Cognition 184 (2019) 28–43

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

# Gaze allocation in face-to-face communication is affected primarily by task structure and social context, not stimulus-driven factors

Roy S. Hessels[a,b,*], Gijs A. Holleman[a,b], Alan Kingstone[c], Ignace T.C. Hooge[a], Chantal Kemner[a,b,d]

[a] Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, the Netherlands
[b] Developmental Psychology, Utrecht University, Utrecht, the Netherlands
[c] Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada
[d] Brain Center Rudolf Magnus, University Medical Center, Utrecht, the Netherlands

ARTICLE INFO

*Keywords:*
Eye movements
Gaze allocation
Face-to-face
Communication
Faces

ABSTRACT

Gaze allocation to human faces has recently been shown to be greatly dependent on the social context. However, what has not been considered explicitly here, is how gaze allocation may be supportive of the specific task that individuals carry out. In the present study, we combined these two insights. We investigated (1) how gaze allocation to facial features in face-to-face communication is dependent on the task-structure and (2) how gaze allocation to facial features is dependent on the gaze behavior of an interacting partner. To this end, participants and a confederate were asked to converse, while their eye movements were monitored using a state-of-the-art dual eye-tracking system. This system is unique in that participants can look each other directly in the eyes. We report that gaze allocation depends on the sub-task being carried out (speaking vs. listening). Moreover, we show that a confederate's gaze shift away from the participants affects their gaze allocation more than a gaze shift towards them. In a second experiment, we show that this gaze-guidance effect is *not* primarily stimulus-driven. We assert that gaze guidance elicited by the confederate looking away is related to the participants' sub-task of monitoring the confederate for when they can begin speaking. This study exemplifies the importance of both task structure and social context for gaze allocation during face-to-face communication.

## 1. Introduction

The world as it is relevant to a human being[1] is, for a large part, social. Social partners inhabit our worlds, and there are often times when one can or needs to interact with another. These interactions contain a constant flow of information between the interacting partners (or agents), which can be conceived as a continuous cycle of perceiving and acting upon that which is perceived. For example, one may perceive the emotional expression of another and act upon it by approaching and consoling or retreating (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), or one may perceive the direction of another's gaze and act upon it by changing gaze direction as well (Langton, Watt, & Bruce, 2000; Frischen, Bayliss, & Tipper, 2007). In sharp contrast to the interactive nature of social behavior, the dominant approach in psychological research has been to separate perception from action (insofar this is possible; Heft (2001)), and focus on the individual passive agent observing the faces of others (Risko, Richardson, & Kingstone, 2016). Recently, however, a movement of researchers has arisen advocating the study of attention, perception or cognition in interaction (Caruana, McArthur, Woolgar, & Brock, 2017; Cole, Skarratt, & Kuhn, 2016; De Jaegher, Di Paolo, & Gallagher, 2010; Pfeiffer, Vogeley, & Schilbach, 2013; Risko, Laidlaw, Freeth, Foulsham, & Kingstone, 2012; Risko et al., 2016; Schilbach et al., 2013; Schilbach, 2015).

One exemplar field of research in which the interactive approach has been brought to fruition is that of social attention – here conceptualized as the study of attentional processes in relation to social partners of the agent. A paramount discovery in this field has been that when a social partner is physically present, attentional processes such as the allocation of gaze are modulated compared to when static pictures of that partner are observed. For example, Laidlaw, Foulsham, Kuhn, and Kingstone (2011) showed that a confederate who was physically present in a waiting room was not often fixated by observers, whereas a video feed of that confederate was. This indicates that the

* Corresponding author. at: Heidelberglaan 1, 3584CS Utrecht the Netherlands.
*E-mail address:* royhessels@gmail.com (R.S. Hessels).

[1] The world as it is relevant to a (human) being is best captured by the term *Umwelt*, as coined by Jakob Von Uexkull (Umwelt und Innenwelt der Tiere, Berlin: Springer, 1909).

https://doi.org/10.1016/j.cognition.2018.12.005
Received 25 June 2018; Received in revised form 30 November 2018; Accepted 6 December 2018
Available online 14 December 2018

social context influences gaze behavior. Moreover, gaze allocation to the face of a person was increased by an interviewer making eye contact versus not making eye contact, but only when that interviewer was physically present compared to when the interviewer was pre-recorded (Freeth, Foulsham, & Kingstone, 2013). It has even been shown that gaze behavior to facial features (eyes, nose, and mouth) was modulated by the fact that a person of lower or higher rank would watch a video of the observer later on (Gobel, Kim, & Richardson, 2015). Specifically, the eyes of higher ranked social partners would be looked at for a shorter duration when observers knew that they might be watched by the other later on, whereas the eyes of lower ranked social partners were looked at for a longer duration. The studies by Freeth et al. (2013), Gobel et al. (2015), and Laidlaw et al. (2011) highlight that gaze allocation (where people look at a given time) does not only serve visual information uptake, but may also serve as a social signal to others with whom one may interact (see also Jarick & Kingstone, 2015). This fact is exemplified in conversational settings, where gaze allocation may facilitate turn-taking (Argyle, 1972; Ho, Foulsham, & Kingstone, 2015).

Although the social attention literature has shown that the *social context* (with whom is being interacted and in what setting – live, video-mediated, and so forth) affects gaze behavior in face-to-face interaction, no models thus far exist that predict gaze allocation in face-to-face interaction based on the setting, task being carried out, or stimulus factors. The fact that gaze allocation may serve as a social signal in interaction, means that models of gaze allocation have to go beyond mere visual saliency (e.g. Itti & Koch, 2000) and incorporate contextual factors (Birmingham, Bischof, & Kingstone, 2009). Here we are concerned with the feasibility of such a model, and investigate what factors are relevant when modeling gaze allocation in interaction. Such models are important for understanding social interaction in general, atypical development of social behavior (e.g. Guillon, Hadjikhani, Baduel, & Rogé, 2014), and have wide applicability to, for example, the study of student-teacher interaction in educational settings and patent-physician interaction in medical settings (e.g. Asan, Young, Chewning, & Montague, 2015; MacDonald, 2015). Besides the social attention literature, we draw on perspectives from the study on the task control of eye movements, and observational studies on face-to-face communication to ascertain what factors may be relevant when modeling eye movement in interaction. We start, however, with the assumptions inherent in much of the eye movement research in psychological science. The study of gaze allocation in face-to-face interaction is founded on these assumptions, yet they are often implicit.

### *1.1. Assumptions of eye-movement research in psychological science*

The investigation of eye movements in psychological science is predicated foremost on the assumption that the allocation of gaze is in some way reflective of psychological processes. Humans are foveated animals, and make saccades to consecutively bring areas of the visual world onto the fovea (the part of the retina with the highest acuity) for further scrutiny. The location of fixation (or point of regard) is generally assumed to coincide with the spatial location of the focus of attention. Moreover, it is assumed that information around the fixation location is processed. That is, if one fixates the eye region of a face, it is assumed that there is some information at that location that may be relevant for the observer. These assumptions are mostly implicit in the social attention literature, although they readily apply to it. Researchers in the social attention literature, however, supplement these assumptions by further assuming a potential communicative role for gaze allocation.

An important question is whether and when these assumptions hold. Is, for example, the allocation of gaze to a certain area necessary for completing a particular task? Does the eye region contain information that is critical for e.g. recognizing who's face it is (e.g. Henderson, Williams, & Falk, 2005)? And if so, does the eye region need to be fixated, or is the relevant information accessible from peripheral vision when fixating the nose? As Tatler, Hayhoe, Land, and Ballard (2011) state, "It is probably a mistake to think that every fixation must have an identifiable purpose" (p. 8). Is it possible to identify when the allocation of gaze has a function or when it does not? Such questions are already being tackled in the literature on task control of eye movements and although it is not aimed at gaze allocation in face-to-face interaction per se, it may provide a useful perspective on our problem.

### *1.2. Task control of eye movements*

Already in the 60's of the last century, Yarbus (1967) showed that task instructions influence gaze behavior. Since then, Land, Mennie, and Rusted (1999) and Hayhoe (2000) have shown that the gaze allocation of an individual can be tightly coupled to the task that one is carrying out. Land et al. (1999), for example, had a participant make tea while their eye movements were recorded. The allocation of gaze turned out to serve multiple functions during the task, for example locating an object (e.g. a cup or kettle) that needed to be used at a later time, or checking the status of the water level when filling the kettle. Gaze allocation thus serves the current informational demands for completing a specific task. Indeed, as Land et al. (1999) state: "…although the actions of tea-making are 'automated' and proceed with little conscious involvement, the eyes closely monitor every step of the process. This type of unconscious attention must be a common phenomenon in everyday life." (p. 1311).

Modeling gaze allocation as a function of informational demands during ongoing activities is a daunting task, as it requires an understanding of the moment-to-moment requirements for the task at hand (Hayhoe & Ballard, 2005, 2014). In other words, the following questions need to be considered. What task is a person carrying out? What current sub-task(s) needs to be completed? How can the allocation of gaze support that sub-task by supporting the acquisition of the task-relevant information? These questions have remained largely unanswered in the literature on social interaction. We believe that they need to be considered if one is to fully understand the role of gaze allocation during face-to-face interactions.

Modeling gaze allocation based on the underlying task structure (the sub-task division and its order) has shown promise in understanding gaze allocation in a broader sense than saliency-based models alone (Tatler et al., 2011). However, when one applies the concepts from the task control of eye movements literature in order to better understand gaze allocation in face-to-face interactions, a problem is encountered. In face-to-face interactions, the allocation of gaze is not only affected by informational demands. Social context, that is, the relation between the two interacting partners may play a pivotal role. For instance, it may be that the allocation of gaze subserves information acquisition only, such as when one fixates the mouth of an individual in order to better understand what is being said (e.g. Vatikiotis-Bateson, Eigsti, Yano, & Munhall, 1998; Vø, Smith, Mital, & Henderson, 2012). It may also be the case that the allocation of gaze serves a communicative role, for instance when a speaker finishes talking and looks to the listener to signal that it is now his/her turn to talk (Argyle, 1972; Ho et al., 2015). The problem is further complicated when one considers that the direction of one's eyes may, as part of a highly configurable facial expression, signal various emotions or affect (Crivelli & Fridlund, 2018; Jack & Schyns, 2017). As such, understanding how gaze is deployed in face-to-face communication requires knowledge of how the task structure and the social context affect eye-movement patterns. Insights for this question stem from observational studies on face-to-face communication, and more recent eye-tracking studies that have focused on gaze allocation to faces in relation to social context or task demands. These are discussed in detail below.

### *1.3. Insights for the study of gaze allocation in face-to-face communication*

The role of gaze in face-to-face communication became a topic of empirical investigation in the 70's and 80's of the last century (e.g. Allen & Guy, 1977; Argyle, 1972; Beattie, 1981; Cary, 1978; Duncan & Fiske, 1977; Krantz, George, & Hursh, 1983). Cary (1978), for example, showed that looking at a person in a waiting room precedes the initiation of conversation. Argyle (1972), reported that participants looked longer toward a conversational partner when they were listening, and their looks away were much shorter, than when they spoke. Furthermore, the patterns of looking at each other during conversation were shown to depend on the relation between the partners (e.g. competitive versus cooperative; Foddy, 1978), and was related to personality differences (Libby & Yaklevich, 1973). In more recent observational work, Clark and Krych (2004) have shown that looks at the face of a person giving instructions occurred when the instructed needed to resolve a conflict, attesting to the fact that gaze in face-to-face interaction is likely tied to spoken communication. Although observational studies have yielded valuable insights, the assessment of gaze direction by this method does not, however, have the spatial resolution that modern eye trackers have, and it depends on the subjective assessment of gaze direction (e.g. Beattie & Bogle, 1982).

Using eye-tracking technology, which allows the objective measurement of gaze direction with high spatial resolution, researchers have shown how gaze allocation to faces depends on task demands. Itier, Villate, and Ryan (2007), for example, instructed participants to determine the orientation of the eyes or head in photographs of faces. When told to determine the orientation of the head, more saccades went into the direction in which the head was oriented, compared to the direction in which the eyes were oriented. Based on these findings, they conclude that "these task-modulated behaviors argue against a purely exogenous and automatic orienting-to-gaze mechanism" (p. 1025). Nummenmaa, Hyönä, and Hietanen (2009) instructed participants to decide on which side to skirt an animated person walking towards them. Participants tended to look in the direction in which the animated person was not looking and chose to skirt the person on that side. They conclude that "Gaze following is not always an obligatory social reflex; social-cognitive evaluations of gaze direction can lead to reversed gaze-following behavior" (p. 1454). These eye-tracking studies show that task demands can determine one's gaze direction in relation to other people. However, the investigation of gaze allocation has come at the cost of restraining participants, and precluding them to actually interact with others.

