

From Gaze Perception to Social Cognition: The Shared-Attention System

Lisa J. Stephenson, S. Gareth Edwards, and
Andrew P. Bayliss 

School of Psychology, University of East Anglia

Perspectives on Psychological Science
1–24

© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1745691620953773

www.psychologicalscience.org/PPS



Abstract

When two people look at the same object in the environment and are aware of each other's attentional state, they find themselves in a shared-attention episode. This can occur through intentional or incidental signaling and, in either case, causes an exchange of information between the two parties about the environment and each other's mental states. In this article, we give an overview of what is known about the building blocks of shared attention (gaze perception and joint attention) and focus on bringing to bear new findings on the initiation of shared attention that complement knowledge about gaze following and incorporate new insights from research into the sense of agency. We also present a neurocognitive model, incorporating first-, second-, and third-order social cognitive processes (the shared-attention system, or SAS), building on previous models and approaches. The SAS model aims to encompass perceptual, cognitive, and affective processes that contribute to and follow on from the establishment of shared attention. These processes include fundamental components of social cognition such as reward, affective evaluation, agency, empathy, and theory of mind.

Keywords

gaze leading, joint attention, shared attention, social cognition

Sharing attention with another individual allows people to share their thoughts, intentions, and desires and serves an important affiliative function that contributes to the human ability to maintain such large social groups (Manninen et al., 2017). Indeed, some theorists have suggested that the intent to engage in such interactions is uniquely human (Tomasello, Carpenter, Call, Behne, & Moll, 2005). Successful sharing of attention relies on two people: the initiator of the interaction and the responder (Emery, 2000). The initiator who has deliberately led the gaze of the responder to an object (also known as *gaze leading*, Bayliss et al., 2013) needs to monitor the gaze response that may have been caused.

Conceptualizations of shared attention have greatly advanced over the past 20 years or so thanks to extensive research into gaze following (from the responder's perspective). Over the past decade, the role of the initiator has become a specific research focus as a separate and important component of shared attention, affording the development of a more elaborated picture of the mechanisms responsible for sharing attention.

The emergence of a focus on the initiator of shared attention, with an understanding that gaze leading can complement the advances made toward understanding gaze following, has motivated this review. We acknowledge the great progress the field has made in current approaches to shared attention, but it is timely to develop these ideas further in a way that places the initiator of joint attention as a fully integrated agent into a model of shared attention. It is this key advance that forms the basis for the new model we propose.

In this article, we review the basic mechanisms underpinning gaze processing and how they facilitate joint and shared attention. Of key relevance to most models of shared attention is the developmental trajectory in typical and atypical populations such as autism, of which we give an overview, along with neurophysiological and neuroimaging findings. It is these traditions that have inspired many previous models of shared attention. Our

Corresponding Author:

Andrew P. Bayliss, School of Psychology, University of East Anglia
E-mail: andrew.p.bayliss@uea.ac.uk

proposed neurocognitive model directly follows from these but with some key additions relating to emotional evaluation, sense of self and agency, reward, empathy, theory of mind, social bonding/affiliation, and person and object knowledge together with a first-, second-, and third-order framework.

Gaze Perception and Shared Attention

Eye contact is often a precursor to initiating joint attention as well as a crucial stimulus that people prioritize during social cognition processing and a modulator of behavior (for a review, see Hamilton, 2016). Being the target of another person's attention is a powerful experience that is affectively stimulating. Indeed, there is a clear preference for the amount of time people feel comfortable with gaze being directed toward them, averaging 3.3 s (Binetti, Harrison, Coutrot, Johnston, & Mareschal, 2016). The morphology of the human eye with distinctive white sclera facilitates detection of gaze signals (Kobayashi & Kohshima, 1997, 2001) and discrimination of small angular differences in the direction of observed gaze (Calder, Jenkins, Cassel, & Clifford, 2008; Lawson, Clifford, & Calder, 2011). The discriminability of observed gaze in humans has led to the development of the cooperative-eye hypothesis, which suggests that human eyes have evolved their high-contrast visibility to serve the need for social interactions (Tomasello, Hare, Lehmann, & Call, 2007). In this way, accurate gaze direction discrimination (Calder et al., 2008) acts as a building block to higher level social cognition because gaze-information processing can help people infer other's mental states (for reviews, see Bayliss, Frischen, Fenske, & Tipper, 2007; Emery, 2000; Langton, Watt, & Bruce, 2000). Moreover, gaze following is one of the key cognitive processes that enables people to learn through observation (Frith & Frith, 2007).

Once accurate gaze direction perception has been achieved, the social cognition system can attempt to use this information in some way. One way to use observed averted gaze is to engage in *joint attention*. Joint attention occurs when an individual (the initiator or gaze leader) gazes at an object, causing another individual (the responder or gaze follower) to orient his or her gaze to the same object.

Thus, establishing joint attention engages orienting mechanisms that allow the use of the directional gaze cue to shift spatial attention to the common object (for a review, see Frischen, Bayliss, & Tipper, 2007). *Shared attention* is often referred to as a subtly different state from *joint attention* in which both agents are aware of their common attentional state for shared attention, but this is not necessary for joint attention (Emery, 2000). Note that this definitional distinction is not universally

adopted; these two terms are often used interchangeably, and some researchers use the term *joint attention* to include shared knowledge of attentional focus, whereas others do not (for a more detailed discussion, see Carpenter & Call, 2013). For our present purposes, we argue that it is preferable to use two different terms to make the distinction clear. Although they are tightly related processes, acknowledging the distinction between them allows a more nuanced examination of their underpinning cognitive mechanisms. Therefore, the definitions of *joint attention* and *shared attention* offered by Emery (2000) are adopted here; shared attention requires both parties to know they are attending to the same referent object, whereas joint attention does not. Much of what is already known about the role of joint attention in humans has come from developmental work on the trajectory of infant-mother social gaze behaviors, and so it is to this work that we turn next.

Developmental Trajectory of Social Attention

From birth, human infants preferentially orient toward faces that make eye contact with them (Farroni, Csibra, Simion, & Johnson, 2002). There is some evidence for neonates having an ability to follow eye gaze, at least if they have seen the preceding eye movement (Farroni, Massaccesi, Pividori, & Johnson, 2004), and there is evidence for gaze following ability in 3-month-olds (Hood, Willen, & Driver, 1998). However, there remains debate in the field precisely when infants do meaningfully follow gaze cues, perhaps partly because what constitutes gaze following can vary among studies (for a review, see Del Bianco, Falck-Ytter, Thorup, & Gredebäck, 2018). One longitudinal study found gaze following developed between 2 and 8 months and stabilizes between 6 and 8 months (Gredebäck, Fikke, & Melinder, 2010). For a detailed review of how gaze following develops in infancy, see Del Bianco et al. (2018).

Once gaze following has developed, joint attention can emerge. *Intersubjectivity* is the sharing of experiences between people (Bard, 2009; Trevarthen & Aitken, 2001). During their first year, an infant and the infant's primary caregiver will share experience in their dyad by paying attention to each other (the *primary intersubjective* stage; Bruner & Sherwood, 1975; Terrace, 2013). Toward the end of the first year (from as early as 8 months), infants will follow gaze. Crucially, at around 12 months, infants begin to be able to "check back" toward the person whose gaze they followed (Scaife & Bruner, 1975). Checking back coincides with the primary intersubjective phase moving on to the *secondary intersubjective* phase in the infant's second year (Terrace, 2013). This is when the child and the

caregiver can start to share experiences by sharing attention toward common referent objects.

The low-level processes that rely on perceptual and attentional capabilities to support shared attention appear to be related to the development of higher level social cognition abilities, such as early language competence (Morales et al., 2000; Mundy & Newell, 2007). For example, the frequency of mother–child joint attention is positively correlated with efficiency in word learning (Tomasello & Farrar, 1986), and Brooks and Meltzoff (2015) found in a longitudinal study that infants who gaze followed more at 10.5 months could produce more words associated with mental state at 2.5 years, which, in turn, also correlated with theory of mind ability at 4.5 years. Recently, jointly attending to a film alongside an experimenter was found to increase the chances of 3- to 4-year-olds passing a verbal false-belief task presented in the film (Psouni et al., 2018).

In summary, the critical age for joint-attention development appears to be during the latter part of the first year of life and during the second year; initiating joint attention develops later than responding to joint attention (Mundy et al., 2007). Understanding that gaze is referential with respect to objects and people develops by the end of the first year of life (Hoehl, Wiese, & Striano, 2008), whereas joint-attention initiation develops later, by 18 months for a typically developing child (for a review of joint-attention development, see Happé & Frith, 2014). The early emergence of joint attention, typically within the first 2 years of life, exemplifies its key role not only in the development of language but also in social cognition processes generally.

Autism and Social Attention

The challenges that people with autism can face with social cognition, and specifically an apparent lack of engagement in joint attention, has directly driven a great deal of empirical and theoretical work (for a review, see Mundy, 2018). A key diagnostic element of an autistic spectrum condition (ASC) is a deficit in non-verbal communication, including eye contact abnormalities (American Psychiatric Association, 2013, which uses the term *autistic spectrum disorder*).

Although it will become clear from the studies reviewed below that there are some atypicalities in interactive gaze behavior in the autistic population, what is equally important to note from the research is how intact gaze processing can be among people with this heterogeneous condition. For example, a recent study showed that a subset of an autistic sample displayed no impairment at all in their gaze perception and suggested that idiosyncratic aspects of gaze perception underlie atypicalities in autism (Pantelis & Kennedy,

2017). Another example is that people with autism and high levels of autism-like traits demonstrated typical priors for gaze direction perception such that gaze was more likely to be perceived as direct when noise was added to eye regions (Pell et al., 2016). Likewise, typical sensory adaptation to eye-gaze direction has been reported for autistic adults as is found in nonclinical populations (Palmer, Lawson, Shankar, Clifford, & Rees, 2018). These types of findings make it hard to argue that social interaction difficulties for people with autism arise only from “first-order” differences such as atypical gaze perception.

It is generally thought that people with autism are less likely to initiate joint attention or, at least, may have atypical gaze-leading behavior (Billeci et al., 2016; Mundy & Newell, 2007; Nation & Penny, 2008; however, for an alternative view, see Gillespie-Lynch, 2013). For example, Billeci et al. (2016) found that whereas toddlers with an ASC diagnosis displayed the same eye movements as typically developing control subjects when responding to joint attention, their patterns of fixations when initiating joint attention were distinguishable from the control group (e.g., fixating for longer on the face than the typically developing control subjects and making more transitions from the object to the face).

In another recent study, it was found that recognition memory for pictures was better when children had gaze led to the pictures than when they had been gaze cued to them. Critically, this was found for typically developing children but not for children with an autism diagnosis (Mundy, Kim, Mcintyre, & Lerro, 2016). Most recently, a large study of 338 toddlers (242 twins, 84 nonsibling children, and 99 children with autism) made the revealing finding that when free viewing video scenes, monozygotic twins showed remarkably similar patterns of gaze fixations on the eye regions of faces, $r = .91$, compared with $r = .35$ for dizygotic twins and no correlation for nonsiblings (Constantino et al., 2017). Gaze fixations on the mouth region, similarly, suggested that these gaze behaviors are highly heritable. In addition, Constantino et al. (2017) found that children with autism looked markedly less at eyes and mouth regions of faces than typically developing children. Another recent study showed that typically developing adults and children preferred a set of stuffed animal toys with visible white sclera over those without, whereas participants with a diagnosis of autism did not (Segal, Goetz, & Maldonado, 2016). This is suggestive of the importance of eye gaze to the typical development of social cognition and supports the cooperative-eye hypothesis.

Findings that suggest that people with autism are less likely to spontaneously look at the eye region of faces are important because people with autism may

appear to lack motivation for social interaction (Chevallier, Kohls, Troiani, Brodtkin, & Schultz, 2012) and it is joint-attention initiation that can signal social motivation to interact with others (Mundy & Newell, 2007). Chevallier et al. (2012) argued that suboptimal social cognition in autism arises from motivational deficits rather than vice versa. However, the social motivation theory of autism has increasingly been challenged. For example, a recent systematic review of empirical studies into the social-motivation hypothesis identified that only 57% of reviewed studies supported the idea (Bottini, 2018). Authors of another article challenged the theory strongly, including pointing out that people with autism do not report lack of motivation for social interaction (Jaswal & Akhtar, 2018). It is also noteworthy that implicit social biases, measured using implicit association tests for gender and race, may be relatively typical in people with an autism diagnosis (Birmingham, Stanley, Nair, & Adolphs, 2015). It is therefore questionable whether the social interaction challenges that are associated with an ASC can be explained by differences in social motivation. Nevertheless, differences between how people with ASC diagnoses and people without ASC diagnoses manage eye contact and gaze leading may affect how social interactions unfold.

There may be individual differences in the broader autism phenotype. Edwards, Stephenson, Dalmaso, and Bayliss (2015), across three experiments, found a negative correlation between a gaze-leading effect (attentional orienting toward faces that had just followed participants' gaze) and level of autism-like traits: the greater the autism-like traits, the less attentional capture from faces who followed gaze. This indicates there may be individual differences in joint attention initiation behaviors across the typical population, specifically linked to the personality traits commonly found in people with an ASC.

There are associations between social skills and joint-attention skills. For example, better joint-attention skills in 3-year-old children with an ASC have been associated with better friendships at age 8 (Freeman, Gulsrud, & Kasari, 2015). Lawton and Kasari's (2012) intervention to improve joint-attention initiation in preschool children with an autism diagnosis increased social interaction duration. Other interventions to improve joint-attention interaction in children diagnosed with an ASC have resulted in improved language development, play skills, and social development (for reviews, see Goods, Ishijima, Chang, & Kasari, 2013; Murza, Schwartz, Hahs-Vaughn, & Nye, 2016; Reichow & Volkmar, 2010). However, improvements from joint-attention interventions have often proved short-lived (e.g., Whalen & Schreibman, 2003) or have not been assessed to ascertain whether the improvements are

maintained (for a review, see Stavropoulos & Carver, 2013). All of the evidence for differences in joint attention for people with autism has not only fueled debates about what can be learned about autism more generally (see Chevallier et al., 2012) but also has led to a wealth of studies on the efficacy of joint-attention skills interventions (for a meta-analysis, see Murza et al., 2016). In the field of autism interventions, there has been a growing interest in how technology-based approaches, including virtual reality, can be used (for a meta-analysis, see Grynszpan, Weiss, Perez-Diaz, & Gal, 2014), and the use of assistive robotics specifically is another emerging area (for a review, see Boucenna et al., 2014).

