Visual cognition during real social interaction

Paul A. Skarratt[1]*, Geoff G. Cole[2] and Gustav Kuhn[3]

[1] Department of Psychology, University of Hull, Hull, UK
[2] Centre for Brain Science, University of Essex, Essex, UK
[3] Department of Psychology, Brunel University, Uxbridge, UK

Laboratory studies of social visual cognition often simulate the critical aspects of joint attention by having participants interact with a computer-generated avatar. Recently, there has been a movement toward examining these processes during authentic social interaction. In this review, we will focus on attention to faces, attentional misdirection, and a phenomenon we have termed *social inhibition of return (Social IOR)*, that have revealed aspects of social cognition that were hitherto unknown. We attribute these discoveries to the use of paradigms that allow for more realistic social interactions to take place. We also point to an area that has begun to attract a considerable amount of interest—that of Theory of Mind (ToM) and automatic perspective taking—and suggest that this too might benefit from adopting a similar approach.

Keywords: vision, attention, social cognition, misdirection, eye gaze, social inhibition of return

SOCIAL ATTENTION: THE NEED FOR REAL SOCIAL INTERACTION

Ever since its inception during the late 1960s (e.g., Neisser, 1967), research into human visual attention has moved toward examining the behavior of individuals as they perform tasks alone. In the standard experiment a single observer is seated in front of a visual display and performs a required task. Clearly, this laboratory-based paradigm has been instrumental in uncovering many of the fundamental properties of visual cognition (e.g., Eriksen and Eriksen, 1974; Posner, 1980; Duncan, 1984; Tipper, 1985; Raymond et al., 1992; Watson and Humphreys, 1997; Simons and Rensink, 2005). However, humans are social animals and the majority of people spend some part of each day interacting with others. As this issue of *Frontiers* demonstrates, a growing number of visual cognition studies are beginning to reflect this and examine how attention is deployed when a person interacts with another individual or individuals. In this article we first present our own assessment of why the new field of social neuroscience can be considered as more than an attempt to improve the ecological validity of our experiments. We then go on to show how the social neuroscience method not only informs us about the mental processes involved in social interaction but has also revealed the existence of visual mechanisms that were previously unknown. Indeed, we cite particular cases where the method has revealed effects previously thought not to occur. Examples are drawn from our own research examining attention to faces, attentional misdirection, and a phenomenon we have previously labeled *social inhibition of return (SIOR)*. Finally, we point to one recent debate that could likely benefit from this new approach. That is the question of whether Theory of Mind (ToM) and perspective-taking are automatic processes.

The study of social attention is often considered to have begun in the late 1990s with the first report of the *gaze cueing effect* (Friesen and Kingstone, 1998) in which a person's

attention is oriented on the basis of another person's direction of gaze. However, it is more accurate to say that developmental psychologists have been studying these types of phenomena for decades and, as with the developing field of social neuroscience, their methods involved measuring behavior during interaction between *real people*. An early example (Scaife and Bruner, 1975) involved young infants sitting with a caregiver and looking directly ahead toward an experimenter. The experimenter would then turn their head 90° to the left or right and fixate an object. The infant's propensity for gaze following would then be recorded. This line of work was subsequently placed within the context of ToM in which infants were assessed for their ability to understand others as intentional agents (Tomasello et al., 1993).

The lead that developmental psychologists have taken in studying infant cognition in the real world is beginning to find favor amongst those advocating a *cognitive ethological* approach to adult cognition (Kingstone et al., 2008). Ethology, the study of animal behavior, was developed by a number of naturalists in Europe during the 1930s. Its central position was that animal "routines" and "patterns" should be examined in as natural environment as possible. Ethology's basic philosophy explicitly contrasted with that of the American-led Behaviorists during the same period. Their models of behavior were derived from laboratory studies of animals, usually rats and pigeons. Although many influential models of behavior were developed from the behaviorist approach, the field was often criticized for its lack of ecological validity. In the same way, cognitive ethologists emphasize the importance of ecological factors in human cognition and consider social interaction as being central. Although social neuroscience does not advocate a naturalistic setting per se, the field does employ paradigms that take into consideration the social situations in which cognition occurs.

A number of recent attention researchers have therefore examined adult visual attention in scenarios where participants

perform tasks in conjunction with other individuals. For instance, Brennan et al. (2008) made the point that many everyday situations involve joint visual search, such as when an adult and child look through a picture book together. With this in mind, Brennan et al. examined whether joint visual search might be more efficient than that of a solitary observer. Pairs of participants were asked to search arrays for a target letter appearing amongst distractors, with cursors allowing each participant to see the location of their partner's gaze at any point during the search. Results showed that searches were almost twice as efficient when made jointly than alone. Furthermore, Brennan et al. showed that observers coordinated their search with their partner's without explicit training. The authors concluded that joint gaze can be used spontaneously to minimize collective effort and optimize search success.