More recently, several studies have been conducted using wearable eye-tracking technology, which measures gaze direction while people can walk around or interact in an environment. Although wearable eye trackers allow the objective measurement of gaze direction, they often lack in spatial resolution to allow one to discriminate whether e.g. facial features are looked at. In an interview setting using a wearable eye tracker, Freeth et al. (2013) reported that participants looked less at the face of another person when speaking as compared to when listening (cf. Argyle, 1972). Similarly, Macdonald and Tatler (2013) have shown that individuals gazed more often at the face of an instructor when they were given ambiguous instructions, and Hanna and Brennan (2007) have shown that looking at someone helps to disambiguate statements made in conversation. Macdonald and Tatler (2018) have furthermore shown that looking at an interacting partner when that partner speaks depends on roles they have been given in a cooperative task. These wearable eye-tracking studies show that gaze to the faces of others depends on the task being carried out (e.g. speaking versus listening), is sensitive to the broader task-context (specific role-division in cooperation versus no role-division), and is linked to the spoken communication. However, in these studies, no data were available about gaze allocation to facial features (i.e. eyes, nose, or mouth), only whether faces were looked at or not. Thus, an important question for modeling gaze allocation in face-to-face communication is whether information from facial features is differentially used for completing a task. The eyes and mouth do not convey the same information in conversation (e.g. Võ et al., 2012; Jack & Schyns, 2017), but do these areas need to be looked at in order to extract information from them? Although Hessels, Cornelissen, Hooge, and Kemner (2017a) have recently investigated gaze allocation to facial features in interaction with a high-spatial resolution dual eye-tracking setup, they did not consider task structure or social context specifically.

### 1.4. The present study

We are concerned with the development of a model of gaze allocation in face-to-face interaction, and aim to integrate perspectives from the social attention literature, the literature on task control of eye movements, and observational work on face-to-face communication. In this study, we consider gaze allocation to facial features in a task-structure that is likely to produce different patterns of gaze to faces (speaking vs. listening, see e.g. Argyle, 1972; Freeth et al., 2013; Macdonald & Tatler, 2018). By modulating the behavior of a confederate in the context of this task-structure, we aim to begin to unfold the social context and task-structure influences on gaze allocation to facial features.

The first aim of the present study was to investigate whether gaze allocation to the face and facial features of a confederate was dependent on the task at hand. For this, we employed a conversational setting in which a confederate and participant were asked to tell each other something about themselves on a certain topic. During the conversation, participants could see each other through a state-of-the-art video connection, which allows two people to look each other straight in the eyes. Eye movements of the participant and confederate were recorded using two eye trackers. The participant's task can be conceived to consist of two sub-tasks: first to monitor the confederate for when his story was finished, and then second to talk about their own experience on the topic. Based on the literature (Argyle, 1972; Freeth et al., 2013), we expect that participants will look at the face of the confederate less often and for shorter durations when speaking compared to when listening. Moreover, we will investigate to what degree gaze allocation to the facial features depends on the task of speaking versus listening. This research question sheds further light on the relation between gaze allocation to facial features and task structure in an interactional setting. However, this question does not address the social context in which conversation occurs. Is it, for example, the case that the gaze behavior of the conversational partner affects gaze behavior of the participants as well, and if it does, how does it affect gaze behavior?

The second aim of the present study was therefore to investigate the effect of a confederate's gaze behavior on the gaze behavior of participants in a conversational setting (i.e. with a specific task structure). For this, we instructed the confederate to either look at or away from the participant during different parts of the confederate's story, and investigated whether changes arose in measures of participants' gaze behavior. As the direction of gaze in conversation may signal the end of a speaker's turn (Ho et al., 2015) and gaze allocation is likely to be variant to moment-to-moment changes in the task (Hayhoe & Ballard, 2005, 2014), we particularly investigated changes in participants' gaze behavior in a short time window around the confederate's gaze direction change. The important question here is whether, and to what degree, the gaze behavior of one person in a conversational setting affects the gaze behavior of another.

In discussing the findings from our experiment, we put emphasis on teasing apart which factors may be relevant for gaze allocation in face-to-face communication, which we further investigate in a second experiment. Combined, these experiments pave the way for modeling gaze allocation in human interaction in general.

## 2. Methods experiment 1

### 2.1. Participants

Participants were recruited at the Faculty of Social and Behavioral

Sciences of Utrecht University, and through social media dedicated to (paid) experiments at the faculty. 33 volunteers participated in the experiment (21 female, 12 male). Mean age was 23.19 years (*sd* = 3.90 years). Participants were informed about the study by e-mailing them a full letter of information after applying for the study. Written informed consent was obtained upon arrival in the laboratory. Participation took between 15 and 30 min, and participants were compensated by receiving either 0.5 so-called 'participant hours' (of which Psychology students have to acquire 12 during their studies) or 4€. This research project does not belong to the regimen of the Dutch Act on Medical Research Involving Human Subjects, and therefore there is no need for approval of a Medical Ethics Committee. However, given that participants were deceived in this experiment and that videos of their faces were recorded, we consulted the Ethics Committee of the Faculty of Social and Behavioral Sciences of Utrecht University, whom approved the experiment. The protocol is filed under number FETC17-052.

### 2.2. Apparatus

Eye movements of the participant and confederate were recorded using the dual eye-tracking social interaction setup introduced by Hessels et al. (2017a), which was specifically designed to map eye movements of one participant with high spatial accuracy to the face of another participant while they are interacting. The participant and confederate were placed at either end of the dual eye-tracking setup, and looked at each other through a video connection while their eye movements were recorded. The setup was designed in such a way that participants can look each other in the eyes, even though they are looking at each other through a video connection (by use of half-silvered mirrors behind which cameras are placed). An intuitive comparison may be made by comparing our setup to a Skype call (or any videoconference system). During a Skype call, one can either look at the eyes of the other (i.e. on the computer screen), or look directly in the camera. Only when one looks directly in the camera is the other presented with an image of someone looking straight ahead. Here we've placed the camera exactly at eye height, such that one can make eye contact (so to say) with the interacting partner, making for a more natural conversational setting compared to Skype. Participants were positioned at either end of the table on which the setup was placed. In physical space, the distance between the participants was about 2.3 meters. The presentation of the participants on screen was such that the 'visual' distance was 1.36 meters. As participants were in the same room, they could therefore hear each other, although they saw each other only through the video connection. Importantly, if a participant would move out of the camera image and the setup, (s) he would appear on that side in the visual periphery of the other participant.

Each participant was filmed by a Logitech webcam (recording at 30 Hz at a resolution of 800 by 600 pixels). The live-feed of each participant was concurrently recorded to disk and presented to the other participant at a resolution of 1024 by 768 pixels in the center of a 1680 by 1050 pixels computer screen. Two SMI RED eye trackers were used to record eye movements of the participant and confederate at 120 Hz. A stimulus computer running Ubuntu 12.04 LTS was used to handle the live-feed, and communicated to the eye-tracker computers when to start and stop recording eye movements. For further technical details, see Hessels et al. (2017a). The setup is depicted in Fig. 1. An example movie of the videos and eye-movement data recorded using this setup is available from http://www.royhessels.nl/research.html.

### 2.3. Design

The experiment was set up to investigate (1) the effect of task structure on participants gaze behavior, and (2) the effect of the confederate's gaze behavior on the participant's gaze behavior during conversation. To investigate this, a 26 year-old male confederate conversed live with a participant about their work and holiday experience.

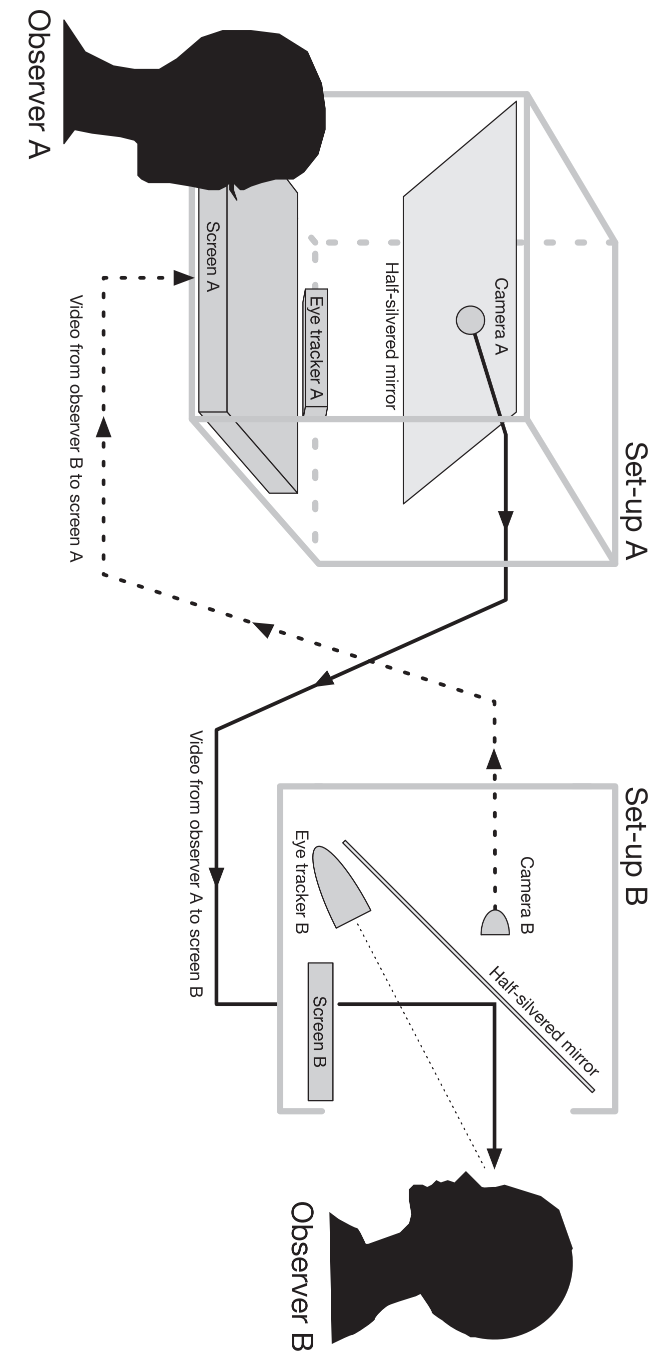


**Fig. 1.** Schematic overview of the dual eye-tracking setup used in the present study. Set-up A and set-up B are depicted at different angles for viewing purposes only. In reality, they were positioned with the backs of the set-ups towards each other. Camera A is depicted in front of the half-silvered mirror for illustrative purposes only. In reality, it was placed behind the half-silvered mirror.

The confederate's story and gaze behavior were scripted. The confederate began talking in each scenario. For each scenario, a script was written that contained four sections based on the autobiographical details of the confederate. For example, in the work scenario, the confederate addressed four jobs he was involved in. During each section of a scenario, the confederate would either look at the participant's face, or would look away from the participant. When the confederate looked away, he looked up and around in various directions, as has been shown to occur during speech more than during listening (Argyle, 1972). The script was not rehearsed word for word, but such that the

**Table 1**
The four possible combinations of confederate gaze behavior in experiment 1. S = section, CF = confederate, P = participant, A = away.

| | Scenario 1 | | | | | Scenario 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | S1 | S2 | S3 | S4 | End | S1 | S2 | S3 | S4 | End |
| CF gazes at | P | A | P | A | P | A | P | A | P | A |
| CF gazes at | P | A | P | A | A | A | P | A | P | P |
| CF gazes at | A | P | A | P | A | P | A | P | A | P |
| CF gazes at | A | P | A | P | P | P | A | P | A | A |

timing was roughly consistent across sections and measurements, and the confederate did not make mistakes in when to look at, or away from, the participant. An example of a translated script is given in Appendix A.

The confederate either began a scenario by looking at, or looking away from, the participant. Moreover, the moment at which the confederate stopped talking was scripted such that the confederate either looked at, or away from, the participant. This gaze direction was scripted to occur at the beginning of the final utterance ("I think that's it."). Previous research (Kendon, 1967) has shown that gaze shifts away or toward a conversational partner increase in the last two seconds prior to utterance ending or beginning. Both were counterbalanced between scenarios for each participant. In which scenario the confederate began looking at the participant, and looked at the participant when the confederate stopped talking, was counterbalanced between participants. The possible combinations of confederate gaze behavior across the two scenarios are given in Table 1. On two occasions, the confederate made an error in following the protocol by looking at the participant instead of looking away during the end of turn. As the eye-tracking data recorded are still analyzable with regard to the confederate's gaze direction when he was speaking, these measurements were not excluded.

