradigm (Friesen & Kingstone, 1998), although the usefulness of such paradigms has also been called into question (Risko, Laidlaw, Freeth, Foulsham, & Kingstone, 2012). This research describes a novel method for examining a crucial and understudied facet of social gaze processing, namely the ability to make
inferences based on gaze location. Importantly, this method allows us to measure the behavior of both parties in an interactive design which reveals how social gaze can serve a communicative function. We begin by placing this research in the context of experimental work in the cognitive sciences and recent trends in social cognitive neuroscience.

1.1. The components of social gaze

When considering the relationship between gaze processing and social behavior, a distinction can be drawn between three components of social attention.

First, there is considerable evidence that typically developing humans preferentially detect and monitor the eyes of other agents. When observers view single faces, complex scenes, or videos, the eyes are looked at early and often (Birmingham & Kingstone, 2009; Foulsham, Cheng, Tracy, Henrich, & Kingstone, 2010; Yarbus, 1967). Selecting the eyes does not seem to occur because they are high contrast or because they are in the center of the head (Birmingham & Kingstone, 2009; Levy, Foulsham, & Kingstone, 2013). Participants with autism, who show a range of impairments in social situations, display reduced eye contact in face-to-face situations and may also avoid looking at the eyes in pictures and video (Klin, Jones, Schultz, Volkmar, & Cohen, 2002). This evidence implies that the eyes are fixated because of the meaning that they convey to the observer.

Second, once the eyes have been selected, successfully attending in a social context requires gaze following. Observers must use the information from the eyes (and also the head and body) to calculate what is being attended. This often results in the observer attending to the same location (i.e., joint attention). Milgram, Bickman, and Berkowitz (1969) observed that pedestrians would follow the gaze of a confederate or group of confederates who stood on a busy street and looked upward (see Gallup et al., 2012; for a recent extension). Humans possess a high-contrast eye that may have evolved to be particularly easy to follow (Kobayashi & Koshima, 1997). Gaze following has been widely studied using a cueing paradigm, which demonstrates that participants automatically shift their attention in response to eyes pointing in a particular direction (Friesen & Kingstone, 1998; Frischen, Bayliss, & Tipper, 2007).

The majority of previous research investigating social attention has focused on these two requirements. However, a third component of attending in a social context involves inferring an agent’s beliefs and feelings about what he or she is looking at. An integral part of our understanding of other people’s mental states—”theory of mind”—concerns perspective taking (Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010). In the context of social attention, observers may take the perspective of a gazing interactant to answer the question of why they are looking at a particular location. Relatively few studies have examined the attributions made after the target of gaze has been established. Children of 4–5 years old can infer which item an actor likes based on differences in gaze (Einav & Hood, 2006). The cognitive and neural mechanisms which support these inferences are likely present in the first year of life (Grossmann & Johnson, 2007; Woodward, 2003). In adults, it has been demonstrated that objects which are looked at are better liked than those that are not looked at (Bayliss, Paul, Cannon, & Tipper, 2006).
In nonhuman primates, the target of gaze is used to infer social hierarchy: Individuals who are looked at frequently are perceived as high in social status (Chance, 1967). Children also use gaze as a cue regarding who is best to learn from in the environment (Chudek, Heller, Birch, & Henrich, 2012).

1.2. Gaze as part of a two-way social interaction

In the laboratory studies of gaze following discussed so far, the behavior of a participant is typically measured as they respond to an image or video representation of a person’s eyes. Of course, an image may not capture the meaning of a real social situation because the people depicted cannot look back (Risko et al., 2012). In real conversation, for example, gaze provides not just a means of selecting pertinent information but also a communicative signal (e.g., regarding whose turn it is to speak next; Argyle & Cook, 1976). Interpreting this signal may play a particular role in grounding ambiguous verbal utterances (e.g., Hanna & Brennan, 2007; Knoeferle & Crocker, 2006).

Several other recent studies confirm that the presence of a real person changes eye movement behavior, and that real interactions may be sensitive to different factors. For example, in Gallup et al.’s (2012) study of gaze following in natural conditions, observers were more likely to follow gaze if they were following from behind, where they could not be seen. The tendency to look at faces and eyes is dramatically reduced when in a room with a real person, compared to an image of the same person (Laidlaw, Foulsham, Kuhn, & Kingstone, 2011). When walking or conversing, other people are attended to differently when they can see you (Foulsham, Walker, & Kingstone, 2011; Freeth, Foulsham, & Kingstone, 2013). In social cognitive neuroscience it has also been acknowledged that studying the mental and neural processes involved in two-way interactions is crucial. Studies of “mentalizing,” which ask people to use their theory of mind, typically measure cognitive processes “offline”: while people observe others and think about interacting with them (Schilbach et al., 2006). In contrast, studying “online” cognition, where participants are engaged in actual interactions with each other, can reveal different results in terms of brain and behavior (Wilms et al., 2010).

1.3. Gaze transfer in applied settings

The research discussed so far indicates that it is important to study both the way in which people interpret gaze and the way in which gaze may change according to interactions with another person.

