Preprint of paper accepted in Frontiers in Psychology on 28th February 2019

The Role of Eye Gaze during Natural Social Interactions in Typical and Autistic People

Roser Cañigueral[1]*, Antonia F. de C. Hamilton[1]

[1]Institute of Cognitive Neuroscience, Division of Psychology and Language Sciences, University College London, London, United Kingdom

*** Correspondence:**
Roser Cañigueral
roser.canigueral.15@ucl.ac.uk

Word count: 11,130

Keywords: eye gaze, audience effect, interpersonal dynamics, social interaction, autism

Abstract

Social interactions involve complex exchanges of a variety of social signals, such as gaze, facial expressions, speech and gestures. Focusing on the dual function of eye gaze, this review explores how the presence of an audience, communicative purpose and temporal dynamics of gaze allow interacting partners to achieve successful communication. First, we focus on how being watched modulates social cognition and behaviour. We then show that the study of interpersonal gaze processing, particularly gaze temporal dynamics, can provide valuable understanding of social behaviour in real interactions. We propose that the Interpersonal Gaze Processing model, which combines both sensing and signalling functions of eye gaze, provides a framework to make sense of gaze patterns in live interactions. Finally, we discuss how autistic individuals process the belief in being watched and interpersonal dynamics of gaze, and suggest that systematic manipulation of factors modulating gaze signalling can reveal which aspects of social eye gaze are challenging in autism.

1 Introduction

In any face-to-face interaction between two people, both agents are continuously exchanging a variety of social signals, such as gaze, gestures or facial expressions. This two-way exchange of social information is possible because they are able to see each other, and consequently both agents can gather and communicate information. Although traditional cognitive research has largely ignored this interactive nature of social encounters, an increasing number of studies are looking at how social behaviour changes in a live interaction, as well as how eye gaze of two individuals coordinates to achieve successful communication, that is, to accurately process incoming signals and send back meaningful signals at a suitable pace.

In the present paper, we explore gaze as a communicative signal in a two-person interaction, considering both patterns of gaze to/from the other person and the interpersonal dynamics of gaze in relation to other behaviours. To explore these issues, we first introduce the dual function of eye gaze and describe two cognitive theories that explain changes in behaviour when being watched. We then consider gaze exchanges during communicative situations, and propose the Interpersonal Gaze Processing model as a framework to study the dynamics of gaze in face-to-face interactions. Finally, we look into the case of autism to discuss how studies on the audience effect and interpersonal dynamics of gaze can shed light on why autistic people find social communication challenging.

2 The dual function of eye gaze

Eye gaze has a dual function in human social interaction – we can both perceive information from others and use our gaze to signal to others (Argyle & Cook, 1976; Gobel, Kim, & Richardson, 2015; Risko, Richardson, & Kingstone, 2016). Simmel (1921) already stated that "the eye cannot take unless at the same time it gives". This contrasts with the auditory modality, where we use our ears to hear, but our mouth to speak. This makes our eyes a powerful tool for social interactions, with a "uniquely sociological function" (Simmel, 1921). For instance, when we see a pair of eyes we can gather information about what other people are looking at (Frischen, Bayliss, & Tipper, 2007), and how they feel or think (Baron-Cohen, Wheelwright, & Jolliffe, 1997). At the same time, we can use our eyes to strategically cue another's attention (Kuhn, Tatler, & Cole, 2009). Depending on the duration and direction of our gaze, we are also able to perceive and signal a variety of meanings, such as desire to communicate (Ho, Foulsham, & Kingstone, 2015), threat and dominance (Ellyson, Dovidio, & Fehr, 1981; Emery, 2000), attractiveness (Argyle & Dean, 1965; Georgescu et al., 2013), or seeking for approval (Efran, 1968; Efran & Broughton, 1966).

The dual function of the eyes has often been ignored in cognitive research studying social interactions. In typical lab studies, participants interact with a monitor that displays pictures or videos of other people, while their gaze or other behaviour is recorded (see Risko, Laidlaw, Freeth, Foulsham, & Kingstone, 2012 for a review). In these experimental settings signals are sent only one-way (from the picture to the participant) and the dual function of gaze is completely lost. Although these traditional approaches allow good experimental control, they are not interactive (Gobel et al., 2015; Risko et al., 2016; Schilbach et al., 2013). Recent research has implemented more ecologically valid approaches that can restore the dual function of gaze. The belief that someone can see us, intrinsic to live interactions, is thought to recruit a range of social cognitive processes that are missing when participants interact with videos or pictures (Risko et al., 2012, 2016; Schilbach et al., 2013). Moreover, in face-to-face interactions communication is multimodal (Vigliocco, Perniss, Vinson, & Vigliocco, 2014): information is exchanged through eye gaze, but also through gestures, facial expressions or speech, and all these signals need to be integrated over time and across agents (Hirai & Kanakogi, 2018; Holler, Kendrick, & Levinson, 2018; Jack & Schyns, 2015).

In the following, we first describe two cognitive theories that explain changes in behaviour when being watched. Then, we discuss why interpersonal dynamics are relevant when studying social eye gaze.

3 Cognitive theories of the audience effect

We behave differently when we are alone or in the presence of others. For instance, when we are with other people our actions become more prosocial (Izuma, Matsumoto, Camerer, & Adolphs, 2011; Izuma, Saito, & Sadato, 2009), our memory improves (Fullwood & Doherty-Sneddon, 2006), and we smile more (Fridlund, 1991). Triplett first introduced this idea 120 years ago, when he showed that cyclists were faster when competing against each other than against a clock (Triplett, 1898). To explain this effect, he suggested that the "bodily presence of another" causes changes in the behaviour of participants, which makes them more competitive when racing against others.

However, previous research has shown that there is more than one way in which the presence of another person can change our behaviour.

On the one hand, social facilitation refers to a change in behaviour caused by the presence of a conspecific who may or may not be watching us (Zajonc, 1965). This effect is present in humans but also in a wide range of species (e.g. cockroaches, rats and monkeys), suggesting that it relies on a simple mechanism like arousal. Zajonc further claimed that an increase in arousal in the presence of others would facilitate dominant behaviours (i.e. responses that are elicited most quickly by a stimulus). For instance in an easy task the dominant response is usually the correct one, while in a difficult task the dominant response is usually the incorrect one. Zajonc and Sales found that, in the presence of a conspecific, participants performed better on a verbal recognition task with familiar items (easy task), and worse on the same task with unfamiliar items (hard task) (Zajonc & Sales, 1966). This effect has been found in a range of tests on both mental (Geen, 1985) and physical skills (Strauss, 2002). Blascovich and colleagues replicated these findings and also showed that, in the presence of others, the cardiovascular system is differently triggered depending on the task: in a difficult task the cardiovascular response fits a threat-like pattern, whereas in an easy task the cardiovascular response fits a challenge-like pattern (Blascovich, Mendes, Hunter, & Salomon, 1999). This suggests that the facilitation of different dominant responses in the presence of others is mediated by different arousal patterns.

On the other hand, the audience effect is a change in behaviour specifically caused by the belief that someone else is watching me. It builds on mechanisms which process the perceptual state of the other, known as perceptual mentalising (Teufel, Fletcher, & Davis, 2010). Perceptual mentalising modulates the processing of social information from the eyes in a variety of ways. For example, seeing a live-feed of a person with transparent glasses (who can see) leads to a larger gaze cuing effect than a matched stimulus of a person with opaque glasses (who cannot see) (Nuku & Bekkering, 2008; Teufel, Alexis, Clayton, & Davis, 2010), and similar results are seen in tests of visual perspective taking (Furlanetto, Becchio, Samson, & Apperly, 2016). This demonstrates that even basic social processing is influenced by the knowledge that another person can see something. The audience effect takes this one step further, considering how our social cognition is affected by the knowledge that another person can see us.

Audience effects differ from social facilitation in that social facilitation could occur if another person is present but looking away, whereas audience effects are specific to the case when another person is believed to be watching (even from another location). When people believe they are being watched, they typically change their behaviour to maintain a positive public image. This has been described in terms of self-presentation theory (Bond, 1982), which claims that people modulate their performance in front of others to maintain a good public image and increase their self-esteem. Bond (1982) further showed that making errors while being observed translates into decreased self-esteem and poor performance, regardless of task difficulty.

The audience effect and the dual function of gaze are closely linked in that both require someone who can see us. In line with this, recent evidence suggests that being watched modulates gaze patterns directed at the face of the observer, because in this context direct gaze acquires a social meaning that an individual may or may not wish to signal to someone else. These studies show that in a live interaction people look less to the other person than in a pre-recorded interaction (Gobel et al., 2015; Laidlaw, Foulsham, Kuhn, & Kingstone, 2011). This change in gaze patterns is further modulated by several factors, such as the observer's social status (high rank or low rank; Gobel et al., 2015) or role in the interaction (speaker or listener; Freeth, Foulsham, & Kingstone, 2013; Ho et al., 2015). Thus, when being watched eye gaze is adjusted to send appropriate signals to the observer, rather than to only gather information from the environment.

In the following, we strictly focus on changes in social behaviour that derive from audience effects, that is, from the belief in being watched. To explain these changes, two main cognitive theories have been proposed: the Watching Eyes model (Conty, George, & Hietanen, 2016) and reputation management theory (Emler, 1990; Resnick, Zeckhauser, Swanson, & Lockwood, 2006; Tennie, Frith, & Frith, 2010). Both theories give plausible explanations about the relationship between an individual and an observer, but they have different focus. The Watching Eyes model concentrates on how an observer influences cognitive processing within individuals (self-focus), beyond self-esteem effects proposed by self-presentation theory.Reputation management theory explains how individuals manipulate the observer's beliefs to their advantage (other-focus) in an updated version of the self-presentation theory. Below we describe each of these theories in more detail.

3.1 Watching Eyes model

A pair of eyes watching us are an ostensive communicative cue (Csibra & Gergely, 2009) that rapidly captures our attention (Senju & Hasegawa, 2005). Early work on gaze processing proposed various mechanisms how direct gaze modulates our attention and behaviour. For instance, Baron-Cohen (1995) suggested that there is a specialised Eye Direction Detector module in the brain. This module rapidly identifies whether we are the target of someone else's attention by processing the direction of other people's eyes relative to us. The detection of direct gaze will in turn trigger mentalising processes that allow us to interpret the other person's mental states (Baron-Cohen & Cross, 1992; Baron-Cohen et al., 1997). Later, Senju & Johnson (2009) coined the term "eye contact effect" to describe changes in cognitive processing following perception of direct gaze, and introduced the Fast-track Modulator model of gaze processing. This model suggests that detection of direct gaze is implemented by a fast subcortical route involving the pulvinar and amygdala, and is modulated by higher cortical regions that depend on social context and task demands. The recently proposed Watching Eyes model (Conty et al., 2016) builds up on these models and suggests that audience effects are due to the "self-referential power of direct gaze".

Similar to the Fast-track Modulator model by Senju & Johnson (2009),the Watching Eyes model proposes two stages in the processing of direct gaze. In the first stage, direct gaze captures the beholder's attention by a subcortical route. This seems to be an automatic effect of direct gaze (Senju & Hasegawa, 2005), and is thought to be triggered by the detection of low-level visual cues in eye gaze (e.g. luminance distribution in the eye; Kobayashi & Kohshima, 2001; von Grünau & Anston, 1995). Then, the subcortical route engages mentalizing brain areas (medial prefrontal cortex and temporo-parietal junction) that process the perceptual state of the observer, that is, the belief that s/he is or is not watching us. In the second stage, if the observer can see us, then direct gaze will elicit self-referential processing, and the sense of self-involvement in the interaction will increase. This will lead to the Watching Eyes effects, causing a change in behaviour in various ways, such as enhancement of self-awareness (Baltazar et al., 2014; Hazem, George, Baltazar, & Conty, 2017; Pönkänen, Alhoniemi, Leppänen, & Hietanen, 2011) or promotion of prosocial actions (Izuma et al., 2011, 2009).