### 2.4. Procedure

The participant and confederate were welcomed into the laboratory. Up until the debriefing, the confederate was asked to complete the same actions as the participant so as not to give him away as a confederate. The participant and confederate were first asked if there were any remaining questions after reading the information letter. Then informed consent was obtained and a short demographical questionnaire was presented to the participants.

The participant and confederate were positioned at either end of the dual eye-tracking setup, at a distance of 70 cm from the eye tracker and their eyes at the height of the camera positioned behind the one-way mirrors. A 5-point calibration sequence was then started, followed by a 4-point calibration validation. We aimed for a systematic error (validation accuracy) below 1° in both the horizontal and vertical direction (they are returned separately by iViewX). If we could not achieve a sufficiently low systematic error after several attempts, the participant was excluded. This resulted in the exclusion of 3 participants. Achieving sufficiently low systematic offset for the confederate was never a problem. Mean systematic offset for the included participants and the confederates is reported in the *Results*. When the eye trackers were calibrated for both the participant and the confederate, instructions were given about the experiment. This was done in Dutch, but the translated instruction was as follows.

> *Scenario 1*: You are seated in front of a Skype-like video connection. In a moment I will switch on the video connection, and you will get to see each other. I'd like to ask you to tell each other about your work experience. What kind of work have you done? What were your duties? When did you do this work?
>
> [Confederate], you can start this time. As soon as you see [participant], you can start telling your story. As soon as you, [confederate], are done, you, [participant], can start your story. You do not need to wait for a signal from me, as soon as [confederate] is done, you can just start. As soon as [participant] is done, I will stop the video connection. Any questions?
>
> *Scenario 2:* This time I would like to ask you to tell something about your holiday experiences. Where have you been? With whom? And why (if there was a specific reason)?
>
> Let's see. [Confederate], you can start again this time. As soon as you see [participant], you can start telling your story. As soon as you, [confederate], are done, you, [participant], can start your story. Again, you do not need to wait for a signal from me, as soon as [confederate] is done, you can start. As soon as [participant] is done, I will stop the video connection. Any questions?

Eye-tracking data were saved following each scenario. After both scenarios were completed, the operator asked the participant and confederate how they experienced the experiment and whether they noticed anything in particular. Then the participant was debriefed about the role of the confederate and the purpose of the study. Finally, participants were asked if they suspected the confederate's role. Answers to this question were recorded.

### 2.5. Data analysis

#### 2.5.1. Confederate behavior

In order to determine whether the gaze behavior of participants differed as a function of task structure (listening or speaking) and confederate behavior, we needed to know at which periods in time the confederate spoke or listened and looked at or away from the participant. As the timing of the confederate was slightly different for each participant, videos of the participants were coded by the first author for a number of time points. The times at which the confederate switched from looking at or away from the participant was coded by finding the video frame in which the eye movement had just been made and the pupil was visible again. Originally, we wanted to code the start of the gaze shift by recording the frame number at which the pupil starts moving. However, in some cases a blink occurred at the time of the gaze shift, making it impossible to code the onset of the movement of the pupil. Two alternatives were available: (1) coding the start of the pupil movement when it was visible and the start of the blink when the pupil was not visible when the gaze shift started or (2) coding the end of the gaze shift when the pupil was visible again. We chose the latter option for two reasons. First, it meant that we only needed one criterion (when is the pupil visible in its final position?) instead of two (when does the pupil move, or when does the blink start?). Second, by aligning the gaze shifts by their end and assuming that the gaze shifts are approximately the same duration on every occasion (Bahill, Clark, & Stark, 1975), they are also aligned by their start. When a blink occurred, it was not always clear when exactly the gaze shift itself started. The coded time thus marks the completion of the gaze shift. In the scenarios where the confederate looked at the participant during the last section of the confederate's script, and had to look at the participant at the end of turn, there was no eye movement. Here, the beginning of the confederate's last sentence ("that was it"), was coded, as this co-occurred with the gaze shift away from the participant in the scenarios when the confederate did have to change gaze direction.

#### 2.5.2. Eye-tracking data reduction

The eye trackers report a time series for each participant with a gaze coordinate each 8.33 ms. In order make this time series analyzable in terms of periods of looking at a facial feature (termed dwells), the eye-tracking data were reduced by the following procedure:

1. Using automatic software in MATLAB, eye-tracking data were automatically trimmed to the start and the end of the video based on

the timestamps from the computer controlling the videos stream.

2. We wanted to know where and when the participant gazed on the face of the confederate. However, as the confederate talked and moved slightly during the conversation, the face does not remain in a fixed location in the video. Moreover, the face rotates and slightly deforms during the conversation. To avoid manual coding, we used a method that automatically constructs Areas-of-Interest (AOIs) for the left eye, right eye, nose and mouth. The method is based on the OpenFace toolbox (Baltrušaitis, Robinson, & Morency, 2013; Baltrušaitis, Robinson, & Morency, 2016) and the Limited-Radius Voronoi tesselation (LRVT) method (Hessels, Kemner, van den Boomen, & Hooge, 2016). The method is validated extensively and described in Hessels, Benjamins, Cornelissen, and Hooge (2018), and is at least as good as manual construction of AOIs. The radius of the LRVT method specifies how large the AOIs will be. A radius of 2°, for example, will mean that a gaze coordinate is assigned to the closest AOI, but only if it does not exceed a 2° distance to the AOI center. In previous work (Hessels et al., 2016), we have shown that a large radius results in measures of gaze behavior that are most robust to noise. As such, the radius was set to 4.0°. The average distance between the confederate's eyes was around 3.03° and the AOI span (Hessels et al., 2016) was 1.96°. An example of the AOIs for two frames of one video is given in Fig. 2.
3. As we were interested in the allocation of gaze to the facial features, and the visual information presented on each screen was only updated with 30 Hz, eye-tracking data were downsampled to 30 Hz (the refresh rate of the video). This was done by averaging the 2 preceding and following samples thereby increasing the signal-to-noise ratio by $\sqrt{4}$. Each sample was subsequently assigned to one of the AOIs. When a gaze position was reported by the eye tracker, but was not on any of the four facial AOIs, it was assigned to the "other" AOI. As visible from Fig. 2, the "other" AOI encompasses the background, the upper body of the confederate, and a small part of the top of the head. The "other" AOI is boundless: looking outside the screen will be counted as "other" as well, although the eye tracker cannot track far outside of the screen.
4. Dwells were defined as a minimum of four consecutive video frames (120 ms) in which the gaze position was on the same AOI.

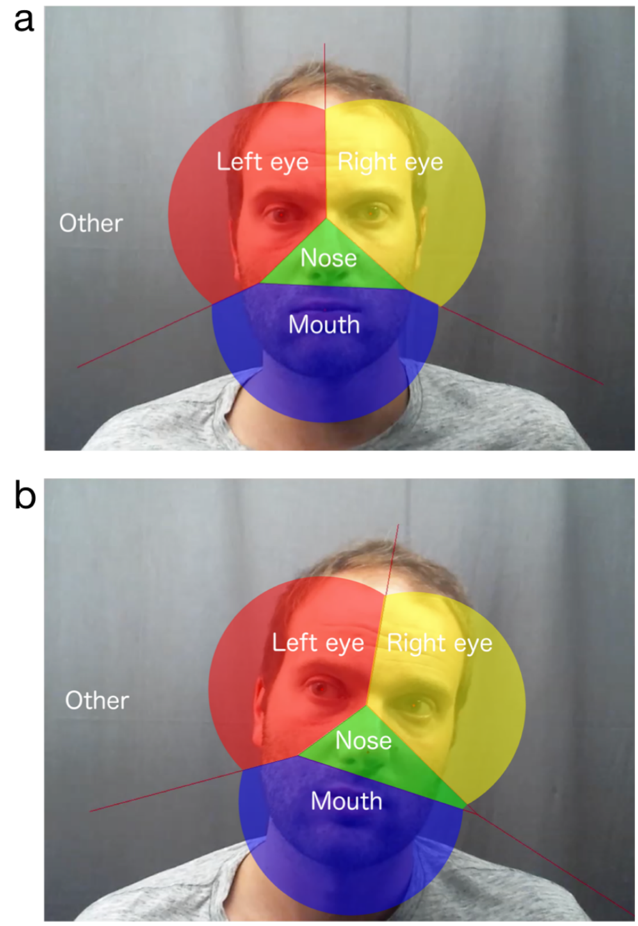


**Fig. 2.** Example of computer-generated Areas of Interest (AOIs) for two frames from one video of the confederate. AOIs for the left eye, right eye, nose and mouth are given in color, and are based on the Voronoi-tesselation method with a bounding radius. These colors are used throughout the data analysis when depicting eye-tracking measures from the AOIs. The Limited-radius Voronoi-tesselation (LRVT) method divides the entire area into regions that are closest to the center of the facial features. The area for the left eye, for example, is the area that is closest to the left eye, but does not exceed the LRVT-radius of 4° used in this study. (a) Example when the confederate is upright and looking straight ahead. (b) Example when the confederate's head is tilted and looking away from the participant. These examples serve to show that our automatic AOI technique was capable of tracking the confederate's face through varying head poses and expressions. Note that the image is mirrored with regard to how it was presented, such that the "left" and "right" eye-AOIs actually correspond neatly to the left and right eyes of the confederate. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

*2.5.3. Statistical analysis of gaze behavior*

Gaze behavior of the participants to the confederate's face was analyzed in two ways, using aggregate measures of gaze behavior and time-dependent measures. For the aggregate measures, total dwell time, dwell time, and number of dwells to the left eye, right eye, nose, mouth and other-AOIs were computed using MATLAB R2017a. We use the terminology and definitions by Holmqvist et al. (2011). Total dwell time was calculated as a relative measure by determining the proportion of time spent looking at an AOI in an episode (e.g. while the participant listened or spoke). Dwell time was calculated as the average time spent looking at an AOI from entry to exit in milliseconds. The number of dwells was calculated by dividing the total number of dwells to an AOI by the duration of the episode in seconds, yielding a number of dwells per second which is irrespective of the duration of the episode. These measures were (1) computed separately for when the participant was listening or speaking and (2) computed for the periods when the confederate looked at the participant or away from the participant (note, this only included the portion of the scenario when the confederate spoke).

Linear mixed-effects models (LMEMs) were used to analyze the data, and were implemented using the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2017). We opted for LMEMs, as they are particularly suited for analyzing data with repeated measurements and are superior compared to, for example, ANOVAs when dealing with unbalanced and missing data (Baayen, Davidson, & Bates, 2008). For the average dwell time measure, for example, we knew beforehand that missing data was going to be an issue. If a participant does not look at the mouth during one of the sections of the confederate's script, we cannot compute an average dwell time measure for that AOI. In standard statistical analyses, this would lead to the exclusion of that entire participant from the analyses. LMEMs were constructed for each measure using AOI (left eye, right eye, nose, mouth, other) and state (participant listening vs. speaking, or the confederate looking at vs away from the participant) as well as the interaction between state and AOI as predictors. For the total dwell time models, "data loss" was added as an AOI. This encompasses the relative time during which no gaze coordinate could be reported, e.g. when a speaker turns away from the screen. All predictors were dummy coded: The reference level was set to the left eye for the AOI predictor and to the 'participant listening'-episode or the 'confederate looking at the participant'-episode for the State predictor. Models were constructed with random intercepts only, and with random slopes for AOI by participant. As the more complex models did not yield additional insights, the models with random intercepts only were further analyzed. p-Values for all fixed effects are reported using Satterthwaite approximations to degrees of freedom in the lme4-package in R. The alpha level was set at 0.05.

Aggregate measures of gaze behavior are useful when capturing changes in biases of looking at facial features as a function of the behavior of the participant or confederate. However, it may be the case that there are no 1-to-1 relations between the gaze allocation of a participant and the gaze behavior of a confederate, but that there are higher-order patterns in the gaze allocation of the participant group. Is it the case, for example, that a shift of gaze away from the participants' face by the confederate predicts the gaze direction of the participants at a short time scale?