The acknowledgment that psychological models can benefit from ecological consideration can be seen in other areas of vision research, as in the case of vision and action. The study of visual perception has been dominated by the view that the function of vision is to generate a representation of the external environment. That is, to provide a percept. However, a number of authors have pointed out that vision is most often accompanied by action (e.g., Milner and Goodale, 1995; Prinz, 1997; Jeannerod, 1999; Hommel, 2009) and paradigms have therefore been developed with this in mind. For instance, an abundance of work has shown that the appearance of a new object is particularly effective in attracting attention (e.g., Cole et al., 2004; Cole and Liversedge, 2006; Cole et al., 2007; Davoli et al., 2007; Cole and Kuhn, 2009, 2010; Yantis and Jonides, 1984). However, Welsh and Pratt (2006) demonstrated that the propensity with which new onsets capture attention is influenced by the type of action an individual makes when responding to them. More specifically, the authors showed that task-irrelevant offsets interfere with new object capture when a standard keyboard press is required but do not when a reaching response is made. Thus the act of reaching toward an object enables attention to be focused more effectively. As is the case with social neuroscience, proponents of the vision-for-action perspective argue that a phenomenon in question may be better elucidated when consideration is given to its functional significance in the real world.

Social neuroscience is clearly grounded in the notion that humans are social animals and this ought to be reflected in our experimental paradigms. In the field of visual attention the social aspect of this process has typically been implemented by presenting participants with social stimuli in the form of static images or video clips of people, and then measuring their effects on visual attention. The use of these often well controlled, yet rather reductionist depictions of real world stimuli offer a valuable tool to investigate social attention in the laboratory. This is particularly the case in the field of neuroscience, where experimental protocols are often limited by the logistical constraints of the apparatus (e.g., MRI, EEG; but for recent developments see Guionnet et al., 2012; Guionnet et al., and Schippers et al., 2010). Progress in overcoming these constraints has allowed face-to-face interaction between people by channeling a live video feed inside a scanner. In one such experiment, Redcay et al. (2010) recorded functional MRI data as participants interacted with an experimenter via a video screen in one of three cooperative scenarios: a "live" interaction, a recording of one of their earlier interactions, or a recording of the experimenter's interaction with a different participant. Hence, all three conditions contained identical visual information but differed according to the mental state imputed to the experimenter's actions. Results showed that the live feed elicited greater activation in the ventral striatum, amygdala, and anterior cingulate cortex (ACC), areas activated in studies of social reward (e.g., Walter et al., 2005), and the right posterior superior temporal sulcus (rpSTS), a region implicated in social perception and social cognition (Allison et al., 2000; Saxe, 2006). That these brain regions are differentially activated on the basis of the *authenticity* of face-to-face interactions leads one to enquire as to their importance. Are such areas central or peripheral to the processes we as experimenters attempt to measure in our studies of social visual cognition? How does their involvement impact upon processes occurring elsewhere in the brain? Do they scale up to produce measurable differences in behavior? And the question most pertinent to the current review: Are the processes that mediate perception of these social stimuli the same as those involved in perceiving a real person? Hence, it is not merely a matter of improving the ecological validity of our experiments; it is about the extent to which the findings from social attention studies translate to real person interaction. If these processes differ, it is likely to have serious implications for our understanding of social attention, and social neuroscience in general. In the next section we compare examples of what we have learned about social attention from classical behavioral studies and from those in which real people interact. We find that the former can yield very different results from the latter.

ATTENTIONAL ORIENTING TOWARD OTHER PEOPLE

Rather than processing all of the available sensory input, the visual system selects only that which is likely to be behaviorally important. Metaphors such as the attentional spotlight (Posner et al., 1980; Broadbent, 1982) or zoom-lens (Eriksen and St. James, 1986) describe the way in which attention is oriented around the field of vision, selecting for further processing any objects or locations falling within the "illuminated" boundary. Attentional orienting can be overt (i.e., where people look), or covert (i.e., where people attend without moving their eyes or head). Given the importance that attention plays in mediating what we see, it is not surprising that there has been much interest in determining how we select the information deemed to be "important" (Henderson, 2003). Some have argued that this selection process is largely driven by bottom-up stimulus features (e.g., Itti and Koch, 2000). According to these models, certain stimulus features, such as luminance contrasts, are particularly salient and thus automatically capture attention. Detailed computational models are remarkably accurate in predicting people's eye movements as they view images of natural scenes (Itti and Koch, 2001). However, due to the complexity of the processes they attempt to simulate, these models necessarily simplify humans as passive observers of the world (Findlay and Gilchrist, 2003). In reality, however, vision is an active process that enables us to carry out multifarious tasks in which required objects might not be the most salient aspect of the visual scene (Land, 2006).