Recently, Hietanen and Hietanen (2017) have directly tested the Watching Eyes model of self-referential processing. To measure self-referential processing they used the foreign-language task, where participants read sentences in a language that they do not understand and need to match underlined words with pronouns in their native language. In this task, more use of first person singular pronouns is thought to be related to more self-referential processing. Participants completed this task but they watched a video-clip of a person with direct or averted gaze before each sentence was presented. Results showed no effect of eye gaze direction on the pronouns used. Then, a second group of participants completed the same task while they watched live faces with direct or averted face. They found that participants in the direct gaze group used more first person singular pronouns than the averted gaze group. In line with this, a recent study on bodily self-awareness (Hazem et al., 2017) has found that participants are more accurate in rating the intensity of a physiological signal when they believe they are in online connection with someone wearing clear sunglasses (the observer can see through) rather than someone wearing opaque sunglasses (the observer cannot see through). Taken together, these findings show evidence in favour of the Watching Eyes model: to trigger self-reference and self-awareness it is not enough to see a pair of eyes directly gazing at us – the belief that these pair of eyes can see us is also required.

Yet, it is important to consider that different tasks measure different forms of self-reference and self-awareness. This means that different tasks are likely to engage different self-related cognitive processes, which might have different sensitivity to the belief in being watched. For instance, the pronoun-selection task used by Hietanen and Hietanen (2017) is rather intuitive and has been shown to be sensitive to manipulations of self-awareness (Davis & Brock, 1975). However, it could be that other tasks which elicit more complex self-referential cognitive processes (e.g. self-referential effect memory task; Craik & Tulving, 1975; Lombardo, Barnes, Wheelwright, & Baron-Cohen, 2007) are not as sensitive to this top-down modulation. It is equally important to distinguish between different forms of self-awareness, such as bodily self-awareness (accuracy in reporting physiological signals; Cameron, 2001) and metacognitive self-awareness (accuracy in judging performance in a task; Fleming & Dolan, 2012). Thus, it remains to be seen whether direct gaze and the belief in being watched modulate all forms of self-referential processing and self-awareness or not.

3.2 Reputation management theory

Reputation is a social construct that emerges from the desire to cultivate good self-impressions in others (Silver & Shaw, 2018). It is based on how we think others see us, and it changes over time depending on our actions (Cage, 2015; Izuma, 2012). People can gain approval from others and increase their own reputation in various ways, such as acting for the benefit of other people or behaving according to social norms. To maintain or manage reputation, individuals need to think about what others think of them, care about how others see them, and have the desire to foster positive impressions in others (Cage, 2015; Izuma, 2012). Thus, mentalizing and social motivation have a central function in reputation management (Cage, 2015; Izuma, 2012; Saito et al., 2010; Tennie et al., 2010). In line with this, neuroimaging studies have shown that mentalizing and reward brain areas are engaged during different phases of reputation management, such as processing what others think of them (e.g. medial prefrontal cortex; Frith & Frith, 2006; Izuma, Saito, & Sadato, 2010) or anticipating positive reputation (e.g. ventral striatum; Izuma et al., 2009, 2010) respectively.

One strategy that people use to maintain a good reputation in front of others is to behave in a more prosocial fashion (Bradley, Lawrence, & Ferguson, 2018; Smith & Bird, 2000). A way to measure prosocial behaviour in the lab is by using economic games. Because they usually have repeated trials, this facilitates reputation building between participants in the game (Bradley et al., 2018; T. Pfeiffer & Nowak, 2006). For instance, Filiz-Ozbay and Ozbay (2014) used the Public Goods game and found that people invest more effort to contribute to public, but not private, goods when someone is observing them. Izuma and colleagues (2011) used the Dictator game (Guala & Mittone, 2010; Kahneman, Knetsch, & Thaler, 1986) as a donation task, where participants receive a sum of money and must decide on repeated trials whether to accept a proposal to share the money with a charity, or reject it and keep all the money. Results showed that in the presence of a confederate who pretended to monitor the answers, participants decided to accept the proposed sharing more often than when they were alone in the room. These findings clearly illustrate how participants manipulate the beliefs of the observer to maintain their good reputation.

Several factors modulate how strong the audience effect is on prosocial behaviour (Bradley et al., 2018), such as the identity of the observer (experimenter, other participants, stranger) or whether decisions of participants are consequential. For instance, Cage and colleagues (Cage, Pellicano, Shah, & Bird, 2013) also used the Dictator game in the presence and absence of a confederate, but additionally contrasted two conditions: one in which participants believed the recipient of the sharing arrangement was an individual who could later reciprocate (consequential decision), and one in which the recipient could not reciprocate (non-consequential decision). They found that participants accepted the chance to share money most frequently in the presence of a confederate and when the confederate could later reciprocate. This shows that the context associated with the observer (e.g. can s/he reciprocate or not?) also modulates the extent to which being watched affects behaviour.

Another strategy used to maintain reputation is to behave according to social norms. Social norms can be of various kinds, such as saying thank you or holding a door for someone after you. A more subtle type of social norm is civil inattention (Goffman, 1963), which proposes that the amount of gaze directed to strangers "should be enough to acknowledge their presence but not so much as to indicate that they are of special interest". Multiple studies have used eye-tracking to test if social attention is modulated according to social norms of eye gaze. For instance, Laidlaw and colleagues (Laidlaw et al., 2011) found that participants sitting in a waiting room would look more to a confederate in a video-clip than to the same confederate present in the room. The authors claimed that this change in gaze patterns is due to a social norm whereby it is not polite to stare at someone, which in turn translates into active disengagement.

Some of these studies also show that gaze patterns in live contexts are modulated by a number of factors that do not have any effect when participants watch video-clips. Gobel et al. (2015) found that participants spend more time gazing at video-clips of a low rank confederate and less time gazing at video-clips of a high rank confederate, but only when they believe the confederate will later see their gaze recording. These two gaze behaviours, direct and averted gaze, have been associated with signalling of dominance and submission, respectively (Ellyson et al., 1981; Emery, 2000). In another study, Foulsham and colleagues (Foulsham, Walker, & Kingstone, 2011) showed that participants gaze less to close pedestrians than distant pedestrians to avoid appearing as an interaction partner to strangers (see also Argyle & Dean, 1965; Gallup, Chong, & Couzin, 2012). These studies indicate that, when an observer is watching, eye gaze acquires a signalling function and this will subtly modulate gaze patterns to send appropriate signals to the observer. Moreover, the social skills of participants and their looking behaviour are correlated in live but not lab settings (Laidlaw et al., 2011). This suggests that individuals who successfully interact with other people are those who can modulate social behaviour according to requirements of the social context.

So far, we have discussed how the presence of an observer modulates an individual's cognitive processing, both self-focused (Watching Eyes model) and other-focused (reputation management theory). However, the studies presented above have a major limitation: confederate and participant are not expected (and do not intend) to interact, verbally or physically, with each other. This means that there is no explicit communicative exchange between them. In the same way that social behaviour changes when participants watch a video-recorded person or a live person, it could be that it also differs between a situation where there is potential for an interaction and a situation where there is an actual interaction with explicit communicative exchanges (henceforth communicative encounter; Foulsham et al., 2011; Macdonald & Tatler, 2018; Wu, Bischof, & Kingstone, 2013). Focusing on the particular case of eye gaze, in the next section we argue that interpersonal gaze dynamics have a key role in modulating social behaviour during communicative encounters.

4 Interpersonal dynamics of eye gaze

Original studies about the role of eye gaze during communicative encounters date back to the 60s, when Argyle and colleagues (Argyle & Cook, 1976; Argyle & Dean, 1965) put forward the intimacy equilibrium model, which is the first account on the relationship between "looking and liking": they showed that gaze directed at other people serves to control the level of intimacy or affiliation with the partner, and that it compensates with other behaviours (e.g. physical proximity) to achieve an equilibrium level of intimacy (see also Loeb, 1972). Furthermore, Watzlavick and colleagues (Watzlawick, Helmick Beavin, & Jackson, 1967) proposed the idea that "one cannot not communicate", since the lack of response is a response in itself (e.g. not looking at someone signals lack of interest in the interaction; Goffman, 1963).

Recent studies show that direct gaze can act as an ostensive communicative signal (Csibra & Gergely, 2009). During face-to-face interactions, where individuals exchange information with communicative purpose through a variety of channels (e.g. gaze, gestures, facial expressions, speech), direct gaze helps to integrate and coordinate auditory and visual signals (Bavelas, Coates, & Johnson, 2002). Moreover, it has been shown that to successfully produce and detect gestures with communicative purpose, information conveyed by gaze signals (e.g. direct gaze) is preferentially used over information conveyed by kinematics of the gesture (Trujillo, Simanova, Bekkering, & Özyürek, 2018). Thus, eye gaze has a core function in leading social interactions up to successful communicative exchanges, where there is efficient transmission of information between sender and receiver.

In the studies presented in the previous section, the authors claim that changes in eye gaze when participants are being watched respond to demands of social norms (Foulsham et al., 2011; Gobel et al., 2015; Laidlaw et al., 2011). The context of those studies does not require participants to explicitly communicate with the confederate, but only look (or not) at each other. Moreover, the confederate is usually a complete stranger to the participant. It is therefore not surprising that this awkward interaction without communicative purpose leads participants to modulate eye gaze in compliance with social norms (Wu et al., 2013). However, in communicative encounters (e.g. conversation) gaze patterns need to coordinate with other verbal and non-verbal signals to successfully receive and send signals (Bavelas et al., 2002; Trujillo et al., 2018). In studying such communications, we must consider not just the average pattern of gaze (towards/away from the face) but also the dynamics of gaze behaviour in relation to other social events (speech, turn taking, facial expressions, etc.). This means that to succeed during communicative exchanges, eye gaze needs not only modulation by social norms, but also constant adjustments to keep pace with interpersonal dynamics that emerge as the interaction develops.

In the following, we first describe the main social functions that eye gaze has during communicative interactions. Then, we focus on the temporal dynamics of gaze as a key mechanism that enables meaningful interpersonal exchanges during communication, as well as successful progression of the interaction.

4.1 Social functions of eye gaze during conversation

During communicative encounters, such as conversations, the eyes of both agents are generally very active. In a seminal study on gaze direction during conversation, Kendon identified asymmetrical gaze behaviour between speakers and listeners (Kendon, 1967): while listeners gazed at speakers most of the time, speakers shifted their gaze toward and away from listeners. More recently, Rogers and colleagues (Rogers, Speelman, Guidetti, & Longmuir, 2018) found that during a 4 min conversation participants spent on average 60% of the time directing their gaze towards the face of the other person (only 10% of the time it was directed specifically to the eyes), and that these events were approximately 2.2 s long (for direct eye contact events were 0.36 s long). The brief duration of these events supports Kendon's original findings, because it indicates that participants are constantly alternating their gaze between face or eyes of their partner and other regions. There has been much debate about the meaning of these rapid and subtle changes in eye gaze direction and duration. Kendon originally suggested that they give rise to three main social functions of gaze (Kendon, 1967). Note that, although the gaze patterns described below allow us to send signals to another person, these signals are sent implicitly and without awareness.

First, he proposed that eye gaze has a regulatory function during conversation, because it allows individuals to modulate transitions between speaker and listener states (i.e. turn-taking). In line with this, it has been found that speakers use averted gaze when they begin to talk and during hesitation (probably to indicate that they want to retain their role as speakers), but they use direct gaze to the listener when they are about to end an utterance (probably to signal that their turn is ending and that the listener can take the floor) (Cummins, 2012; Duncan & Fiske, 1977; Ho et al., 2015; Kendon, 1967; Sandgren, Andersson, Weijer, Hansson, & Sahlén, 2012). However, as noted by Ho and colleagues (Ho et al., 2015) conversation is a two-way process and this means that the listener is also responsible to regulate in turn-taking. For instance, it has been shown that listeners make more gestures, head shifts and gaze shifts before speaking, probably to indicate to the speaker that they want to take the turn (Harrigan, 1985).

Second, Kendon suggested that eye gaze has a monitoring function: it allows each participant to track attentional states and facial displays of the partner to ensure mutual understanding and seek social approval from others (Efran, 1968; Efran & Broughton, 1966; Kleinke, 1986). Indeed, speakers try to gain more information about what listeners think by engaging in brief periods of mutual eye gaze, which elicit back-channelling (i.e. listener's brief responses showing comprehension of what the speaker is saying) (Bavelas et al., 2002). Rogers et al. (2018) have also proposed that brief and rapid gaze shifting between gaze directed to the eyes and to other facial regions (e.g. mouth, eyebrows) may serve to scan facial features and pick subtle cues that help interpreting the meaning of what is being said. The monitoring function of gaze can also have high cognitive costs. For instance, when participants are asked to look at the face of the experimenter, they perform worse than participants who can avert their gaze naturally (Beattie, 1981), or who are asked to fixate on other static or dynamic stimuli (Markson & Paterson, 2009). Thus, Kendon also claimed that speakers avert their gaze partly to reduce the costs associated with monitoring a face.