The interaction setup that we employ allows us to compute such measures. As there are large differences in the bias that participants have in looking at the eyes, nose or mouth (e.g. Arizpe, Walsh, Yovel, & Baker, 2017; Kanan, Bseiso, Ray, Hsiao, & Cottrell, 2015; Mehoudar, Arizpe, Baker, & Yovel, 2014; Peterson & Eckstein, 2013), we wanted a measure that can accurately captures changes in the distribution of the participant groups' gaze location on the face of the confederate. A prime candidate for such a measure is entropy (Shannon, 1948), which is a measure of disorder and has previously been applied to eye-tracking data (e.g. Allsop & Gray, 2014; Hessels et al., 2017a; Hooge & Camps, 2013). Entropy can be used to quantify the nature of the spatial distribution of gaze allocation of the participants. If the spatial distribution is heterogeneous, entropy is low. If the spatial distribution is homogeneous, entropy is maximal. In this study, this translates into the following: If all participants gaze at one location only (e.g. the eyes or the mouth), entropy is low. If, on the other hand, participants' gaze is equally divided over the AOIs, entropy (or disorder) is high, and the gaze behavior of the participants is different from one another. Such a measure can be calculated over time, and time-locked to the gaze shifts by the confederate. Entropy was calculated as follows:

1. Dwells to the left eye and right eye were combined into an "eyes" AOI, as we were not interested in whether participants looked at the left or right eye when they did so in response to the confederate's behavior.
2. Dwells were time-locked to the start of (1) the periods during which the confederate gazed at the participant's face and (2) the periods during which the confederate gazed away from the participant.
3. Entropy was computed for ± 10 s around the event with 33 ms interval using the equation:

$$-\sum_{i=1}^{n} p(i)\log_2 p(i) \tag{1}$$

   where $p(i)$ is the proportion of recorded dwells on the $i$th AOI, and $n$ is the number of AOIs (4 in this case: eyes, nose, mouth, and other).
4. As the absolute values of entropy are dependent on the average bias in gaze location of the participant group, a differential signal was computed by subtracting the entropy signals of the confederate looking at or away from the participant from each other. This differential signal captures the relative strength of gaze guidance by the confederate's gaze direction at its onset.
5. The mean and standard deviation of the differential signal in the entire ± 10 s range were calculated as a benchmark. We will investigate whether the differential entropy exceeds the mean value around the time of the gaze shift by the confederate.

## 3. Results experiment 1

None of the participants reported that they suspected the conversational partner to be a confederate. As such, all measurements were considered for further analysis.

### 3.1. Eye-tracking data quality

Two measures of eye-tracking data quality were computed: the systematic error (or validation accuracy) and the proportion of lost data. Data loss may occur when, for example, participants blink or look away from the screen. However, it can also occur due to technical difficulties in tracking the participants' eyes (e.g. Hessels, Andersson, Hooge, Nyström, & Kemner, 2015). The systematic error[2] was on average 0.58° ($sd$ = 0.18°) for the confederate, and 0.70° ($sd$ = 0.24°) for the participants. Proportion of time without eye-tracking data (data loss) for participants while they were listening was 0.03 on average ($sd$= 0.03). As the proportion of data loss did not exceed 0.13, no participants or scenarios were excluded. We did not consider the proportion of data loss when participants spoke as an exclusion criteria, as we expect people to look away from the eye tracker when speaking, and thereby causing data loss not related to technical difficulties in tracking gaze. Values for the systematic error were well below the inter-AOI distance and data loss was low. The quality of the eye-tracking data was therefore considered to be adequate to allow further analysis.

### 3.2. Aggregate measures of gaze behavior

In order to test whether participants[3] looked less at the facial features when speaking compared to listening, we compared the total dwell time (how long do participants look at the facial features overall as the proportion of time), mean dwell time (how long does a visit to a facial feature last), and number of dwells per second (how often do participants look at a facial feature) between period of speaking and listening using linear mixed-effects models. Estimates and significance of the fixed effects of the six LMEM are given in Table 2.

The aggregate gaze measures are depicted in Fig. 3. In accordance with the generally accepted results, we found that not every facial feature was looked at equally often, nor equally long (see Table 2). When participants listened, they looked at the facial features (left eye, right eye, nose, and mouth) of the confederate for a much longer total duration (panel A), for a longer average duration (panel B), and more often per second (panel C), as opposed to when participants' spoke. When participants spoke, they looked longer at the other-AOI (encompassing all areas on and off screen without the face), and away from the screen completely (as evidenced from the "data loss" category). This was confirmed statistically by the significant State by AOI interaction for the speak/listen models. The fact that the estimates for the State × right eye, nose and mouth AOIs are non-significant (and very small), indicates that the relative pattern of which facial features are preferentially fixated was preserved when participants spoke compared to when they listened. Participants looked more often and for a longer total duration to the eyes and nose than to the mouth when they did look at the confederate's face (see also Supplementary Figs. 2 and 3, which depict total dwell time for the AOIs calculated as the proportion of all dwells, and all dwells on the face, respectively). As visible from panels D-F, and confirmed by the statistical results, the aggregate measures did not differ significantly between the confederate looking at or away from the participant. In other words, aggregate measures of participants' gaze behavior were dependent on whether the participants were listening or speaking, but not dependent on whether the confederate looked at or away from the participants.

### 3.3. Entropy of participants' gaze location

The entropy of the participants' gaze location time-locked to the

[2] The 2-d systematic error was calculated using Pythagoras' theorem on the validation accuracy as reported by the SMI iViewX software in the horizontal and vertical direction.

[3] see Supplementary Fig. 1 for the confederate's gaze behavior while listening

**Table 2**
Estimates for the predictors in the linear mixed-effect models on participants' gaze behavior as a function of participant listening/speaking, or the confederate looking at or away from the participant. The estimates are relative to the left eye Area of Interest (AOI) when the participant was listening or when the confederate looked at the participant. Standard errors for the estimates are given in parentheses. TDT = total dwell time (proportion of episode), DT = dwell time (ms), ND = number of dwells per second. Significance is indicated by: *** < .001, ** < .01, * < .05 using Satterthwaite approximations to degrees of freedom in the lme4-package in R.

| Fixed effects | Model | | | | | |
|---|---|---|---|---|---|---|
| | Listen/Speak TDT | Listen/Speak DT | Listen/Speak ND | At/Away TDT | At/Away DT | At/Away ND |
| Intercept | 0.26 (0.02)*** | 514.05 (27.89)*** | 0.51 (0.03)*** | 0.25 (0.01)*** | 515.38 (36.64)*** | 0.49 (0.03)*** |
| State | −0.12 (0.02)*** | −164.81 (35.37)*** | −0.14 (0.04)*** | 0.03 (0.02) | −19.93 (42.02) | 0.03 (0.03) |
| Right eye AOI | −0.06 (0.02)* | −60.22 (35.04) | −0.08 (0.04) | −0.04 (0.02)* | −43.20 (42.53) | −0.06 (0.03) |
| Nose AOI | −0.04 (0.02) | −119.87 (35.04)*** | 0.02 (0.04) | −0.02 (0.02) | −128.51 (41.77)** | 0.06 (0.03) |
| Mouth AOI | −0.12 (0.02)*** | −57.99 (35.20) | −0.22 (0.04)*** | −0.11 (0.02)*** | −76.26 (42.87) | −0.22 (0.03)*** |
| Other AOI | −0.22 (0.02)*** | −176.40 (36.63)*** | −0.41 (0.04)*** | −0.21 (0.02)*** | −153.90 (48.07)** | −0.39 (0.03)*** |
| Data loss | −0.23 (0.02)*** | | | −0.21 (0.02)*** | | |
| State × Right eye AOI | 0.03 (0.04) | 35.50 (49.89) | 0.02 (0.06) | −0.04 (0.03) | −7.36 (59.76) | −0.03 (0.05) |
| State × Nose AOI | 0.03 (0.04) | 88.69 (49.79) | −0.05 (0.06) | −0.03 (0.03) | 40.43 (58.90) | −0.07 (0.05) |
| State × Mouth AOI | 0.04 (0.04) | 22.84 (50.22) | 0.02 (0.06) | −0.02 (0.03) | 45.01 (60.00) | −0.00 (0.05) |
| State × Other AOI | 0.29 (0.04)*** | 365.18 (50.91)*** | 0.40 (0.06)*** | −0.03 (0.03) | 12.01 (67.28) | −0.04 (0.05) |
| State × Data loss | 0.34 (0.04)*** | | | −0.04 (0.03) | | |

onset of the confederate's gaze direction toward or away from the participant is depicted in Fig. 4. As visible from the black line, there is a large peak in the differential signal around the time of (and slightly before) the onset of the confederate's gaze direction towards or away from the participant. The entropy is lower when the confederate looks away compared to when the confederate looks towards the participant. The maximum difference is >3.4 standard deviations away from the mean entropy difference. Moreover, a peak exceeding ± 3 standard deviations does not occur in the entire −10 to +10 s relative to the episode onset during which the entropy was calculated.

The fact that entropy is much lower when the confederate looks away from compared to towards the participant, is a result of the underlying distribution of participants' gaze over the eyes, nose, mouth and other-AOI, which we decided to investigate more closely. Fig. 5 depicts the proportion of dwells on each AOI relative to the onset of confederate's gaze direction toward or away from the participant. When the confederate looked towards the participant, there was a dip in the proportion of dwells to the eyes, whereas a peak is observed when the confederate looked away from the participant. Moreover, the reversed pattern is observed for the other-AOI; a peak when the confederate looked at the participant, versus a dip when the confederate looked away from the participant.

## 4. Discussion experiment 1

The present study was meant to assess (1) whether gaze behavior to the face and facial features of a confederate was dependent on the task for the participant – i.e. whether the participant had to speak or listen, and (2) whether the gaze behavior of a confederate affected participants' gaze behavior. For this, a confederate and a participant engaged in conversation during which both had to speak for a period of time, and listen for a period of time, while their eye movements were monitored using a state-of-the-art dual eye-tracking setup.

As hypothesized, participants looked more often, for longer durations, and for a longer total duration at the facial features (eyes, nose, and mouth) when they listened compared to when they spoke. Even though participants looked less often at the face of the confederate when they spoke (and for shorter durations), we observed only small changes in the distribution of gaze over the facial features when participants did look at the face. These changes were a slight decrease in proportion of looking time to the mouth when participants spoke as compared to when participants listening, and vice versa for the eyes. The changes were on the order of a few percent in total looking time, consistent with earlier reports (cf. Rogers, Speelman, Guidetti, & Longmuir, 2018). This indicates that gaze allocation in face-to-face communication is task-dependent, and suggests that gaze allocation to the faces of others should always be considered in the context of the specific task being carried out.

The confederate's gaze behavior (i.e. either looking at or away from the participants) did not affect aggregate measures of participants' gaze behavior. This means that there were no changes in the overall biases to look at the eyes, nose, or mouth when the confederate looked at, or away from, the participants. However, we observed a large difference in the entropy of the distribution of participants' gaze location around the time when the confederate looked towards or away from the participant. When the confederate looked toward the participants entropy was near its maximum, meaning that the differences between the participants' gaze allocation were nearly as great as they could possibly be. When the confederate looked away from the participants, the gaze allocation of the participants was more similar. This could be seen as the confederate's gaze shift away from the participant guiding the participants' gaze direction. In terms of the underlying gaze location of the participants, more participants looked at the eyes and less at the other-AOI when the confederate looked away versus when the confederate looked towards the participant.

Note that the peak in entropy seems to start before the change in gaze direction, which seems counterintuitive if the change in gaze direction (i.e. away from or toward the participant) differentially affects participants' gaze behavior. As stated, however, the gaze direction of the confederate was coded at the moment when the pupil was visible again, as in some cases blinks masked the onset of the gaze shifts. When no blink occurred, or when the blink occurred halfway through the eye movement, some rotation of the eye was visible even though the pupil was not yet visible, which may have been usable information for the participants. It may also be the case, that there were additional cues available to the participants that accompanied the gaze shift and signaled an upcoming end to a sentence. Such cues may have been linguistic, e.g. prosody (Cutler & Pearson, 1986), or non-linguistic, e.g. head movement (Maynard, 1987). Note, however, that we obtain a differential entropy signal here, which reflects the relative change in entropy when the confederate looked away from versus toward the participant. Assuming that any additional cues are available to the participant regardless of confederate gaze direction, the differential entropy measure should not be affected by them.