Consequently, others have argued that eye movements are driven by our top-down goals rather than by salient aspects of the visual scene (Land et al., 1999; Hayhoe and Ballard, 2005; Land, 2006). This reinforces the view outlined in the previous section that vision and action are intricately linked.

In addition to task goals and bottom-up salience, it has become clear that the attention system is strongly influenced by social factors. Some of the earliest eye tracking studies by Yarbus (1967) have shown that whilst viewing images of social situations our eyes are particularly attracted by the people in the scene. More recent studies have replicated this observation and shown further that attention is strongly drawn toward faces, and in particular the eyes (Yarbus, 1967; Kuhn and Land, 2006; Birmingham et al., 2008b, 2009; Fletcher-Watson et al., 2008; Kuhn et al., 2009). Indeed, such is the appeal of eyes that observers still tend to look at them even when faces are presented in isolation (Walker-Smith et al., 1977; Pelphrey et al., 2002; Itier et al., 2007).

However, to what extent do these findings generalize to more complex social interactions? Most studies of face perception involve faces presented in isolation and are, as a consequence, already attended (Walker-Smith et al., 1977; Pelphrey et al., 2002). Hence these studies may demonstrate a preference for the eyes because these are often the most complex or salient component of a pre-selected face. The true measure of eye-preference is whether they are able to summon attention when faces are embedded within a complex scene. Evidence has shown that this is indeed the case (Birmingham et al., 2008a,b, 2009). Although those studies used static scenes, others have investigated where observers look in dynamic scenes. For instance, Kuhn et al. (2009) had participants observe a magician performing a magic trick. In spite of the trick involving the magician's hands, the proportion of fixations on his head and eyes was close to 70%. Likewise, when participants were asked to watch videos of other students engaging in conversation, 77% of fixations were directed to the people in the clips (Foulsham et al., 2010).

Much empirical work has therefore demonstrated that as humans, we generally prioritize other humans, their faces and, in particular, their eyes when viewing natural scenes. Whilst these studies vary in terms of their ecological validity, there remain questions as to whether these studies capture the true nature of social cognition. Indeed, social cognition involves more than passively observing images of people; it involves *interaction* with real people. Interestingly, there is evidence that even the *potential* for real social interaction can influence behavior. For example, people will often meet the gaze of an approaching stranger that is depicted in an image (Henderson et al., 2005; Itier et al., 2007) but will avoid direct eye contact when the same event occurs in real life (Ellsworth et al., 1972). Hence the presence of a real person clearly elicits a different behavioral response. The main difference between these settings is that the latter involves the potential for social interaction whilst the former does not. A recent study by Laidlaw et al. (2011) directly examined the effect of those two scenarios. In the study, participants' eye movements were monitored as they sat in a waiting room. The crucial manipulation in the experiment was whether participants were joined by a real confederate posing as another research participant or the confederate appearing on a video screen in the waiting room.

Results showed that whereas participants frequently looked at the confederate on the video screen, they rarely did so in person. Moreover, ratings of the participants' social skills correlated positively with the amount of time spent looking at the live confederate, yet did not in the video condition. Similar conclusions concerning the difference between real and artificial social interactions have been drawn from studies examining eye movements in response to social cues in autism (e.g., Nation and Penny, 2008).

In sum, it is clear that attempts to measure aspects of social cognition can yield different results depending on whether the social context is real or merely simulated. The majority of research demonstrates that our willingness to look at others is strongly influenced by whether or not they are physically present. Although traditional, controlled, computer-based tasks are important in examining some of the mechanisms involved in social attention, its underlying mechanisms may only be fully understood in more naturalistic settings that take into account how we interact with other people.

ATTENTIONAL ORIENTING AWAY FROM OTHER PEOPLE: GAZE FOLLOWING

Eyes are not only highly effective in attracting attention, but also in orienting attention to other parts of the visual field indicated by their gaze direction. This orienting response to where other people look has been termed *gaze following* or *gaze cueing* and has been studied extensively since the late 1990s (Friesen and Kingstone, 1998; Driver et al., 1999). In these experiments, participants are typically presented with a face in the centre of a display with its eyes and/or head directed to the left or right. A target is then presented at either the gazed-at location or the opposite hemifield. The characteristic results are that response times are reduced for targets appearing in the gaze-indicated position, a facilitatory effect arising from the gaze cue having automatically shifted the observers' attention. In the years since this discovery, many variations of this paradigm have been developed to determine the parameters of gaze cueing and its underlying neural bases (Williams et al., 2005; Frischen et al., 2007; Materna et al., 2008).