One context in which this has previously been studied concerns how people can use explicit representations of others’ gaze. For example, Litchfield, Ball, Donovan, Manning, and Crawford (2010) presented novice and expert radiographers with the dynamic gaze pattern made by someone searching for a nodule in a chest X-ray. The new observers, particularly the novices, found the search targets more quickly when viewing another person’s eye movements. Viewing another person’s gaze can also guide problem solving, by triggering shifts of attention consistent with the solution (Grant & Spivey, 2003;
Litchfield & Ball, 2011), and change biases in perceptual judgments (Wu, Shimojo, Wang, & Camerer, 2012). In other settings, it might help performance of a shared task if one could view the gaze of an interacting partner. For example, in Brennan, Chen, Dickinson, Neider, and Zelinsky (2008), participants collaborating in a search task were sometimes shown their partners’ gaze position. These participants strategically changed their looking patterns, dividing their labor, and allowing them to perform better than solitary searchers.

These studies presented the gaze locations of others as a model for learning, and they demonstrate that this can be a useful tool for improving performance. However, as discussed recently by Müller, Helmert, Pannasch, and Velichkovsky (2013), the process of interpreting a partner’s gaze can lead to costs as well as benefits in collaborative tasks. During a puzzle task where one partner knew the solution, displaying the gaze of this person induced uncertainty, leading to response delays and increased verbal effort relative to a condition which displayed the partner’s mouse cursor. Participants in such tasks must interpret potentially ambiguous gaze signals, and this interpretation presumably requires an appreciation of why the partner is looking in this location—a theory of mind. However, although “gaze transfer” has been exploited in numerous studies, these studies are typically tied to specific applications and do not investigate the interpretation of gaze in a two-way interaction in detail.

1.4. This study

This study describes a simple but novel paradigm designed to investigate both how people interpret gaze signals and how these signals change through social interaction. Specifically we aimed to (a) confirm that people can spontaneouly use the target of gaze to make inferences about the thoughts and feelings of another person and (b) examine how attention can be flexibly deployed to mislead these inferences. Participants were eye-tracked while they chose which of a set of stimuli they liked the most. While we could have used any task, a preference task was particularly suitable because of reported relationships between looking and liking (e.g., Shimojo, Simion, Shimojo, & Scheier, 2003) and because it provided a range of idiosyncratic choices. In the second part, observers were shown a replay of the eye movements of the previous participant and asked to guess how that participant had responded during the preference task. Thus, they were provided with information about what was being looked at and had to make an inference about that person’s response.

Observers subsequently repeated the preference task but with new instructions to try to hide their decision from the guesser. The case of deception is an important one because it requires an appreciation of another person’s (false) beliefs. There has been extensive interest in determining the degree to which infants and nonhuman primates practice deception (Chandler, Fritz, & Hala, 1989; Whiten & Byrne, 1988). Moreover, folk wisdom suggests that the eyes are particularly revealing during deception, causing, for example, poker players to wear sunglasses in the hope of masking their intentions. Indeed, the ability to hide one’s gaze direction may have exerted evolutionary pressure on the development of the primate eye, causing most nonhuman primates to have a darkly
colored sclera (Kobayashi & Koshima, 1997). The human white sclera sacrifices this camouflage in favor of communication. In this study, we did not require participants to follow gaze from the eye, but presented a gaze cursor (as used in studies of gaze transfer). Can we instead camouflage our preferences by changing our gaze behavior? By comparing how well people are able to guess the preference response from eye movements in the truthful block and from the deceptive block, we aimed to determine whether naive observers can intuit a response based on fixations and whether attention can be flexibly deployed to mislead.

2. Experiment 1

2.1. Method

2.1.1. Participants
Twenty volunteers from the University of Essex community took part in the experiment (15 male; mean age = 27.8 years). Due to the nature of the design, which required participants to react to another person’s behavior, an additional pilot participant was tested to provide preference responses for the first experimental participant. Their fixation data were not included in the analysis.

2.1.2. Stimuli and apparatus
A set of computer-generated fractals was chosen as stimuli from freely available collections (http://www.nahee.com/spanky/index.html). Fractals are mathematically derived patterns with no explicit meaning, and we therefore expected them to elicit a range of idiosyncratic preferences. One hundred and forty-four fractals were used in total, and these were randomly assigned to groups of four (see Fig. 1). In each trial, four fractals were displayed in a $2 \times 2$ arrangement on a white background.

The stimuli were presented on a 19-inch color monitor subtending approximately 31˚ by 24˚ at a fixed viewing distance of 60 cm. Stimuli and eye tracking were controlled by the

![Fig. 1. The experimental procedure, with example stimuli. Participants chose one of four patterns. In Block 1, they chose the one they preferred. In Block 3, the preference task was repeated with the instruction to mislead the next guesser. In the intervening block they guessed the pattern the previous participant preferred after watching an animated replay of their fixations via a circular cursor which moved around the display.](image-url)
EyeLink system (SR Research Ltd., Mississauga, Canada) via Experiment Builder software. Eye movements were recorded using a desktop-mounted EyeLink 1000 which recorded eye position from the pupil and corneal reflection at 1,000 Hz. Head movements were restricted by a chin rest. Responses were made by pressing one of four buttons on a gamepad.