Several explanations might be posited for the observed differences in the distribution of participants' gaze location as a result of the confederate looking at, or away from, the participant. Researchers from an experimental background or those familiar with saliency-based models of gaze allocation might assert that the observed differences are stimulus-driven. Researchers studying the task control of eye movements

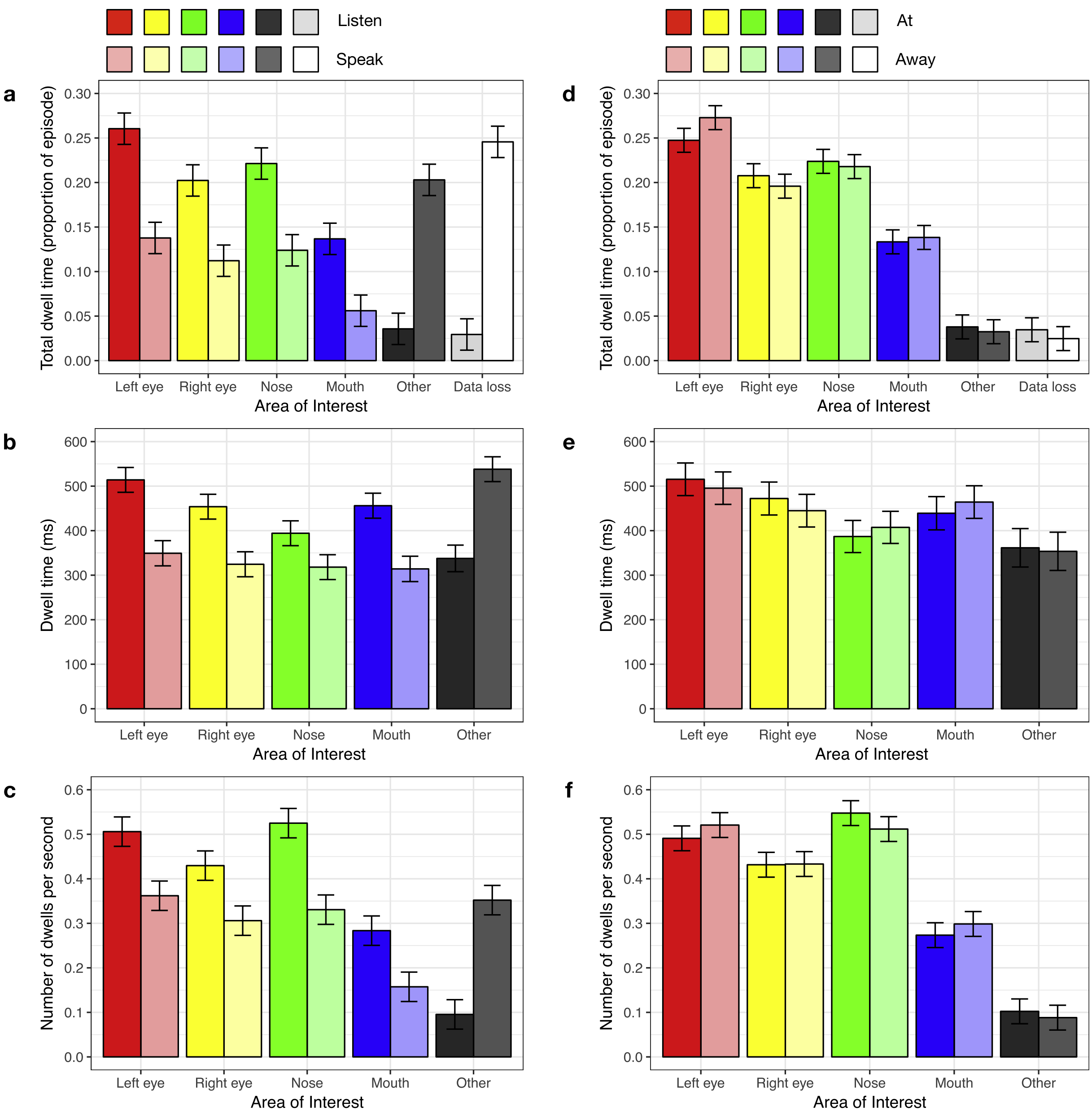


**Fig. 3.** Aggregate measures of participants' gaze behavior (total dwell time as a proportion of the episode, dwell time in milliseconds, and number of dwells per second) while the participant listened or spoke (panels A-C) and when the confederate looked at or away from the participant (panels D-F). Error bars indicate the standard error of the linear mixed-effects models' intercept.

might assert that the observed differences are task-driven, and researchers from the social attention literature may assert that they are driven by the social context. We discuss each possible explanation in turn.

With stimulus-driven, we mean that the confederate looking away from the participants guides the participants' gaze direction more than the confederate looking toward the participant, irrespective of any specific task structure or contextual factors. In other words, the change in gaze direction of the confederate in itself causes more participants to look toward the eye region. Although previous research has shown that task demands can influence gaze behavior to the faces of others (Itier et al., 2007; Nummenmaa et al., 2009), it is important to consider whether stimulus factors alone could have influenced the participants' gaze behavior, or whether they occur in a specific task-structure or social context.

With task-driven, we mean that the observed differences in the participants' gaze distributions are specific to the task posed upon the participants. In this case, each participant had to monitor the confederate for when it was the participant's turn to start speaking. A speaker looking away from someone during a conversation is more likely to signal that the speaker intends to preserve their turn to speak (Argyle, 1972; Ho et al., 2015). A speaker looking toward someone while the end of a sentence is uttered is more likely to signal that it is the other's turn to speak. We found that when the confederate looked away, the distribution of participants' gaze location was less uniform over the AOIs than when the confederate looked towards the participant. This might have been because the confederate's gaze direction away from the participant sends a clear signal that it remains the confederate's turn to speak, which makes it unnecessary for the participants to figure out whether the confederate's turn is about to end by, for instance, looking at the mouth and monitoring the end of a sentence.

With social context-driven, we mean that the important factor is the presence of an interacting partner in itself (regardless of the specific task structure). As many authors have stated (e.g. Argyle & Cook, 1976;

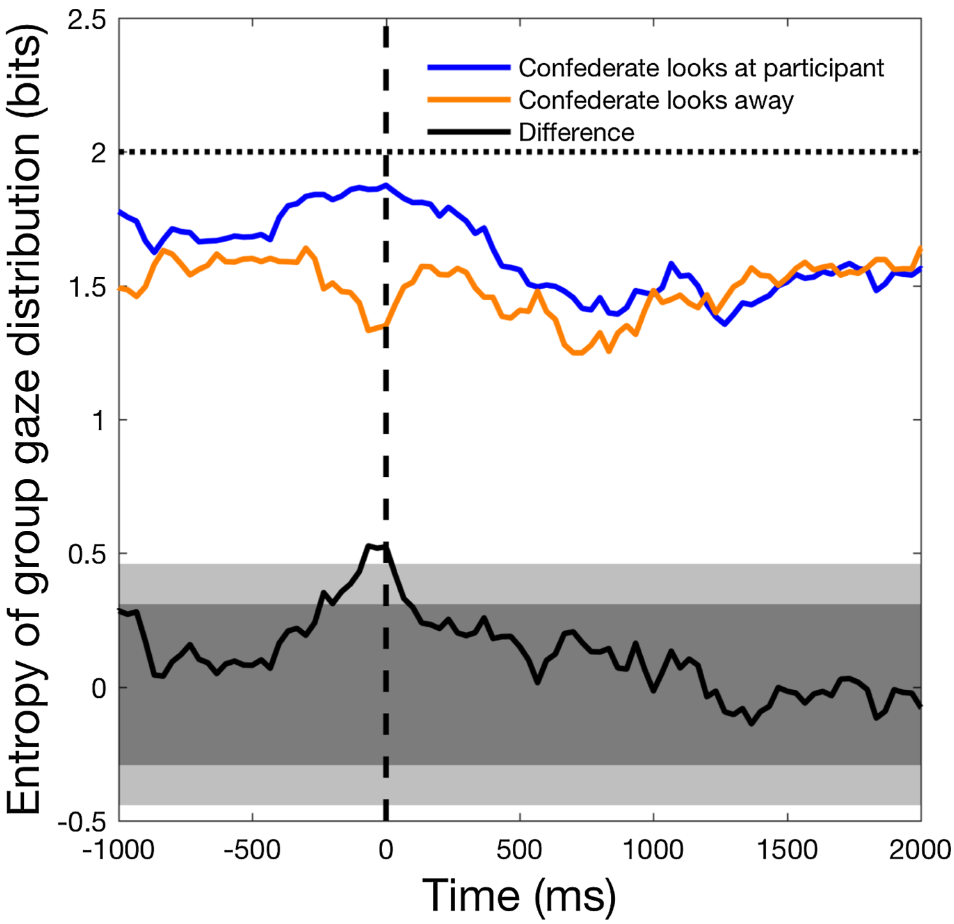


**Fig. 4.** Entropy of the distribution of participants' gaze location, time-locked to the start of an episode of the confederate looking at or away from the participant in experiment 1. The dashed line indicates the start of the episode. The dotted line indicates the maximum entropy for four Areas of Interest (eyes, nose, mouth, and other). The black line indicates the entropy difference between the two confederate behaviors. As it is a differential signal, it can be negative. The dark grey block indicates ± 2 standard deviations around the mean entropy difference calculated from −10 to +10 s around the episode onset. The light grey block indicates ± 3 standard deviations around the mean entropy difference.

Gobel et al., 2015; Risko et al., 2016), gaze has a "dual function" during face-to-face communication: gaze allocation may be for visual information uptake, but may also signal information to an interacting partner. Indeed, it has long been known in biology that the direction of one's gaze (e.g. whether eye contact is made or not) is a crucial component in regulating social interaction, for example in primates (Emery, 2000; de Waal, 2009). Furthermore, Senju and Johnson (2009) point out that the perception of eye contact in humans modulates many psychological processes. In the present study, the confederate either looked directly at the participants or averted his gaze. It might be that the confederate's direct gaze evoked the perception of eye contact in the participants – eye contact being defined as the subjective experience of a person looking directly into one's eyes. This might have caused some participants to avoid looking at the eyes and thereby causing the peak in the entropy signal. One reason why not all participants need necessarily do this, is that some may be more socially anxious than others, which we have previously linked to the maintenance and avoidance of looking at the eyes in face-to-face interaction (Hessels, Holleman, Cornelissen, Hooge, & Kemner, 2018). Why this would then only occur on a short time scale and is not visible in the aggregate measures of gaze behavior is, however, unclear. The confederate always maintained direct or averted gaze direction for a longer period. Moreover, we do not know when participants might have perceived eye contact and for how long that experience lasted, as perceived eye contact need not necessarily coincide with being looked at in the eyes (Honma, Tanaka, Osada, & Kuriyama, 2012).

Differentiating between the task- and social context-driven explanations is not trivial, particularly as the communicative role of gaze behavior in face-to-face communication has received little attention in previous eye-tracking research. We return to this in the *General Discussion*. However, whether the observed entropy peak is stimulus-driven – caused primarily by the gaze shift away from the participant in itself – can be investigated. It is particularly important to rule out a stimulus-driven account, if we are to understand how gaze behavior is affected by social contextual and task-related factors. Before attributions to particular instances of gaze behavior are made (e.g. to a communicative role), we need to be sure they cannot be attributed to stimulus-factors. We therefore conducted a second experiment in which videos were recorded of a confederate acting out their part of the scenarios from experiment 1. Participants were recruited to observe these videos, without it having been posed as a task of conversing with the confederate. Should the same effect of the confederate's gaze direction on the participants' distribution of gaze over the facial features be observed, it is likely that the eye movement in itself guides participants' gaze direction. In other words, the change in the participants' distribution of gaze direction would be stimulus-driven.