Even though gaze cues are intended to represent real human faces, it has been argued that the paradigm may not necessarily capture the true nature of social attention (Kingstone et al., 2003, 2008; Kingstone, 2009). Although some researchers have tried to address this concern by improving the realism of the images used in their experiments (e.g., Hermens and Walker, 2010), the very nature of simulating social interaction via a computer monitor is questionable (Kingstone, 2009). Hence other researchers have begun to examine gaze following in more naturalistic settings whereby target locations are cued by real people. For instance, Gallup et al. (2012a) used a hidden video camera to record the number of glances received by an attractive stimulus as pedestrians walked by. The critical measure concerned whether a pedestrian's gaze would increase the likelihood of other passers-by glancing toward the stimulus. This was indeed the case. Moreover, this likelihood was greater for those who walked behind the pedestrian than for those who approached from the front. This finding is consistent with the notion of

gaze avoidance by approaching strangers (Ellsworth et al., 1972), and demonstrates that the effectiveness of a visual cue in directing attention can be modulated by the social context (see also Gallup et al., 2012b, for gaze following in large crowds of people). Kuhn and colleagues have adopted a similar ethological approach to examining visual cognition by recording the eye movements of observers as they watch magic tricks (for reviews see Kuhn et al., 2008; Macknik et al., 2008; Kuhn and Martinez, 2012). Magicians are highly skilled in directing—and misdirecting—the attention of observers. Social cues play a crucial role in misdirection, and numerous studies have now demonstrated that gaze cues are instrumental in successfully achieving this (Kuhn and Land, 2006; Tatler and Kuhn, 2007; Kuhn et al., 2009; but see Cui et al., 2011). For instance, Kuhn et al. (2009) found that the magicians' gaze influenced where people looked, and consequently the likelihood of successful detection (see also Tatler and Kuhn, 2007). The advantage of this over the standard gaze cueing paradigm is not only are the cues generated in a more naturalistic way, but that they also compete against other salient features in the visual scene as well as the participant's intention to discover the trick. These paradigms therefore offer a significant step toward investigating attention in a more realistic social context. Importantly however, attempts have been made to improve the ecological validity still further by comparing the likelihood of trick detection when observed on a video or in a live setting (Kuhn and Tatler, 2005; Tatler and Kuhn, 2007). These have indicated that misdirection experienced during a face-to-face interaction is more effective, suggesting that social cues are stronger when presented by a real person. Moreover, the instructions concerning what participants would expect to see in the face-to-face scenario did not influence their eye movement behavior, nor did it improve their detection of the trick (Kuhn and Tatler, 2005). However, when viewed on a computer monitor, prior instructions influenced both detection as well as eye movement behavior (Kuhn et al., 2008).

In sum, whilst the gaze cueing paradigm has been immensely valuable in investigating different aspects of social attention, there is a clear difference in the way attention operates in the presence of real people as opposed to simulated people. We now turn to further evidence of this in the next section.

SOCIAL INHIBITION OF RETURN (SIOR)

Another example of how the presence of others influences attention is the way *inhibition of return* (IOR) is expressed during individual and joint visual search tasks. Indeed, the differences are such that Skarratt et al. (2010) have proposed that IOR and its social counterpart *social IOR* may even be independent processes rather than facets of the same processes as revealed in social and solitary search contexts.

IOR refers to the slowing of responses to targets appearing in previously attended locations (Kingstone and Pratt, 1999; Taylor and Klein, 2000; Godijn and Theeuwes, 2002). It has been proposed as having evolved as a means of expediting visual search (Klein and MacInnes, 1999). To this end, inhibitory mechanisms serve to bias attention from returning to previously inspected locations, and to discourage successive eye-movements being programmed to the same spatial location (Rafal et al., 1989). One can imagine the utility of such mechanisms during a search for

a friend in a crowd of people. The search is less likely to yield a successful outcome if attentional and saccadic resources are repeatedly realigned with spatial locations that have recently been searched. However, as social animals, we are likely to have carried out many of our predatory and defensive search behaviors in conjunction with other individuals. This raises the interesting question of whether one might inhibit a spatial location knowing that another person has previously searched there. This very rationale motivated Welsh and colleagues (Welsh et al., 2005) to examine whether IOR can be socially "transferred" between different individuals. This was investigated by having pairs of participants sit across a table from one another. Each took turns at reaching out to one of two targets as they appeared on the workspace. The basic social IOR phenomenon is the observation that participants are slower to initiate a reaching action to a location previously responded to by a partner. That is to say, one inhibits a location on the basis that another person has searched there. As such, it can be said that IOR can indeed be "transferred" between two individuals. This effect is clearly a visual phenomenon based on real social interaction. However, not only has this paradigm revealed information concerning such interaction but, as the following two sections show, the procedure also reveals aspects of visual attention previously unknown.