Third, Kendon proposed that eye gaze has an expressive function, which allows participants to regulate the level of arousal in the interaction. He found that some participants tended to avert their gaze at moments of high emotion, and that the amount of eye contact was inversely related to the frequency of smiling. He suggested that averting gaze at this highly emotional moments could be interpreted as a 'cut off' act to express embarrassment and reduce arousal. Moreover, the expressive function of mutual eye gaze has been associated with affiliation and attraction (Argyle & Cook, 1976; Argyle & Dean, 1965; Georgescu et al., 2013), with dominance and power (Ellyson et al., 1981; Emery, 2000; Gobel et al., 2015), and more recently with expressing response preference to polar questions (Kendrick & Holler, 2017).

It is important to bear in mind that the social functions of gaze are only meaningful during face-to-face interactions, where both partners can see each other. It is only in this context that eye gaze has a dual function and both agents can perceive and signal information (Gobel et al., 2015; Risko et al., 2016). Moreover, gaze signals are not isolated: speakers need to shift their gaze toward or away from the listener at specific time points during speech, listeners need to coordinate gaze direction with facial expressions to indicate preference or reduce arousal, and speakers and listeners need to engage in brief mutual gaze periods to exchange turns or elicit back-channelling. Thus, to succeed in communicative encounters social signals need to be coordinated within and across conversation partners over time.

4.2 Temporal dynamics of gaze

Successful communication requires that both agents involved in the interaction process incoming signals and send back meaningful signals at a suitable pace. Since these signalling exchanges (specially for eye gaze) happen very quickly, timing becomes a critical factor to enable successful progression of the interaction. The need for timed coordination gives rise to patterns of gaze behaviour, that is, temporal dependencies that emerge between gaze and other social signals. For instance, using gaze cueing paradigms (e.g. Posner's paradigm; Posner, 1980) it has been shown that averted gaze results in reflexive gaze following behaviour, which is key to build joint attention (U. J. Pfeiffer, Vogeley, & Schilbach, 2013). Similarly, there could be a systematic relationship between gaze and speech within an individual (e.g. direct gaze at others when finishing an utterance, but avert gaze when hesitating; Ho et al., 2015), or between the gaze direction of two conversation partners (e.g. establish mutual eye gaze to elicit back-channelling; Bavelas et al., 2002). The presence and direction of these temporal dependencies at different time points can contribute to identifying which social cognitive processes modulate gaze behaviour in the course of the interaction.

Experimentally manipulating temporal dynamics of eye gaze in the lab can be challenging, because it requires some degree of control over gaze patterns for at least one of the agents. Virtual reality and humanoid robot avatars offer an efficient alternative to this issue, because their behaviour can be meticulously controlled while participants respond with comparable social behaviours as in interactions with real human beings (U. J. Pfeiffer et al., 2013). With the aim of studying interactions in a truly reciprocal context, Wilms and colleagues (Wilms et al., 2010) created the now widely used gaze-contingent eye-tracking paradigm (see also Bayliss et al., 2012; Edwards, Stephenson, Dalmaso, & Bayliss, 2015; Kim & Mundy, 2012). In this paradigm, participants wearing an eye-tracker interact with an avatar whose gaze is controlled by the real-time gaze data collected from the participant. Thus, the avatar becomes a gaze-contingent stimulus that responds to the participant's gaze behaviour. Using this paradigm in the context of joint attention, it has been shown that avatars are perceived as more human-like (U. J. Pfeiffer, Timmermans, Bente, Vogeley, & Schilbach, 2011) and more likeable (Grynszpan, Martin, & Fossati, 2017; Willemse, Marchesi, & Wykowska, 2018) if they follow the gaze of participants to achieve joint attention. Another study has shown that participants are quicker to assume that the avatar understands their instructions when there is contingent gaze following (Frädrich, Nunnari, Staudte, & Heloir, 2018). At the neural level, joint attention has been linked to activation in brain areas related to gaze direction (superior temporal sulcus), processing of rewards (ventral striatum) and mental states (medial prefrontal cortex, temporo-parietal junction) (Caruana, Brock, & Woolgar, 2015; Pelphrey, Viola, & McCarthy, 2004; U. J. Pfeiffer et al., 2013; Schilbach et al., 2010).

Some attempts have also been made to study the nature of temporal dynamics of gaze in real human-to-human interactions. For instance, Lachat and colleagues (Lachat, Hugueville, Lemaréchal, Conty, & George, 2012) designed a joint attention task where dyads of participants engaged in joint and no-joint attention periods respectively. They found that during joint attention periods mu rhythms in centro-parietal regions were supressed for both leaders and followers, which has been previously associated with interpersonal coordination processes (Naeem, Prasad, Watson, & Kelso, 2012). In another study, participants completed a structured interview with a pre-recorded or live confederate, whose gaze was directed at them or averted (Freeth et al., 2013). They found that participants gazed more to the confederate's face if her gaze was directed at them than if her gaze was averted, but only in the live condition. This means that participants' gaze was adjusted according to the looking behaviour of the confederate only when their gaze acquired a signalling function (i.e. they were in a live interaction), thus creating a reciprocal social signal. Recently, a dual eye-tracking study (Macdonald & Tatler, 2018) has also shown that pairs of participants who are given specific social roles in a collaborative task align their gaze quicker than pairs who have no social role. This indicates that eye gaze adjusts to the communicative purpose embedded in different social contexts.

Gaze dynamics are fundamental to efficiently communicate with other people, that is, to enable information transfer between individuals. It has recently been suggested that brain-to-brain coherence (i.e. synchronisation of neural activity between two brains) provides a marker of the success of a communication between two people (Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012), and several hyperscanning studies show that mutual gaze triggers neural coherence between partners. For instance, mutual gaze mediates neural coupling between parents and infants, which has been associated with appropriate use of communicative signals according to each social context later in development (Piazza, Hasenfratz, Hasson, & Lew-Williams, 2018). Neural coherence between parents and infants has been shown to be stronger in live versus pre-recorded interactions (Leong et al., 2017). Moreover, in a joint attention task through a video-feed, moments of eye contact were characterised by increased synchronisation of frontal brain activity between participants (Saito et al., 2010). Hirsch and colleagues (Hirsch, Zhang, Noah, & Ono, 2017) have also shown that only when partners in a dyad make eye contact (compared to when both partners look at a photograph of a face) brain-to-brain coherence between partners increases in regions associated with processing of social information (temporo-parietal and frontal regions). These findings suggest that direct gaze acts as a signal that enhances the temporal alignment of two brains (Gallotti, Fairhurst, & Frith, 2017; Hasson et al., 2012), thus facilitating the sharing of information.

All these studies show that temporal coordination of gaze patterns are characteristic of human interactions (U. J. Pfeiffer et al., 2011; Willemse et al., 2018), and that they have beneficial effects for the interacting partners, such as increasing the reward value of the interaction (Schilbach et al., 2010), or facilitating social coordination (Frädrich et al., 2018; Freeth et al., 2013; Lachat et al., 2012) and information transfer (Hirsch et al., 2017; Leong et al., 2017; Saito et al., 2010). They also highlight that gaze is a dynamic and interpersonal signal which changes over time depending on the social situation and communicative purpose. However, there is no cognitive model of gaze processing that takes into account these interactive factors. We believe that in the current context of social cognitive research, which has a strong focus on ecologically valid approaches (Risko et al., 2016; Schilbach et al., 2013), there is an urgent need to build up a cognitive model of eye gaze in live interactions. With this aim, in the next section we introduce the Interpersonal Gaze Processing model, which tries to makes sense of gaze dynamics during face-to-face interactions.

5 Interpersonal Gaze Processing model: active sensing and social signalling

The dual function of the eyes means that our gaze both gains information form the environment and signals information to others. Early cognitive research already described how the visual system gains information from the environment in non-social contexts (Itti & Koch, 2001; Koch & Ullman, 1985). However, to our knowledge there is no cognitive model of gaze processing in social contexts. Here we draw on two distinct frameworks, from motor control (active sensing; Yang, Wolpert, & Lengyel, 2016) and from animal communication (signalling theory; Grafen, 1990; Zahavi, 1975), to introduce the Interpersonal Gaze Processing model. This model considers how these two frameworks can be combined in the domain of social gaze to take into account both its sensing and signalling functions. In the following, we describe how active sensing and signalling theory are useful to explain gaze behaviour.

5.1 Active sensing in eye gaze

Active sensing is a key process in our interaction with the world, since it allows our sensors to be directed to the environment in order to extract relevant information (Yang et al., 2016). Gaze behaviour (i.e. deciding where to look) can be considered a form of active sensing in that we choose to move our eyes to specific locations to sample useful information from a visual scene. Since our visual system only gains high-resolution information for items falling in the fovea, the motor system needs to move our eyes to orient the fovea to different locations of interest. Thus, our motor actions shape the quality of the sensory information we sample (Yang et al., 2016).

The active sensing framework provides a mathematical account of how we can sample the world with our eyes to get useful information. Because we can only direct our eyes at one location at a time, each eye movement (i.e. saccade) comes at some opportunity cost. For instance, in Figure 1a, looking at the woman and child on the bottom means we might lose the chance to get information about the house in the centre or the woman and child on the left. Similarly, in Figure 1b, looking at the landscape on the right means we will lose information about the blue car on the left or the speedometer. Active sensing suggests that saccades are planned to maximise the information we sample depending on the goal of the task at hand.

To understand how sampled information is maximised it is useful to consider the concept of saliency maps. A saliency map is "an explicit two-dimensional topographical map that encodes stimulus conspicuity, or saliency, at every location in the visual scene" (Itti & Koch, 2001). It results from the combination of different topographical or feature maps, each representing a single visual feature (Itti, Koch, & Niebur, 1998; Koch & Ullman, 1985; Veale, Hafed, & Yoshida, 2017), such as intensity or colour. A saliency map is a pre-attentive computation, in the sense that at this stage all locations are competing for representation in the visual cortex (Itti & Koch, 2001). Only the location that is most salient will gain further access in downstream visual areas and the oculomotor nerve, and guide the next eye movement so as to deploy attention in that specific location (Itti & Koch, 2001; Kastner & Ungerleider, 2000; Koch & Ullman, 1985; Veale et al., 2017) (see Figure 1c[1-3]).

Early models of saliency maps only included static features of visual scenes (e.g. colour, orientation, intensity, center-surround difference; Itti & Koch, 2001; Koch & Ullman, 1985), but later proposals have suggested saliency maps that also integrate dynamic features (Jeong, Ban, & Lee, 2008; Milanese, Gil, & Pun, 1995). For instance, the integrated saliency map by Jeong and colleagues (2008) considers dynamic features such as rotation, expansion, contraction or planar motion. These dynamic features are especially effective in attracting visual attention, and have been associated with an alerting mechanism that rapidly detects moving objects (Milanese et al., 1995). Both static and dynamic features generate a bottom-up bias on the saliency map.

However, saliency maps can also be modelled by a top-down bias emerging from affective features (e.g. preference or dislike for the visual stimuli; Itti et al., 1998; Jeong et al., 2008; Olshausen, Anderson, & Van Essen, 1993; Tsotsos et al., 1995; Veale et al., 2017) (see Figure 1c[4]). Affective features are mainly associated with the goal of the task at hand, and are integrated with bottom-up information in associative visual areas (extrastriate cortex) (Veale et al., 2017). For instance, as shown on Figure 1d, different search goals will model different priority maps derived from the same saliency map. Recent evidence has also found that when participants view social naturalistic scenes low-level salient features are less important, and participants primarily fixate on the faces and eyes of people in the scene (End & Gamer, 2017; Nasiopoulos, Risko, & Kingstone, 2015; Rubo & Gamer, 2018). This suggests that there is an implicit preferential bias to attend to others in social scenes to obtain information about them (Nasiopoulos et al., 2015). In the same way that non-social task goals (e.g. search for the cell phone) model different priority maps, implicit social task goals (e.g. identify feelings of an actress in a movie) will model different sensing maps. This top-down bias is particularly important in the context of active sensing, since the task goal will modify the reward value of each location in the visual scene and, in turn, determine which information needs to be maximized (Jeong et al., 2008; Yang et al., 2016).