### 2.1.3. Procedure and design

Following a 9-point eye-tracker calibration, the experiment consisted of three blocks (Fig. 1). In each trial of the first and third block participants performed a preference task requiring them to choose which of the four fractals they preferred by pressing one of four buttons. In the intervening block, participants performed the “Guess” task, which required them to guess which item the previous observer had chosen.

Each preference trial began with a central fixation point and after the onset of the fractals participants could take as long as they liked to make their choice. Eye movements were recorded throughout and the location and duration of each fixation were written to disk. Each preference block had 18 trials, selected at random without replacement. In the first block participants were told only to choose which pattern they liked the most. No other instructions were given and, because at this point in the study participants were not aware of the other parts of the task, we consider responses in the block to represent natural behavior. Thus, we refer to this as the “Truth” block.

In the final, “Lie” block, participants repeated the preference task with the additional instruction to deceive the next participant by “hiding your decision or misleading the guesser.” It was stressed that participants should still respond accurately, indicating which item they truly preferred.

The eye movements recorded from each Truth and Lie trial formed the basis of the Guess block for the next participant. In each Guess trial observers were shown a replay of any eye movements before being asked to “guess which image that participant had chosen.” Four fractals were displayed in the same way as before and an animated cursor moved over the images representing the location and duration of the fixations made by the previous participant when viewing those images. The cursor was a red, filled circle with a diameter subtending approximately 1°. Participants were given the option of replaying the gaze animation as many times as they liked before making their judgment. The Guess block continued for 36 trials, divided equally into three conditions, which were shown in a random, interleaved order. In the Truth and Lie conditions eye movements from the corresponding block of preference trials were replayed. In the third, “Control,” condition, no eye movements were shown. In Control trials, the fractals were presented with the instruction that participants should make their best guess as to how the participant responded. Therefore, this provided a baseline measure of participants’ ability to guess with no extra information. Across all participants each particular set of stimuli was equally likely to appear in all conditions.

### 2.2. Results

The responses made during the Guess block were compared with the actual choice made by the previous participant. Thus, we asked whether guessing accuracy was greater
than chance and whether it varied between conditions. We then examined the eye movements made in each preference block to quantify how overt attention was moderated by the instruction to mislead another observer. In each case, participant means were compared using repeated-measures ANOVA.

2.2.1. Guessing performance

Guessing participants took an average of 6.13 s ($SD = 0.37$ s) per trial to make their response. This time includes the time spent playing and replaying any eye movements and deciding on a guess. Although they had the option of replaying the animation, they tended to do this only rarely (on only 6.7% of trials, on average; $SD = 3.3$). Fig. 2 summarizes the proportion of correct responses during the guess block. As there were four possible responses, a completely random guessing strategy would give 25% correct.

Condition had a reliable effect on accuracy, $F(2,38) = 16.1$, $p < .001$, $\eta_p^2 = .46$. Guesses to trials with eye movements from the Truth block achieved a mean accuracy of over 60%, much higher than chance and significantly different from the other conditions (both $t > 5$, $p < .001$, $d > 1.2$). Control trials (shown without eye movements) were guessed less accurately and Lie trials elicited the poorest performance (these conditions were not significantly different, $p = .16$). Only on Lie trials were observers unable to guess the previous participant’s response at a level reliably greater than chance.

It is worth noting that guessing performance was greater than chance in the Control block. This indicates that participants had intuitions about which fractals were likely
to be preferred, and that these intuitions were sometimes correct. For example, they may have chosen the more colorful image, or the one that they themselves would have picked, and thus this level of accuracy is not surprising. If we consider performance in Control trials as the true baseline, it is interesting that the Lie block resulted in performance below chance. Thus, participants were able to mislead with their gaze.

2.2.2. Eye movements

Participants took several seconds to make their choice during the preference block, and this did not differ reliably between Truth and Lie blocks ($M_s = 5.3$ s and 5.7 s, respectively; $SD_s = 0.46$ s and 0.41 s; $t(19) < 1$). By examining the eye movements made in each trial we can ask how gaze was related to the most preferred item, and how this relationship changed in the Lie block. We began by looking at the spatial distribution of fixations, before looking at the time course of attention.

Fig. 3. (a) Fixation position distributions for Truth and Lie trials. Maps show the fixations from all participants, with brighter points being fixated more frequently, and the location of the four items (the preferred item, $P$, and the three distractors). (b) Mean proportion of fixations on each item, with standard error bars. Distractors are labeled A–C, clockwise from the preferred item (e.g., see labels in [a]).
To examine where people looked during the preference task, we plotted all fixations relative to the item that was ultimately chosen (Fig. 3a). Fixation coordinates were rotated around the center to align all trials to a common reference frame, with the preferred item in the top left. From this figure, it is clear that participants modulated their attention in the Lie block. In Truth trials participants spent longest fixating the preferred item, whereas in Lie trials attention was distributed more evenly.