SOCIAL IOR AND NEW INSIGHTS INTO HUMAN COGNITION
THE ROLE OF VISIBLE TRANSIENTS

Skarratt et al. (2010) sought to investigate the extent to which social IOR is generated on the basis of social information rather than the visual information carried by another person's responses. If it is the former, then it ought to occur when participants simply *know* where their partner has responded without having seen it take place. To address this possibility, the view each participant had of their partner was restricted to a central portion measuring 12° across. All peripheral information was occluded, meaning participants could not see their partner's targets, response buttons and, consequently, the completion of their responses. In other words, all the visual information that could give rise to IOR at the response location was eliminated. White noise also masked the sound of the response buttons being pressed, thus precluding the likelihood of IOR occurring due to auditory stimulation (cf. Spence and Driver, 1998). This meant that participants could infer a response location only from their partner's eye gaze (signaling their intention to respond), or their initial hand movement toward the target. Results showed that social IOR emerged even under these restricted viewing conditions. Moreover, it was the same magnitude as the corresponding IOR effect observed under free-viewing conditions, to which all the sensory information had contributed. Thus, simply *knowing* where a person had responded was as effective as seeing the complete reaching response.

The implication of these findings is that a visual cognition effect, mediated in this case by inhibitory processes, is initiated by *inferred events* occurring in the external world. Importantly, this contrasts markedly with what was previously assumed about IOR from classical precueing studies in which participants perform alone. For instance, Cole et al. (2011a) showed that even when observers are aware that an occluded visual event has taken

place in a spatial location, they do not inhibit it. The experiment involved participants having to detect a target appearing in a cued or uncued location. On some trials, a luminance cue indicated the possible target location. On others, a pattern mask briefly occluded the cue onset, but the indicated target location was revealed after the mask was removed. Thus, in both cases a precue indicated the potential target location, but only in one case did the participant see the cue generated. Results showed that IOR emerged only when the cue transient was visible, indicating that localized sensory input is required for inhibition to occur. Indeed, these findings concur with a great deal of evidence suggesting that local transients are necessary for attention to be marshaled at all. Using a similar method to occlude lower level visual transients, Franconeri et al. (2005) examined whether new perceptual objects can capture attention without an abrupt visual onset. To this end, they used a standard irrelevant singleton paradigm (Egeth and Yantis, 1997) comparing search slopes yielded by new versus already present targets. An annulus shape was presented around the perimeter of the array, which then contracted during the course of each trial. As the shape contracted, the new object appeared behind and was then revealed in a location previously seen to be unoccupied. As with Cole et al. (2011a), participants were aware that a scene change had taken place, but had not seen the accompanying transient that signaled its arrival. Results showed that new objects failed to capture attention under these conditions, yet they did attract attention when the annulus was seen to move *behind* the search items thus rendering a visible onset (but see Cole et al., 2011a, Experiment 6). In a similar vein, Skarratt et al. (submitted) have shown that attention is captured by objects that loom towards or recede away from the observer (see also Skarratt et al., 2009). These objects began their motion paths in far and near depth planes, respectively, before moving into alignment with objects remaining static throughout the trial. However, when the motion sequence was replaced with a blank frame, thus removing the transients associated with the objects' movement path, these objects were no longer capable of attracting attention. In the case of social IOR, however, spatial locations undergo inhibitory tagging on the basis of knowing rather than seeing that a stimulus event has taken place.

DO CENTRAL CUES ELICIT IOR?

That social IOR occurs under restricted viewing conditions indicates that peripheral locations can be inhibited on the basis of centrally presented information. This contrasts with IOR, whose emergence in a peripheral location requires a localized peripheral cue, and which is not reliably observed when a peripheral location is indicated by a central arrow cue. Like peripheral cues, central arrows can facilitate processing when they precede a target by a short interval (Ristic et al., 2002; Tipples, 2002) yet they do not give rise to later IOR (Posner and Cohen, 1984; Abrams and Dobkin, 1994). Hence there is a clear difference in the effects of central cueing in a conventional precueing paradigm and in a social IOR paradigm. Indeed, this discrepancy becomes even more apparent when examining the effects of gaze cues. As we described earlier, attention can be oriented by the gaze direction of centrally presented faces. They too result in prolonged facilitation that rarely gives way to subsequent IOR (McKee et al.,

2007; Greene et al., 2009). As far as we are aware, only two studies have shown IOR in response to gaze cues. According to Frischen et al. (2007; see also Frischen and Tipper, 2004), gaze cues do give rise to IOR but at much later cue-target intervals than can be observed with peripheral cues (around 2400 ms rather than 300 ms), and only when attention is disengaged from the gazed-at location prior to target presentation. The highly specific circumstances in which gaze cues elicit IOR are in contrast with those in which social IOR occurs. For instance, Skarratt et al. (2010; Experiment 3) ensured that each participant saw only their partner's face as they performed the alternating response task. This provided a very close approximation of the classic gaze-cueing method in that the partner looked toward their response location after which the participant's own target appeared in the same or opposite location. The results showed reliable social IOR occurring much earlier (i.e., between 1300–1700 ms) than the IOR effect found by Frischen et al., and without a controlled attempt to remove attention from the gazed-at location. These findings suggest that the mechanisms underlying attention and inhibition respond differently to real and simulated biological behavior.