Recently, the social-interaction-mismatch hypothesis of autism has been proposed (Bolis & Schilbach, 2018; Schilbach et al., 2013). This hypothesis emphasizes that people with high autistic traits do perceive social information in their environment in a similar way to people without a diagnosis of autism but are less likely to update their beliefs on the basis of this perception; this, in turn, influences their social behavior. We will consider this theory further in the light of our model in a later section.

Sense of Agency and Social Attention

Above we reviewed some of the extensive research on the basic mechanisms underpinning shared attention in humans. Here we introduce an important factor in social behaviors such as social attention—the sense of agency. Sense of agency is experienced when a person causes or generates actions and, through them, feel that they have controlled events around them (Gallagher, 2000). It is obvious that sensing causality over the actions of a partner with whom one is performing a joint activity, whether it be playing tennis or moving furniture, is necessary to achieve the joint enterprise. However, it is just as critical, but perhaps less obvious, that social interactions may benefit from a sense of agency when coordinating social attention. Specifically, detecting that someone has shifted his or her own gaze to coordinate and share attention with a person because of that person's own initial action of shifting his or her overt attention to a new location could be one source of evidence that they are in a new social interaction or that an ongoing interaction is continuing successfully. This is a relatively unexplored area but one we feel is important because one's own role in the world and the actions one elicits in others are key elements of social cognition. There are two studies we know of that have begun to explore agency and joint attention.

In one study, it was suggested that having one's gaze followed generates a sense of agency in the initiator. Pfeiffer et al. (2012) had participants make a saccade

to an object and subsequently an on-screen face would display a gaze shift toward or away from that object in response. Participants self-reported that having their gaze followed, compared with not, made the on-screen gaze shift feel more “related” to their own eye movement. In their experiments, Pfeiffer et al. varied the time delay between the participant’s fixation of the object and the on-screen responding gaze shift toward that object. Participants reported maximal relatedness with a 400-ms delay, with a linear decrease thereafter up to 4,000 ms. However, this temporal sensitivity to the response of the on-screen face was limited to instances in which the face would definitely “look” toward the same referent object as the participant: In one experiment in which the face could look at the same or different location to the participant, there was little effect of latency (the time between participant gaze shift and the on-screen gaze shift response) on reported relatedness (Pfeiffer et al., 2012). Therefore, the optimal temporal range within which a response to shared-attention initiation feels naturalistic remains a subject to be explored in future further research. Such information could help inform the interventions that seek to improve social skills for people with autism that were discussed above.

In another study examining sense of agency in joint gaze contexts, findings showed that the temporal-binding effect—often used as an index of an implicit sense of agency (Haggard, 2017)—occurs when a participant’s own change in fixation location results in an on-screen gaze shift to the same location (Stephenson, Edwards, Howard, & Bayliss, 2018). This temporal-binding effect (often termed *intentional binding*) is a subjective compression of the perceived passage of time when a person causes an action outcome and the absence of this effect when an external cause generates the outcome (David, Newen, & Vogeley, 2008). Thus, this effect evidences an implicit sense of agency over congruent gaze shifts elicited by gaze leaders during a simulated social interaction (Stephenson et al., 2018). Previously, most temporal-binding effects had been demonstrated for manual button presses (Moore & Obhi, 2012), and so this study was the first to reveal the same effects can be elicited by oculomotor actions that take place during joint attention.

Electrophysiological Correlates of Gaze Perception

Because of its high temporal resolution, electroencephalography (EEG) is an experimental technique ideally placed to uncover the basic mechanisms underpinning gaze perception and social attention. Sharing attention, responding to perceptual inputs with consequent shifts

of attention and so on, is an example of an unfolding sequence of processing stages (albeit across two individuals). Therefore, time-sensitive measurements during this unfolding interaction may be especially revealing. In the next section, we give an overview of electrophysiological studies that looked at gaze perception and a small number of studies that have attempted to investigate higher-order aspects of social attention such as shared or joint attention.

Early Event-Related Potentials: The N170, EDAN, and N2pc

Early perceptual components are useful for examining the initial processing of stimuli within a gaze interaction. The N170 has been the subject of a large body of work showing its involvement in face processing, but it has also been implicated specifically in gaze processing for observed gaze shifts toward or away from the participant (for a review, see Itier & Batty, 2009). Indeed, it has long been known that observing eyes alone elicits an N170 greater in amplitude than the onset of a whole face (Bentin, Allison, Puce, Perez, & McCarthy, 1996). More recently, gaze-contingent studies have demonstrated that the sensitivity of the N170 to gaze stimuli can be detected when participants fixate on the eyes of a face, not just when eye regions are presented in isolation (Itier & Preston, 2018; Nemrodov, Anderson, Preston, & Itier, 2014; Parkington & Itier, 2018, 2019). In children, the N170 has been shown to be larger for eyes than for faces and to reflect an adult-like event-related potential (ERP) profile by the age of 11, suggesting the N170 may be driven largely by eyes and mature earlier than the N170 for complete faces, which continues into adulthood (Taylor, Edmonds, McCarthy, & Allison, 2001). The importance of the eyes to early face-related processing is confirmed by patient intracranial N200 ERP recordings, which have been consistent with N170 studies, showing face-specific neural processing (Allison, Puce, Spencer, & McCarthy, 1999; McCarthy, Puce, Belger, & Allison, 1999; Puce, Allison, & McCarthy, 1999). Engell and McCarthy (2014) added to these earlier findings, with intracranial ERPs demonstrating that the human cortex has more eye-selective cells than face-sensitive cells, specifically, in the lateral occipitotemporal cortex.

The sensitivity of the N170 for particular gaze directions has also been investigated and resulted in somewhat mixed findings in adults. Some studies have shown greater N170 elicited for observed gaze shifts away from participants compare with gaze shifts toward participants (Latinus et al., 2014; Rossi, Parada, Latinus, & Puce, 2015), some the opposite effect (e.g., Conty, N’Diaye, Tijus, & George, 2007), and others no modulation at all (e.g.,

Myllyneva & Hietanen, 2016). More recently, researchers using combined EEG and magnetoencephalography (MEG) found that the N170 was greater for observed direct gaze than for averted gaze only in the right hemisphere for both frontal and deviated head presentations (Burra, Baker, & George, 2017).

Contextual factors such as task and stimulus types could explain these disparate findings, and because many N170 studies have employed only one task, it is not possible to disentangle stimulus effects from task effects. One within-subjects study (McCrackin & Itier, 2019) demonstrated that task type (face emotion, direction of attention, and gender discrimination) elicited different ERP activation for direct and averted gaze from around 220 ms to 290 ms. Because this study also used mass-univariate analyses rather than classic analyses, it is worth noting that the analysis choices that researchers subject their ERP data to could also contribute to disparate findings within the field. In another study, Latinus et al. (2014) used different tasks but the same stimuli—the same participants performed a social task (observing gaze shifts either away or toward the participants) and a nonsocial task (observing left/right gaze shifts)—and found that a greater N170 for gaze aversion was elicited in the right hemisphere for the nonsocial task but not the social task. Stephenson, Edwards, Luri, Renoult, and Bayliss (2020) also investigated the impact of shared-attention context as a modulator of the N170. They found that both when the task was gaze-related and gaze-unrelated, participants showed an enhanced N170 to the onset of averted gaze when the gaze stimulus followed the participant's antecedent saccade to a peripheral location compared with when the observed gaze direction did not establish shared attention. Stephenson et al. noted that the N170 was small and a little delayed because of the peripheral presentation of the gaze onset, but their finding further supports the notion that contextual factors relating to gaze can modulate the N170 and in particular, relation to shared attention.

Clearly the sensitivity and modulation of the N170—traditionally conceived as a face-specific response—by observed eye gaze direction—demonstrates that gaze processing affects, or even drives, some aspects of face processing in the cortex. What is not fully resolved, however, is how stimulus or task contexts modulate the N170 or the extent to which different analysis techniques of the ERP signal can account for mixed findings. Understanding these factors will elucidate further the role of gaze processing in the elicitation of the N170 and explain the, sometimes, inconsistent findings.

Rapid perceptual processing of observed gaze direction triggers attention processes. A clear candidate ERP for investigating attention shifts in response to gaze cues is, therefore, the early-directing attention negativity

(EDAN). This ERP component is elicited in tasks that involve spatial attention (Harter, Miller, Price, LaLonde, & Keyes, 1989). The EDAN was found to be modulated in response to spatial cues of attention from arrows but surprisingly not from eye gaze (Hietanen, Leppänen, Nummenmaa, & Astikainen, 2008). An explanation for this could be that gaze cueing and arrow cueing are controlled by different systems. Brignani, Guzzon, Marzi, and Miniussi (2009) reported a curious reverse EDAN-like effect from eye gaze shifts and a more typical EDAN effect from arrows, again supporting the notion of different attentional systems supporting social and nonsocial cueing (for a discussion, see Frischen et al., 2007). The EDAN's role in attentional orientating has been debated, specifically whether it reflects processing of the stimulus itself or orienting attention on the basis of the directional cue being given (for further discussion on this point, see Van Velzen & Eimer, 2003; Woodman, Arita, & Luck, 2009). Recently, Kirk Driller, Stephani, Dimigen, and Sommer (2019) found a large EDAN was elicited when participants counted the number of gaze shifts they observed. This supports the idea that the EDAN reflects attention orienting because the direction of gaze was not task relevant.

The N2pc's role is usually explored in visual-search paradigms (e.g., Grubert & Eimer, 2015). The N2pc ERP component is characterized by greater negative activity at the posterior sites that are contralateral to the side on which the stimuli are presented, implicated in spatial attentional shifting (Galfano et al., 2011). Galfano et al. (2011) used the N2pc as an index of spatial attention reorientation to the target needed when incongruent gaze cueing was observed. Galfano et al. predicted, and found, greater N2pcs elicited from incongruent gaze cueing than congruent. These findings do suggest that perhaps at the completion of a shift of attention, attention systems responsible for gaze cueing are operating in a similar manner compared with how other cues are processed. We note that it is not yet clearly established whether the EDAN and the N2pc reflect overlapping processes rather than distinct processes.

These early perceptual components (N170, EDAN, and the N2pc) are useful for examining the processing of stimuli per se. However, EEG also affords researchers a method for studying higher level processes involved in social interactions. The establishment of a sometimes brief, but stable, "state" is relevant not only for studying early perceptual processes but also for the subsequent social interaction of shared attention, and so later ERP components and the study of EEG oscillations could hold promise. For example, conflict processing/expectation of congruence later than 300 ms could, therefore, be revealing in joint attention. In addition, the role of alpha rhythms can offer insight because they can code a shared/nonsocial momentarily lingering state of shared

or joint attention rather than a delimited stimulus processing event, as ERPs offer by definition offer. We, therefore, now review some studies that have looked at later components that emerge after 300 ms and some that have examined alpha waves because both could be revealing for the study of joint and shared attention.

Electrophysiological Correlates of Dynamic Joint Gaze States: The Anterior Directing Attention Negativity, N330, P3, and Oscillatory Studies

The anterior directing-attention negativity (ADAN) is a later ERP occurring at around 300 ms to 600 ms contralateral to a cued location (Eimer, Van Velzen, & Driver, 2002). Lassalle and Itier (2013, 2015) found both the EDAN and ADAN were elicited by gaze cueing, reflecting the orienting of attention and holding of attention to a cued location, respectively. These studies were the first to show the EDAN and ADAN for gaze cues, having previously been elicited only by arrow cues (Hietanen et al., 2008).

Greater occipito-temporal negativity (N330 ERP component) has been demonstrated in response to incongruent gaze shifts away from an object, compared with congruent gaze shifts (Senju, Johnson, & Csibra, 2006). The suggested explanation was that the N330 reflected the greater effort required to process the violation of the expectancy that gaze would be shifted to an object. In addition, the N330 was believed to reflect activity in the posterior superior temporal sulcus (pSTS) because corresponding functional MRI (fMRI) data showed increased activity in response to incongruent gaze shifts (also see Pelphrey, Viola, & McCarthy, 2004). Tipples, Johnston, and Mayes (2013) also found an enhanced negative occipito-temporal ERP (occurring slightly earlier at N300) for incongruent gaze shifts. In addition, Tipples et al. found an enhanced N300 when arrows provided the directional shifts of attention, suggesting a domain general mechanism for detecting and processing unexpected events, perhaps not limited to gaze shifts. Therefore, a little is already known about ERP correlates when participants observe a face looking toward or away from an object.

Although the studies reviewed above focused on observing gaze shifts toward or away from objects or toward or away from participants, some studies have specifically examined the neural time course of processing responses to participants who have initiated joint attention. Caruana, de Lissa, and McArthur (2015) found that an enhanced central parietal P3 ERP (reported as a “P350”) occurred when participants’ joint-attention bids were ignored (an averted gaze shift resulted) over when successfully reciprocated. Caruana

et al. found no such effect when another group of participants undertook a similar task that replaced eye gaze responses with arrows. In another article, Caruana et al. (2017) found that the P350 was not modulated by averted gaze or congruent gaze shifts when participants were expressly told that they were engaging with a computer program rather than being told that the gaze shifts they observed were being controlled by a real human. More research is needed to build on these findings, but they offer evidence of a specific social evaluation of the outcome of a joint-attention bid.