To quantify these trends, we examined the proportion of fixations during the Truth and Lie blocks that landed on each of the items (Fig. 3b). The four regions of interest were classified as the preferred item (chosen by that viewer during the preference task) or a distractor (labeled A–C, clockwise from the preferred item). In a two-way, repeated-measures ANOVA, there was an interaction between condition and item, demonstrating that items were fixated differently in Truth and Lie trials ($F(3,57) = 28.8$, $p < .001$, $\eta_p^2 = .60$). In the Truth condition, there were significantly more fixations on the preferred item than on any of the distractors (simple main effect: $F(3,17) = 33.5$, $p < .001$, $\eta_p^2 = .85$; all Bonferroni-corrected pairwise comparisons, $t$s(19) > 8.5, $p$s < .001, $d$s > 1.9).

In the Lie condition, fixations were distributed between the four images and the preferred item was looked at less often than in the Truth block ($t(19) = 5.9$, $p < .001$, $d = 1.8$). However, participants still showed a systematic pattern, even in the Lie trials ($F(3,17) = 15.7$, $p < .001$, $\eta_p^2 = .73$). The distractor diagonally opposite the preferred item was looked at most frequently when trying to mislead (reliably different from the other distractors, both $t$s(19) > 5.4, $p$s < .001, $d$s > 1.2), and the preferred item was looked at slightly more often than the two other distractors (but there were no other significant differences: all $p$s > .16). These results are averaged across many trials, and so they could mask a variety of different misleading strategies in the Lie trials. For example, participants might spend a disproportionately long time on a single distractor, to falsely indicate this as their chosen item. Alternatively, they might try to look at all four items more equally (e.g., spending 25% of fixations on each). On average, the most looked at item in the Lie trials was selected on 34% of fixations, compared to 41% in the Truth block ($t(19) = 2.9$, $p < .01$). Thus, attention in individual trials was more evenly distributed in the Lie block.

Next, we investigated the time course of looks toward the different items. A central fixation point constrained gaze to start in the center of the display. We computed the proportion of subsequent fixations on each item as a function of ordinal fixation number. Trials ranged in length, with a mean number of 16.2 fixations per trial (SD = 5.6) and no reliable difference between the number of fixations in each block ($t(19) < 1$). Because of the variation in trial length, we calculated the gaze distribution separately for the last fixation, which indicates where people were fixating when they made their manual response. The results are shown in Fig. 4.

These results mirror those of the trial overall. In the Truth block, fixations were biased to the preferred item. This was confirmed by an item by time ANOVA, where there was a main effect of item, $F(3,57) = 29.8$, $p < .001$, $\eta_p^2 = .61$, and the pattern was clear even from the second fixation (i.e., following the first free saccade; Preferred
Fig. 4. The time course of fixations to preferred items and the distractors (labeled A–C), for trials in the Truth and Lie blocks. Data points show the mean proportion of fixations in each condition across participants, along with the value for the last fixation in the trial.
item different from all others, $ts(19) > 4.1$, $ps < .005$, $ds > 1.4$). There were also significantly more final fixations on the preferred item than on the other fractals in the Truth trials ($ts(19) > 4.0$, $ps < .005$, $ds > 1.3$). In the Lie block, there was much less of a difference between fixations to the four items over the first 10 fixations. There was a smaller main effect of item, $F(3, 57) = 6.0$, $p < .01$, $\eta_p^2 = .24$, and the chosen item was only preferentially fixated on the 9th fixation (significantly different from item C, $t(19) = 3.8$, $p = .001$, $d = 1.2$; no other significant Bonferroni-adjusted comparisons). As well as this slight tendency to fixate the ultimately preferred item, there was also a bias toward the item placed diagonally from it (labeled distractor “B”). Interestingly, this bias was very strong on the last fixation, where participants were much more likely to be looking at the diagonally placed distractor than at any other item ($ts(19) > 2.1$, $ps < .05$, $ds > 0.8$).

2.2.3. Modeling inference from gaze

The eye movement data show that participants looked more at the preferred item during the Truth block. This would provide a cue which the guessers could have used to infer preferences. To further investigate such cues, we computed a simple model to classify preference based on gaze using multinomial logistic regression with maximum likelihood estimation. The model was trained on one third of the trial data from the Truth block, equivalent to the trials in the Control condition. The model was then tested on the remaining Truth and Lie trials, just as the human guessers were.

The resulting model had four continuous predictor variables: the proportion of gaze on each of the four items. The outcome variable was the chosen item, coded A–D, clockwise from the top left of the display. This model was significantly better than a model with no predictors ($\chi^2(12) = 190.8$, $p < .001$). Adding a nominal predictor variable which coded for the last item being fixated did not improve the model fit significantly. Inspection of the parameter estimates confirmed that the proportion of fixations on an item predicted the odds of that item being preferred. For example, Table 1 shows the results of the multinomial logistic regression for one outcome, the choice of item B, compared to the reference category A. As we would expect, the proportion of fixations on B was a significant positive predictor of choosing this item. There was also a negative effect of fixations on A (the reference category), such that fixations on A made it less likely to choose B (and more likely A). These trends persisted across Truth trials where other items were chosen, with coefficients for the percentage of fixations on the chosen item ranging from 0.064 to 0.143 (Wald $\chi^2 s(1) > 2.8$, $p$ values from .003 to .093) and odds ratios from 1.07 to 1.15. In the training set, classification accuracy was 74.2%.