REAL VERSUS ANIMATED BIOLOGICAL BEHAVIOR

This point is further underscored by our attempts to induce social IOR using a realistic animation of a person's response behavior. In this experiment (Skarratt et al., 2010, Experiment 1), individual participants performed the alternating response task in conjunction with an animated partner. This was achieved by projecting a movie of a male partner onto a screen such that he appeared to be seated opposite. In keeping with our other experiments, we disambiguated the social and visual information conveyed by the partner's response by manipulating the participants' view of him. Results showed the inhibitory effect occurred only when participants had an unrestricted view of the animated partner's responses, indicating that IOR was elicited by the associated lower level visual transients. The absence of social IOR in the restricted viewing condition suggests that the visuomotor system is less sensitive to simulated biological behavior than it is to the same behavior performed by a real person. Hence the critical factor in the generation of social IOR is that the observed behavior must demonstrate *agency*. This view is supported by several recent findings revealing different neural substrates for the perception of real and virtually real biological behavior. For instance, Gobbini et al. (2011) compared the BOLD responses of participants observing either human or robot faces performing basic emotional expressions. Perhaps unsurprisingly, both face types activated face-specialized regions such as the fusiform gyrus and the superior temporal sulcus. More interesting, however, was that human faces evoked stronger activations in the medial prefrontal and the anterior temporal cortices, and the right amygdala. The latter system has long been associated with emotion (Breiter et al., 1996; Morris et al., 1996), suggesting participants were less sensitive to automated displays of emotion, whilst the former regions are thought to be involved in the representation of others' mental states and ToM (Leibenluft et al., 2004; Mitchell et al., 2005; Amodio and Frith, 2006; Frith and Frith, 2006). These findings can be interpreted as observers empathizing more with sentient than with automated beings. This claim

is corroborated with the observation that robotic faces elicited stronger activation in three gyri associated with the perception of inanimate objects and automated motion—the medial fusiform, the lingual, and mid temporal gyri (see Beauchamp et al., 2002). In a similar study by Perani et al. (2001), positron emission tomography (PET) was used to record the neural responses of participants whilst they observed scenes of a hand grasping various geometrical objects. Responses to scenes involving a real hand were compared to those evoked when the same scenes were rendered in 3D virtual reality or 2D movie clips. Results showed that observation of the real and virtual hands was associated with greater activation in the inferior temporal cortices and the right inferior parietal cortex. These are regions that have been implicated, respectively, in the perceptual representation of actions and motor planning (e.g., Decety, 1996; Decety et al., 1997), and these stronger activations may reflect greater sensitivity to more realistic depictions of behavior. Finally, when the same brain regions are activated by live-action and computer-animated behavior, overall activation is stronger for live-action images (Mar et al., 2007).

FUTURE DIRECTIONS

The findings described above demonstrate that classical attention paradigms can not only underestimate effects but may also fail to reveal aspects of human cognition. Throughout this review we have pointed out cases in which important theoretical advances in visual cognition have been made by implementing social interaction into experimental manipulations. In the remainder of this article, we focus on one such phenomenon that until recently has been the exclusive domain of developmental researchers, but which is now enjoying increased interest within our own discipline. We propose that a recent development concerning the ToM phenomenon is perfectly suited to experiments employing real social interaction.

Theory of Mind is the ability to impute mental states to oneself and to others. Although the concept was originally employed in the context of animal cognition (Premack and Woodruff, 1978), a number of developmental psychologists applied the idea to human infants (e.g., Wimmer and Perner, 1983). Indeed, ToM has now been applied within various contexts including, for instance, schizophrenia (Harrington et al., 2005), autism (Baron-Cohen, 2000), Alzheimer's disease (Gregory et al., 2002), decision-making (Torralva et al., 2007), and evolutionary psychology (Povinelli and Preuss, 1995). Real person interaction studies have been employed in work on ToM. For instance, Stuss et al. (2001) examined the ability of frontal lobe patients to infer the visual experience of others; that is, to perspective take. Rather than depict an individual on a computer monitor, the patients were asked to consider the perspective of a real person who was seated opposite. Using both depicted and real person interaction, Rilling et al. (2004) examined whether economical decision making in conjunction with another individual is subserved by cortical areas known to be involved in ToM (e.g., anterior paracingulate cortex and posterior superior temporal sulcus). Not only did Rilling et al. find this to be the case but the activation observed was greater when the decision maker was interacting with a real person.