## 5. Methods experiment 2

### 5.1. Participants

Participants were recruited during a programming course in the Psychology bachelor at Utrecht University. 65 students (37 female, 27 male) participated in the study. The gender of one participant was not recorded, as the participant did not fill in the demographic questionnaire. Additionally, four participants filled in 2017 as their birth year, causing us to be unable to compute an accurate age for five participants. Mean age of the other 60 participants was 21.8 years ($sd$ = 3.38 years). One participant was excluded due to severe pendular nystagmus. This participant was allowed to complete the experiment for his own interest. The experiment was conducted in accordance with the Declaration of Helsinki. Participation was voluntary, took between 5

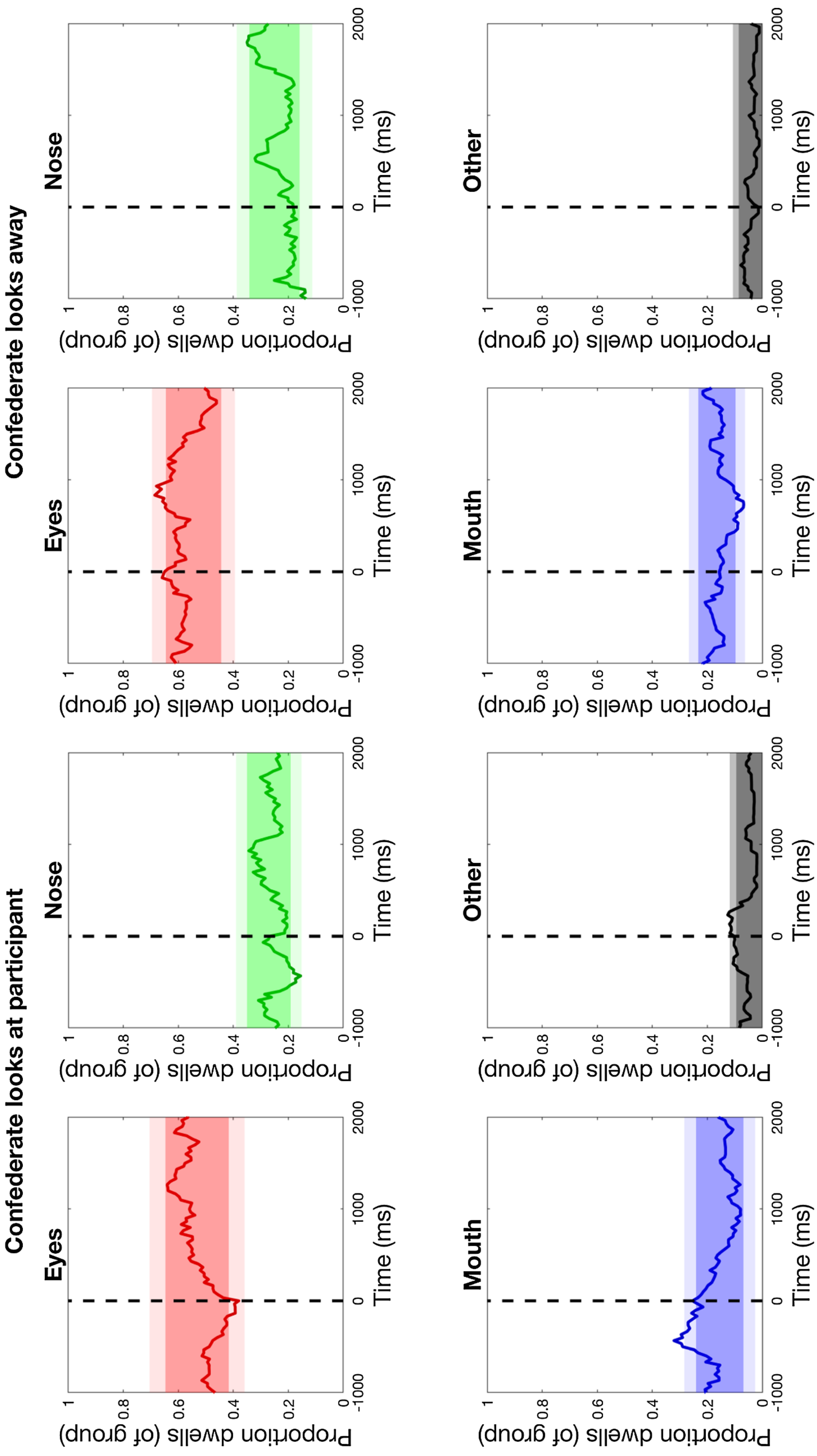


**Fig. 5.** Proportion of recorded dwells from to the eyes, nose, mouth and other Areas of Interest, time-locked to the start of an episode of the confederate looking at or away from the participant in experiment 1. Left panels indicate proportion of dwells for when the confederate looked towards the participant, right panels when the confederate looked away from the participant. The dashed line indicates the start of the episode. The dark colored blocks indicate ± 2 standard deviations around the mean proportion of dwells for that AOI calculated from −10 to +10 s around the episode onset. The light colored blocks indicate ± 3 standard deviations around the mean proportion of dwells.

and 10 min, and participants were compensated by receiving 0.5 participant hours. Participants were informed about the study by one of the operators, after which written informed consent was obtained.

### 5.2. Apparatus and stimuli

Eye movements were recorded using a Tobii TX300, placed in a moveable booth (see Fig. 6). A chin rest was mounted on the end of the booth, in order to control participants' head movement and improve

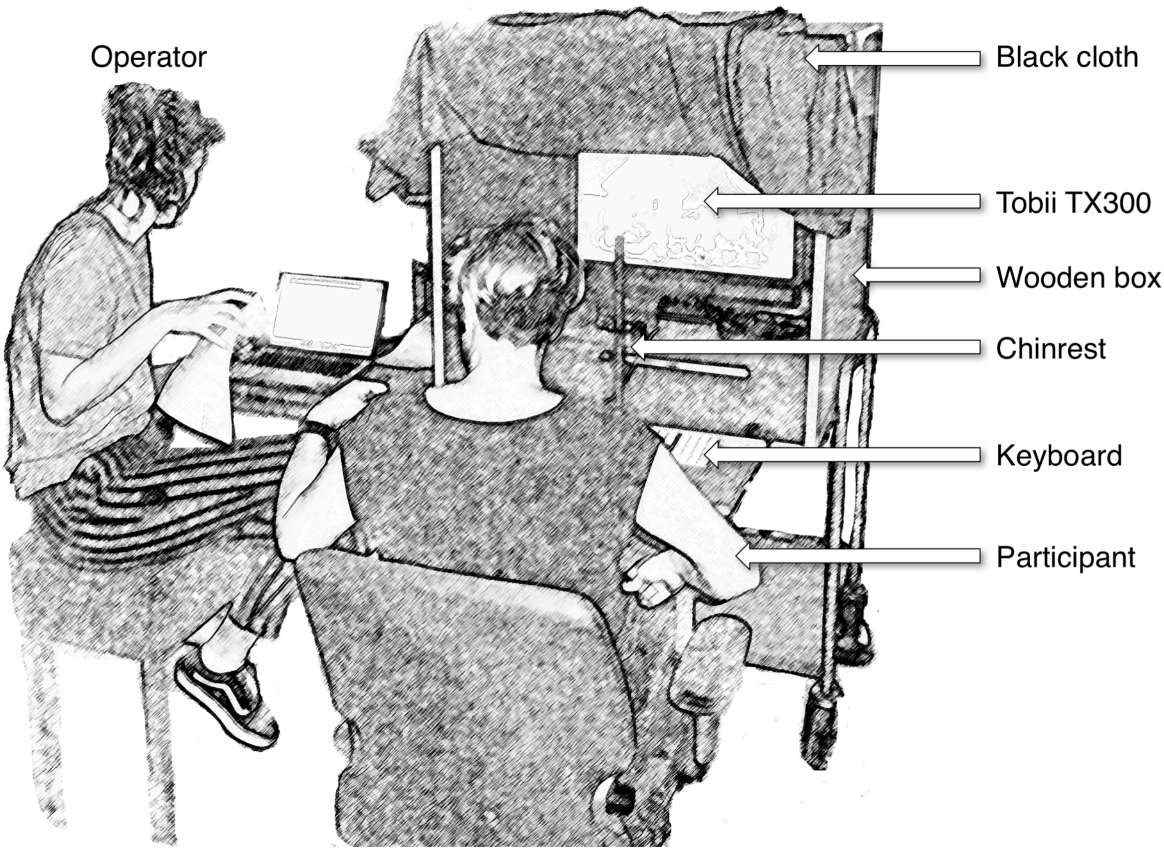


**Fig. 6.** Tobii TX300 placed in a moveable booth. The participant was positioned in the chin rest, while the operator sat on the side initiating calibration and monitoring the measurement. A black curtain was drawn over the participants to ensure identical lighting conditions. Reprinted from Hooge et al. (2018) with permission.

eye-tracking data quality (Niehorster, Cornelissen, Holmqvist, Hooge, & Hessels, 2017). A black curtain was drawn around the participant to ensure equal lighting conditions across participants. A MacBook Pro (2.8 GHz i7 processor, OS X 10.9.5) running MATLAB R2012b and the Tobii SDK was used for stimulus presentation and communication with the eye tracker. Stimulus presentation was done on a 23″, 1920 by 1080 pixels, computer screen attached to the Tobii TX300.

Each participant received two videos from one of two possible confederates. Confederate 1 was a 26 year-old male, the same confederate as in experiment 1 and the second author. Confederate 2 was a 20 year-old female. Two confederates were included, such that we could ascertain whether the effects of the confederate's gaze shift was specific to our initial confederate, or generalizes across confederates (e.g. male versus female confederates). Note that this is only applicable, should the same effect be observed as in experiment 1. For both confederates, video recordings were made of all the possible combinations of looking at, or away from, the participant at the beginning of the experiment, and when the confederate stopped talking (see Table 1). Videos were recorded by having the confederate carry out their script while looking at the first author via a one-way mirror. In this manner, the story is told to another person and not to a camera. We conjectured that this would lead to a conversation-like viewing pattern of the confederate when looking at the camera. A Nikon D3200 DSLR-camera was placed behind this mirror, and recorded the confederate at 1280 by 720 pixels. Videos were trimmed to the start and end of the script, and cropped to 960 by 720 pixels. There were presented at 1024 by 768 pixels in the center of the computer screen. The duration of the videos was between 1:09 and 1:40 min. Two example videos are available from https://osf.io/k32vz/.

### 5.3. Procedure

Participants were first informed about the study, after which they gave written informed consent and filled in a short demographic questionnaire. Then participants were seated on an office chair in front of the booth containing the eye tracker. Participants positioned themselves in the chin rest, which was placed such that the distance between the participants' eyes and the eye tracker was about 65 cm, the optimal tracking distance for the Tobii TX300. Then a 9-point calibration procedure was performed. Participants were presented with an instruction screen stating: "You are about to view a videoclip of [confederate]. [Confederate] has been asked to tell something about his/her work experience". A spacebar press initiated the first video. After the first video was finished, a second instruction screen was presented stating "You are about to view a videoclip of [confederate]. [Confederate] has been asked to tell something about where (s) he has been on holiday". A spacebar press again initiated the video. After the second video was finished, a 9-point validation sequence was presented.

### 5.4. Data analysis

#### 5.4.1. Confederate behavior

Confederate behavior was coded as in experiment 1.

#### 5.4.2. Eye-tracking data reduction

Eye-tracking data were reduced to dwells by the following procedure:

1. As the automatic AOI-construction method used in experiment 1 had trouble with detecting the position of the nose, due to the lighting conditions in the videos in experiment 2, we applied the semi-

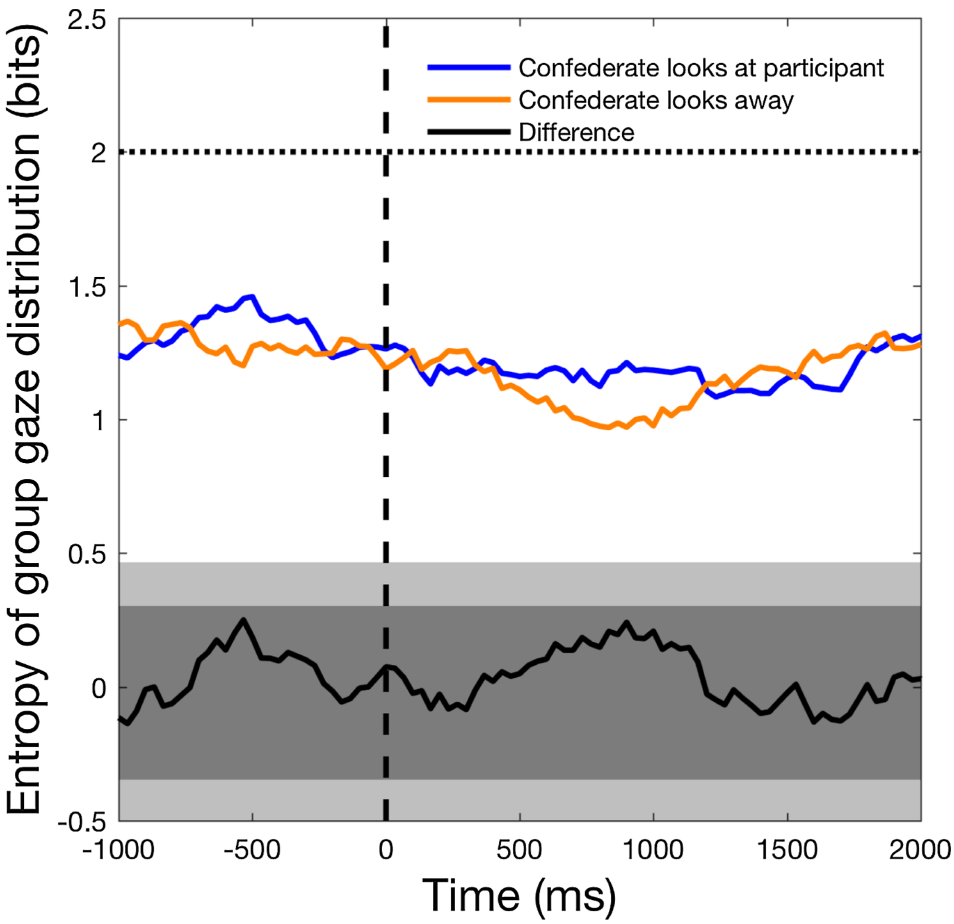


**Fig. 7.** Entropy of the distribution of participants' gaze location, time-locked to the start of an episode of the confederate looking at or away from the participant in experiment 2. The dashed line indicates the start of the episode. The dotted line indicates the maximum entropy for four Areas of Interest (eyes, nose, mouth, and other). The black line indicates the entropy difference between the two confederate behavior. As it is a differential signal, it can be negative. The dark grey block indicates ± 2 standard deviations around the mean entropy difference calculated from −10 to +10 s around the episode onset. The light grey block indicates ± 3 standard deviations around the mean entropy difference.

automatic method used by Hessels et al. (2017a) which has been shown to perform as well as the fully-automatic method used in experiment 1 (Hessels, Benjamins, et al., 2018). As in experiment 1, AOIs were constructed for the left eye, right eye, nose, and mouth. The radius for the LRVT method was set to 4.0°. The average distance between the confederate's eyes was around 3.55° and the AOI span (Hessels et al., 2016) was 2.20°.