Active sensing provides a useful framework to understand how eye movements are planned to process non-social stimuli (e.g. objects or landscapes), as well as social stimuli in pictures or videos. In both cases, the saccade planner combines bottom-up and top-down features in a priority or sensing map to maximise information relevant for the task and decide where gaze is next directed (Yang et al., 2016). However, in the case of face-to-face interactions, our gaze not only needs to maximise the information gained but also send signals to another person (i.e. dual function of eyes; Argyle & Cook, 1976; Gobel et al., 2015; Risko et al., 2016).

5.2 Social signalling and eye gaze

Research on animal communication has explored in detail the question of what behaviour counts as a social signal and what message (if any) is sent (Stegmann, 2013). A cue is a behaviour or feature that can be used by another creature to guide its behaviour; for example, mosquitos use the increased carbon dioxide in exhaled air as a cue to find people to bite, but there is no benefit here to those sending the cue. In contrast, the mating call of a bird that attracts a mate acts as a signal because it benefits both sender and receiver (Stegmann, 2013). A key way to distinguish between these is that signals are sent with the purpose of having an effect on another individual, which means they are more likely to be sent when they can be received. In the context of human interaction, signals are sent when another person is present (an audience effect) but should not be sent when a person acts alone. A stronger definition of explicit and deliberate signalling might require sending a signal repeatedly or elaborating on the signal until it is received. However, based on animal communication models (Stegmann, 2013), we will use a minimal definition of communication where signals are sent implicitly.

As described above, our eyes can act both as a cue to our current thoughts (e.g. if I am looking at my watch, I want to know the time) and as a signal to another person (e.g. I ostentatiously stare at my watch to signal to my friend that we must leave the party) (Argyle & Cook, 1976; Gobel et al., 2015; Risko et al., 2016). As Watzlavick's axiom "one cannot not communicate" (Watzlawick et al., 1967) suggests, even in a waiting room where two people are not intended to communicate and avoid engaging in eye contact, they are sending a signal that means "I do not want to interact with you" (Foulsham et al., 2011). This means that, in line with signalling theory, in face-to-face interactions our eye movements are constantly planned so as to send signals to others, and not just to gain information from the world. We propose that the signalling function of gaze creates a signalling map in the brain equivalent to the sensing map generated by the sensing function. In the same way that sensing maps show where to look to gain information, we hypothesize that signalling maps are computed in the brain to show where to look to send an appropriate signal to another person. In the following, we argue that the signalling map is computed by taking into account three key factors: communicative purpose, other's gaze direction, and coordination with other social signals.

First, the value of each gaze target in the signalling map will vary depending on the communicative purpose, that is, the type of message we wish to send. Just as saliency maps incorporate the task goal to create priority or sensing maps of visual attention, signalling maps need to take into account the communicative purpose. Imagine a waiting room with two people, where one person (A) wants to engage in an interaction, but the other person (B) does not. For person A, the optimal signalling behaviour is to direct gaze to person B in order to send the message "I want to engage in an interaction with you". However, person B should avert gaze to efficiently signal "I do not want to interact with you". Thus, the signalling map will be different for person A and B, depending on the message they want to send.

Second, the signalling map will change according to the direction of the other person's gaze. The relationship between other's gaze direction and the signalling map lies in the fact that my signal will be received depending on whether the other person is gazing at us or not. Let's go back to the case of the waiting room with person A and B. For person A, who wishes to interact with person B, the optimal signalling behaviour is to direct her gaze when person B is also looking at her, in order to disclose interest in the interaction. Directing her gaze when B is not looking has little benefit, because the signal will not be received. Equally, for person B the optimal signalling behaviour is to avert gaze specifically when A is looking at her. This illustrates how the values associated with each location in the signalling map changes on a moment-by-moment basis, contingent on the gaze direction of the other person and in relation to communicative purpose.

Finally, the signalling map depends on the need to coordinate with other social signals that are sent in multimodal communication, such as speech or gestures (Hirai & Kanakogi, 2018; Ho et al., 2015; Holler et al., 2018; Jack & Schyns, 2015; Trujillo et al., 2018; Vigliocco et al., 2014). This is particularly relevant for explicit communicative encounters. Imagine that person A and B in the waiting room are now engaged in a lively conversation: to signal interest in keeping the conversation going, the choice of direct or averted gaze will vary depending on the role of each partner in the conversation, as well as the time-course of speech itself. For instance, when person A starts speaking, she may avert gaze every now and then to signal she still has more things to say (Ho et al., 2015; Kendon, 1967). While person B is listening, her gaze may be directed towards person A in order to signal interest in what A is saying (Ho et al., 2015; Kendon, 1967). However, when person A is finishing the utterance, she may look towards person B to signal that she can take the floor (Ho et al., 2015; Kendon, 1967). Thus, the coordination with other social signals also modulates the optimal location in the signalling map on a moment-by-moment basis.

Signalling theory provides a framework to understand how the communicative function of gaze shapes the planning of eye movements during face-to-face interactions. In the following, we propose a model where both active sensing and social signalling are combined to make sense of gaze patterns in human-to-human communication.

5.3 The Interpersonal Gaze Processing model

The Interpersonal Gaze Processing model considers how gaze transitions from one state to the other (i.e. how eye movements are planned) when presented with social stimuli (Figure 2 and 3). This model distinguishes between two situations that differ in the belief in being watched: one where the social stimulus is a picture or video (i.e. cannot see us), and one where the social stimulus is a real person in front of us (i.e. can see us).

In the first case, where the stimulus is a picture or video of another person, there is no need to send a signal because it will not be perceived. Thus, the planning of eye movements only responds to active sensing, which aims to gain maximal information from the stimulus (Yang et al., 2016). The Interpersonal Gaze Processing model considers that gaze patterns derived from active sensing correspond to baseline gaze behaviour. When the goal is to get social information from the picture or video (e.g. what is the man in the picture feeling?) gaze patterns will be mostly influenced by sensing maps (see Figure 2 and 3a). This baseline sensing map reveals how people use gaze to gain different types of social information during interactions. For example, in a noisy environment where it is hard to hear, they will look more to the centre of the face to help with speech comprehension; conversely, to recognise emotions they will look more to the eyes (Buchan, Paré, & Munhall, 2007, 2008; Lewkowicz & Hansen-Tift, 2012). This also demonstrates how task goals (e.g. speech comprehension or emotion recognition) translate in different eye movements depending on the information that needs to be maximised.

In the second case, where the stimulus is a real person in front of us, our eyes will be sending a signal to the other person. Here, the Interpersonal Gaze Processing model proposes that gaze patterns result from a trade-off between sensing maps and signalling maps (see Figure 2 and 3b). This means that the planning of eye movements combines the maximal gain of information from a particular location in the sensing map (e.g. eyes of the other person), together with the optimal benefit of gazing to that location in the signalling map. Figure 4 illustrates how different possible gaze targets on the face of the man can provide various types of information to the woman (sensing function), but also can send different signals to the man (signalling function). Comparing baseline gaze behaviour in a video to gaze behaviour in a matched real-life interaction, can provide a measure of the signalling components of eye gaze. For example, some studies show that people direct gaze to the eyes of a stranger in a video, but not to the eyes of a live stranger: this indicates that averting gaze from the real person has a meaningful signalling value, since it expresses no desire to affiliate with the stranger and reduces the intensity of the interaction (Argyle & Dean, 1965; Foulsham et al., 2011; Laidlaw et al., 2011). This example considers the case of watching a stranger with a rather neutral face, but another interesting situation is that where partners show emotional facial expressions. Although this scenario has not yet been tested, it would give further insight on how sensing maps and signalling maps are integrated during gaze planning. Moreover, we acknowledge there may also be changes in arousal in association with being watched by a live person (Lyyra, Myllyneva, & Hietanen, 2018; Myllyneva & Hietanen, 2015; Zajonc, 1965), but these effects are not included in our model because of non-specific predictions on sensing and signalling maps.

Thus, the Interpersonal Gaze Processing model proposes that, moment-by-moment, the gaze control systems in the brain must evaluate both the information gained and the signalling potential of a saccade, to determine where to look next. This model and other theories of the audience effect (i.e. Watching Eyes model and reputation management theory) are linked because they are all modulated by the belief in being watched. The Watching Eyes model and reputation management theory explain how the presence of an *observer* modulates an individual's self- and other-focused cognitive processing, but they do not attempt to explain the dynamics of eye gaze in live *communicative exchanges*. By contrast, the Interpersonal Gaze Processing model places special emphasis on communicative purpose and coordination with other social signals (e.g. other's gaze direction, speech, facial expressions): while comunicative purpose (together with the belief in being watched) is key to define the signalling map, the coordination with other social signals modulates this map on a moment-by-moment basis. Future studies on gaze processing should try to elucidate how each of these factors modulates gaze sensing and signalling during communication, as well as if and how these maps are computed and integrated in the brain.

6 Gaze processing in autism

Autism Spectrum Condition (ASC) is a developmental condition characterized by difficulties in interpersonal interaction and communication, as well as the presence of restricted and repetitive patterns of behavior (*Diagnostic and Statistical Manual of Mental Disorders 5th Ed.*, 2013). Since eye gaze has a critical role in regulating social interactions and enabling successful communicative exchanges, it is not surprising that the presence of abnormal gaze patterns is one of the most used diagnostic criteria for ASC from early infancy (Zwaigenbaum et al., 2005). Although research into gaze behaviour in autistic adults has identified some general patterns, it has also yielded some inconsistent findings: some studies using pictures and videos suggest that they avoid looking at the eyes, whereas others indicate that they have typical gaze patterns (Chita-Tegmark, 2016; Falck-Ytter & Von Hofsten, 2011; Frazier et al., 2017). Some of these discrepancies may be a consequence of the wide spectrum in autistic individuals, but in line with the second-person neuroscience framework (Schilbach, 2016; Schilbach et al., 2013), it has been suggested that they could also be a consequence of the lack of experimental paradigms for studying eye gaze in real social interactions (Chevallier et al., 2015; Drysdale, Moore, Furlonger, & Anderson, 2018; Von dem Hagen & Bright, 2017). Moreover, a recent qualitative study highlights that self-declared autistic adolescents and adults struggle with the appropriate use and timing of eye gaze during face-to-face interactions (Trevisan, Roberts, Lin, & Birmingham, 2017). These findings suggest that to fully understand autistic social cognition it is necessary to examine how they process social signals in real dynamic interactions.

6.1 Audience effects in autism

We have previously presented two distinct cognitive theories to explain audience effects: the Watching Eyes model (Conty et al., 2016) and reputation management theory (Emler, 1990; Resnick et al., 2006; Tennie et al., 2010). Both theories involve mentalizing and distinction between self-beliefs and other-beliefs, either to process the perceptual state of the observer (Teufel, Fletcher, et al., 2010) or to further infer what the observer thinks of us (Cage, 2015; Izuma et al., 2010). This means that mentalizing is a key cognitive component of audience effects (Hamilton & Lind, 2016). Difficulties in processing mental states of others is one of the hallmarks of autism: they have trouble inferring beliefs and intentions of other people (Happé, 1994; White, Happé, Hill, & Frith, 2009), as well as attributing a social meaning to eye gaze (Baron-Cohen et al., 1997), especially when they need to do so spontaneously (Senju, Southgate, White, & Frith, 2009). Thus, impaired mentalizing in autistic people implies that being watched will elicit less self-related processing and reputation management, and they will show reduced audience effects (Hamilton & Lind, 2016).

To our knowledge, no studies have directly tested the Watching Eyes model on autistic individuals, but instead have looked at differences in self-referential processing between typical and autistic populations. Lombardo and colleagues (Lombardo et al., 2007) used a task measuring self-referential memory and found that high-functioning autistic individuals as well as with Asperger Syndrome (two similar subgroups within the autism spectrum) had smaller self-referential bias compared to typical individuals. Moreover, from early infancy autistic individuals show reduced orienting to their name, which is a salient stimulus uniquely related to oneself (Nadig, Lee, Singh, Kyle, & Ozonoff, 2010; Werner, Dawson, Osterling, & Dinno, 2000). These studies suggest that autistic people have a general impairment in processing self-related information as distinct from other-related information, already when they are in a non-interactive environment. Interestingly, it has recently been suggested that autistic people might have a narrower cone of direct gaze (i.e. the range of gaze directions that an individual judges as being directed to oneself), which means that they might be less likely to perceive that an observer is watching them (Gianotti, Lobmaier, Calluso, Dahinden, & Knoch, 2018). Thus, a plausible prediction is that autistic individuals will fail to process self-relevant signals in interactive environments, such as the belief in being watched (Conty et al., 2016). Studies directly testing effects of being watched on self-referential processing will be needed to clarify this question.