Using a face-to-face joint attention interaction, Lachat, Hugueville, Lemaréchal, Conty, and George (2012) found that the act of engaging in joint attention suppressed signal power in the α and μ frequency band in centro-parietal-occipital scalp regions relative to non-joint-attention episodes. The authors suggested this modulation was consistent with joint attention’s association with social coordination, mutual attention, and attention mirroring. More recently, an infant EEG study looked specifically at oscillations when infant joint-attention initiation was reciprocated with a congruent gaze shift to an object and found greater α band suppression for congruent gaze shift responses than incongruent (Rayson, Bonaiuto, Ferrari, Chakrabarti, & Murray, 2019). This is evidence that infants as young as 6.5 to 9.5 months old detect whether their gaze has been followed.

It can be concluded from this review of EEG studies that there is electrophysiological evidence of gaze perception in three types of observed gaze behaviors. First, there is ERP modulation when an observed face gaze shifts toward or away from an object. Second, modulation occurs when an observed face shifts gaze toward or away from the participant and third, when the participant’s own direction of gaze is followed to establish joint attention. In addition, the EEG methodologies used have enabled some useful information to emerge about the rapid detection of gaze signals and the timing involved, which is likely an important moderator in the social interaction. In sum, the EEG studies reviewed above support the notion that gaze is detected rapidly, then attention processing is engaged, and the outcome is evaluated relative to expectation all within half a second. This rapid processing is highly context dependent and supports the establishment of ongoing iterative interactions that rely on social attention.

Current Theories and Models of Shared Attention

Two influential models of social attention come from the disparate traditions of developmental psychology and neurophysiology of high-level vision. Nevertheless,

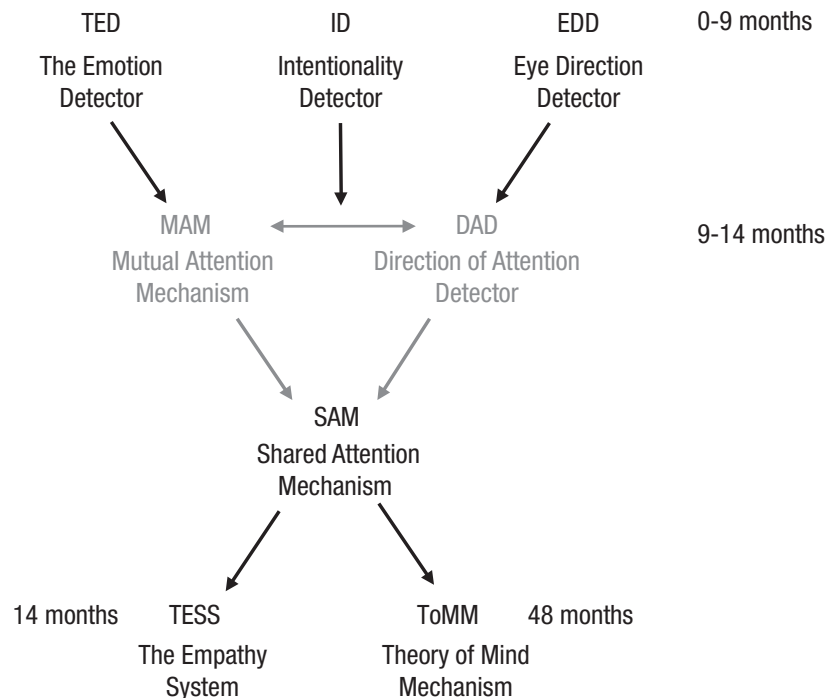


Fig. 1. The empathizing system adapted from Baron-Cohen (2005) and Perrett and Emery (1994). All mechanisms and developmental age indications are from Baron-Cohen's model, whereas the mechanisms in gray text are those proposed by Perrett and Emery. The latter's specific addition was to propose that shared attention (via a shared-attention mechanism, or SAM) arises from the coupling of a mutual-attention mechanism (MAM) and a direction-of-attention director (DAD). This figure is our interpretation of how to combine the two models.

these two models proposed during the 1990s bear a number of similarities, including being evolutionarily based and involving dedicated cognitive modules. Baron-Cohen's (1994, 2005) developmentally inspired model proposed evolved mechanisms to facilitate "mind-reading" and empathizing from gaze perception. Baron-Cohen (1994, 2005) hypothesized an eye-direction detector, an intentionality detector, the emotion detector, a shared-attention mechanism, the empathizing system, and a theory-of-mind mechanism. The eye-direction detector has subsequently found support from the ERP research reviewed earlier, which suggests a specific eye-region detector (e.g., Engell & McCarthy, 2014; Parkington & Itier, 2019). Based partly on the findings of complex response profiles in monkey superior temporal sulcus (STS) to social-attention signals, Perrett and Emery (1994) theorized a direction-of-attention detector and a mutual-attention mechanism additional to those in Baron-Cohen's model. The mechanisms proposed by both Baron-Cohen and Perrett and Emery are summarized in Figure 1 by schematically integrating the two models.

It is important to understand how multiple attention cues are integrated during shared attention, and this

was something Perrett, Hietanen, Oram, Benson, and Rolls (1992) was able to shed light on using neurophysiological evidence. For example, Perrett et al.'s single-cell recordings showed that the system included inhibitory mechanisms to support the prioritization of more reliable information about the direction of observed attention when they are available, such that information from the eyes is always prioritized over head and body orientations (for a more detailed discussion, see Langton et al., 2000). However, there is evidence that rather than being simply inhibitory, the system may allow integration of the information from eye and head orientation, providing an attenuated effect of head information if the eye information conflicts (Langton et al., 2000).

Our review is primarily about human social cognition, but within the gaze body of research, there are debates about whether sharing attention is exclusive to humans (see e.g., Carpenter & Call, 2013; Leavens & Racine, 2009; Van der Goot, Tomasello, & Liszkowski, 2014). In addition, because earlier models (Baron-Cohen, 2005; Perrett & Emery, 1994) acknowledged the role of evolution, we will briefly consider some nonhuman species and their neurophysiology for processing gaze. For example, rhesus macaques and chimpanzees

follow gaze direction of conspecifics (Tomasello, Hare, & Fogleman, 2001). Chimpanzees use gaze and head direction cues (Tomasello et al., 2007) and also exhibit checking-back behaviors (Bräuer, Call, & Tomasello, 2005; for a detailed review, see also Carpenter & Call, 2013). In rhesus macaques, single-cell recordings have shown that there is a neural network that supports gaze-direction encoding (Perrett et al., 1985, 1992). This work revealed a hierarchical system in the monkey anterior STS that codes—in order of priority—direction of gaze, head, and body orientations.

More recently, recordings from the macaque amygdala suggest eye contact is detected in social interactions (Mosher, Zimmerman, & Gothard, 2014) and that face-responsive neurons discriminate gaze and head orientation and integrate that information, with greater responses to direct than averted gaze (Tazumi, Hori, Maior, Ono, & Nishijo, 2010). As Nummenmaa and Calder (2009) pointed out, there is no equivalent evidence that a hierarchical system exists in humans, but it would seem reasonable for such a system to exist given the eyes offer the best clues for social attention. Evidence that human neurons are dedicated to separate coding of gaze, head, and body orientation have been shown repeatedly, and Nummenmaa and Calder offered a succinct review of the adaptation paradigms used to explore this separate coding system (e.g., Calder et al., 2008; Lawson et al., 2011). There are also behavioral studies that show both eye gaze and head direction are integrated to inform perceived gaze direction (e.g., Todorović, 2006, 2009).

The Baron-Cohen (2005) and Perrett and Emery (1994) models are mainly concerned with perception and attention and how these processes give rise to higher level, broader consequences of shared attention. Likewise, Mundy and Newell's (2007) developmental-attention-systems model of joint attention and social cognition focused on attentional states, with self-other attention processing identified as a necessary step toward social cognition, and highlighted early joint attention (before 9 months) as a contributing factor in social cognition development. Mundy and Newell's influential model emphasized the important distinction between initiating and responding to joint attention and how these processes interact.

Thus far, we have demonstrated how existing models provide elegant and detailed descriptions of the perception and attention processes at work and how they can lead to higher level processes, but this can be taken a step further to examine how all these processes really feed into the broader ability of humans to cooperate with one another. This is something Shteynberg (2015) addressed by reviewing behavioral shared attention studies, mainly from the field of social psychology. The

review includes studies that looked at effects of sharing attention in online social networks, encompassing any studies in which participants believe that they are jointly attending, and so goes beyond the much more narrow definition of shared attention in this review that is between two people who are in a face-to-face interaction. However, Shteynberg's review of behavioral studies does demonstrate that the increased recruitment of cognitive resources seems to be one result of sharing attention. Shteynberg's model lists five empirically demonstrated effects of sharing attention: enhanced memory, stronger motivation, more extreme judgments, higher affective intensity, and greater behavioral learning. The model also postulates a shared-attention mechanism that facilitates group coordination toward mutual goals.

With a focus on the broader consequences of sharing attention, one relevant behavioral outcome of initiating joint attention is better memory for images participants gaze led to over those they, themselves, responded to in response to gaze cueing (Kim & Mundy, 2012). Another recent finding was that jointly attending to the same side of a computer screen with a social partner increased ratings on a social bonding scale, whether or not there was a shared goal (W. Wolf, Launay, & Dunbar, 2016). This indicates that people feel connected when jointly attending, and this could be built on by exploring whether this sense of closeness is enhanced more by initiating the joint-attention interaction rather than responding to it.

The models we have described above all provide insight into how joint attention operates and interacts with other mechanisms involved with social cognition and beyond. It is remarkable to consider how these models aim to account for the same phenomenon from diverse perspectives, from perception and neurophysiological traditions (Perrett & Emery, 1994), developmental perspective (Baron-Cohen, 2005; Mundy & Newell, 2007), and social psychological, interactionist, and third-person approaches (Shteynberg, 2015). We therefore have a rich research base with multifarious approaches aiming to offer different levels of explanation. This allows us to seek to integrate these perspectives within one new model.

A Shared-Attention System and Associated Neural Mechanisms

The findings reviewed above enable the formulation of a novel model, building on previous work, to capture all the processes at work during a joint and shared-attention interaction and the neural regions involved. Our shared-attention system (SAS) aims to capture both how people in a joint-attention interaction have to

coordinate their behavior and how this leads to a state of shared attention, which, in turn, facilitates a number of subsequent social cognitive processes (in addition to initial gaze detection and perception).

A depiction of SAS can be found in Figure 2 and shows how two agents during joint attention need to coordinate their attention and how the act of reorienting attention to another's object of attention leads to the higher level state of sharing attention. This facilitates key components of social cognition: emotional evaluation, sense of self and agency, reward, empathy, theory of mind, social bonding/affiliation, and person knowledge. After two parties become aware that joint attention has been initiated and responded to, only then does their deliberately coordinated gaze achieve shared attention, facilitating the cognitive outcomes in the model. Previous models and reviews have been proposed for direct gaze and the ensuing cognitive processes that flow from that (see Conty, George, & Hietanen, 2016; Hamilton, 2016; Senju & Johnson, 2009). Our model has a wider framework than direct gaze effects, with joint and shared attention at its heart, but we acknowledge that these previous models have captured the importance of direct gaze within social interaction and that the attentional state that direct gaze allows is a necessary process in building to a state of shared attention and also in maintaining that state.

The SAS model is structured around first-, second-, and third-order processes to reflect that when engaging in social interactions, the social cognitive system needs to solve a series of problems. The first is to derive meaning from the social signals available from others (first-order). The second is to generate an appropriate response (second-order). A critical next step takes account of the fact that as an interacting individual, one may effect changes in the mental, emotional, or physical state of another person. A third-order problem is to determine and process the social outcomes that one's own actions have had. Therefore, our model is structured around these first-, second-, and third-order processes, beginning with basic first-order gaze detection and ending with higher order third-person processes such as theory of mind.

How the SAS model complements previous approaches

Previous seminal research has been critically influential in conceptualizing the mechanisms at work—for example, shared-attention mechanism (Baron-Cohen, 1994) and mutual-attention mechanism (Perrett & Emery, 1994) and some of the outcomes of gaze processing, such as theory of mind and empathy (Baron-Cohen, 2005). Our model adds to these by including not only the mechanisms needed to engage in gaze interactions

but also both the initiator's and responder's roles, how these need to be coordinated, which brain regions are implicated, and all of the social cognition outcomes that flow from sharing attention. These outcomes number seven compared with the two (empathizing and theory of mind) included in Baron-Cohen's (2005) influential model. Our model is also informed by more recent advances that have occurred since the formulation of these previous models, for example, from EEG research (e.g., Lachat et al., 2012; Rayson et al., 2019), imaging studies (e.g., Cavallo et al., 2015; Koike et al., 2016), and behavioral studies (e.g., Stephenson et al., 2018).

We note here that the first-order processes of detection of gaze and encoding of gaze direction might arise by either covert attention or overt attention. In other words, the gaze signal may result in a gaze shift in which overt attention is deployed or not, if the gaze cue is just detected through covert attention. The research on gaze cueing establishes that attention is shifted in the observed direction of gaze, but this is a covert attentional shift, without necessarily the accompanying eye movement that gaze following involves (for a review, see Frischen et al., 2007). For clarity, our model assumes there has been an overt attentional gaze shift, as is assumed by definitions of joint attention (see e.g., Emery, 2000).

The field has been remarkably successful in developing and testing theories of how first-order, second-order, and fully interactive dyadic communication is achieved. Take, for example, Perrett et al.'s (1992) model of the hierarchical encoding of gaze direction in the STS of the macaque. Likewise, researchers have a very good idea about the mechanisms underpinning gaze following (see e.g., Frischen et al., 2007). These are first- and second-order processes, respectively, and help social cognitive scientists build models of the fundamental building blocks of social cognition (e.g., Frith & Frith, 2012; Happé, Cook, & Bird, 2017). What our model does is speak to third-order social cognition too by taking into account the iterative and continuous aspects of social interactions by including both agents in the interaction, the mechanisms that need to be coordinated between them, and more of the social outcomes that can follow than have previously been encapsulated in one model. For example, detecting gaze and decoding gaze direction are first-order processes. Reorienting attention is an example of a second-order process, and examples of third-order processes in the model are gaze coordination and cognitive outcomes.