The model can be validated by comparing the model-predicted choice with both the actual chosen item and the guess made by the human guesser. In the test set, the model classified 78% of Truth trials correctly, much better than chance and also better than the human guessers (who guessed 62% of Truth trials, see Fig. 2). The accuracy of model predictions and human guesses were strongly associated ($\chi^2$ test of association, $\chi^2(1) = 35.5$, $p < .001$): They tended to get the same trials right or wrong and predicted
the same item on 67% of Truth trials. Indeed, even when both the model and the guesser made an incorrect classification they made the same error in a majority of cases (62%).

Perhaps most interesting, we can use this model, trained on data from the Truth trials, to predict performance in Lie trials, in the same way that the human guessers were asked to do. The model classifier, just like the guessers in the experiment, performed poorly when classifying the Lie trials. Overall, only 31% of trials were correctly classified. Human and model accuracy on a trial-by-trial basis were associated ($\chi^2 (1) = 22.9$, $p < .001$). The same preference was inferred by both the guessing participant and the statistical model on 48% of Lie trials.

In sum, modeling the inference from gaze confirms that eye movements changed in the Lie condition, and that guesser accuracy and the particular errors made can be accounted for by a systematic response to the observed gaze behavior. The change in eye movements caused lower guessing accuracy when the observer was trying to mislead.

2.2.4. Preference patterns across Truth and Lie blocks

Participants in this experiment were instructed to make their true preference choice, but in the Lie block they were also encouraged to mislead the later observer. It is possible that some participants might have not followed this instruction and moved their eyes normally, while giving a dishonest preference response. Moreover, because items were fixated differently during the Lie block, we might expect participants’ preferences to change. For these reasons, it is useful to consider the pattern of preference judgments in the two conditions. Specifically, if the preference expressed by participants is altered by the instructions or a change in their attention, the items that tend to be liked or disliked will be different in the Lie block.

To examine this, we asked a new and heterogeneous sample of participants to choose their preferred item via the crowdsourcing website Amazon Mechanical Turk. Seventy-four respondents rated the same sets of fractals that were seen by the

<table>
<thead>
<tr>
<th>Coefficients from a multinomial logistic regression predicting one of the outcomes: the choice of item B relative to the reference category A</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Odds Ratio (with 95% confidence interval)</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Percentage of fixations on this item</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
</tbody>
</table>

Note. **$p < .001$, *$p < .01$.**
eyetracked participants. We then conducted an item analysis, comparing the likelihood that each fractal would be preferred in the Truth and Lie blocks and the online survey. If participants were choosing likeable items, even in the Lie block, then there will be a positive correlation between the probabilities of a particular fractal being chosen in each dataset. Conversely, if participants are misleading subsequent observers by choosing unlikeable items, there will be a negative correlation.

The correlation between the probability of a fractal being preferred in the Truth block and the online survey was 0.64 ($p < .01$), demonstrating that there was significant inter-observer agreement. However, in the Lie block the same correlation was much weaker: $r = 0.12$, $p = .44$. On the one hand, this shows that participants in the Lie block did indeed choose different items. On the other hand, this correlation is not negative, so it does not appear to be the case that participants chose deliberately unlikable fractals.

Importantly, the tendency to reduce gaze toward the preferred item in the Lie block does not seem to be caused only by the different items that were chosen. We repeated the eye movement analysis above, restricting it to only those trials where participants chose the item which was most preferred by the neutral, online raters. In these trials we can be more confident that the eyetracked participant’s manual response was a true reflection of preference. The eye movement results remained unchanged in these trials, showing again that participants were most likely to look at the chosen item in Truth trials (on 45% of all fixations, on average), and much less likely to do so in Lie trials (25% of fixations).

2.3. Discussion

This experiment aimed to introduce a two-way social interaction into a laboratory experiment, allowing an investigation of how inferences can be made from eye gaze. Gaze following provides information about where people are attending, information that in our experiment was provided by the replayed fixations. The results confirmed that participants in the Guess block could spontaneously use this information to infer the chosen item. This result is consistent with findings from developmental psychology (Einav & Hood, 2006), but for task and stimuli that were arguably more complex, abstract, and dynamic. Human guesses, both correct and erroneous, could be statistically modeled as a response to the gaze distribution.

Although Experiment 1 demonstrates that participants can recognize preferences from eye movements, the guessers’ ability to do this may have been enhanced because they themselves had just completed the task. Moreover, during the guess block, these participants did not yet know of the possibility that they were being deceived. Can guessers also distinguish Truths from Lies? Answering this question will give further insights into whether participants are using theory of mind mechanisms in this task, as well as into how well participants can mislead with their eyes. In Experiment 2, a new group of naïve guessers took part who did not have the experience of completing the preference task.
These participants were first asked to infer the chosen item. They were then asked to identify Truth and Lie trials.