In contrast, a body of research has investigated reputation management in autism. Using the donation task, it has been found that the frequency of donations of autistic participants is not affected by the presence or absence of a confederate who is watching them (Cage et al., 2013; Izuma et al., 2011). It is worth noting that Izuma et al. (2011) found a social facilitation effect in autistic participants on a perceptual task, which indicates that autistic people have specific difficulties with reputation management processes. Cage and colleagues (2013) further showed that, while typical participants donated more frequently when the observer could reciprocate, autistic participants had reduced expectation of reciprocity. Moreover, autistic children do not engage in flattery behaviour towards others (Chevallier, Molesworth, & Happé, 2012) and do not use strategic self-promotion when describing themselves in front of an audience (Scheeren, Begeer, Banerjee, Meerum Terwogt, & Koot, 2010). These findings demonstrate that autistic people are less inclined to manipulate beliefs of observers to maintain their reputation, either due to mentalizing impairments (U. Frith, 2012) or to social motivation deficits (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2013).

However, it is not clear how social norms of eye gaze (i.e. civil inattention; (Goffman, 1963) are implemented in autism, since no study has directly contrasted gaze patterns of autistic individuals in live versus pre-recorded non-communicative interactions. A study by von dem Hagen and colleagues approached this question in typical individuals with high and low autistic traits (Von dem Hagen & Bright, 2017, Experiment 1). Participants were shown videos of a confederate and were deceived to believe that the videos were either pre-recorded or a live video-feed. They found that people with low autistic traits decreased the amount of gaze directed to the face of the confederate in the live video-feed condition, but no reduction was found in the group with higher autism traits. This finding indicates some degree of insensitivity to the belief in being watched and, consequently, to social norms associated with social behaviour towards strangers. However, it remains to be seen whether these findings are true for individuals with an ASC diagnosis.

6.2 Interpersonal dynamics of gaze in autism

Few studies have looked at how gaze patterns differ between typical and autistic groups during interactions with communicative purpose, and the evidence is mixed. For instance, when asked to actively engage in an interaction (Q&A task) over a video-feed, individuals with high autistic traits looked less towards the face of the confederate than individuals with low autistic traits (Von dem Hagen & Bright, 2017, Experiment 2). Using a similar Q&A task in a face-to-face interaction, it was found that high amount of autistic traits was not associated to reduced looking to the face, but to reduced visual exploration (Vabalas & Freeth, 2016). However, in a study testing a sample with autism diagnosis, no differences in visual exploration were found between typical and autistic groups

(Freeth & Bugembe, 2018). It is worth noting that in all these studies they found no between-group differences in gaze patterns during speaking and listening periods (i.e. typical/low autistic traits and high autistic traits behave equally), which suggests that to some extent social functions of gaze are preserved in autism (e.g. regulating turn-taking during conversation).

We previously argued that in communicative encounters (direct) eye gaze needs to coordinate with other verbal and non-verbal signals, within and between agents, to successfully exchange information (Bavelas et al., 2002; Trujillo et al., 2018). Several studies indicate that autistic individuals do not use direct gaze as a signal to coordinate intra- and inter-personal social behaviour in the same way that typical participants do. Using non-interactive stimuli, it has been shown that autistic adults do not follow gaze after eye contact as much as typical participants (Böckler, Timmermans, Sebanz, Vogeley, & Schilbach, 2014). Moreover, while in typical individuals direct gaze reduces reaction times to generate an action (Schilbach, Eickhoff, Cieslik, Kuzmanovic, & Vogeley, 2012) or to mimic an action (Forbes, Wang, & Hamilton, 2017), this effect is not found in ASC. Similarly, when participants interact with a virtual avatar that displays contingent gaze patterns, autistic children show less gaze following (Little, Bonnar, Kelly, Lohan, & Rajendran, 2017) and individuals with high autistic traits engage in less facial mimicry following joint attention than individuals with low autistic traits (Neufeld, Ioannou, Korb, Schilbach, & Chakrabarti, 2016). These findings suggest that reduced coordination between eye gaze and other social behaviour may have an impact on the successful progression of the interaction.

A reason why autistic people show poor coordination of social behaviour could stem from difficulties in appropriately adjusting gaze to the dynamics of communication. It has been found that infants at high risk for ASC alternate less between initiating and responding to joint attention compared to infants at low risk (Thorup, Nyström, Gredebäck, Bölte, & Falck-Ytter, 2018), and that they preferentially orient towards a person that always responds in the same way over a person that can show variable responses (Vernetti, Senju, Charman, Johnson, & Gliga, 2017). This means that, since early infancy, individuals at high risk for ASC experience less dynamic social contexts and less variety in gaze-contingent events. Using a gaze-contingent eye-tracking paradigm with virtual avatars, Caruana and colleagues (Caruana et al., 2017) have found that autistic adults are less accurate and take a longer time than typical adults to respond to joint attention. In line with this, Freeth & Bugembe (2018) have found that when a confederate directly gazes at participants during a Q&A task, autistic adults look less at the confederate’s face than typical adults. These findings suggest that difficulties in adjusting eye contact make it hard for autistic individuals to keep pace with rapid and spontaneous face-to-face interactions.

It has been suggested that a lack of exposure to contingent eye gaze in infancy can impact the specialisation of brain areas related to gaze processing (Vernetti et al., 2018). Indeed, a study using live video-feed found that some regions in the social neural network (superior temporal sulcus and dorsomedial prefrontal cortex) are equally engaged during periods of joint attention and periods of no joint attention in ASC (Redcay et al., 2012). This is corroborated by previous studies using non-interactive stimuli, where they found abnormal activation of the social neural network (e.g. superior temporal sulcus, right temporo-parietal junction) when autistic adults processed social information conveyed by eye gaze (Georgescu et al., 2013; Pelphrey, Morris, & McCarthy, 2005; Philip et al., 2012; Zilbovicius et al., 2006). Moreover, a hyperscanning study using live video-feed (Tanabe et al., 2012) found that inter-brain coherence (in frontal regions) during eye contact was lower in autistic-typical dyads compared to typical-typical dyads, which might reflect difficulties in processing and integrating social signals in ASC. Thus, these studies suggest that atypical intra- and inter-individual patterns of neural activity in response to direct gaze may underlie difficulties in detecting, processing and sending social signals in autism.

Overall, these findings indicate that autistic individuals have difficulties with social dynamics of gaze in real interactions. However, current research is not enough to clearly distinguish which cognitive components of eye gaze processing are most disrupted in autism. In this sense, the Interpersonal Gaze Processing model (Figure 2 and 3) provides common ground where studies manipulating various gaze-related factors can come together. We previously suggested that comparing gaze patterns in a video versus a matched real-life interaction provides a measure of the signalling components of eye gaze. If autistic people do not engage in social signalling, the Interpersonal Gaze Processing model predicts that their gaze patterns in live and video conditions should be similar, which is in line with recent evidence (Von dem Hagen & Bright, 2017). Future research should try to systematically study which factors modulating gaze signalling make interpersonal gaze processing challenging in autism.

7 Conclusions

Natural social interactions are characterised by complex exchanges of social signals, so achieving successful communication can be challenging. This paper aimed to review research manipulating three key factors that modulate eye gaze processing during social interactions: the presence of an interacting partner who can perceive me, the existence of communicative purpose, and the development of interpersonal and temporal dynamics.

Current findings indicate that the belief in being watched has a strong impact on other-focused social cognition (both on prosocial behaviour and social norms of eye gaze), but evidence is less clear for self-focused cognition: future studies should clarify to what extent being watched affects different forms of self-related processes. We also find that, to achieve successful communication, eye gaze needs to coordinate with verbal and non-verbal social signals, both within and between interacting partners. We propose the Interpersonal Gaze Processing model as a framework where gaze sensing and signalling are combined to determine where the eyes will look next in a live interaction. In this model, the belief in being watched and the communicative purpose of the interaction are key to define the gaze signalling map, while the contingencies between different signalling modalities (e.g. gaze, speech) are critical in changing this map on a moment-by-moment basis. Systematic manipulation of these factors could help elucidate how they relate to each other to enable successful communicative encounters, as well as how signalling maps are computed in the brain.

Finally, research on autistic individuals reveals that they are less sensitive to the belief in being watched, but more studies are needed to clarify how the presence of an audience impacts self-related processing in autism. Although evidence on interpersonal dynamics is mixed, it is agreed that autistic individuals have difficulties with social dynamics of eye gaze during real interactions. We argue that the Interpersonal Gaze Processing model provides a framework for future studies to systematically characterise which aspects of gaze communication are most challenging for autistic people.

8 Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

9 Author Contributions

RC wrote the initial draft of the manuscript and prepared the figures. AH made critical revisions to the original draft. RC and AH approved the final version of the manuscript.

10 Funding

This work was supported by University College London Open Access Team. RC acknowledges financial support from "la Caixa" Foundation (ID 100010434, grant code LCF/BQ/EU16/11560039).

11 References

Argyle, M., & Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge, UK: Cambridge University Press.

Argyle, M., & Dean, J. (1965). Eye-Contact, Distance and Affilitation. *Sociometry*, *28*(3), 289–304. http://doi.org/10.1111/j.2044-8295.1970.tb01257.x

Baltazar, M., Hazem, N., Vilarem, E., Beaucousin, V., Picq, J. L., & Conty, L. (2014). Eye contact elicits bodily self-awareness in human adults. *Cognition*, *133*(1), 120–127.

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

Baron-Cohen, S., & Cross, P. (1992). Reading the eyes: evidence for the role of perception in the development of a Theory of Mind. *Mind & Language*, *7*(1–2), 172–186.

Baron-Cohen, S., Wheelwright, S. J., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition*, *4*(3), 311–331.

Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener Responses as a Collaborative Process: The Role of Gaze. *Journal of Communication*, (September), 566–580.

Bayliss, A. P., Murphy, E., Naughtin, C. K., Kritikos, A., Schilbach, L., & Becker, S. I. (2012). Gaze leading: Initiating simulated joint attention influences eye movements and choice behavior. *Journal of Experimental Psychology: General*, *142*(1), 76–92. http://doi.org/10.1037/a0029286

Beattie, G. W. (1981). A further investigation of the cognitive inference hypothesis of gaze patterns during conversation. *British Journal of Social Psychology*, *20*, 243–248.

Blascovich, J., Mendes, W. B., Hunter, S. B., & Salomon, K. (1999). Social "faciliation" as challenge and threat. *Journal of Personality and Social Psychology*, *77*(1), 68–77.

Böckler, A., Timmermans, B., Sebanz, N., Vogeley, K., & Schilbach, L. (2014). Effects of observing eye contact on gaze following in high-functioning autism. *Journal of Autism and Developmental Disorders*, *44*(7), 1651–1658. http://doi.org/10.1007/s10803-014-2038-5

Bond, C. F. J. (1982). Social Facilitation: A Self-Presentational View. *Journal of Personality and Social Psychology*, *42*(6), 1042–1050.

Bradley, A., Lawrence, C., & Ferguson, E. (2018). Does observability affect prosociality? *Proceedings of the Royal Society B: Biological Sciences*, *285*(20180116). http://doi.org/10.1098/rspb.2018.0116

Buchan, J. N., Paré, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience*, *2*(1), 1–13. http://doi.org/10.1080/17470910601043644

Buchan, J. N., Paré, M., & Munhall, K. G. (2008). The effect of varying talker identity and listening conditions on gaze behavior during audiovisual speech perception. *Brain Research*, *1242*, 162–171. http://doi.org/10.1016/j.brainres.2008.06.083

Cage, E. A. (2015). *Mechanisms of social influence: Reputation management in typical and autistic individuals.*

Cage, E. A., Pellicano, E., Shah, P., & Bird, G. (2013). Reputation management: Evidence for ability but reduced propensity in autism. *Autism Research*, *6*(5), 433–442.