Although adult humans are adept at considering others' mental states when required to do so, a number of authors have recently argued that ToM attributions occur *automatically*. That is, without conscious effort. Evidence for automatic ToM has come from a number of different paradigms including gaze cueing (Nuku and Bekkering, 2008; Teufel et al., 2009, 2010a,b). Gazing agents have been employed in the context of ToM because when a person looks to a location, a mental state such as intention can be assumed to be occurring. As Calder et al. (2002) point out, gaze implies that the person may have some intention or goal toward a fixated object. Similarly, Nuku and Bekkering (2008) argued that gaze cueing occurs because the observer infers that the agent is physically able to attend to the target. They based their conclusion on results from a gaze cueing procedure in which the targets would or would not be visible to the agent from his vantage point. The authors found larger cueing effects when the targets were visible to the agent. This clearly suggests that inferring the agent's mental state (i.e., seeing versus not seeing) influenced the degree to which the agent shifted the observer's attention.

Langton (2009), however, has urged caution in concluding that mental state attribution modulates gaze cueing. For instance, objects that have no mental state (e.g., a glove) but which incorporate a pair of eyes are effective in shifting attention toward the "looked-at" direction (e.g., Quadflieg et al., 2004). Moreover, Cole et al. (2011b) found that gaze cueing was not influenced according to whether the inducing agent had their view of a peripheral target blocked or not. These findings suggest that gaze cueing is largely controlled by bottom-up mechanisms with little contribution from higher processes that are responsible for mental state attribution.

Apperly et al. (2006) have also examined whether ToM can occur automatically. Adult participants were shown a video sequence in which an agent marked a box that she knew contained an object. After she was seen to leave the room, a second agent then secretly placed the marker on a different box, meaning that when she returned, the first agent would hold a false belief about which box contained the object. Participants were then given true/false statements assessing their own perspective (e.g., "the object is in the left box") or occasionally that of the female agent ("she thinks the object is in the left box"). Apperly et al. reasoned that if participants automatically infer and encode another's perspective then judgments about the agent's beliefs should be made as quickly as are judgments about their own. Results showed, however, that participants were relatively slow to indicate the agent's belief when unexpectedly asked to do so. This therefore challenges the notion that ToM can occur automatically. By contrast, German et al. (2004) have provided support for the automatic ToM hypothesis using neuroimaging. They found that brain areas known to be concerned with inferring another person's intentions (medial prefrontal, inferior frontal, and temporoparietal cortex; Frith and Frith, 2006) are also recruited when participants view videos of social situations but are not required to make judgments about mental states; a phenomenon the authors refer to as *automatic engagement of the intentional stance*. Given that the issues relating to the automaticity of ToM are, by definition, concerned

with the attribution of mental states to real people, we suggest that the real-person interaction paradigm we have advocated in this review would appear particularly suited to its investigation. Furthermore, the present review shows how much more sensitive the paradigm can be to cognitive phenomena.

CONCLUSIONS

In this review we have emphasized how real-person social interaction research can yield very different results when compared with paradigms in which the social context is merely depicted. Indeed, new information concerning visual cognition is being derived from the method. It is clear from the work described in this review that a fresh insight into human cognitive abilities can be gained from experiments that allow for more realistic social interaction.

The development of such paradigms is particularly timely given the burgeoning interest in issues such as ToM and automatic perspective taking. Indeed, we suggest that if any debate within cognition could benefit from real-person interaction paradigms it is this. In the same way as processes underlying attentional orienting and IOR can be elucidated during social interaction with other people, one might hypothesize that those underlying perception of others' thoughts, intentions, goals and actions might also be better understood. The adoption of such an approach can only increase our understanding of these fascinating processes.

REFERENCES

Abrams, R. A., and Dobkin, R. S. (1994). Inhibition of return: effects of attentional cuing on eye movement latencies. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 467–477.

Allison, T., Puce, A., and McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* 4, 267–278.

Amodio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.

Apperly, I. A., Riggs, K. J., Simpson, A., Samson, D., and Chiavarino, C. (2006). Is belief reasoning automatic? *Psychol. Sci.* 17, 841–844.

Baron-Cohen, S. (2000). Theory of mind and autism: a review. *Int. Rev. Res. Ment. Retard.* 23, 169–184.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., and Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron* 34, 149–159.

Birmingham, E., Bischof, W. F., and Kingstone, A. (2008a). Gaze selection in complex social scenes. *Vis. Cogn.* 16, 341–356.

Birmingham, E., Bischof, W. F., and Kingstone, A. (2008b). Social attention and real world scenes: the roles of action, competition, and social content. *Q. J. Exp. Psychol. (Hove)* 61, 986–999.

Birmingham, E., Bischof, W. F., and Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Res.* 49, 2992–3000.

Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Buckner, P. L., Strauss, M. M., Hyman, S. E., and Rosen, B. R. (1996). Response and habituation of the human amygdala during visual processing of facial expressions. *Neuron* 17, 875–887.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., and Zelinsky, G. J. (2008). Coordinating cognition: the costs and benefits of shared gaze during collaborative search. *Cognition* 106, 1465–1477.