3. Experiment 2

3.1. Method

3.1.1. Participants
Twenty volunteer students took part in exchange for payment or course credit. There were 12 males and participants had a mean age of 21.2 years. All the participants were naïve to the purpose of the experiment and had not taken part in Experiment 1. Each participant in this experiment was yoked to two different participants from Experiment 1, and across all participants all the eye movement data were replayed in both blocks of Experiment 2.

3.1.2. Stimuli and apparatus
Participants watched replays of eye gaze in exactly the same way as during the guess block of Experiment 1. The same stimuli and monitor were used, but eye movements were not recorded in this experiment.

3.1.3. Procedure and design
At the beginning of the session, the participants were told about the preference task and given an example of the eye movement animations. They then completed a block of 36 trials, under the same instructions as the Guess block in Experiment 1. Participants were asked to make their best judgment as to which of the four items was chosen by the original observer. As in Experiment 1, the 36 Guess trials comprised trials displaying eye fixations from either the Truth or Lie condition, or they displayed no eye movements (Control condition). Condition order was randomized and interleaved.

After completing this block, participants completed a short questionnaire which asked them two open-ended questions about the aim of the experiment (“What do you think this experiment was about?”) and their performance (“How did you make your judgment?”), as well as asking them to rate the difficulty of the task. Only after this questionnaire had been completed were participants told that there would be a second task.

In the second block, it was explained that some of the time participants had been asked to mislead the person who was guessing. Participants then saw a second block of 36 trial animations, all of which displayed eye movements from a participant who had taken part in Experiment 1 (and who was different from the source of the trials shown in the first block). The trials were drawn equally from the Truth and Lie blocks and shown in a random order. After watching each animation as many times as they wanted, guessers were instructed to judge (a) whether the trial was a Truth or a Lie and (b) what the actual preferred item had been. At the end of the experiment, guessers completed a second short questionnaire asking how they had made their judgments.
3.2. Results

The manual guessing responses were analyzed to see how well participants could guess the response and condition from Experiment 1. Because the eye fixations being replayed were exactly the same as those in that experiment, the first question was whether the same pattern of guessing accuracy would be seen in a new group of naïve participants.

3.2.1. Block 1 guessing accuracy

The mean proportion of trials in which participants correctly guessed the chosen fractal is shown in Fig. 5a. These data represent a partial replication of the findings from the Guess block in Experiment 1. Participants were most accurate at guessing based on Truth trials, and they performed at chance levels when responding to Lie or Control trials. There was a significant main effect of condition on guessing accuracy, $F(2,38) = 57.9$, $p < .001$, $\eta_p^2 = .75$, and accuracy in guessing Truth trials was reliably greater than in the other conditions (both $t(19) > 8.9$, $p < .001$, $d > 2$). Lie and Control trials were not significantly different from each other, and neither were they different from the chance level of 25% accuracy (all $t(19) < 1$).

3.2.2. Lie detection accuracy

In the second block, participants made a two-alternative forced choice as to whether each trial was a Truth or a Lie. A guessing strategy here would result in 50% accuracy. On average, lie detection accuracy was not significantly different from chance (one-sample $t$-test against 50%, $t(19) < 1$). Looking at the accuracy for correctly identifying Truths and Lies separately (Fig. 5b), there was no evidence for a bias toward responding with one category and thus sensitivity was essentially zero.

Participants were also asked again to guess which item the observer had chosen after he or she had made a Truth/Lie judgment. In those Truth trials which were correctly recognized as such, mean guessing accuracy was 55% ($SEM = 5\%$). While lower than the guessing accuracy in the first block, this remained much greater than the 1 in 4 expected by chance. In contrast, guessing accuracy when responding to correctly detected Lie trials was no better than chance ($M = 25\%$, $SEM = 5\%$). Thus, even when Lie trials were correctly detected, participants were unable to overcome the deception of the original observer.

3.2.3. Participant self-reports

Questionnaire responses following the Guess block confirmed that participants did not know about the Truth/Lie instructions beforehand. Participants rated guessing the item chosen to be of moderate difficulty (mean rating 3.4 of 5, where 1 was “very easy” and 5 was “very hard”). The most commonly reported strategies were choosing the item that was looked at the most (mentioned by 70% of respondents) and choosing the item looked at last (also stated by 70%).
Participants found discriminating lies more difficult (average rating, 3.7). Four of 20 participants stated that they were guessing, whereas most others claimed to recognize lies where eye movements were too “quick” or “erratic.”

Fig. 5. (a) The mean accuracy of participants in Experiment 2 when guessing which item was preferred. (b) The mean accuracy when identifying Truth or Lie trials. Error bars show ± 1 SEM. Chance performance is shown by the horizontal dashed lines.
4. General discussion

This study tested whether participants could guess the preference of another person based on their overt attention. In addition, it tested whether people could manipulate this attention to deceive an observer.