Cameron, O. G. (2001). Interoception: The inside story - A model for psychosomatic processes. *Psychosomatic Medicine*, *63*(5), 697–710. http://doi.org/10.1097/00006842-200109000-00001

Caruana, N., Brock, J., & Woolgar, A. (2015). A frontotemporoparietal network common to initiating and responding to joint attention bids. *NeuroImage*, *108*, 34–46. http://doi.org/10.1016/j.neuroimage.2014.12.041

Caruana, N., Stieglitz Ham, H., Brock, J., Woolgar, A., Kloth, N., Palermo, R., & McArthur, G. (2017). Joint attention difficulties in autistic adults: An interactive eye-tracking study. *Autism*, 136236131667620. http://doi.org/10.1177/1362361316676204

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2013). The social motivation theory of autism. *Trends in Cognitive Neuroscience*, *16*(4), 231–239. http://doi.org/10.1016/j.tics.2012.02.007.The

Chevallier, C., Molesworth, C., & Happé, F. (2012). Diminished social motivation negatively impacts reputation management: Autism spectrum disorders as a case in point. *PLoS ONE*, *7*(1), 1–6. http://doi.org/10.1371/journal.pone.0031107

Chevallier, C., Parish-Morris, J., McVey, A., Rump, K. M., Sasson, N. J., Herrington, J. D., & Schultz, R. T. (2015). Measuring social attention and motivation in Autism Spectrum Disorder using eye-tracking: stimulus type matters. *Autism Research*, *8*(5), 620–628.

Chita-Tegmark, M. (2016). Research in Developmental Disabilities Review article Social attention in ASD : A review and meta-analysis of eye-tracking studies. *Research in Developmental Disabilities*, *48*, 79–93. http://doi.org/10.1016/j.ridd.2015.10.011

Conty, L., George, N., & Hietanen, J. K. (2016). Watching Eyes effects: when others meet the self. *Consciousness and Cognition*, *45*, 184–197.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294.

Cramer, D. (1994). Books Reconsidered: Gaze and Mutual Gaze. *British Journal of Psychiatry*, *165*, 848–850. http://doi.org/10.1017/S0007125000073980

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. http://doi.org/10.1016/j.tics.2009.01.005

Cummins, F. (2012). Gaze and blinking in dyadic conversation: A study in coordinated behaviour among individuals. *Language and Cognitive Processes*, *27*(10), 1525–1549. http://doi.org/10.1080/01690965.2011.615220

Davis, D., & Brock, T. C. (1975). Use of first person pronouns as a function of increased objective self-awareness and performance feedback. *Journal of Experimental Social Psychology*, *11*(4), 381–388.

Diagnostic and Statistical Manual of Mental Disorders 5th Ed. (2013). Washington, DC: American Psychological Association.

Drysdale, B. M., Moore, D. W., Furlonger, B. E., & Anderson, A. (2018). Gaze Patterns of Individuals with ASD During Active Task Engagement: a Systematic Literature Review. *Review Journal of Autism and Developmental Disorders*, *5*(1), 1–14. http://doi.org/10.1007/s40489-017-0119-z

Duncan, S., & Fiske, D. W. (1977). *Face-to-face interaction: Research, methods, and theory.* (L. Erlbaum). New Jersey.

Edwards, S. G., Stephenson, L. J., Dalmaso, M., & Bayliss, A. P. (2015). Social orienting in gaze leading: A mechanism for shared attention. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1812), 1–8. http://doi.org/10.1098/rspb.2015.1141

Efran, J. S. (1968). Looking for Approval: Effects on Visual Behavior of Approbation from Persons Differing in Importance. *Journal of Personality and Social Psychology*, *10*(1), 21–25.

Efran, J. S., & Broughton, A. (1966). Effect of expectancies for social approval on visual behavior. *Journal of Personality and Social Psychology*, *4*(1), 103–107. http://doi.org/10.1037/h0023511

Ellyson, S. L., Dovidio, J. F., & Fehr, B. J. (1981). Visual Behavior and Dominance in Women and Men. In C. Mayo & N. M. Henley (Eds.), *Gender and Nonverbal Behavior* (pp. 63–81). New York, NY: Springer-Verlag. http://doi.org/10.1007/978-1-4612-5953-4

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, *24*(6), 581–604.

Emler, N. (1990). A social psychology of reputation. *European Review of Social Psychology*, *1*(1), 171–193.

End, A., & Gamer, M. (2017). Preferential processing of social features and their interplay with physical saliency in complex naturalistic scenes. *Frontiers in Psychology*, *8*, 1–16. http://doi.org/10.3389/fpsyg.2017.00418

Falck-Ytter, T., & Von Hofsten, C. (2011). *How special is social looking in ASD: A review*. *Progress* *in Brain Research* (1st ed., Vol. 189). Elsevier B.V. http://doi.org/10.1016/B978-0-444-53884-0.00026-9

Filiz-Ozbay, E., & Ozbay, E. Y. (2014). Effect of an audience in public goods provision. *Experimental Economics*, *17*(2), 200–214. http://doi.org/10.1007/s10683-013-9363-y

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 1338–1349.

Forbes, P. A. G., Wang, Y., & Hamilton, A. F. de C. (2017). STORMy Interactions: Gaze and the Modulation of Mimicry in Adults on the Autism Spectrum. *Psychonomic Bulletin & Review*, *24*(2), 529–535. http://doi.org/10.3758/s13423-016-1136-0

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, *51*(17), 1920–1931.

Frädrich, L., Nunnari, F., Staudte, M., & Heloir, A. (2018). (Simulated) listener gaze in real-time spoken interaction. *Computer Animation and Virtual Worlds*, *29*(3–4), 1–11. http://doi.org/10.1002/cav.1831

Frazier, T. W., Strauss, M., Klingemier, E. W., Zetzer, E. E., Hardan, A. Y., Eng, C., & Youngstrom, E. A. (2017). A Meta-Analysis of Gaze Differences to Social and Nonsocial Information Between Individuals With and Without Autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(7), 546–555. http://doi.org/10.1016/j.jaac.2017.05.005

Freeth, M., & Bugembe, P. (2018). Social partner gaze direction and conversational phase; factors affecting social attention during face-to-face conversations in autistic adults? *Autism*, 1–11.

Freeth, M., Foulsham, T., & Kingstone, A. (2013). What Affects Social Attention? Social Presence, Eye Contact and Autistic Traits. *PLoS ONE*, *8*(1). http://doi.org/10.1371/journal.pone.0053286

Fridlund, A. J. (1991). Sociality of Solitary Smiling: Potentiation by an Implicit Audience. *Journal of Personality and Social Psychology*, *60*(2), 229–240.

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. http://doi.org/10.1037/0033-2909.133.4.694.Gaze

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, *50*(4), 531–534. http://doi.org/10.1016/j.neuron.2006.05.001

Frith, U. (2012). Why we need cognitive explanations of autism. *The Quarterly Journal of Experimental Psychology*, *65*(11), 2073–2092.

Fullwood, C., & Doherty-Sneddon, G. (2006). Effect of gazing at the camera during a video link on recall. *Applied Ergonomics*, *37*(2), 167–175. http://doi.org/10.1016/j.apergo.2005.05.003

Furlanetto, T., Becchio, C., Samson, D., & Apperly, I. A. (2016). Altercentric interference in level 1 visual perspective taking reflects the ascription of mental states, not submentalizing. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(2), 158–163.

Gallotti, M., Fairhurst, M. T., & Frith, C. D. (2017). Alignment in social interactions. *Consciousness and Cognition*, *48*, 253–261. http://doi.org/10.1016/j.concog.2016.12.002

Gallup, A. C., Chong, A., & Couzin, I. D. (2012). The directional flow of visual information transfer between pedestrians. *Biology Letters*, *8*(4), 520–522. http://doi.org/10.1098/rsbl.2012.0160

Geen, R. G. (1985). Evaluation apprehension and response withholding in solution of anagrams. *Personality and Individual Differences*, *6*(3), 293–298. http://doi.org/10.1016/0191-8869(85)90052-2

Georgescu, A. L., Kuzmanovic, B., Schilbach, L., Tepest, R., Kulbida, R., Bente, G., & Vogeley, K. (2013). Neural correlates of "social gaze" processing in high-functioning autism under systematic variation of gaze duration. *NeuroImage: Clinical*, *3*, 340–351.

Gianotti, L. R. R., Lobmaier, J. S., Calluso, C., Dahinden, F. M., & Knoch, D. (2018). Theta resting EEG in TPJ/pSTS is associated with individual differences in the feeling of being looked at. *Social Cognitive and Affective Neuroscience*, *13*(2), 216–223. http://doi.org/10.1093/scan/nsx143

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

Goffman, E. (1963). *Behavior in Public Places*. Simon and Schuster.

Grafen, A. (1990). Biological Signals as Handicaps. *Journal of Theoretical Biology*, *144*, 517–546. http://doi.org/10.1016/S0022-5193(05)80088-8

Grynszpan, O., Martin, J. C., & Fossati, P. (2017). Gaze leading is associated with liking. *Acta Psychologica*, *173*, 66–72. http://doi.org/10.1016/j.actpsy.2016.12.006

Guala, F., & Mittone, L. (2010). Paradigmatic experiments: The Dictator Game. *Journal of Socio-Economics*, *39*(5), 578–584. http://doi.org/10.1016/j.socec.2009.05.007

Hamilton, A. F. de C., & Lind, F. (2016). Audience effects: what can they tell us about social neuroscience, theory of mind and autism? *Culture and Brain*, *4*(2), 159–177. http://doi.org/10.1007/s40167-016-0044-5

Happé, F. (1994). An advanced test of Theory of Mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, *24*(2), 129–154. Retrieved from http://link.springer.com/10.1007/BF02172093%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/8040158

Harrigan, J. A. (1985). Listener's Body Movements and Speaking Turns. *Communication Research*, *12*(2), 233–250.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-Brain coupling: A mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, *16*(2), 114–121. http://doi.org/10.1016/j.tics.2011.12.007

Hazem, N., George, N., Baltazar, M., & Conty, L. (2017). I know you can see me: Social attention influences bodily self-awareness. *Biological Psychology*, *124*, 21–29.

Hietanen, J. O., & Hietanen, J. K. (2017). Genuine eye contact elicits self-referential processing. *Consciousness and Cognition*, *51*, 100–115.

Hirai, M., & Kanakogi, Y. (2018). Communicative Hand-Waving Gestures Facilitate Object Learning in Preverbal Infants. *Developmental Science*, (April), e12787. http://doi.org/10.1111/desc.12787

Hirsch, J., Zhang, X., Noah, J. A., & Ono, Y. (2017). Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. *NeuroImage*, *157*(January), 314–330. http://doi.org/10.1016/j.neuroimage.2017.06.018

Ho, S., Foulsham, T., & Kingstone, A. (2015). Speaking and listening with the eyes: Gaze signaling during dyadic interactions. *PLoS ONE*, *10*(8), 1–18. http://doi.org/10.1371/journal.pone.0136905

Holler, J., Kendrick, K. H., & Levinson, S. C. (2018). Processing language in face-to-face conversation: Questions with gestures get faster responses. *Psychonomic Bulletin and Review*, *25*(5), 1900–1908. http://doi.org/10.3758/s13423-017-1363-z

Itti, L., & Koch, C. (2001). Computational Modelling of Visual Attention. *Nature Reviews Neuroscience*, *2*(March), 194–203.

Itti, L., Koch, C., & Niebur, E. (1998). A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259. http://doi.org/10.1109/34.730558

Izuma, K. (2012). The social neuroscience of reputation. *Neuroscience Research*, *72*(4), 283–288.

Izuma, K., Matsumoto, K., Camerer, C. F., & Adolphs, R. (2011). Insensitivity to social reputation in autism. *Proceedings of the National Academy of Sciences*, *108*(42), 17302–17307.

Izuma, K., Saito, D. N., & Sadato, N. (2009). Processing of the incentive for social approval in the ventral striatum during charitable donation. *Journal of Cognitive Neuroscience*, *22*(4), 621–631.

Izuma, K., Saito, D. N., & Sadato, N. (2010). The roles of the medial prefrontal cortex and striatum in reputation processing. *Social Neuroscience*, *5*(2), 133–147.

Jack, R. E., & Schyns, P. G. (2015). The Human Face as a Dynamic Tool for Social Communication. *Current Biology*, *25*(14), R621–R634. http://doi.org/10.1016/j.cub.2015.05.052

Jeong, S., Ban, S. W., & Lee, M. (2008). Stereo saliency map considering affective factors and selective motion analysis in a dynamic environment. *Neural Networks*, *21*(10), 1420–1430. http://doi.org/10.1016/j.neunet.2008.10.002

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the Assumptions of Economics. *The Journal of Business*, *59*(4).