The dynamics of the SAS model

The model shows how the second-order gaze processes, in particular, are dynamic and iterative and have bidirectional influences on each other; for example, both

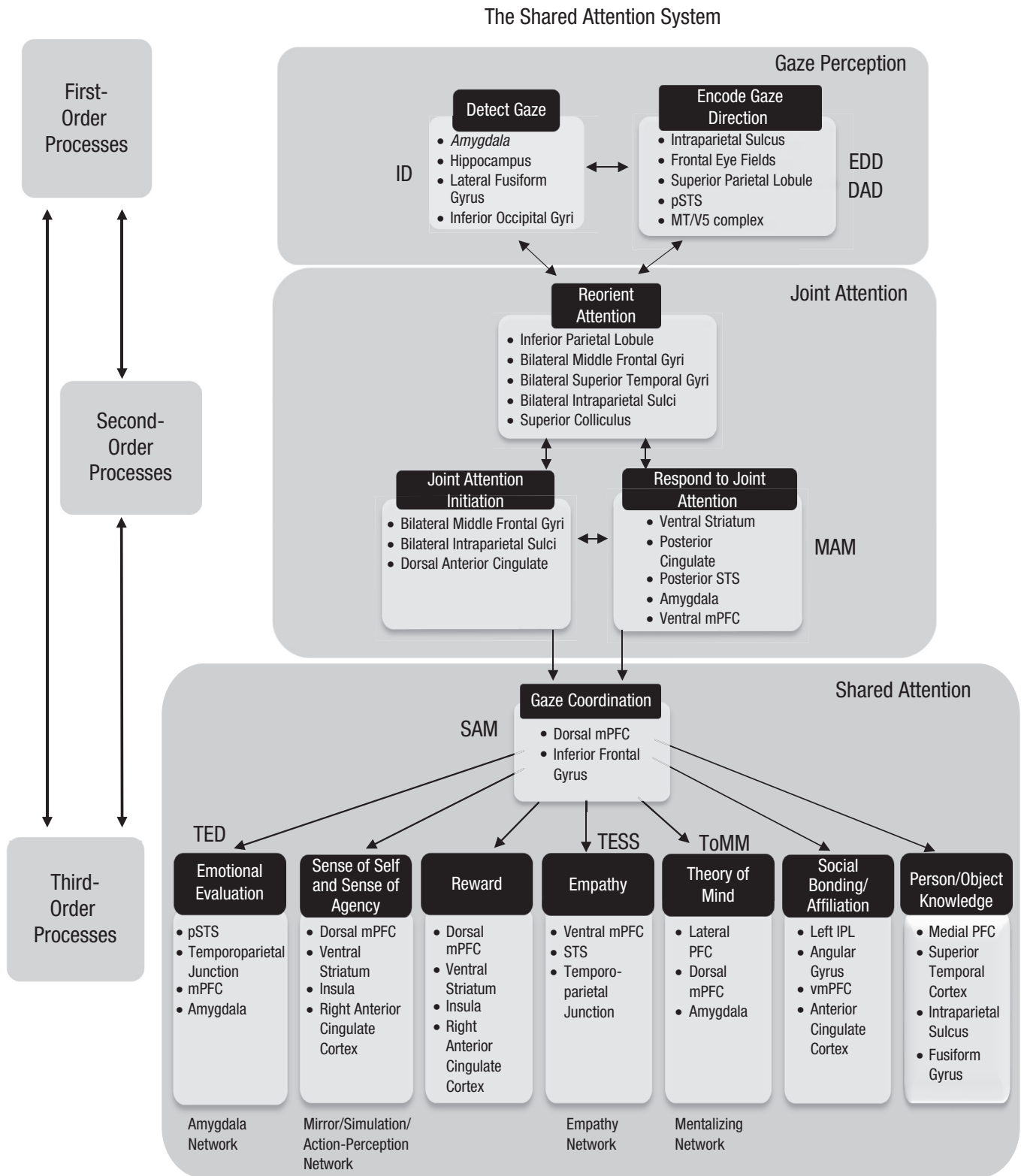


Fig. 2. The shared-attention system. This figure shows how first-, second-, and third-order processes are involved in facilitating gaze detection, leading to the act of joint attention, which is a necessary precursor to shared attention when both agents are aware of each other's attentional state, which can lead to important social cognitive outcomes: emotional evaluation, sense of self and agency, reward, empathy, theory of mind, social bonding/affiliation, and person knowledge. In our model, we included all of Baron-Cohen's (1994, 2005) conceptualized mechanisms (ID = intentionality detector; EDD = eye-direction detector; SAM = shared-attention mechanism; TED = the emotion detector; TESS = the empathizing system; ToMM = theory-of-mind mechanism) and also those of Perrett and Emery (1994; DAD = direction-of-attention detector; MAM = mutual-attention mechanism). We included subserving brain regions for each of the processes in the model, and in addition, we included the "social brain" networks highlighted by Stanley and Adolphs (2013)—the amygdala, empathy, mentalizing, and mirror/simulation/action-perception networks. See The Brain Regions in the Model section for the research that informs the involvement of these regions. STS = superior temporal sulcus; IPL = inferior parietal lobule; PFC = prefrontal cortex; m = medial; p = posterior; vm = ventromedial.

gaze initiator and responder must reorient their attention after detecting each other's gaze (either direct gaze or direction of gaze) to achieve gaze coordination, when third-order processes can begin because that is when the social outcome of one's own gaze behaviors on another person can be determined. We included arrows between first-, second-, and third-order processes in Figure 2 because we recognize that the third-order process of shared attention and the ensuing cognitive outcomes may have a top-down impact on gaze detection and encoding and, similarly, on second-order processes such as initiating joint attention and responding to joint attention. This reflects the interconnected nature of gaze interactions in which several bidirectional cognitive processes continually send feedback to one another.

To summarize the novelty of our SAS model, it is not just an overview of the processes at work when people share attention and the associated neural regions, although it does aim to encapsulate these together, building on subsequent work since earlier models were formulated. The model makes clear the roles of both parties in the shared attention episode by placing the initiator of joint attention as a fully integrated agent alongside the gaze responder. In addition, the model aims to identify all of the possible social cognitive outcomes that could flow from sharing attention. However, we do not claim that all of the processes in the model are necessary or sufficient for each and every social cognitive outcome. The model offers a novel first-, second-, and third-order framework within which to conceptualize the processes at work.

Application of the SAS model to theory. The extensive network subserving gaze processing, initiating joint attention, responding to joint attention, and shared attention consequences are summarized in the model. The theoretical framework offered is that initiating shared attention has fundamental benefits for the initiator; people are motivated to share attention as part of the human capacity for social cognition and intergenerational transmission of culture, including language. Just as gaze following allows access to mentalizing about other's intentions, beliefs, and expected behavior, so does initiating shared attention allow us to share our thoughts and experiences with others. The motivation to share thoughts and experiences with others, supported by joint and shared attention, was identified by Tomasello et al. (2005) and argued to be what sets humans apart as a species and facilitates shared intentionality and, critically, allows culture to evolve. The model includes the mechanisms conceptualized by previous theorists (Baron-Cohen, 1994, 2005; Perrett & Emery, 1994) and the "social brain" networks reviewed by Stanley and Adolphs (2013), building and expanding on those to create a more complete picture of all the processes at work in sharing attention.

A sense of agency for causing eye gaze shifts in others is captured in the model (Stephenson et al., 2018). Specifically, experiencing agency over gaze shifts one causes in others may be necessary to correctly attribute the outcome as a response to one's action and could facilitate coordinated gaze during the ongoing social interaction, which, in turn, leads to the sociocognitive outcomes, including sense of self and agency over other's actions. A recent theory, termed *sociomotor action control* (Kunde, Weller, & Pfister, 2017), is consistent with this and with third-order social cognition. In the SAS model, the joint-attention initiator needs to detect the response to his or her gaze-leading action in another person to coordinate gaze and lead to ongoing cognitive outcomes. It is this detection of one's action outcomes on other people's behavior that is captured by the idea of sociomotor action control, which is that the responses elicited in another's behavior feedback inform further action control (Kunde et al., 2017). We add to this idea that social responses from other people are much less predictable than typical action-outcome in inanimate objects manipulated by oneself. Arguably, the variance in possible outcomes from another person whose behavior can change on a whim is far greater than the variance in expected outcomes from inanimate objects. Therefore, people need to be particularly flexible in their assessment of feedback from social outcomes; the system must be capable of processing a huge range of responses. For example, in the context of shared attention, possible outcomes include being ignored in one's gaze-leading bid and having to reestablish eye contact and repeat the gaze-leading saccade. We also note that, as shown in the model, incidental joint gaze may occur because one's gaze can be followed without any deliberate intent to establish shared attention, which a person would, nevertheless, benefit from noticing and monitoring.

The model captures the neural mechanisms of the gaze-detection process, the coordination needed between both initiator and responder, and the potential resulting cognitive and affective processes (emotional evaluation, sense of self and agency, reward, empathy, theory of mind, social bonding/affiliation, and person/object knowledge), which are integral to the way people interact as human beings. People's motivation to engage with others is facilitated through shared attention, which is adaptive to functioning in social groups and the shared intentionality people can engage in that makes humans so successful as a species (Tomasello & Herrmann, 2010).

The SAS model contributes to wider theories of social cognition. For example, the model can lend support to Frith's "we-mode" theory that when agents are interacting, they engage in a collective mode of cognition and tend to corepresent actions of social partners

(Gallotti & Frith, 2013). This is supported by studies showing activation of the inferior frontal gyrus when engaging in mutual gaze, specifically, coordinating gaze with a social partner (Cavallo et al., 2015; Koike et al., 2016), which is the same region in which evidence for human mirror neurons has been offered (Kilner, Neal, Weiskopf, Friston, & Frith, 2009). Future research could investigate whether repetition suppression is found both when executing a repeated joint-attention initiation to an object and when observing another person repeating a joint-attention initiation. Such research would enable exploration of whether there is evidence of mirror neurons within the inferior frontal gyrus specifically for joint-attention bids, which would support the idea both of a human mirror neuron system and the overlapping theory of corepresenting a social partner's actions. This would be consistent with the computational model proposed by Triesch, Jasso, and Deák (2007) in which a new class of mirror neurons was postulated that is specific to looking behaviors. There is some empirical evidence for such neurons in the monkey; Shepherd, Klein, Deaner, and Platt (2009) showed that lateral intraparietal neurons fire during attention orienting and observing other monkeys shifting attention.

Koike et al. (2016) used hyperscanning fMRI in which two participants shared attention and found synchronization of neural activity of the inferior frontal gyrus (IFG) during mutual gaze and also IFG activation during both initiating and responding to joint attention. Furthermore, eye blinks were shown to be coordinated during a joint attention task. Taken together, this is further evidence of a shared representational state during shared attention that facilitates theory of mind and other key elements of social cognition.

Social motivation to share experiences with others as a driver of gaze behaviors has been identified and encapsulated in the cooperative-eye hypothesis, which is that the morphology of human eyes evolved to facilitate the cooperation needed in social interactions involving joint attention (Emery, 2000; Kobayashi & Kohshima, 2001; Tomasello et al., 2007). It has also been proposed that social interactions can be conceptualized as a form of mental alignment, which does not necessarily rely on a shared common goal but does involve an exchange of information (Gallotti, Fairhurst, & Frith, 2017). The third-order outcomes in our model flowing from sharing attention all potentially speak to these ideas. The outcomes in the model allow for the exchange of information postulated by Gallotti et al. (2017). We would add that all of the third-order outcomes in our model are socially motivating in one way or another (e.g., the reward value of interactions, social bonding, or access to other's mental states).

Another recent influential theoretical perspective in the field of social neuroscience is that a second-person approach is needed, that is, that social cognition is crucially different when engaging with another in an interaction compared with passive observation (Schilbach et al., 2013). This second-person approach places emotional engagement and the act of social interaction itself as the keys to accessing other's mental states and emphasizes that this takes place in an iterative fashion. Our model, similarly, recognizes the iterative nature of the gaze dynamics between the gaze initiator and gaze responder. Unlike Schilbach et al.'s (2013) second-person approach, our model does not offer a conceptual framework for the whole of social cognition but focuses on the main processes that set up and then flow on from shared attention. Our approach to shared attention is a natural consequence of, and is consistent with, the second-order framework and applied here specifically to joint and shared attention.

Finally, because the model is structured around first-, second-, and third-order social cognition, it is worth briefly considering how the model might speak to a recent theory of autism mentioned in a previous section. The social mismatch hypothesis of autism (Bolis & Schilbach, 2018; Schilbach et al., 2013) acknowledges some first-order typicalities in social cognition for people with autism while also seeking to account for second- and third-order difficulties in social interactions. The social-mismatch hypothesis is that it is the integration of perception and action-based processes that are critical to social interactions and that only by studying two-person real-time interactions (rather than simply passive-observation paradigms) can light to be shed on how the social difficulties arise (Schilbach et al., 2013). Our model is consistent with this emphasis on the importance of looking at the processes at work on the different levels of first-, second-, and third-order cognition and how they might interact and the central importance of both gaze leader and gaze responder being engaged in the interaction.

Developmental trajectory of the SAS. The processes in the model follow a progression from first-order to third-order social cognition, and this can map onto human development of these processes. The first-order processes of detecting gaze and encoding gaze direction are likely to develop early, possibly from birth, given behavioral evidence of orienting to targets that others are looking at (Farroni et al., 2004), but the very early ontology of this has not yet been fully established (for a review, see Del Bianco et al., 2018).

The developmental trajectory of second-order processes such as reorienting attention through gaze following and engaging in joint attention have more

support in the research. Spontaneous gaze following itself seems to develop from 3 to 6 months of age (Del Bianco et al., 2018); and then responding to joint attention, which involves gaze following, develops by around 1 year; and finally, initiating joint attention develops somewhat later during the second year (for a review, see Happé & Frith, 2014).