4.1. Making inferences from gaze

The results confirmed that people were able to read the minds of someone in the preference task by watching the person’s eye movements. Guessers shown the fixations were much more accurate than expected by chance. This is particularly interesting given that the representation of fixations was novel for all the observers and no training or feedback was given. The results were replicated in Experiment 2, when participants had no previous experience with the task. It is significant that guesses in the Truth condition also outperformed the control trials, where no fixations were shown. In Experiment 1, these control trials—blind guesses—were accurate about 40% of the time and more so than the ¼ predicted by chance. This margin likely reflects between-subjects similarity in preferences: Participants could sometimes guess correctly because some fractals were actually more likeable. Inspection of the choices made in particular stimulus displays confirmed that they were idiosyncratic, which was important for our design and hence performance in the control condition was not even higher (as it would be if everyone had chosen the same item). Displaying (truthful) eye movements significantly enhanced any intuitions guessers had about the previous participant, in both experiments. We propose that the same mechanisms are used to make inferences from gaze during humans’ everyday social interactions.

During Truth trials—where participants behaved naturally—the preferred item was fixated most frequently. This finding is in agreement with those from previous studies showing the link between fixations and choice (e.g., Shimojo et al., 2003). Detailed inspection of the time course of fixations in our multi-item preference task showed that the bias toward the preferred item emerged early, and that it persisted, with final fixations also most likely to land on this item. However, most of the nonchosen items were also fixated at least once, confirming that participants were evaluating the different stimuli before making their choice. Holmes and Zanker (2012) recently reported that combining oculomotor measures such as those used here could successfully predict aesthetic. Of course, this study cannot distinguish between the tendency to look at items one finds aesthetically pleasant and the tendency to look at something one is going to choose. Models of choice such as the gaze cascade model (Glaholt & Reingold, 2009) and the drift-diffusion model (Krajbich, Armel, & Rangel, 2010) can account for the bias seen toward the chosen item, and recent evidence suggests that it is decision processes and not aesthetic value which modulates fixation (Isham & Geng, 2013). Importantly, the present results show that naïve participants know and can spontaneously use this link between gaze and choice to guide their inferences.
4.2. Manipulating attention to mislead

Regarding our second aim, there was also evidence that people could consciously change their eye movements to mislead. The large bias toward looking at the to-be-chosen item changed to a more uniform distribution. An important question for decision-making researchers interested in gaze is the degree to which the looking behavior associated with a choice is under conscious control or reflects an automatic process. The present results demonstrate that gaze during the choice task was affected, top–down, by a change in the task instructions.

More specifically, behavior in the Lie trials provides strong evidence that people can moderate their gaze behavior according to the demands of the social situation—in this case to mislead an observer. When doing so, they decreased their looking at the preferred item and successfully caused a naïve observer to guess incorrectly. The ability to deceive is considered a hallmark of theory of mind (e.g., Leslie, 1987). Participants in the Lie block had to intuit how their attention was being deployed and how the guesser would interpret this behavior. Further research could manipulate the feedback and information available to both parties—the deceiver and the guesser—in order to probe the theory of mind and perspective-taking necessary to perform the task. The simplicity of the current procedure makes it well suited for such investigations. For example, it has been proposed that adults possess two systems for performing theory of mind type tasks: a fast, inflexible system which is used by young infants and a more cognitively demanding but flexible system which develops later (Apperly & Butterfill, 2009). The faster and more primitive of these systems, perhaps relying on automatization or learned behavioral associations, could conceivably explain the gaze interpretation shown in the current task. To determine this, further experiments with this paradigm could investigate how cognitively demanding both gaze interpretation and gaze regulation are, perhaps through use of a dual task.

It is important to note that behavior in the Lie condition did not show the same bias as the natural task but neither was it completely random. This confirms what the participants told us—that they responded with their actual preference but tried to disguise this choice by changing their looking behavior. Although participants successfully misled the subsequent observer during the Lie block—reducing guessing performance to chance levels—their eye movements still provided clues to their choice. Participants were not given any instructions about how to change their fixations. However, they showed a consistent pattern to look more often at the diagonally opposite item in the display, particularly on the last fixation. It is not clear why this pattern emerges. One possibility is that it reflects a desire to look further away from the to-be-chosen item. There was also a tendency to look slightly more at the preferred item, which again indicates that participants were not responding randomly.

At this point, it is important to consider an alternative explanation of the change in guessing accuracy for Lie trials. During these trials, despite being asked to genuinely indicate the most preferred item, participants might have behaved normally, fixating their chosen item, but then deceived by indicating a different, nonpreferred item with their manual response. It is possible that this happened on some of the trials, and this is one
potential explanation of the fact that preferences seemed to change in the Lie block. If participants were acting in this way, it would still show sophisticated awareness of one’s own gaze position and how another party would interpret this gaze position (“I prefer X, but the guesser would be able to infer that by the fact that I looked at it frequently, therefore I better choose Y”).

However, a deceptive manual response at the end of the Lie trials cannot provide a full account of the results, for several reasons. First, anecdotal self-reports from participants in Experiment 1 suggest that most participants were aware of changing their gaze behavior during the Lie block, and not their manual response. Second, if observers were behaving normally during the trial and then making a deceptive response, we would expect eye movements during Truth and Lie trials to be the same. However, even apart from the tendency to look or not look at the preferred item, there was evidence for different strategies in each block. Participants in the Lie block tended to focus less on one particular item than those in the Truth block. The nonrandom relationship between gaze and choice, even when trying to mislead, also counts against this possibility. Third, because participants are exposing themselves to stimuli to different degrees in the Lie trials, preferences are likely to change relative to truthful participants. However, if we consider inter-observer agreement as a reflection of a “true” preference, there were still many Lie trials where participants chose a fractal that naïve observers agreed was most likeable. In these trials the eye movement results were identical in that participants continued to avoid looking at the chosen item. Nevertheless, future research could examine how gaze can be regulated in different judgment tasks, and how changes in looking behavior might actually alter future choices (as in Shimojo et al., 2003).