2. As an eye tracker was used with a higher sampling rate compared to experiment 1, we first classified periods during which the gaze was directed at roughly the same location as fixations. This was done using the Identification by 2-Means Clustering algorithms by Hessels, Niehorster, Kemner, and Hooge (2017b). This algorithm has been shown to be most robust to variations in variable error (i.e. precision) that may occur in the eye-tracking data.
3. Fixations were subsequently assigned an AOI, using the AOI positions at fixation onset. When a gaze position was recorded, but was not on any of the four facial AOIs, it was assigned to the "other" AOI.
4. Dwells were defined as consecutive fixations on the same AOI.

*5.4.3. Statistical analysis of gaze behavior*

Aggregate measures of gaze behavior and the entropy of group gaze behavior were computed as in experiment 1. As participants were only required to listen to the confederate and not speak themselves, a comparison between participant speaking and listening was not possible.

## 6. Results experiment 2

### *6.1. Eye-tracking data quality*

Mean systematic error was 0.7°. Mean variable error was 0.28° (*sd* = 0.11°). Mean proportion of data loss was 0.07 (*sd*= 0.05). Data were excluded on a scenario basis, if the variable error exceeded the mean plus 2 standard deviations or proportion of data loss exceeded 0.25. This led to the exclusion of eye-tracking data from 17 scenarios; for 7 participants eye-tracking data from 2 scenarios were excluded (i.e. the entire measurement), for 3 participants eye-tracking data from only 1 scenario were excluded.

### *6.2. Aggregate measures of gaze behavior*

LMEMs for the effect of the confederate looking toward or away from the participant on total dwell time (proportion of episode), dwell time (ms), and number of dwells per second yielded the same overall results as in experiment 1 (see Supplementary Fig. 4). These results are therefore not discussed any further.

### *6.3. Entropy of participants' gaze location*

The entropy of the participants' gaze location time-locked to the onset of the confederate's gaze direction toward or away from the participant is depicted in Fig. 7. As opposed to experiment 1, there is no peak in the difference in the entropy signal between when the confederate looked toward or away from the participant. This was the case both when the confederates were investigated separately and when they were combined. Note however, that the overall entropy is slightly lower in this experiment, as more participants were likely to look at the eyes compared with experiment 1.

## 7. General discussion

In two experiments, we investigated which factors may be relevant for understanding the allocation of gaze to facial features in face-to-face interaction. We investigated the role of task structure and gaze behavior of an interacting partner on the gaze allocation of an agent. In the first experiment, we engaged participants and a confederate in a conversational setting in a state-of-the-art dual eye-tracking interaction setup. This setup is unique in that it allows us to investigate gaze to facial features of two individuals, as well as their coupling, at a high temporal and spatial resolution. Here we have shown (1) that gaze behavior to facial features depends on the sub-task being carried out (speaking vs. listening), and (2) that the location of gaze to facial features is more similar between participants within 500 ms of the confederate looking away rather than toward the participants. We interpret this as the confederate's gaze shift away from the participants guiding gaze more than a gaze shift toward the participants.

In the second experiment we investigated whether the gaze-guidance effect of a confederate's gaze shift depends on the social context and task structure (i.e. the fact that a participant and confederate are engaged in face-to-face conversation), or reflects stimulus-driven factors. We find that the gaze shift of a confederate, without the conversational setting, does not guide the gaze of participants. Stimulus factors alone – specifically, a change in gaze direction toward or away from the participant – are unlikely to have produced the observed effects in experiment 1. These findings support the proposition (e.g. Risko et al., 2016) that an individual passively observing the faces of others is not representative for an agent involved in face-to-face communication with a specific task structure.

Our study combines perspectives from a growing literature on the importance of social context for gaze behavior (Freeth et al., 2013; Gobel et al., 2015; Laidlaw et al., 2011; Macdonald & Tatler, 2018; Richardson et al., 2012; Risko & Kingstone, 2011), the literature on the role of task structure on gaze allocation (Hayhoe, 2000; Hayhoe & Ballard, 2005, 2014; Land et al., 1999), and observational studies on face-to-face communication (Argyle, 1972; Duncan & Fiske, 1977). Previous research has focused primarily on how the (implied) presence of others (e.g. Gobel et al., 2015) or the potential for interaction (e.g. Gregory & Antolin, 2018) affects gaze behavior to faces. However, this has thus far not resulted in a model of gaze behavior in face-to-face interaction. Here we investigated which factors might be relevant for such models, by relating gaze behavior to faces to the task structure, and manipulating the behavior of an interacting agent. Both the present study, as well as previous studies, show that models of gaze allocation to facial features need to emphasize task-structure and social context above and beyond traditional visual saliency (e.g. Itti & Koch, 2000), at least insofar as other people are involved.

In accordance with previous research, we have shown that individuals look less often, for shorter average durations, and for a total shorter duration at the face (Argyle, 1972; Ho et al., 2015) and facial features (Rogers et al., 2018) when speaking compared to when listening. Although individuals look less at the face and for shorter durations, we report that individuals do not look proportionally longer at the mouth when listening than while speaking, and do not look proportionally longer at the eyes when speaking than while listening when they *do* look at the face (Supplementary Fig. 3). Rogers et al. (2018), however, reported changes of around 5% in the proportion of total dwell time to the eyes and mouth when listening versus speaking. If the differences in looking at facial features as a proportion of looking at the face, are small (Rogers et al., 2018) to non-existent (the present study), what does this entail for the interpretation of gaze allocation to the facial features during speaking or listening, and for the relation between gaze and task structure in interaction in general? As our time-locked analyses indicate, gaze allocation may be sensitive to minute changes in informational needs given a particular task. For example, gaze allocation in conversational settings may be quite predictable when a speaker signals the end of a turn (by fixating on the eyes of the interacting partner), or when one tries to follow a low-volume speaker (looking at the mouth of the interacting partner; Vatikiotis-Bateson et al. (1998)), but not at other times during conversation. As stated before, "It is probably a mistake to think that every fixation must have an identifiable purpose" (Tatler et al., 2011, p. 8). Aggregate measures of gaze behavior are unlikely to capture such task-dependent gaze allocation, and we urge researchers to consider the allocation of gaze to facial features on a finer timescale. Given the fact that other means of communication (e.g. language) operate on finer timescales too and are related to gaze allocation (Clark & Krych, 2004; Hanna & Brennan, 2007; Macdonald & Tatler, 2013), a fine-grained investigation of gaze allocation in face-to-face interaction would benefit the field.

As previously discussed, the observed change in the similarity of participants' gaze behavior by the confederate's gaze direction could be explained as task-dependent or social context-dependent. It could be task-driven in the sense that looking toward or away from the participant has meaning in the context of the participants' putative sub-task of monitoring the confederate for when they could start speaking. As stated, looking toward or away from a person in conversation may signal the end of, or the preservation of, one's turn, respectively (e.g. Ho et al., 2015). While in conversations there may be a quicker back and forth than we simulated in the present study (although at least some linguistic interaction occurred in the form of vocal back channels, see Supplementary Table 1), we assume that monitoring a conversational partner for when to speak is at least part of the task-structure being carried out. It is likely, however, that more sub-tasks may be identified in conversation that may guide the allocation of gaze (e.g. memorizing another's words, or deciding upon an answer to a question posed by the conversational partner). On the other hand, the effect might be driven according to the social context in the sense that the gaze direction of the confederate toward the participant may evoke a perception of eye contact and a corresponding avoidance of the eye region in some participants but not others, which does not occur in the absence of an interacting partner (therefore 'social context'). However, why this might then occur on a short time scale only, and is not visible in aggregate measures of gaze behavior (the confederate maintained direct or avoided gaze for longer periods) is unclear. As such, we are inclined to assert that the task-driven explanation is more likely. Future research may elucidate the independent contributions of task structure and social context on the allocation of gaze to facial features during face-to-face communication, or whether these two can be separated at all.

A limitation of the present study is that participants were placed at either end of a table with a video and eye-tracking setup between them. This precludes participants from referencing to, and working with, objects in a shared workplace. However, the video and eye-tracking setup allows us to measure eye-movements with high spatial resolution, and to collect video data from the participants for automatic data analysis. This is generally not possible with alternative eye-tracking equipment, such as wearable eye trackers. As wearable eye-tracking technology advances, one may be able to integrate these methods to better understand gaze allocation beyond dyadic interaction, towards triadic and dyad-object interaction. Furthermore, it should be noted that we investigated the effect of a confederate's gaze direction on the similarity of gaze allocation by the entire participant group. Ideally, one might want to investigate such effects at the level of an individual. However, to prevent participants from finding out that the interacting partner was a confederate carrying out a specific set of actions (looking at/away from the participants) at specific points in time, we engaged participants in a short conversation. This meant that there were too few systematic changes in the gaze behavior of the confederate to investigate the effects on the individual participant's gaze behavior. Additionally, scan patterns of faces are highly idiosyncratic (e.g. Arizpe et al., 2017; Kanan et al., 2015; Mehoudar et al., 2014; Peterson & Eckstein, 2013; Rogers et al., 2018), making it likely that task-dependent changes in gaze allocation are not easily picked up at a group level. Longer conversational settings may be better suited for capturing the effects of a confederate's gaze behavior on the gaze behavior of an individual.

Future research may benefit from studies focusing more on modeling the allocation of gaze in face-to-face communication of individuals based on their idiosyncratic scan patterns, combined with stimulus-, task-, and contextual-factors, as opposed to mere biases or preferences in aggregate gaze behavior. Models of gaze allocation incorporating stimulus-, task-, and contextual-factors are likely to be of benefit to several research fields outside of the social attention literature. For example, the study of atypical gaze behavior of individuals with Autism Spectrum Disorder (ASD) has been full of inconsistent reports on the supposed diminished bias of looking at the eyes of others (Guillon et al., 2014). Therefore, a benchmark of typical gaze allocation in face-to-face communication may open up new avenues for the investigation of atypical gaze allocation, for example in ASD. Likewise, studies on student-teacher or patient-physician interaction (e.g. Asan et al., 2015) may benefit from models on gaze allocation in face-to-face interaction, and how gaze allocation may facilitate the transmission of information. Finally, the study of human-robot interaction may benefit from such benchmarks in order to establish what gaze allocation a robot should portray in order to be perceived as a comfortable interaction partner (e.g. Mutlu, Shiwa, Kanda, Ishiguro, & Hagita, 2009).

## 8. Conclusion

We have shown that gaze behavior in face-to-face communication is dependent on the task at hand and influenced by the behavior of the interacting partner. Our findings reinforce and extend the limitations of saliency-based models of gaze allocation, and indicate that both task structure and contextual factors play an important role. We particularly highlight the importance of the gaze behavior of an interacting partner in relation to the task structure, for example monitoring when to take turns in conversation. Future studies on gaze behavior in face-to-face communication will benefit from emphasizing and including contextual factors and task-specific informational demands on the allocation of gaze.