Moving on to third-order sociocognitive outcomes, the development of these follows on later, and their precise developmental trajectory has been the subject of some debate in large bodies of research (for reviews, see Baillargeon, Scott, & He, 2010; Rakoczy, 2012). In general, we observe that first-order processes will emerge in development before second- and third-order cognition. For example, Schilbach et al. (2013) argued convincingly that development of awareness of other's minds depends on first experiencing minds directed toward the self and that the developmental gaze research evidences this well, with infants showing strong preferences (e.g., Farroni et al., 2002) and neural responses (e.g., Grossmann, Johnson, Farroni, & Csibra, 2007) for self-directed gaze.

Neural systems relevant to the model. The putative brain regions are listed for each process in the model to offer a summary of all of the regions associated with gaze processing, joint attention, and the cognitive outcomes. It is beyond the scope of our model to try to identify how all these regions may interconnect, but see the helpful review of brain regions in typical and autistic neurodevelopment by Mundy (2018) for a more extensive summary. Stanley and Adolphs (2013) identified one current view of a social brain composed of a set of functional networks that each are thought to subservise various social cognition processes. These comprise the amygdala, mentalizing, empathy, and mirror/simulation/action-perception networks. We have included these networks in our model, which are shown in the shared attention, third-order part of our model, alongside the cognitive outcomes, as a guide to how the brain regions identified may relate to established social brain networks. The current view of the social brain will inevitably evolve further, as postulated by Stanley and Adolphs, informed by more extensive modeling frameworks from future imaging studies.

We now offer a summary of the research that has informed the brain regions identified in our model as key contributors to shared attention. The regions involved in detecting gaze presence are the amygdala (Adolphs, 2008; Adolphs & Spezio, 2006; Gamer, Schmitz, Tittgemeyer, & Schilbach, 2013; Kawashima et al., 1999), the hippocampus and lateral fusiform gyrus, and the inferior occipital gyri (reviewed in Nummenmaa & Calder, 2009). A recent study by Mormann et al. (2015) took recordings from neurons in the human amygdala of neurosurgical

patients and found that gaze direction itself was not encoded, although face identity was. The authors suggested that the human amygdala may have selectivity for gaze information but that might be used for information such as identity and emotion rather than for processing gaze direction itself. The right pSTS also has a causal role in orienting to eye gaze, as demonstrated by less orienting to eye regions when inhibitory transcranial magnetic stimulation was applied to this region (Saitovitch et al., 2016). More recently, the ventromedial prefrontal cortex too has been shown to play a role in driving attention to the eye region given that this is impaired in people with lesions to this region (R. C. Wolf, Philippi, Motzkin, Baskaya, & Koenigs, 2014). A review by George and Conty (2008) summarized how perceiving direct gaze has been associated with face and eye movement encoding within 200 ms of perception, while also triggering social brain processes such as emotion and theory of mind, through fMRI, MEG, and EEG methods. These studies show that perception of direct gaze elicits greater fusiform responses than averted gaze and also elicits early recruitment of social brain networks.

Following gaze detection, the encoding of gaze direction has been implicated in the intraparietal sulcus (Hoffman & Haxby, 2000), frontal eye fields (O'Shea, Muggleton, Cowey, & Walsh, 2004), superior parietal lobule and pSTS (Calder et al., 2007), and MT/V5 complex (Watanabe, Kakigi, Miki, & Puce, 2006). Once a gaze shift is detected, the responder reorients attention toward the initiator's gaze-cued location, which involves the inferior parietal lobule (Calder et al., 2007; Perrett et al., 1985, 1992), the bilateral middle frontal gyri, the bilateral superior temporal gyri, the bilateral intraparietal sulci (Thiel, Zilles, & Fink, 2004, 2005), and the superior colliculus (Furlan, Smith, & Walker, 2015). While the initiator is already attending to the referent object, it has been demonstrated that the face of the responder has an attentional capture effect for the initiator, and so reorienting, at least, covertly toward the responder is part of the process for the initiator (Edwards et al., 2015). This gaze-leading effect is theorized to be a mechanism that facilitates the state of joint attention to move onto the higher level sociocognitive state of shared attention because it enables the initiator to monitor the behavior of the responder (Edwards et al., 2015). In addition, people who cooperatively follow gaze leading produce less of a gaze-cueing effect in a person when the person subsequently reencounters them (Dalmaso, Edwards, & Bayliss, 2016). These studies may indicate that shared attention is affected by previous interactions and is not exclusively an automatic process but subject to contextual influences.

One intriguing neural correlate is that shown by Schilbach et al. (2010), who demonstrated enhanced

ventral striatum activity for initiating joint attention, suggesting this is a rewarding experience. This activity also correlated with self-reported subjective feelings of pleasantness. The greater the activity change in the ventral striatum was, the greater the sense of pleasantness reported for looking at objects with another person was. In this case, the other “person” was an on-screen face, but participants were told that the on-screen face was controlled by a real person. Seeking to examine online social interactions rather than offline in this way has become more numerous and influential in recent years (for a review, see Pfeiffer, Vogeley, & Schilbach, 2013).

A further study from the same research group showed that gaze-based behaviors with another person activated the ventral striatum, and it did not matter whether the participants believed their partner had a shared goal (Pfeiffer et al., 2014). Another study found increased striatum activity when initiating joint attention was reciprocated with gaze following compared with an averted gaze response (Gordon, Eilbott, Feldman, Pelphrey, & Vander Wyk, 2013). Finally, the ventral striatum was activated more even when participants simply passively observed actors in a video clip engaging in a shared purpose than when the actors were simply acting in parallel (Eskenazi, Rueschemeyer, de Lange, Knoblich, & Sebanz, 2015).

A revealing recent fMRI study went further and identified functional connections between the visual and dorsal attention networks as initiating joint attention develops in toddlers in a large sample of 1-year-olds ($n = 116$) and 2-year-olds ($n = 98$), 37 of whom provided behavioral and imaging data at both age points (Eggebrecht et al., 2017). Infants were assessed for their initiating joint-attention abilities. Then, brain functional connectivity was measured while the infants slept so that correlations between joint-attention initiation abilities and brain functional connectivity between regions of interest identified by the work in adults could be examined (e.g., Redcay, Kleiner, & Saxe, 2012). Broadly, the findings were that initiating joint-attention abilities was most strongly associated with connectivity between the visual and dorsal attention networks and between the visual network and posterior cingulate default mode network (Eggebrecht et al., 2017).

Cognitive and affective outcomes from the shared attention system. Seven key cognitive and affective outcomes from shared attention are identified in the model: emotional evaluation, sense of self and agency, reward, empathy, theory of mind, social bonding/affiliation, and person/object knowledge. We will consider each of these in turn. Before we do, we should explain that we recognize that most of these processes will overlap in time and have different rise times to activation. We recognize that the processes on the left side of the model diagram are

likely to give rise to behavioral outcomes sooner than those on the right given that the outcomes range from lower-order to higher-order cognition.

First, making emotional evaluations (which are often inferred by decoding a facial expression of emotion) can result from shared attention and may be the first of the outcomes in our model to occur in time. The precise timing of making emotional evaluations is not yet established, although ERP studies have attempted to identify the time course. There have been some mixed findings in using the N170 to attempt to index emotional evaluation. For example, Eimer and Holmes (2002) found the N170 was not affected by facial emotion (using fearful and neutral faces), but fearful faces elicited a frontocentral positivity within 120 ms and a sustained positivity after 250 ms. However, the N170 was modulated by emotion in another study (Williams, Palmer, Liddell, Song, & Gordon, 2006) in which both fear and happiness modulated the N170 relative to neutral faces. Such discrepancies in findings might be accounted for by different treatment of ERP data. For example, Rellecke, Sommer, and Schacht (2013) found that emotional modulations recorded from N170-typical electrodes were less pronounced when a mastoid reference was used than an average one and concluded that N170 modulations were really accounted for by early posterior negativity, which is emotion sensitive. The earliest intracranial recordings from the amygdala for neutral and fearful faces are at 150 ms to 200 ms (Pourtois, Spinelli, Seeck, & Vuilleumier, 2010), whereas what is being recorded on the scalp is thought to be due to modulation of the face system by amygdala top-down effects.

Part of the variability in the research on the time course of emotional evaluation may come from the use of a single task type in most studies. This is illustrated by daSilva, Crager, and Puce (2016), who used both implicit and explicit tasks with respect to emotion in a within-subjects design and showed that the early potential N170 (and even P1) can be modulated by task effects. In addition, daSilva, Crager, Geisler, et al. (2016) showed that changes in local luminance and contrast (e.g., the presence or absence of teeth in an expression) can also have a modulatory effect. So, early ERP effects in studies dealing with emotion may occur because of task in addition to low-level stimulus feature changes. For a more detailed explanation of this, see Puce et al. (2015). Whatever the explanations for the mixed findings, Calvo and Nummenmaa (2016), in a detailed review, concluded that emotional expression is implicitly identified as different from neutral expressions sometime between 150 ms and 300 ms.

Gaze cues and facial expression cues need to be integrated to convey qualitative affective information

of objects in the environment. Some ERP studies and gaze-cuing studies have suggested this occurs at 300 ms (e.g., Graham, Kelland Friesen, Fichtenholtz, & LaBar, 2010; Klucharev & Sams, 2004). More recent studies suggest that this integration might start before or around 200 ms (Conty et al., 2016; McCrackin & Itier, 2018). In an EEG experiment in which gaze and emotion cues were temporally dissociated, responses to gaze only occurred first, those to emotion occurred subsequently, and the interaction between the two occurred still later—at around 700 ms (Ulloa, Puce, Hugueville, & George, 2014).

Turning to associated neural regions, the process of emotional evaluations can be split into evaluations about oneself and evaluations of others' emotions. Those regions involved in one's own emotional evaluations are the insula and the right anterior cingulate cortex (ACC), whereas those used when evaluating other's emotions are the STS and the temporo-parietal junction (TPJ), and those used for both types of evaluation are the amygdala, the lateral prefrontal cortex, and the dorsal medial prefrontal cortex (for a review, see Lee & Siegle, 2009).

Second, initiating shared attention has been associated with the dorsomedial area of the prefrontal cortex (PFC; Schilbach et al., 2010), which has been implicated in processing self-referential information (Bergström, Vogelsang, Benoit, & Simons, 2014; Schmitz & Johnson, 2007) and with regions associated with processing reward such as the ventral striatum, insula, and right ACC (Redcay et al., 2010; Schilbach et al., 2010). Because these two processes of reward and self-referential information implicate the same brain regions, we have placed those processes beside one another in the model. The dorsolateral PFC appears to play a role in sense of agency by monitoring the fluency of action selection processes, whereas the angular gyrus may be involved in detecting agency violations, although there is some inconsistency in findings of the involvement of the angular gyrus (e.g., Beyer, Sidarus, Fleming, & Haggard, 2018; Chambon, Wenke, Fleming, Prinz, & Haggard, 2012; Kühn, Brass, & Haggard, 2013).

Third, the brain regions implicated both by joint attention and the processing of empathy are the pSTS and the TPJ, whereas another overlapping region involved is the ventromedial PFC (vmPFC; for a review, see Bernhardt & Singer, 2012; for an activation likelihood activation meta-analysis, see Bzdok et al., 2012). Fourth, sharing attention with another person facilitates the human attribute of theory of mind and its accompanying potential for cooperation, teaching, control, and communication. The neural mechanisms of theory of mind have been identified as the pSTS, the TPJ, and the medial PFC (Saxe, 2006; for a meta-analysis, see Schurz, Radua, Aichhorn, Richlan, & Perner, 2014). Engaging in

shared attention can lead to affiliative processes, a sense of social inclusion, and rapport (mutual attentiveness, positivity, and coordination, as defined by Tickle-Degnen & Rosenthal, 1990), as demonstrated, for example, by R. C. Wolf et al. (2014). Brain regions associated with being in synchrony with a social partner are the left intraparietal lobule, angular gyrus, vmPFC, and the ACC (Cacioppo et al., 2014).

Finally, there are findings that gaze encounters with a person influence subsequent gaze behaviors when reencountering that person (e.g., Bayliss, Naughtin, Lipp, Kritikos, & Dux, 2012; Edwards & Bayliss, 2019) and influence later personality judgments such as trustworthiness (Bayliss et al., 2007). Brain regions involved in person knowledge are the medial prefrontal cortex, superior temporal cortex, intraparietal sulcus, and the fusiform gyrus (Mitchell, Heatherton, & Macrae, 2002). There may be object knowledge and working memory effects too. For example, Kim and Mundy (2012) demonstrated that gaze leading toward pictures enhanced recognition memory for those stimuli compared with when responding to gaze-cued pictures, even controlling for picture-viewing time. Effects of gaze following on working memory were also reported by Gregory and Jackson (2017), who demonstrated that being gaze cued to stimuli can enhance working memory for cued stimuli compared with when there was no joint attention on the items to be recalled. In addition, affective consequences have been demonstrated for gaze interaction with objects, with increased preferences for objects that have been gazed at by others in a gaze-cueing or gaze-leading context (Bayliss et al., 2013; Bayliss, Paul, Cannon, & Tipper, 2006). Furthermore, objects looked at with happy expressions are preferred over those looked at with disgust (Bayliss et al., 2007).

To summarize, the SAS model captures the neural mechanisms of gaze processing and subsequent social cognitive outcomes, the coordination needed between both joint attention initiator and responder, and the potential resulting cognitive processes (emotional evaluation, sense of self and agency, reward, empathy, theory of mind, social bonding/affiliation, and person/object knowledge), which are integral to the way people interact as human beings. People's motivation to engage with others is facilitated through joint and shared attention, which is adaptive to functioning in social groups and the shared intentionality people can engage in that makes humans so successful as a species (Tomasello & Herrmann, 2010).

Future Directions

The majority of experiments reviewed used a computer-simulated joint-attention paradigm with a virtual other rather than a real-world interaction between dyads. This

facilitates tightly controlled, measurable variables. The field recognizes how grappling with these issues has resulted in compromises to ecological validity in much of the social-gaze-based research, but the field is, nevertheless, increasingly working toward attempts to develop other useful paradigms (Kingstone, Smilek, Ristic, Kelland Friesen, & Eastwood, 2003; Schilbach, 2015).