Much of the previous scientific interest in deception has been directed at improving observers’ ability to spot liars (DePaulo et al., 2003). In Experiment 2, we asked whether participants could detect whether the replayed eye movements came from a Lie or a Truth trial. Performance was statistically indistinguishable from chance, showing that, although they intuited the relationship between gaze and preference, participants could not determine the lies. There was also no sign of a “truth bias,” which is commonly seen when participants in lie detection experiments are biased toward identifying truths (Zuckerman, DePaulo, & Rosenthal, 1981). Even when they did detect a lie, participants were unable to determine the true preference. Nevertheless, it is possible that the remaining systematic eye movement patterns could be used to improve guesses in the Lie trials. Recently eye movements, along with pupil dilation, have been touted as a possible implicit sign of guilty knowledge in a lie detection scenario (Schwedes & Wentura, 2012) and in game-theoretic biased-transmission tasks (Wang, Spezio, & Camerer, 2010). For example, Schwedes and Wentura (2012) adapted the “concealed information test” to see whether participants asked to withhold their recognition response to a previously seen face would nonetheless fixate it differently. The findings showed that the eyes did indeed reveal recognition, making overt attention a potentially useful measure for detecting deception. Steptoe, Steed, Rovira, and Rae (2010), meanwhile, showed that incorporating realistic eye movements into avatar-based teleconferencing makes it easier to decide on the truthfulness and trustworthiness of verbal statements. There are many other contexts
where detecting deception involves the reading of subtle cues. For example, Sebanz and Shiffrar (2009) showed that participants could interpret the nonverbal cues of basketball players attempting to deceive with fake passes. This study suggests that such research should look at the interaction between both the “faker” and the judge in these, inherently social, situations.

Humans transmit their gaze direction very clearly, unlike most primates whose sclera may camouflage where they are looking (Kobayashi & Koshima, 1997). In this study, the selection of the eyes and following of gaze direction was not required and fixation location was displayed symbolically. However, the spontaneous way in which participants interpreted and regulated fixation, as well as previous research showing effective gaze transfer via a cursor, implicate gaze-following and theory of mind mechanisms. Our simple paradigm could be used to test the extent to which interpreting a gaze cursor activates the neural structures known to be involved in more traditional gaze cueing tasks. Importantly, our results highlight a flipside of the human ability to transmit and receive gaze cues: The eyes can also tell lies and be used to mislead observers.

4.3. Implications

It is clear that detecting and following eye gaze is important for behaving in a social context. This study shows that, when provided with information about the focus of attention, participants are adept at using this information to predict preferences. However, they are also able to change their attention depending on the social situation and communicate their preferences, even when such communication is deceptive. This novel paradigm provides an innovative method for investigating the reading and misleading of inferences based on social attention. Because participants are involved in an interaction with the guesser, this paradigm offers rich possibilities for studying attention and inferences in this context (possibilities which are limited in paradigms such as gaze cueing, see Risko et al., 2012).

The results also have implications for the use of gaze as a model in learning, expertise, and problem solving. In several studies across different domains it has been shown that participants in skilled visual search tasks can improve if shown the eye movements of an expert (Litchfield et al., 2010; Nalanagula, Greenstein, & Gramopadhye, 2006). Moreover, showing an expert model’s eye movements has the potential to improve learning of a perceptual task (van Gog & Scheiter, 2010). These studies, along with eye movement cued retrospective think aloud procedures, rely on observers making inferences from gaze, and thus the present results support their validity. Our results also emphasize that the model’s eye movements may change according to the socio-communicative context (e.g., if they know that they are a model; see Brennan et al., 2008; for a similar insight). Although we have focused on how gaze can mislead, it is clear that we often regulate our attention in a more positive fashion to enhance communication. Indeed, experts in gaze transfer tasks may make their fixations deliberately informative, and problem solvers learn quicker when deliberate, “didactic” scanpaths are replayed (Litchfield & Ball, 2011).
4.4. Conclusion

The current experiments address two main aims. First, they demonstrate that participants can use the point of regard (e.g., from gaze following) to infer a participant’s choice or preference. Second, they show that this inference is part of a two-way communication in that participants can manipulate their eye movements to transmit misleading information. Experimental psychology and cognitive neuroscience have frequently studied the way that individuals attend to social stimuli, but this study is an example of using a more interactive design to look at bilateral gaze interpretation in a controlled setting.

Acknowledgments

We thank Damien Wright for assistance with data collection. We are also grateful for the helpful comments of Damien Litchfield and an anonymous reviewer, and for the input of presenters and attendees at the symposium on “Social Gaze” held during the European Conference on Eye Movements, 2013, in Lund, Sweden.

References