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of Visual Attention in the Human Cortex.

Annual Reviews Neuroscience, *23*, 315–341.

Kendon, A. (1967). Some Functions of Gaze-Direction in Social Interaction. *Acta Psychologica*, *26*, 22–63.

Kendrick, K. H., & Holler, J. (2017). Gaze Direction Signals Response Preference in Conversation. *Research on Language and Social Interaction*, *1813*(February), 1–21. http://doi.org/10.1080/08351813.2017.1262120

Kim, K., & Mundy, P. (2012). Joint Attention, Social-Cognition, and Recognition Memory in Adults. *Frontiers in Human Neuroscience*, *6*(June), 1–11. http://doi.org/10.3389/fnhum.2012.00172

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, *100*(1), 78–100.

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of Human Evolution*, *40*(5), 419–435. http://doi.org/10.1006/jhev.2001.0468

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219–227.

Kuhn, G., Tatler, B. W., & Cole, G. G. (2009). You look where I look! Effect of gaze cues on overt and covert attention in misdirection. *Visual Cognition*, *17*(6–7), 925–944. http://doi.org/10.1080/13506280902826775

Lachat, F., Hugueville, L., Lemaréchal, J.-D., Conty, L., & George, N. (2012). Oscillatory Brain Correlates of Live Joint Attention: A Dual-EEG Study. *Frontiers in Human Neuroscience*, *6*(June), 1–12. http://doi.org/10.3389/fnhum.2012.00156

Laidlaw, K., 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.

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences*, *114*(50), 13290–13295.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences*, *109*(5), 1431–1436. http://doi.org/10.1073/pnas.1114783109

Little, G. E., Bonnar, L., Kelly, S. W., Lohan, K. S., & Rajendran, G. (2017). Gaze contingent joint attention with an avatar in children with and without ASD. In *2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics, ICDL-EpiRob 2016* (pp. 15–20). http://doi.org/10.1109/DEVLRN.2016.7846780

Loeb, B. K. (1972). *Mutual eye contact and social interaction and their relationship to affiliation.* University of Montana.

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in austism. *PLoS ONE*, *2*(9).

Lyyra, P., Myllyneva, A., & Hietanen, J. K. (2018). Mentalizing eye contact with a face on a video: Gaze direction does not influence autonomic arousal. *Scandinavian Journal of Psychology*, *59*(4), 360–367. http://doi.org/10.1111/sjop.12452

Macdonald, R. G., & Tatler, B. W. (2018). Gaze in a real-world social interaction: A dual eye-tracking study. *The Quarterly Journal of Experimental Psychology*, *71*(10), 2162–2173. http://doi.org/10.1177/1747021817739221

Markson, L., & Paterson, K. B. (2009). Effects of gaze-aversion on visual-spatial imagination. *British Journal of Psychology*, *100*(3), 553–563. http://doi.org/10.1348/000712608X371762

Milanese, R., Gil, S., & Pun, T. (1995). Attentive mechanisms for dynamic and static scene analysis. *Optical Engineering*, *34*(8), 2428–2434.

Myllyneva, A., & Hietanen, J. K. (2015). The dual nature of eye contact: to see and to be seen. *Social Cognitive and Affective Neuroscience*, *1*, 29–32. http://doi.org/10.1017/CBO9781107415324.004

Nadig, A. S., Lee, I., Singh, L., Kyle, B., & Ozonoff, S. (2010). How does the topic of conversation affect verbal exchange and eye gaze? A comparison between typical development and high-functioning autism. *Neuropsychologia*, *48*(9), 2730–2739. http://doi.org/10.1016/j.neuropsychologia.2010.05.020

Naeem, M., Prasad, G., Watson, D. R., & Kelso, J. A. S. (2012). Electrophysiological signatures of intentional social coordination in the 10-12Hz range. *NeuroImage*, *59*(2), 1795–1803. http://doi.org/10.1016/j.neuroimage.2011.08.010

Nasiopoulos, E., Risko, E. F., & Kingstone, A. (2015). Social Attention, Social Presence, and the Dual Function of Gaze. In A. Puce & B. I. Bertenthal (Eds.), *The Many Faces of Social Attention* (p. 139). New York, NY: Springer.

Neufeld, J., Ioannou, C., Korb, S., Schilbach, L., & Chakrabarti, B. (2016). Spontaneous Facial Mimicry is Modulated by Joint Attention and Autistic Traits. *Autism Research*, *9*(7), 781–789. http://doi.org/10.1002/aur.1573

Nuku, P., & Bekkering, H. (2008). Joint attention: inferring what others perceive (and don't perceive). *Consciousness and Cognition*, *17*(1), 339–349.

Olshausen, B. A., Anderson, C. H., & Van Essen, D. C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *The Journal of Neuroscience*, *13*(11), 4700–4719. http://doi.org/10.1.1.66.2555

Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2005). Neural basis of eye gaze processing deficits in autism. *Brain*, *128*(5), 1038–1048. http://doi.org/10.1093/brain/awh404

Pelphrey, K. A., Viola, R. J., & McCarthy, G. (2004). When Strangers Pass: Processing of Mutual and Averted Social Gaze in the Superior Temporal Sulcus. *Psychological Science*, *15*(9), 598–

603. http://doi.org/10.1111/j.0956-7976.2004.00726.x

Pfeiffer, T., & Nowak, M. A. (2006). All in the game. *Nature*, *441*(June), 583–584. http://doi.org/10.1038/441583a

Pfeiffer, U. J., Timmermans, B., Bente, G., Vogeley, K., & Schilbach, L. (2011). A non-verbal turing test: Differentiating mind from machine in gaze-based social interaction. *PLoS ONE*, *6*(11). http://doi.org/10.1371/journal.pone.0027591

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 and Biobehavioral Reviews*, *37*(10), 2516–2528. http://doi.org/10.1016/j.neubiorev.2013.07.017

Philip, R. C. M., Dauvermann, M. R., Whalley, H. C., Baynham, K., Lawrie, S. M., & Stanfield, A. C. (2012). A systematic review and meta-analysis of the fMRI investigation of autism spectrum disorders. *Neuroscience and Biobehavioral Reviews*, *36*(2), 901–942. http://doi.org/10.1016/j.neubiorev.2011.10.008

Piazza, E. A., Hasenfratz, L., Hasson, U., & Lew-Williams, C. (2018). Infant and adult brains are coupled to the dynamics of natural communication. *BioRxiv*, 359810. http://doi.org/10.1101/359810

Pönkänen, L. M., Alhoniemi, A., Leppänen, J. M., & Hietanen, J. K. (2011). Does it make a difference if I have an eye contact with you or with your picture? An ERP study. *Social Cognitive and Affective Neuroscience*, *6*(4), 486–494.

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. http://doi.org/10.1080/00335558008248231

Redcay, E., Dodell-Feder, D., Mavros, P. L., Kleiner, M., Pearrow, M. J., Triantafyllou, C., … Saxe, R. (2012). Atypical brain activation patterns during a face-to-face joint attention game in adults with Autism Spectrum Disorder. *Human Brain Mapping*, *34*(10), 2511–2523.

Resnick, P., Zeckhauser, R., Swanson, J., & Lockwood, K. (2006). The value of reputation on eBay: A controlled experiment. *Experimental Economics*, *9*(2), 79–101.

Risko, E. F., Laidlaw, K., 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*(May), 1–11.

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

Rogers, S. L., Speelman, C. P., Guidetti, O., & Longmuir, M. (2018). Using dual eye tracking to uncover personal gaze patterns during social interaction. *Scientific Reports*, *8*(1), 1–9. http://doi.org/10.1038/s41598-018-22726-7

Rubo, M., & Gamer, M. (2018). Social content and emotional valence modulate gaze fixations in dynamic scenes. *Scientific Reports*, *8*(1), 1–11. http://doi.org/10.1038/s41598-018-22127-w

Saito, D. N., Tanabe, H. C., Izuma, K., Hayashi, M. J., Morito, Y., Komeda, H., … Sadato, N. (2010). "Stay Tuned": Inter-Individual Neural Synchronization During Mutual Gaze and Joint Attention. *Frontiers in Integrative Neuroscience*, *4*(November), 1–12. http://doi.org/10.3389/fnint.2010.00127

Sandgren, O., Andersson, R., Weijer, J. Van De, Hansson, K., & Sahlén, B. (2012). Timing of gazes in child dialogues: A time-course analysis of requests and back channelling in referential communication. *International Journal of Language and Communication Disorders*, *47*(4), 373–383. http://doi.org/10.1111/j.1460-6984.2012.00151.x

Scheeren, A. M., Begeer, S., Banerjee, R., Meerum Terwogt, M., & Koot, H. M. (2010). Can you tell me something about yourself? *Autism*, *14*(5), 457–473. http://doi.org/10.1177/1362361310366568

Schilbach, L. (2016). Towards a second-person neuropsychiatry. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*, 20150081.

Schilbach, L., Eickhoff, S. B., Cieslik, E. C., Kuzmanovic, B., & Vogeley, K. (2012). Shall we do this together? Social gaze influences action control in a comparison group, but not in individuals with high-functioning autism. *Autism*, *16*(2), 151–162. http://doi.org/10.1177/1362361311409258

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*(4), 393–414.

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., … Vogeley, K. (2010). Minds Made for Sharing: Initiating Joint Attention Recruits Reward-related Neurocircuitry. *Journal of Cognitive Neuroscience*, *22*(12), 2702–2715. http://doi.org/10.1162/jocn.2009.21401

Senju, A., & Hasegawa, T. (2005). Direct gaze captures visuospatial attention. *Visual Cognition*, *12*(1), 127–144.

Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences*, *13*(3), 127–134. http://doi.org/10.1016/j.tics.2008.11.009

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind Eyes: An Absence of Sponteneous Theory of Mind in Asperger Syndrome. *Science*, *219*(August), 883–885.

Silver, I. M., & Shaw, A. (2018). Pint-Sized Public Relations: The Development of Reputation Management. *Trends in Cognitive Sciences*, *22*(4), 277–279. http://doi.org/10.1016/j.tics.2018.01.006

Simmel, G. (1921). Sociology of the Senses: Visual Interaction. In R. E. Park & E. W. Burgess (Eds.), *Introduction to the Science of Sociology* (pp. 356–361). Chicago, IL: University of Chicago Press.

Smith, E. A., & Bird, R. L. B. (2000). Turtle hunting and tombstone opening: public generosity as costly signaling. *Evolution and Human Behavior*, *21*(4), 245–261. http://doi.org/10.1016/S1090-5138(00)00031-3

Stegmann, U. E. (2013). *Animal Communication Theory: Information and Influence*. Cambridge University Press.

Strauss, B. (2002). Social facilitation in motor tasks: A review of research and theory. *Psychology of Sport and Exercise*, *3*(3), 237–256. http://doi.org/10.1016/S1469-0292(01)00019-X

Tanabe, H. C., Kosaka, H., Saito, D. N., Koike, T., Hayashi, M. J., Izuma, K., … Sadato, N. (2012). Hard to "tune in": neural mechanisms of live face-to-face interaction with high-functioning autistic spectrum disorder. *Frontiers in Human Neuroscience*, *6*(September), 1–15. http://doi.org/10.3389/fnhum.2012.00268

Tennie, C., Frith, U., & Frith, C. D. (2010). Reputation management in the age of the world-wide web. *Trends in Cognitive Sciences*, *14*(11), 482–488.

Teufel, C., Alexis, D. M., Clayton, N. S., & Davis, G. (2010). Mental-state attribution drives rapid, reflexive gaze following. *Attention, Perception & Psychophyics*, *72*(3), 695–705.

Teufel, C., Fletcher, P. C., & Davis, G. (2010). Seeing other minds: Attributed mental states influence perception. *Trends in Cognitive Sciences*, *14*(8), 376–382. http://doi.org/10.1016/j.tics.2010.05.005

Thorup, E., Nyström, P., Gredebäck, G., Bölte, S., & Falck-Ytter, T. (2018). Reduced Alternating Gaze During Social Interaction in Infancy is Associated with Elevated Symptoms of Autism in Toddlerhood. *Journal of Abnormal Child Psychology*, 1–15. http://doi.org/10.1007/s10802-017-0388-0

Trevisan, D. A., Roberts, N., Lin, C., & Birmingham, E. (2017). How do adults and teens with self-declared Autism Spectrum Disorder experience eye contact? A qualitative analysis of first-hand accounts. *PLoS ONE*, *12*(11), 1–22.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *The American Journal of Psychology*, *9*(4), 507–533.