To address more naturalistic social interactions, many joint-attention researchers have progressed from static to interactive paradigms over the years, and such developments are likely to continue (Pfeiffer et al., 2013), and calls have been made to study real interactions in tandem with experimental, lab-based work (Heerey, 2015). Going further than computer-simulated interactive paradigms, Redcay et al. (2010) had participants interact live via video feed with the experimenter. Koike et al. (2016) deployed hyperscanning fMRI, enabling collection of imaging data from two interacting individuals, and others have employed live joint-attention paradigms to explore electrophysiological effects (Lachat et al., 2012). Of course, portable scanning would enable more naturalistic paradigms to be developed. There have been recent strides toward this with the development of wearable MEG systems (Boto et al., 2018) and the use of virtual reality (Pan & Hamilton, 2018), which may herald more naturalistic interactions becoming the norm in the future.

Another area for future research is to explore EEG data further to seek to evidence the rapid neural detection of gaze responses one elicits following gaze leading and how quickly this happens. The neural time course involved has begun to be addressed (Caruana et al., 2015, 2017). This can be built on to explore whether the face processing of people who follow one's gaze is prioritized in the visual and person knowledge systems. One way to find this out is to examine the electrophysiological signature of visual face processing, the ERP known as the N170 (Bentin et al., 1996), to see whether this is modulated according to gaze response. Such research would build on the highly informative work already undertaken for observing averted and direct gaze shifts, demonstrating rapid detection of eye-gaze signals and alpha wave modulations (Conty et al., 2016; Lachat et al., 2012). Furthermore, another key ERP that could be examined is the N250r that seems specific to faces, increasing on repeated exposure to individuals (Schweinberger & Neumann, 2016). The N250r component seems to index some form of person knowledge and, therefore, could contribute to learning from social-gaze signals. In addition, future behavioral studies could explore whether social learning about faces from repeated encounters with responses to gaze leading results in affective consequences or any effects on face memory.

Summary

The new neurocognitive model of shared attention seeks to capture three main strands. First, the model captures the many findings and advances made over the past 20 years about the brain regions involved in the different processes that make up sharing attention using eye-gaze signals. Second, the model shows how these processes relate to one another in an iterative fashion and, third, identifies the key outcomes that can result from sharing attention: emotional evaluation, sense of self and agency, reward, empathy, theory of mind, social bonding/affiliation, and person/object knowledge. Human motivation to engage with others is facilitated through the shared-attention system that is adaptive to people's shared intentionality and functioning in social groups.

Transparency

Action Editor: Aina Puce

Editor: Laura A. King


Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

The work was supported by Leverhulme Trust Project Grant RPG-2016-173 (to A. P. Bayliss).

ORCID iD

Andrew P. Bayliss  <https://orcid.org/0000-0003-4810-7758>

References

- Adolphs, R. (2008). Fear, faces, and the human amygdala. *Current Opinion in Neurobiology*, *18*, 166–172. doi:10.1016/j.conb.2008.06.006
- Adolphs, R., & Spezio, M. (2006). Role of the amygdala in processing visual social stimuli. *Progress in Brain Research*, *156*, 363–378. doi:10.1016/S0079-6123(06)56020-0
- Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex*, *9*, 415–430. doi:10.1093/cercor/9.5.445
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*, 110–118. doi:10.1016/j.tics.2009.12.006
- Bard, K. A. (2009). Social cognition: Evolutionary history of emotional engagements with infants. *Current Biology*, *19*, R941–R943. doi:10.1016/j.cub.2009.09.037
- Baron-Cohen, S. (1994). The mindreading system: New directions for research. *Current Psychology of Cognition*, *13*, 724–750.

- Baron-Cohen, S. (2005). The empathizing system: A revision of the 1994 model of the mindreading system. In B. Ellis & D. Bjorklund (Eds.), *Origins of the social mind: Evolutionary psychology and child development* (pp. 468–492). New York, NY: Guilford.
- Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, *104*, 644–653. doi:10.1016/j.cognition.2006.07.012
- Bayliss, A. P., Murphy, E., Naughtin, C. K., Kritikos, A., Schilbach, L., & Becker, S. I. (2013). “Gaze leading”: Initiating simulated joint attention influences eye movements and choice behavior. *Journal of Experimental Psychology: General*, *142*, 76–92. doi:10.1037/a0029286
- Bayliss, A. P., Naughtin, C. K., Lipp, O. V., Kritikos, A., & Dux, P. E. (2012). Make a lasting impression: The neural consequences of re-encountering people who emotate inappropriately. *Psychophysiology*, *49*, 1571–1578. doi:10.1111/j.1469-8986.2012.01481.x
- Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review*, *13*, 1061–1066.
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565. doi:10.1162/jocn.1996.8.6.551
- Bergström, Z. M., Vogelsang, D. A., Benoit, R. G., & Simons, J. S. (2014). Reflections of oneself: Neurocognitive evidence for dissociable forms of self-referential recollection. *Cerebral Cortex*, *25*, 2648–2657. doi:10.1093/cercor/bhu063
- Bernhardt, B. C., & Singer, T. (2012). The neural basis of empathy. *Annual Review of Neuroscience*, *35*, 1–23. doi:10.1146/annurev-neuro-062111-150536
- Beyer, F., Sidarus, N., Fleming, S., & Haggard, P. (2018). Losing control in social situations: How the presence of others affects neural processes related to sense of agency. *eNeuro*, *5*, Article ENEURO.0336-17.2018. doi:10.1523/ENEURO.0336-17.2018
- Billeci, L., Narzisi, A., Campatelli, G., Crifaci, G., Calderoni, S., Gagliano, A., . . . Muratori, F. (2016). Disentangling the initiation from the response in joint attention: An eye-tracking study in toddlers with autism spectrum disorders. *Translational Psychiatry*, *6*, Article e808. doi:10.1038/tp.2016.75
- Binetti, N., Harrison, C., Coutrot, A., Johnston, A., & Mareschal, I. (2016). Pupil dilation as an index of preferred mutual gaze duration. *Royal Society Open Science*, *3*, Article 160086. doi:10.1098/rsos.160086
- Birmingham, E., Stanley, D., Nair, R., & Adolphs, R. (2015). Implicit social biases in people with autism. *Psychological Science*, *26*, 1693–1705. doi:10.1177/0956797615595607
- Bolis, D., & Schilbach, L. (2018). Observing and participating in social interactions: Action perception and action control across the autistic spectrum. *Developmental Cognitive Neuroscience*, *29*, 168–175. doi:10.1016/j.dcn.2017.01.009
- Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S. S., . . . Bestmann, S. (2018). Moving magnetoencephalography towards real-world applications with a wearable system. *Nature*, *555*, 657–661.
- Bottini, S. (2018). Social reward processing in individuals with autism spectrum disorder: A systematic review of the social motivation hypothesis. *Research in Autism Spectrum Disorders*, *45*, 9–26. doi:10.1016/j.rasd.2017.10.001
- Boucenna, S., Narzisi, A., Tilmont, E., Muratori, F., Pioggia, G., Cohen, D., & Chetouani, M. (2014). Interactive technologies for autistic children: A review. *Cognitive Computation*, *6*, 722–740.
- Bräuer, J., Call, J., & Tomasello, M. (2005). All great ape species follow gaze to distant locations and around barriers. *Journal of Comparative Psychology*, *119*, 145–154. doi:10.1037/0735-7036.119.2.145
- Brignani, D., Guzzon, D., Marzi, C. A., & Miniussi, C. (2009). Attentional orienting induced by arrows and eye-gaze compared with an endogenous cue. *Neuropsychologia*, *47*, 370–381. doi:10.1016/j.neuropsychologia.2008.09.011
- Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology*, *130*, 67–78. doi:10.1016/j.jecp.2014.09.010
- Bruner, J. S., & Sherwood, V. (1975). Early rule structure: The case of peekaboo. In J. S. Bruner, A. Jolly, & K. Sylva (Eds.), *Play: Its role in evolution and development* (pp. 277–285). New York, NY: Penguin Books.
- Burra, N., Baker, S., & George, N. (2017). Processing of gaze direction within the N170/M170 time window: A combined EEG/MEG study. *Neuropsychologia*, *100*, 207–219. doi:10.1016/j.neuropsychologia.2017.04.028
- Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, *217*, 783–796. doi:10.1007/s00429-012-0380-y
- Cacioppo, S., Zhou, H., Monteleone, G., Majka, E. A., Quinn, K. A., Ball, A. B., . . . Cacioppo, J. T. (2014). You are in sync with me: Neural correlates of interpersonal synchrony with a partner. *Neuroscience*, *277*, 842–858. doi:10.1016/j.neuroscience.2014.07.051
- Calder, A. J., Beaver, J. D., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., & Henson, R. N. A. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Current Biology*, *17*, 20–25. doi:10.1016/j.cub.2006.10.052
- Calder, A. J., Jenkins, R., Cassel, A., & Clifford, C. W. (2008). Visual representation of eye gaze is coded by a non-opponent multichannel system. *Journal of Experimental Psychology: General*, *137*, 244–261. doi:10.1037/0096-3445.137.2.244
- Calvo, M. G., & Nummenmaa, L. (2016). Perceptual and affective mechanisms in facial expression recognition: An integrative review. *Cognition and Emotion*, *30*, 1081–1106. doi:10.1080/02699931.2015.1049124

- Carpenter, M., & Call, J. (2013). How joint is the joint attention of apes and human infants? In J. M. Metcalfe & H. S. Terrace (Eds.), *Agency and joint attention* (pp. 49–61). New York, NY: Oxford University Press.
- Caruana, N., de Lissa, P., & McArthur, G. (2015). The neural time course of evaluating self-initiated joint attention bids. *Brain and Cognition*, *98*, 43–52. doi:10.1016/j.bandc.2015.06.001
- Caruana, N., de Lissa, P., & McArthur, G. (2017). Beliefs about human agency influence the neural processing of gaze during joint attention. *Social Neuroscience*, *12*, 194–206. doi:10.1080/17470919.2016.1160953
- Cavallo, A., Lungu, O., Becchio, C., Ansuini, C., Rustichini, A., & Fadiga, L. (2015). When gaze opens the channel for communication: Integrative role of IFG and MPFC. *NeuroImage*, *119*, 63–69. doi:10.1016/j.neuroimage.2015.06.025
- Chambon, V., Wenke, D., Fleming, S. M., Prinz, W., & Haggard, P. (2012). An online neural substrate for sense of agency. *Cerebral Cortex*, *23*, 1031–1037. doi:10.1093/cercor/bhs059
- Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*, 231–239. doi:10.1016/j.tics.2012.02.007
- Constantino, J. N., Kennon-McGill, S., Weichselbaum, C., Marrus, N., Haider, A., Glowinski, A. L., . . . Jones, W. (2017). Infant viewing of social scenes is under genetic control and is atypical in autism. *Nature*, *547*, 340–344. doi:10.1038/nature22999
- Conty, L., George, N., & Hietanen, J. K. (2016). Watching eyes effects: When others meet the self. *Consciousness and Cognition*, *45*, 184–197. doi:10.1016/j.concog.2016.08.016
- Conty, L., N'Diaye, K., Tijus, C., & George, N. (2007). When eye creates the contact! ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia*, *45*, 3024–3037. doi:10.1016/j.neuropsychologia.2007.05.017
- Dalmaso, M., Edwards, S. G., & Bayliss, A. P. (2016). Re-encountering individuals who previously engaged in joint gaze modulates subsequent gaze cueing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 271–284.
- daSilva, E. B., Crager, K., Geisler, D., Newbern, P., Orem, B., & Puce, A. (2016). Something to sink your teeth into: The presence of teeth augments ERPs to mouth expressions. *NeuroImage*, *127*, 227–241. doi:10.1016/j.neuroimage.2015.12.020
- daSilva, E. B., Crager, K., & Puce, A. (2016). On dissociating the neural time course of the processing of positive emotions. *Neuropsychologia*, *83*, 123–137. doi:10.1016/j.neuropsychologia.2015.12.001
- David, N., Newen, A., & Vogeley, K. (2008). The “sense of agency” and its underlying cognitive and neural mechanisms. *Consciousness and Cognition*, *17*, 523–534. doi:10.1016/j.concog.2008.03.004
- Del Bianco, T., Falck-Ytter, T., Thorup, E., & Gredebäck, G. (2018). The developmental origins of gaze-following in human infants. *Infancy*, *24*, 433–454. doi:10.1111/infa.12276
- Edwards, S. G., & Bayliss, A. (2019). Seeing eye-to-eye: Social gaze interactions influence gaze direction identification. *Attention, Perception & Psychophysics*, *81*, 2755–2765. doi:10.3758/s13414-019-01671-1
- 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*, Article 20151141. doi:10.1098/rspb.2015.1141
- Eggebrecht, A. T., Elison, J. T., Feczko, E., Todorov, A., Wolff, J. J., Kandala, S., . . . Estes, A. M. (2017). Joint attention and brain functional connectivity in infants and toddlers. *Cerebral Cortex*, *27*, 1709–1720. doi:10.1093/cercor/bhw403
- Eimer, M., & Holmes, A. (2002). An ERP study on the time course of emotional face processing. *NeuroReport*, *13*, 427–431.
- Eimer, M., Van Velzen, J., & Driver, J. (2002). Cross-modal interactions between audition, touch, and vision in endogenous spatial attention: ERP evidence on preparatory states and sensory modulations. *Journal of Cognitive Neuroscience*, *14*, 254–271. doi:10.1162/089892902317236885
- Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, *24*, 581–604. doi:10.1016/S0149-7634(00)00025-7
- Engell, A. D., & McCarthy, G. (2014). Face, eye, and body selective responses in fusiform gyrus and adjacent cortex: An intracranial EEG study. *Frontiers in Human Neuroscience*, *8*, Article 642. doi:10.3389/fnhum.2014.00642
- Eskenazi, T., Rueschemeyer, S.-A., de Lange, F. P., Knoblich, G., & Sebanz, N. (2015). Neural correlates of observing joint actions with shared intentions. *Cortex*, *70*, 90–100. doi:10.1016/j.cortex.2015.05.027
- Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, USA*, *99*, 9602–9605. doi:10.1073/pnas.152159999
- Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy*, *5*, 39–60. doi:10.1207/s15327078in0501_2
- Freeman, S. F. N., Gulsrud, A., & Kasari, C. (2015). Brief report: Linking early joint attention and play abilities to later reports of friendships for children with ASD. *Journal of Autism and Developmental Disorders*, *45*, 2259–2266. doi:10.1007/s10803-015-2369-x
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*, 694–724. doi:10.1037/0033-2909.133.4.694
- Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology*, *17*, R724–R732. doi:10.1016/j.cub.2007.05.068
- Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, *63*, 287–313.
- Furlan, M., Smith, A. T., & Walker, R. (2015). Activity in the human superior colliculus relating to endogenous saccade preparation and execution. *Journal of Neurophysiology*, *114*, 1048–1058. doi:10.1152/jn.00825.2014