## Acknowledgements

This work was supported by the Consortium on Individual Development (CID). CID is funded through the Gravitation program of the Dutch Ministry of Education, Culture, and Science and the NWO (Grant No. 024.001.003 awarded to CK). The authors would like to thank Nina Haukes and Thei Bongers for valuable help with data collection.

## Appendix A. Example script

“Uhmm…Well…When I was young I used to go on summer holiday with my parents, my sister and my brother. We used to go camping in Belgium, France or Italy. …For example, in the Ardennes, the Dordogne, or Tuscany. These were my parents’ favorite places. This was always a lot of fun…

Uhmm.. *(confederate switches gaze direction)*

When…I finished high school, I wanted to take a year off. I saved up money and decided to go on a big backpacking adventure in Australia and South-East Asia. I spent 4 months in Australia and another two months in Thailand, Laos and Cambodia. This was my first big trip by myself and it was fantastic. I made many friends and visited many beautiful places. I would love to do something like that again.

Well…*(confederate switches gaze direction)*

I also love to go on winter holidays to do snowboarding. I used to go to Switzerland with my family because my uncle lives there, but last year I went together with four friends to Austria. We rented a little apartment for one week and went snowboarding every day. That was a lot of fun. This year I will go with my girlfriend to France for the winter holiday.

Let’s see, what else…*(confederate switches gaze direction)*

I also like to go on little city trips…Recently, I have visited my sister in Copenhagen, she lives there now…I also recently visited a friend of mine in Oxford, who is currently studying there, and my next trip is going to be Malaga. I will visit several friends of mine who live there. I don’t have any other holiday plans for this summer yet.

Well…I think that’s it…*(confederate gazes gaze at or away from participant)*

## Appendix B. Supplementary material

The raw data and statistical analyses underlying the findings reported in this article are available from https://osf.io/k32vz/ (Hessels, Holleman, Kingstone, Hooge, & Kemner, 2018).

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.cognition.2018.12.005.

## References

Allen, D. E., & Guy, R. F. (1977). Ocular breaks and verbal output. *Sociometry, 40*(1), 90–96.
Allsop, J., & Gray, R. (2014). Flying under pressure: Effects of anxiety on attention and gaze behavior in aviation. *Journal of Applied Research in Memory and Cognition, 3*(2), 63–71.
Argyle, M. (1972). Eye-contact and the direction of gaze. *The psychology of interpersonal behavior*. Harmondsworth, Middlese: Penguin books.
Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze.* Cambridge University Press.
Arizpe, J., Walsh, V., Yovel, G., & Baker, C. I. (2017). The categories, frequencies, and stability of idiosyncratic eye-movement patterns to faces. *Vision Research, 141*, 191–203.
Asan, O., Young, H. N., Chewning, B., & Montague, E. (2015). How physician electronic health record screen sharing affects patient and doctor non-verbal communication in primary care. *Patient Education and Counseling, 98*(3), 310–316.
Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412.
Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences, 24*, 191–204.
Baltrušaitis, T., Robinson, P., & Morency, L. -P. (2013). Constrained local neural fields for robust facial landmark detection in the wild. In *2013 IEEE International Conference on Computer Vision Workshops (ICCVW)* (pp. 354–361). IEEE.
Baltrušaitis, T., Robinson, P., & Morency, L. -P. (2016). OpenFace: An open source facial behavior analysis toolkit. In *IEEE winter conference on applications of computer vision* (pp. 1–10).
Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The ”Reading the Mind in the Eyes” test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241–251.
Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.
Beattie, G. W. (1981). A further investigation of the cognitive interference hypothesis of gaze patterns during conversation. *British Journal of Social Psychology, 20*, 243–248.
Beattie, G. W., & Bogle, G. (1982). The reliability and validity of different video-recording techniques used for analysing gaze in dyadic interaction. *British Journal of Social Psychology, 21*, 31–34.
Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Research, 49*(24), 2992–3000.
Caruana, N., McArthur, G., Woolgar, A., & Brock, J. (2017). Simulating social interactions for the experimental investigation of joint attention. *Neuroscience & Biobehavioral Reviews, 74*(Part A), 115–125.
Cary, M. S. (1978). The role of gaze in the initiation of conversation. *Social Psychology, 41*(3), 269–271.
Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50*(1), 62–81.
Cole, G. G., Skarratt, P. A., & Kuhn, G. (2016). Real person interaction in visual attention research. *European Psychologist, 21*, 141–149.
Crivelli, C., & Fridlund, A. J. (2018). Facial displays are tools for social influence. *Trends in Cognitive Sciences, 22*(5), 388–399.
Cutler, A. & Pearson, M. (1986). On the analysis of prosodic turn-taking cues. In C. Johns-Lewis (ed.), *Intonation in discourse* (pp. 139–155). London, UK.
De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences, 14*(10), 441–447.
de Waal, F. B. M. (2009). *Peacemaking among primates.* Harvard University Press.
Duncan, S., & Fiske, D. W. (1977). *Face-to-face interaction. Research, methods, and theory.* Hillsdale, New Jersey: Lawrence Erlbaum.
Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews, 24*, 581–604.
Foddy, M. (1978). Patterns of gaze in cooperative and competitive negotiation. *Human Relations, 31*(11), 925–938.
Freeth, M., Foulsham, T., & Kingstone, A. (2013). What affects social attention? Social presence, eye contact and autistic traits. *PLOS One, 8*(1), e53286.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133*(4), 694–724.

Gobel, M. S., Kim, H. S., & Richardson, D. C. (2015). The dual function of social gaze. *Cognition, 136*, 359–364.

Gregory, N., & Antolin, J. (2018). Does social presence or the potential for interaction reduce social gaze in online social scenarios? Introducing the Live Lab paradigm. *The Quarterly Journal of Experimental Psychology*.

Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience & Biobehavioral Reviews, 42*, 279–297.

Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of memory and language, 57*(4), 596–615.

Hayhoe, M. (2000). Vision using routines: A functional account of vision. *Visual Cognition, 7*(1–3), 43–64.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9*(4), 188–194.

Hayhoe, M., & Ballard, D. (2014). Modeling task control of eye movements. *Current Biology, 24*(13), R622–R628.

Heft, H. (2001). *Ecological psychology in context.* Lawrence Erlbaum.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33*(1), 98–106.

Hessels, R. S., Andersson, R., Hooge, I. T. C., Nyström, M., & Kemner, C. (2015). Consequences of eye color, positioning, and head movement for eye-tracking data quality in infant research. *Infancy, 20*(6), 601–633.

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

Hessels, R. S., Cornelissen, T. H. W., Hooge, I. T. C., & Kemner, C. (2017a). Gaze behavior to faces during dyadic interaction. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 71*(3), 226–242.

Hessels, R. S., Holleman, G. A., Cornelissen, T. H. W., Hooge, I. T. C., & Kemner, C. (2018). Eye contact takes two – autistic and social anxiety traits predict gaze behavior in dyadic interaction. *Journal of Experimental Psychopathology,* (April-June), 1–17.

Hessels, R. S., Holleman, G. A., Kingstone, A., Hooge, I. T. C., & Kemner, C. (2018). Gaze allocation in face-to-face communication. https://doi.org/10.17605/OSF.IO/K32VZ.

Hessels, R. S., Kemner, C., van den Boomen, C., & Hooge, I. T. C. (2016). The area-of-interest problem in eyetracking research: A noise-robust solution for face and sparse stimuli. *Behavior Research Methods, 48*(4), 1694–1712.

Hessels, R. S., Niehorster, D. C., Kemner, C., & Hooge, I. T. C. (2017b). Noise-robust fixation detection in eye movement data: Identification by two-means clustering (I2MC). *Behavior Research Methods, 49*(5), 1802–1823.

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

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

Honma, M., Tanaka, Y., Osada, Y., & Kuriyama, K. (2012). Perceptual and not physical eye contact elicits pupillary dilation. *Biological Psychology, 89*(1), 112–116.

Hooge, I., & Camps, G. (2013). Scan path entropy and arrow plots: Capturing scanning behavior of multiple observers. *Frontiers in Psychology, 4*(996), 1–12.

Hooge, I. T. C., Holleman, G. A., Haukes, N. C., & Hessels, R. S. (2018). Gaze tracking accuracy in humans: One eye is sometimes better than two. *Behavior Research Methods,* 1–10.

Itier, R. J., Villate, C., & Ryan, J. D. (2007). Eyes always attract attention but gaze orienting is task-dependent: Evidence from eye movement monitoring. *Neuropsychologia, 45*(5), 1019–1028.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489–1506.

Jack, R. E., & Schyns, P. G. (2017). Toward a social psychophysics of face communication. *Annual Review of Psychology, 68*(1), 269–297.

Jarick, M., & Kingstone, A. (2015). The duality of gaze: Eyes extract and signal social information during sustained cooperative and competitive dyadic gaze. *Frontiers in Psychology, 6*(1423), 1–22.

Kanan, C., Bseiso, D. N. F., Ray, N. A., Hsiao, J. H., & Cottrell, G. W. (2015). Humans have idiosyncratic and task-specific scanpaths for judging faces. *Vision Research, 108*, 67–76.

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

Krantz, M., George, S. W., & Hursh, K. (1983). Gaze and mutual gaze of preschool children in conversation. *The Journal of Psychology, 113*, 9–15.

Laidlaw, K. E. W., 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.

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception, 28*, 1311–1328.

Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences, 4*(2), 50–59.

Libby, W. L., & Yaklevich, D. (1973). Personality determinants of eye contact and direction of gaze aversion. *Journal of Personality and Social Psychology, 27*(2), 197–206.

MacDonald, K. (2015). Patient-clinician eye contact: Social neuroscience and art of clinical engagement. *Postgraduate Medicine, 121*(4), 136–144.

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), 1–12.

Macdonald, R. G., & Tatler, B. W. (2018). Gaze in a real-world social interaction: A dual eye-tracking study. *Quarterly Journal of Experimental Psychology, 43*(1) 174702181773922-12.

Maynard, S. K. (1987). Interactional functions of a nonverbal sign. *Journal of Pragmatics, 11*, 589–606.

Mehoudar, E., Arizpe, J., Baker, C. I., & Yovel, G. (2014). Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of Vision, 14*(7), 1–11.

Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Footing in human-robot conversations: how robots might shape participant roles using gaze cues. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction* (pp. 61–68), New York, New York, USA: ACM Press.

Niehorster, D. C., Cornelissen, T. H. W., Holmqvist, K., Hooge, I. T. C., & Hessels, R. S. (2017). What to expect from your remote eye-tracker when participants are unrestrained. *Behavior Research Methods,* 1–15.

Nummenmaa, L., Hyönä, J., & Hietanen, J. K. (2009). I'll walk this way: Eyes reveal the direction of locomotion and make passersby look and go the other way. *Psychological Science, 20*(12), 1454–1458.

Peterson, M. F., & Eckstein, M. P. (2013). Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychological Science, 24*(7), 1216–1225.

Pfeiffer, U. J., Vogeley, K., & Schilbach, L. (2013). From gaze cueing to dual eye-tracking: Novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience & Biobehavioral Reviews,* 1–13.

R Core Team. (2017). R: A language and environment for statistical computing. Technical report, Vienna, Austria.

Richardson, D. C., Street, C. N. H., Tan, J. Y. M., Kirkham, N. Z., Hoover, M. A., & Cavanaugh, A. G. (2012). Joint perception: Gaze and social context. *Frontiers in Human Neuroscience,* 1–8.

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., Laidlaw, K. E. W., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: Toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience, 6*(1), 143.

Risko, E. F., Richardson, D. C., & Kingstone, A. (2016). Breaking the fourth wall of cognitive science. *Current Directions in Psychological Science, 25*(1), 70–74.

Rogers, S. L., Speelman, C. P., Guidetti, O., & Longmuir, M. (2018). Using dual eye tracking to uncover personal gaze patterns during social interaction. *Scientific Reports,* 1–9.

Schilbach, L. (2015). Eye to eye, face to face and brain to brain: Novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. *Current Opinion in Behavioral Sciences, 3*, 130–135.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences, 36*, 393–462.

Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences, 13*(3), 127–134.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27*, 379–423.

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision, 11*(5), 1–23.

Vatikiotis-Bateson, E., Eigsti, I.-M., Yano, S., & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception & Psychophysics, 60*(6), 926–940.

Võ, M. L. H., Smith, T. J., Mital, P. K., & Henderson, J. M. (2012). Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. *Journal of Vision, 12*(13), 3.

Yarbus, A. L. (1967). *Eye movements and vision.* Plenum Press.