Trujillo, J. P., Simanova, I., Bekkering, H., & Özyürek, A. (2018). Communicative intent modulates production and comprehension of actions and gestures: A Kinect study. *Cognition*, *180*(March), 38–51. http://doi.org/10.1016/j.cognition.2018.04.003

Tsotsos, J. K., Culhane, S. M., Kei Wai, W. Y., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence*, *78*(1–2), 507–545. http://doi.org/10.1016/0004-3702(95)00025-9

Vabalas, A., & Freeth, M. (2016). Brief Report: Patterns of Eye Movements in Face to Face Conversation are Associated with Autistic Traits: Evidence from a Student Sample. *Journal of Autism and Developmental Disorders*, *46*(1), 305–314. http://doi.org/10.1007/s10803-015-2546-y

Veale, R., Hafed, Z. M., & Yoshida, M. (2017). How is visual salience computed in the brain? Insights from behaviour, neurobiology and modeling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*, 20160113. http://doi.org/10.1098/rstb.2016.0113

Vernetti, A., Ganea, N., Tucker, L., Charman, T., Johnson, M. H., & Senju, A. (2018). Infant neural sensitivity to eye gaze depends on early experience of gaze communication. *Developmental Cognitive Neuroscience*, *34*(June 2017), 1–6. http://doi.org/10.1016/j.dcn.2018.05.007

Vernetti, A., Senju, A., Charman, T., Johnson, M. H., & Gliga, T. (2017). Simulating interaction: using gaze-contingent eye-tracking to measure the reward value of social signals in toddlers with and without autism. *Developmental Cognitive Neuroscience*. http://doi.org/10.1016/j.dcn.2017.08.004

Vigliocco, G., Perniss, P., Vinson, D., & Vigliocco, G. (2014). Language as a multimodal phenomenon: implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*, 20130292. http://doi.org/10.1098/rstb.2013.0292

Von dem Hagen, E. A. H., & Bright, N. (2017). High autistic trait individuals do not modulate gaze behaviour in response to social presence but look away more when actively engaged in an interaction. *Autism Research*, *10*(2), 359–368.

von Grünau, M., & Anston, C. (1995). The detection of gaze direction: A stare-in-the-crowd effect. *Perception*, *24*(11), 1297–1313. http://doi.org/10.1068/p241297

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, *19*(9), 1395–1407. http://doi.org/10.1016/j.neunet.2006.10.001

Watzlawick, P., Helmick Beavin, J., & Jackson, D. D. (1967). *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies, and Paradoxes*. New York, NY: W. W. Norton & Company.

Werner, E., Dawson, G., Osterling, J., & Dinno, N. (2000). Brief Report: Recognition of Autism Spectrum Disorder Before One Year of Age: A Retrospective Study Based on Home Videotapes. *Journal of Autism and Developmental Disorders*, *30*(2), 157–162. http://doi.org/10.1023/A:1005463707029

White, S., Happé, F., Hill, E., & Frith, U. (2009). Revisiting the Strange Stories: revealing mentalizing impairments in autism. *Child Development*, *80*(4), 1097–1117.

Willemse, C., Marchesi, S., & Wykowska, A. (2018). Robot faces that follow gaze facilitate attentional engagement and increase their likeability. *Frontiers in Psychology*, *9*(FEB), 1–11. http://doi.org/10.3389/fpsyg.2018.00070

Wilms, M., Schilbach, L., Pfeiffer, U. J., Bente, G., Fink, G. R., & Vogeley, K. (2010). It's in your eyes - using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, *5*(1), 98–107. http://doi.org/10.1093/scan/nsq024

Wu, D. W. L., Bischof, W. F., & Kingstone, A. (2013). Looking while eating: The importance of social context to social attention. *Scientific Reports*, *3*(2356). http://doi.org/10.1038/srep02356

Yang, S. C. H., Wolpert, D. M., & Lengyel, M. (2016). Theoretical perspectives on active sensing. *Current Opinion in Behavioral Sciences*, *11*, 100–108.

http://doi.org/10.1016/j.cobeha.2016.06.009

Zahavi, A. (1975). Mate Selection - A selection for a Handicap. *Journal of Theoretical Biology*, *53*, 205–214. http://doi.org/10.1016/0022-5193(75)90111-3

Zajonc, R. B. (1965). Social Facilitation. *Science*, *149*(3681), 269–274.

Zajonc, R. B., & Sales, S. M. (1966). Social facilitation of dominant and subordinate responses. *Journal of Experimental Social Psychology*, *2*(2), 160–168. http://doi.org/10.1016/0022-1031(66)90077-1

Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, F., Samson, Y., & Boddaert, N. (2006). Autism, the superior temporal sulcus and social perception. *Trends in Neurosciences*, *29*(7), 359–366. http://doi.org/10.1016/j.tins.2006.06.004

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience*, *23*, 143–152.

12 Figure legends

Figure 1. **(A-B)** Sample visual scenes with red circles indicating different locations where gaze can be directed. (A) Photographic reproduction of painting ‘Poppies’ by Claude Monet, and (B) original image published by Max Pixel under the Creative Commons CC0 License. **(C)** Feature, saliency and priority maps (original image published by Veale et al. 2017 under the Creative Commons Attribution License). **(D)** Priority maps for different task goals (original image published by Max Pixel under the Creative Commons CC0 License; maps were obtained with SaliencyToolbox for Matlab (Walther & Koch, 2006)).

Figure 2. Diagram summarising the Interpersonal Gaze Processing model.

Figure 3. The Interpersonal Gaze Processing model in a real social scene. **(A)** Planning gaze when watching a video. **(B)** Planning gaze in a live interaction. Blurbs indicate areas of high saliency depending on the type of map. Original image published by Max Pixel under the Creative Commons CC0 License. Original maps were obtained with SaliencyToolbox for Matlab (Walther & Koch, 2006).

Figure 4. Different sensing and signalling maps may be used in different contexts. Original image published by Max Pixel under the Creative Commons CC0 License. Original maps were obtained with SaliencyToolbox for Matlab (Walther & Koch, 2006).

Figure 1

Figure 2

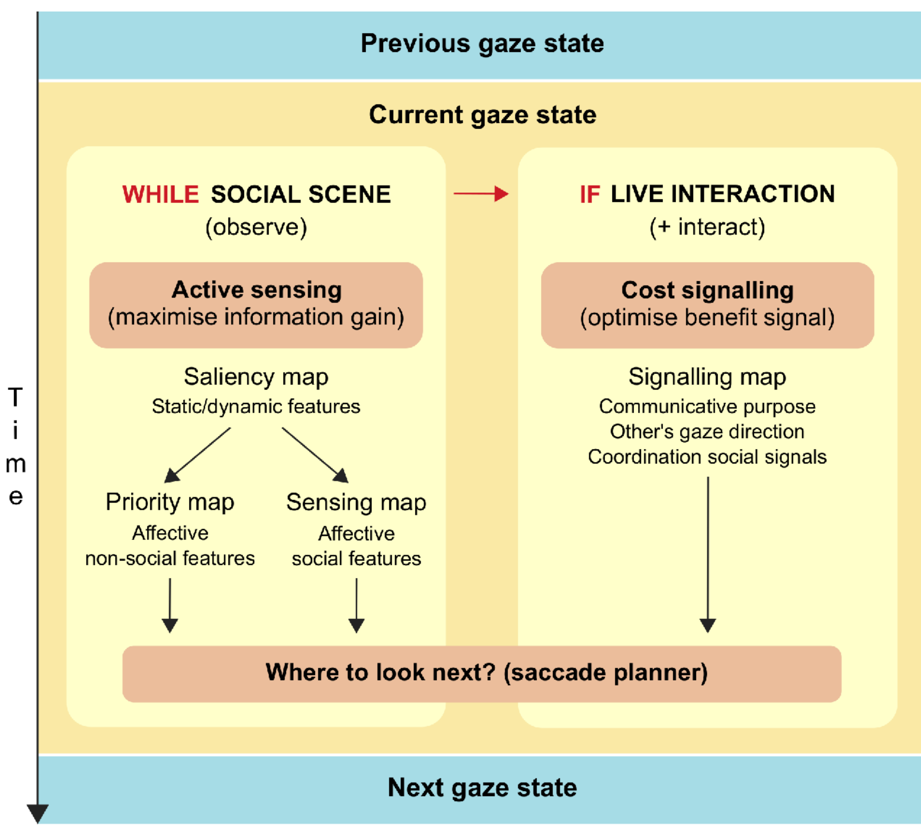

Figure 3

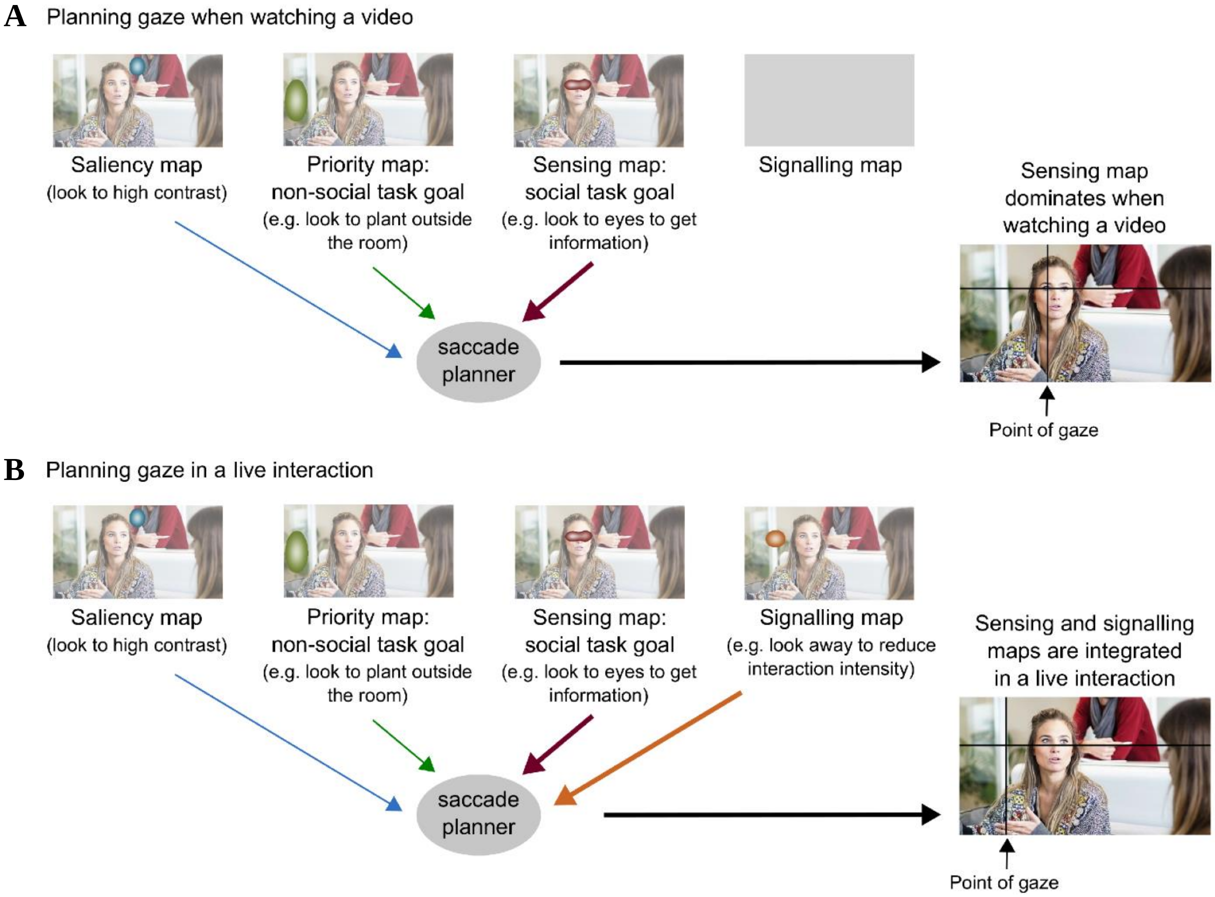

Figure 4

Gaze target	Eyes	Mouth	Background
Sensing function	Emotions Interest in me	Speech Emotions	No social information
Signalling function	Interaction intensity Affiliation Turn-taking	Interest in you	No interest in you Polite gaze aversion Hesitation