- Galfano, G., Sarlo, M., Sassi, F., Munafò, M., Fuentes, L. J., & Umiltà, C. (2011). Reorienting of spatial attention in gaze cuing is reflected in N2pc. *Social Neuroscience*, *6*, 257–269. doi:10.1080/17470919.2010.515722
- Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, *4*, 14–21. doi:10.1016/S1364-6613(99)01417-5
- Gallotti, M., Fairhurst, M. T., & Frith, C. D. (2017). Alignment in social interactions. *Consciousness and Cognition*, *48*, 253–261. doi:10.1016/j.concog.2016.12.002
- Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences*, *17*, 160–165. doi:10.1016/j.tics.2013.02.002
- Gamer, M., Schmitz, A. K., Tittgemeyer, M., & Schilbach, L. (2013). The human amygdala drives reflexive orienting towards facial features. *Current Biology*, *23*, R917–R918. doi:10.1016/j.cub.2013.09.008
- George, N., & Conty, L. (2008). Facing the gaze of others. *Neurophysiologie Clinique/Clinical Neurophysiology*, *38*, 197–207. doi:10.1016/j.neucli.2008.03.001
- Gillespie-Lynch, K. (2013). Response to and initiation of joint attention: Overlapping but distinct roots of development in autism. *OA Autism*, *1*, Article 13. doi:10.13172/2052-7810-1-2-596
- Goods, K. S., Ishijima, E., Chang, Y.-C., & Kasari, C. (2013). Preschool based JASPER intervention in minimally verbal children with autism: Pilot RCT. *Journal of Autism and Developmental Disorders*, *43*, 1050–1056. doi:10.1007/s10803-012-1644-3
- Gordon, I., Eilbott, J. A., Feldman, R., Pelphrey, K. A., & Vander Wyk, B. C. (2013). Social, reward, and attention brain networks are involved when online bids for joint attention are met with congruent versus incongruent responses. *Social Neuroscience*, *8*, 544–554. doi:10.1080/17470919.2013.832374
- Graham, R., Kelland Friesen, C., Fichtenholtz, H. M., & LaBar, K. S. (2010). Modulation of reflexive orienting to gaze direction by facial expressions. *Visual Cognition*, *18*, 331–368. doi:10.1080/13506280802689281
- Gredebäck, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: A longitudinal study of gaze following during interactions with mothers and strangers. *Developmental Science*, *13*, 839–848. doi:10.1111/j.1467-7687.2009.00945.x
- Gregory, S. E. A., & Jackson, M. C. (2017). Joint attention enhances visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 237–249. doi:10.1037/xlm0000294
- Grossmann, T., Johnson, M. H., Farroni, T., & Csibra, G. (2007). Social perception in the infant brain: Gamma oscillatory activity in response to eye gaze. *Social Cognitive and Affective Neuroscience*, *2*, 284–291. doi:10.1093/scan/nsm025
- Grubert, A., & Eimer, M. (2015). Rapid parallel attentional target selection in single-color and multiple-color visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 86–101.
- Grynszpan, O., Weiss, P. L., Perez-Diaz, F., & Gal, E. (2014). Innovative technology-based interventions for autism spectrum disorders: A meta-analysis. *Autism*, *18*, 346–361. doi:10.1177/1362361313476767
- Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*, 196–207. doi:10.1038/nrn.2017.14
- Hamilton, A. F. de C. (2016). Gazing at me: The importance of social meaning in understanding direct-gaze cues. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*, Article 20150080. doi:10.1098/rstb.2015.0080
- Happé, F., Cook, J. L., & Bird, G. (2017). The structure of social cognition: In(ter)dependence of sociocognitive processes. *Annual Review of Psychology*, *68*, 243–267.
- Happé, F., & Frith, U. (2014). Annual research review: Towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry*, *55*, 553–577. doi:10.1111/jcpp.12162
- Harter, M. R., Miller, S. L., Price, N. J., LaLonde, M. E., & Keyes, A. L. (1989). Neural processes involved in directing attention. *Journal of Cognitive Neuroscience*, *1*, 223–237. doi:10.1162/jocn.1989.1.3.223
- Heerey, E. A. (2015). Decoding the dyad: Challenges in the study of individual differences in social behavior. *Current Directions in Psychological Science*, *24*, 285–291. doi:10.1177/0963721415570731
- Hietanen, J. K., Leppänen, J. M., Nummenmaa, L., & Astikainen, P. (2008). Visuospatial attention shifts by gaze and arrow cues: An ERP study. *Brain Research*, *1215*, 123–136. doi:10.1016/j.brainres.2008.03.091
- Hoehl, S., Wiese, L., & Striano, T. (2008). Young infants' neural processing of objects is affected by eye gaze direction and emotional expression. *PLOS ONE*, *3*(6), Article e2389. doi:10.1371/journal.pone.0002389
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*, 80–84. doi:10.1038/71152
- Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*, 131–134. doi.org/10.1111/1467-9280.00024
- Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience & Biobehavioral Reviews*, *33*, 843–863. doi:10.1016/j.neubiorev.2009.02.004
- Itier, R. J., & Preston, F. (2018). Increased early sensitivity to eyes in mouthless faces: In support of the LIFTED model of early face processing. *Brain Topography*, *31*, 972–984. doi:10.1007/s10548-018-0663-6
- Jaswal, V. K., & Akhtar, N. (2018). Being vs. appearing socially uninterested: Challenging assumptions about social motivation in autism. *Behavioral and Brain Sciences*, *42*, Article e82. doi:10.1017/S0140525X18001826
- Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., . . . Nakamura, K. (1999). The human amygdala plays an important role in gaze monitoring: A PET study. *Brain*, *122*, 779–783. doi:10.1093/brain/122.4.779
- Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior

- frontal gyrus. *The Journal of Neuroscience*, *29*, 10153–10159. doi:10.1523/JNEUROSCI.2668-09.2009
- Kim, K., & Mundy, P. (2012). Joint attention, social-cognition, and recognition memory in adults. *Frontiers in Human Neuroscience*, *6*, Article 172. doi:10.3389/fnhum.2012.00172
- Kingstone, A., Smilek, D., Ristic, J., Kelland Friesen, C., & Eastwood, J. D. (2003). Attention, researchers! It is time to take a look at the real world. *Current Directions in Psychological Science*, *12*, 176–180. doi:10.1111/1467-8721.01255
- Kirk Driller, K., Stephani, T., Dimigen, O., & Sommer, W. (2019). Large lateralized EDAN-like brain potentials in a gaze-shift detection task. *Psychophysiology*, *56*, Article 13361. doi:10.1111/psyp.13361
- Klucharev, V., & Sams, M. (2004). Interaction of gaze direction and facial expressions processing: ERP study. *NeuroReport*, *15*, 621–625.
- Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature*, *387*, 767–768. doi:10.1038/42842
- 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*, 419–435. doi:10.1006/jhev.2001.0468
- Koike, T., Tanabe, H. C., Okazaki, S., Nakagawa, E., Sasaki, A. T., Shimada, K., . . . Sadato, N. (2016). Neural substrates of shared attention as social memory: A hyperscanning functional magnetic resonance imaging study. *NeuroImage*, *125*, 401–412. doi:10.1016/j.neuroimage.2015.09.076
- Kühn, S., Brass, M., & Haggard, P. (2013). Feeling in control: Neural correlates of experience of agency. *Cortex*, *49*, 1935–1942. doi:10.1016/j.cortex.2012.09.002
- Kunde, W., Weller, L., & Pfister, R. (2017). Sociomotor action control. *Psychonomic Bulletin & Review*, *25*, 917–931. doi:10.3758/s13423-017-1316-6
- Lachat, F., Hugueville, L., Lemaréchal, J., Conty, L., & George, N. (2012). Oscillatory brain correlates of live joint attention: A dual-EEG study. *Frontiers in Human Neuroscience*, *6*, Article 156. doi:10.3389/fnhum.2012.00156
- Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*, 50–59. doi:10.1016/S1364-6613(99)01436-9
- Lassalle, A., & Itier, R. J. (2013). Fearful, surprised, happy, and angry facial expressions modulate gaze-oriented attention: Behavioral and ERP evidence. *Social Neuroscience*, *8*, 583–600. doi:10.1080/17470919.2013.835750
- Lassalle, A., & Itier, R. J. (2015). Autistic traits influence gaze-oriented attention to happy but not fearful faces. *Social Neuroscience*, *10*, 70–88. doi:10.1080/17470919.2014.958616
- Latinus, M., Love, S. A., Rossi, A., Parada, F. J., Huang, L., Conty, L., . . . Puce, A. (2014). Social decisions affect neural activity to perceived dynamic gaze. *Social Cognitive and Affective Neuroscience*, *10*, 1557–1567. doi:10.1093/scan/nsv049
- Lawson, R. P., Clifford, C. W., & Calder, A. J. (2011). A real head turner: Horizontal and vertical head directions are multichannel coded. *Journal of Vision*, *11*(9), Article 17. doi:10.1167/11.9.17
- Lawton, K., & Kasari, C. (2012). Teacher-implemented joint attention intervention: Pilot randomized controlled study for preschoolers with autism. *Journal of Consulting and Clinical Psychology*, *80*, 687–693. doi:10.1037/a0028506
- Leavens, D., & Racine, T. P. (2009). Joint attention in apes and humans: Are humans unique? *Journal of Consciousness Studies*, *16*, 240–267.
- Lee, K. H., & Siegle, G. J. (2009). Common and distinct brain networks underlying explicit emotional evaluation: A meta-analytic study. *Social Cognitive and Affective Neuroscience*, *7*, 521–534. doi:10.1093/scan/nsp001
- Manninen, S., Tuominen, L., Dunbar, R. I., Karjalainen, T., Hirvonen, J., Arponen, E., . . . Nummenmaa, L. (2017). Social laughter triggers endogenous opioid release in humans. *The Journal of Neuroscience*, *37*, 6125–6131. doi:10.1523/JNEUROSCI.0688-16.2017
- McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex*, *9*, 431–444. doi:10.1093/cercor/9.5.431
- McCrackin, S. D., & Itier, R. J. (2018). Both fearful and happy expressions interact with gaze direction by 200 ms SOA to speed attention orienting. *Visual Cognition*, *26*, 231–252. doi:10.1080/13506285.2017.1420118
- McCrackin, S. D., & Itier, R. J. (2019). Perceived gaze direction differentially affects discrimination of facial emotion, attention and gender—an ERP study. *Frontiers in Neuroscience*, *13*, Article 517. doi:10.3389/fnins.2019.00517
- Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserving person and object knowledge. *Proceedings of the National Academy of Sciences, USA*, *99*, 15238–15243. doi:10.1073/pnas.232395699
- Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition*, *21*, 546–561. doi:10.1016/j.concog.2011.12.002
- Morales, M., Mundy, P., Delgado, C. E., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to joint attention across the 6-through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology*, *21*, 283–298.
- Mormann, F., Niediek, J., Tudusciuc, O., Quesada, C. M., Coenen, V. A., Elger, C. E., & Adolphs, R. (2015). Neurons in the human amygdala encode face identity, but not gaze direction. *Nature Neuroscience*, *18*, 1568–1570. doi:10.1038/nn.4139
- Mosher, C. P., Zimmerman, P. E., & Gothard, K. M. (2014). Neurons in the monkey amygdala detect eye contact during naturalistic social interactions. *Current Biology*, *24*, 2459–2464. doi:10.1016/j.cub.2014.08.063
- Mundy, P. (2018). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *European Journal of Neuroscience*, *47*, 497–514.
- Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development*, *78*, 938–954.

- Mundy, P., Kim, K., McIntyre, N., & Lerro, L. (2016). Brief report: Joint attention and information processing in children with higher functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders, 46*, 2555–2560. doi:10.1007/s10803-016-2785-6
- Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science, 16*, 269–274. doi:10.1111/j.1467-8721.2007.00518.x
- Murza, K. A., Schwartz, J. B., Hahs-Vaughn, D. L., & Nye, C. (2016). Joint attention interventions for children with autism spectrum disorder: A systematic review and meta-analysis. *International Journal of Language & Communication Disorders, 51*, 236–251. doi:10.1111/1460-6984.12212
- Myllyneva, A., & Hietanen, J. K. (2016). The dual nature of eye contact: To see and to be seen. *Social Cognitive and Affective Neuroscience, 11*, 1089–1095. doi:10.1093/scan/nsv075
- Nation, K., & Penny, S. (2008). Sensitivity to eye gaze in autism: Is it normal? Is it automatic? Is it social? *Development and Psychopathology, 20*, 79–97. doi:10.1017/S0954579408000047
- Nemrodov, D., Anderson, T., Preston, F. F., & Itier, R. J. (2014). Early sensitivity for eyes within faces: A new neuronal account of holistic and featural processing. *NeuroImage, 97*, 81–94. doi:10.1016/j.neuroimage.2014.04.042
- Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Sciences, 13*, 135–143. doi:10.1016/j.tics.2008.12.006
- O'Shea, J., Muggleton, N. G., Cowey, A., & Walsh, V. (2004). Timing of target discrimination in human frontal eye fields. *Journal of Cognitive Neuroscience, 16*, 1060–1067. doi:10.1162/0898929041502634
- Palmer, C. J., Lawson, R. P., Shankar, S., Clifford, C. W. G., & Rees, G. (2018). Autistic adults show preserved normalisation of sensory responses in gaze processing. *Cortex, 103*, 13–23. doi:10.1016/j.cortex.2018.02.005
- Pan, X., & Hamilton, A. F. D. C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology, 109*, 395–417. doi:10.1111/bjop.12290
- Pantelis, P. C., & Kennedy, D. P. (2017). Deconstructing atypical eye gaze perception in autism spectrum disorder. *Scientific Reports, 7*, Article 14990. doi:10.1038/s41598-017-14919-3
- Parkington, K. B., & Itier, R. J. (2018). One versus two eyes makes a difference! Early face perception is modulated by featural fixation and feature context. *Cortex, 109*, 35–49. doi:10.1016/j.cortex.2018.08.025
- Parkington, K. B., & Itier, R. J. (2019). From eye to face: The impact of face outline, feature number, and feature saliency on the early neural response to faces. *Brain Research, 1722*, Article 146343. doi:10.1016/j.brainres.2019.146343
- Pell, P. J., Mareschal, I., Calder, A. J., von dem Hagen, E. A., Clifford, C. W., Baron-Cohen, S., & Ewbank, M. P. (2016). Intact priors for gaze direction in adults with high-functioning autism spectrum conditions. *Molecular Autism, 7*, Article 25. doi:10.1186/s13229-016-0085-9
- 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*, 598–603. doi:10.1111/j.0956-7976.2004.00726.x
- Perrett, D. I., & Emery, N. J. (1994). Understanding the intentions of others from visual signals: Neurophysiological evidence. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 13*, 683–694.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., Benson, P. J., & Rolls, E. T. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences, 335*, 23–30. doi:10.1098/rstb.1992.0003
- Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society B: Biological Sciences, 223*, 293–317. doi:10.1098/rspb.1985.0003
- Pfeiffer, U. J., Schilbach, L., Jording, M., Timmermans, B., Bente, G., & Vogeley, K. (2012). Eyes on the mind: Investigating the influence of gaze dynamics on the perception of others in real-time social interaction. *Frontiers in Psychology, 3*, Article 537. doi:10.3389/fpsyg.2012.00537
- Pfeiffer, U. J., Schilbach, L., Timmermans, B., Kuzmanovic, B., Georgescu, A. L., Bente, G., & Vogeley, K. (2014). Why we interact: On the functional role of the striatum in the subjective experience of social interaction. *NeuroImage, 101*, 124–137. doi:10.1016/j.neuroimage.2014.06.061
- Pfeiffer, U. J., Vogeley, K., & Schilbach, L. (2013). From gaze cueing to dual eye-tracking: Novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience & Biobehavioral Reviews, 37*, 2516–2528. doi:10.1016/j.neubiorev.2013.07.017
- Pourtois, G., Spinelli, L., Seeck, M., & Vuilleumier, P. (2010). Temporal precedence of emotion over attention modulations in the lateral amygdala: Intracranial ERP evidence from a patient with temporal lobe epilepsy. *Cognitive, Affective, & Behavioral Neuroscience, 10*, 83–93. doi:10.3758/CABN.10.1.83
- Psouni, E., Falck, A., Boström, L., Persson, M., Sidén, L., & Wallin, M. (2018). Together I can! Joint attention boosts 3-to 4-year-olds' performance in a verbal false-belief test. *Cbild Development, 90*, 35–50. doi:10.1111/cdev.13075
- Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex, 9*, 445–458. doi:10.1093/cercor/9.5.445
- Puce, A., Latinus, M., Rossi, A., Parada, F., Love, S., Ashourvan, A., & Jayaraman, S. (2015). Neural bases for social attention in healthy humans. In A. Puce & I. B. Bertenthal (Eds.), *The many faces of social attention: Behavioural and neural measures* (pp. 93–127). New York, NY: Springer.
- Rakoczy, H. (2012). Do infants have a theory of mind? *British Journal of Developmental Psychology, 30*, 59–74. doi:10.1111/j.2044-835X.2011.02061.x

- Rayson, H., Bonaiuto, J. J., Ferrari, P. F., Chakrabarti, B., & Murray, L. (2019). Building blocks of joint attention: Early sensitivity to having one's own gaze followed. *Developmental Cognitive Neuroscience, 37*, Article 100631. doi:10.1016/j.dcn.2019.100631
- Redcay, E., Dodell-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D. E., & Saxe, R. (2010). Live face-to-face interaction during fMRI: A new tool for social cognitive neuroscience. *NeuroImage, 50*, 1639–1647. doi:10.1016/j.neuroimage.2010.01.052
- Redcay, E., Kleiner, M., & Saxe, R. (2012). Look at this: The neural correlates of initiating and responding to bids for joint attention. *Frontiers in Human Neuroscience, 6*, Article 169. doi:10.3389/fnhum.2012.00169
- Reichow, B., & Volkmar, F. R. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders, 40*, 149–166. doi:10.1007/s10803-009-0842-0
- Rellecke, J., Sommer, W., & Schacht, A. (2013). Emotion effects on the N170: A question of reference? *Brain Topography, 26*, 62–71. doi:10.1007/s10548-012-0261-y
- Rossi, A., Parada, F. J., Latinus, M., & Puce, A. (2015). Photographic but not line-drawn faces show early perceptual neural sensitivity to eye gaze direction. *Frontiers in Human Neuroscience, 9*, Article 185. doi:10.3389/fnhum.2015.00185
- Saitovitch, A., Popa, T., Lemaitre, H., Rechtman, E., Lamy, J.-C., Grévent, D., . . . Samson, Y. (2016). Tuning eye-gaze perception by transitory STS Inhibition. *Cerebral Cortex, 26*, 2823–2831. doi:10.1093/cercor/bhw045
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology, 16*, 235–239. doi:10.1016/j.conb.2006.03.001
- Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*, 265–266.
- Schilbach, L. (2015). Eye to eye, face to face and brain to brain: Novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. *Current Opinion in Behavioral Sciences, 3*, 130–135. doi:10.1016/j.cobeha.2015.03.006
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences, 36*, 393–414. doi:10.1017/S0140525X12000660
- 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*, 2702–2715. doi:10.1162/jocn.2009.21401
- Schmitz, T. W., & Johnson, S. C. (2007). Relevance to self: A brief review and framework of neural systems underlying appraisal. *Neuroscience & Biobehavioral Reviews, 31*, 585–596. doi:10.1016/j.neubiorev.2006.12.003
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews, 42*, 9–34. doi:10.1016/j.neubiorev.2014.01.009
- Schweinberger, S. R., & Neumann, M. F. (2016). Repetition effects in human ERPs to faces. *Cortex, 80*, 141–153. doi:10.1016/j.cortex.2015.11.001
- Segal, N. L., Goetz, A. T., & Maldonado, A. C. (2016). Preferences for visible white sclera in adults, children and autism spectrum disorder children: Implications of the cooperative eye hypothesis. *Evolution and Human Behavior, 37*, 35–39. doi:10.1016/j.evolhumbehav.2015.06.006
- Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences, 13*, 127–134. doi:10.1016/j.tics.2008.11.009
- Senju, A., Johnson, M. H., & Csibra, G. (2006). The development and neural basis of referential gaze perception. *Social Neuroscience, 1*, 220–234. doi:10.1080/17470910600989797
- Shepherd, S. V., Klein, J. T., Deaner, R. O., & Platt, M. L. (2009). Mirroring of attention by neurons in macaque parietal cortex. *Proceedings of the National Academy of Sciences, USA, 106*, 9489–9494. doi:10.1073/pnas.0900419106
- Shteynberg, G. (2015). Shared attention. *Perspectives on Psychological Science, 10*, 579–590. doi:10.1177/1745691615589104
- Stanley, D. A., & Adolphs, R. (2013). Toward a neural basis for social behavior. *Neuron, 80*, 816–826. doi:10.1016/j.neuron.2013.10.038
- Stavropoulos, K. K. M., & Carver, L. J. (2013). Research review: Social motivation and oxytocin in autism—implications for joint attention development and intervention. *Journal of Child Psychology and Psychiatry, 54*, 603–618. doi:10.1111/jcpp.12061
- Stephenson, L. J., Edwards, S. G., Howard, E. E., & Bayliss, A. P. (2018). Eyes that bind us: Gaze leading induces an implicit sense of agency. *Cognition, 172*, 124–133. doi:10.1016/j.cognition.2017.12.011
- Stephenson, L. J., Edwards, S. G., Luri, N., Renoult, L., & Bayliss, A. (2020). The N170 event-related potential differentiates congruent and incongruent gaze responses in gaze leading. *Social Cognitive and Affective Neuroscience, 15*, 479–486.
- Taylor, M. J., Edmonds, G. E., McCarthy, G., & Allison, T. (2001). Eyes first! Eye processing develops before face processing in children. *NeuroReport, 12*, 1671–1676.
- Tazumi, T., Hori, E., Maior, R. S., Ono, T., & Nishijo, H. (2010). Neural correlates to seen gaze-direction and head orientation in the macaque monkey amygdala. *Neuroscience, 169*, 287–301. doi:10.1016/j.neuroscience.2010.04.028
- Terrace, H. S. (2013). Becoming human: Why two minds are better than one. In J. Metcalf & H. S. Terrace (Eds.), *Agency and joint attention* (pp. 11–48). New York, NY: Oxford University Press.
- Thiel, C. M., Zilles, K., & Fink, G. R. (2004). Cerebral correlates of alerting, orienting and reorienting of visuospatial attention: An event-related fMRI study. *NeuroImage, 21*, 318–328. doi:10.1016/j.neuroimage.2003.08.044
- Thiel, C. M., Zilles, K., & Fink, G. R. (2005). Nicotine modulates reorienting of visuospatial attention and neural activity in human parietal cortex. *Neuropsychopharmacology, 30*, 810–820. doi:10.1038/sj.npp.1300633

- Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry, 1*, 285–293. doi:10.1207/s15327965pli0104_1
- Tipples, J., Johnston, P., & Mayes, A. (2013). Electrophysiological responses to violations of expectation from eye gaze and arrow cues. *Social Cognitive and Affective Neuroscience, 8*, 509–514. doi:10.1093/scan/nss024
- Todorović, D. (2006). Geometrical basis of perception of gaze direction. *Vision Research, 46*, 3549–3562. doi:10.1016/j.visres.2006.04.011
- Todorović, D. (2009). The effect of face eccentricity on the perception of gaze direction. *Perception, 38*, 109–132. doi.org/10.1068/p5930
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). In search of the uniquely human. *Behavioral and Brain Sciences, 28*, 721–727. doi:10.1017/S0140525X05540123
- Tomasello, M., & Farrar, M. J. (1986). Object permanence and relational words: A lexical training study. *Journal of Child Language, 13*, 495–505. doi:10.1017/S03050009000684X
- Tomasello, M., Hare, B., & Fogleman, T. (2001). The ontogeny of gaze following in chimpanzees, *Pan troglodytes*, and rhesus macaques, *Macaca mulatta*. *Animal Behaviour, 61*, 335–343. doi:10.1006/anbe.2000.1598
- Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution, 52*, 314–320. doi:10.1016/j.jhevol.2006.10.001
- Tomasello, M., & Herrmann, E. (2010). Ape and human cognition: What's the difference? *Current Directions in Psychological Science, 19*, 3–8. doi:10.1177/0963721409359300
- Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*, 3–48. doi:10.1017/S0021963001006552
- Triesch, J., Jasso, H., & Deák, G. O. (2007). Emergence of mirror neurons in a model of gaze following. *Adaptive Behavior, 15*, 149–165. doi:10.1177/1059712307078654
- Ulloa, J. L., Puce, A., Hugueville, L., & George, N. (2014). Sustained neural activity to gaze and emotion perception in dynamic social scenes. *Social Cognitive and Affective Neuroscience, 9*, 350–357. doi:10.1093/scan/nss141
- Van der Goot, M. H., Tomasello, M., & Liszkowski, U. (2014). Differences in the nonverbal requests of great apes and human infants. *Child Development, 85*, 444–455.
- Van Velzen, J., & Eimer, M. (2003). Early posterior ERP components do not reflect the control of attentional shifts toward expected peripheral events. *Psychophysiology, 40*, 827–831. doi:10.1111/1469-8986.00083
- Watanabe, S., Kakigi, R., Miki, K., & Puce, A. (2006). Human MT/V5 activity on viewing eye gaze changes in others: A magnetoencephalographic study. *Brain Research, 1092*, 152–160. doi:10.1016/j.brainres.2006.03.091
- Whalen, C., & Schreibman, L. (2003). Joint attention training for children with autism using behavior modification procedures. *Journal of Child Psychology and Psychiatry, 44*, 456–468.
- Williams, L. M., Palmer, D., Liddell, B. J., Song, L., & Gordon, E. (2006). The 'when' and 'where' of perceiving signals of threat versus non-threat. *NeuroImage, 31*, 458–467. doi:10.1016/j.neuroimage.2005.12.009
- Wolf, R. C., Philippi, C. L., Motzkin, J. C., Baskaya, M. K., & Koenigs, M. (2014). Ventromedial prefrontal cortex mediates visual attention during facial emotion recognition. *Brain, 137*, 1772–1780.
- Wolf, W., Launay, J., & Dunbar, R. I. M. (2016). Joint attention, shared goals, and social bonding. *British Journal of Psychology, 107*, 322–337. doi:10.1111/bjop.12144
- Woodman, G. F., Arita, J. T., & Luck, S. J. (2009). A cuing study of the N2pc component: An index of attentional deployment to objects rather than spatial locations. *Brain Research, 1297*, 101–111. doi:10.1016/j.brainres.2009.08.011