Broadbent, D. E. (1982). Task combination and selective intake of information. *Acta Psychol. (Amst.)* 50, 253–290.

Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I., and Young, A. W. (2002). Reading the mind from eye gaze. *Neuropsychologia* 40, 1129–1138.

Cole, G. G., Kentridge, R. W., and Heywood, C. A. (2004). Visual salience in the change detection paradigm: the special role of object onset. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 464–477.

Cole, G. G., Kuhn, G., and Skarratt, P. A. (2011a). Non-transient luminance changes do not capture attention. *Atten. Percept. Psychophys.* 73, 1407–1421.

Cole, G. G., Smith, D., and Billing, R.-C. (2011b). "Social attention effects do not depend on mental-state attribution," in *Paper presented at the Psychonomic Society meeting*, (Seattle, WA).

Cole, G. G., Kuhn, G., and Liversedge, S. P. (2007). Onset of illusory figures attenuates change blindness. *Psychon. Bull. Rev.* 14, 939–942.

Cole, G. G., and Kuhn, G. (2009). Appearance matters: attentional orienting by new objects in the precuing paradigm. *Vis. Cogn.* 17, 755–776.

Cole, G. G., and Kuhn, G. (2010). What the experimenter's prime tells the observer's brain. *Atten. Percept. Psychophys.* 72, 1367–1376.

Cole, G. G., and Liversedge, S. L. (2006). Change blindness and the primacy of object appearance. *Psychon. Bull. Rev.* 13, 588–593.

Cui, J., Otero-Millan, J., Macknik, S. L., King, M., and Martinez-Conde, S. (2011). Social misdirection fails to enhance a magic illusion. *Front. Hum. Neurosci.* 5:103. doi: 10.3389/fnhum.2011.00103

Davoli, C., Suszko, J., and Abrams, R. (2007). New objects can capture attention without a unique luminance transient. *Psychon. Bull. Rev.* 14, 338–343.

Decety, J. (1996). Do imagined and executed actions share the same neural substrate? *Brain Res. Cogn. Brain Res.* 3, 87–93.

Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., and Fazio, F. (1997). Brain activity during observation of actions. influence of action content and subject's strategy. *Brain* 120, 1763–1777.

Driver, J., Davies, M., Kidd, P., Maxwell, M., Ricciardelli, P., and Baron-cohen, S. (1999). Gaze perception triggers reflective visiospatial orienting. *Vis. Cogn.* 6, 509–540.

Duncan, J. (1984). Selective attention and the organization of visual information. *J. Exp. Psychol. Gen.* 113, 501–517.

Egeth, H. E., and Yantis, S. (1997). Visual attention: control, representation, and time course. *Annu. Rev. Psychol.* 48, 269–297.

Ellsworth, P. C., Carlsmith, J. M., and Henson, A. (1972). The stare as a stimulus to flight in human subjects: a series of field experiments. *J. Pers. Soc. Psychol.* 21, 302–311.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a non-search task. *Percept. Psychophys.* 16, 143–149.

Eriksen, C. W., and St. James, J. D. (1986). Visual-attention within and around the field of focal attention - a zoom lens model. *Percept. Psychophys.* 40, 225–240.

Findlay, J. M., and Gilchrist, I. D. (2003). *Active Vision*. Oxford, UK: Oxford University Press.

Fletcher-Watson, S., Findlay, J. M., Leekam, S. R., and Benson, V. (2008). Rapid detection of person information in a naturalistic scene. *Perception* 37, 571–583.

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

Franconeri, S. L., Hollingworth, A., and Simons, D. J. (2005). Do new objects capture attention? *Psychol. Sci.* 16, 275–282.

Friesen, C. K., and Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychon. Bull. Rev.* 5, 490–495.

Frischen, A., Bayliss, A. P., and Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychol. Bull.* 133, 694–724.

Frischen, A., Smilek, D., Eastwood, J. D., and Tipper, S. P. (2007). Inhibition of return in response to gaze cues: the roles of time course and fixation cue. *Vis. Cogn.* 15, 881–895.

Frischen, A., and Tipper, S. P. (2004). Orienting attention via observed gaze shift evokes longer term inhibitory effects: implications for social interactions, attention, and memory. *J. Exp. Psychol. Gen.* 133, 516–533.

Frith, C. D., and Frith, U. (2006). How we predict what other people are going to do. *Brain Res.* 1079, 36–46.

Gallup, A. C., Chong, A., and Couzin, I. D. (2012a). The directional flow

of visual information transfer between pedestrians. *Biol. Lett.* doi: 10.1098/rsbl.2012.0160. [Epub ahead of print].

Gallup, A. C., Hale, J. J., Sumpter, D. J. T., Garnier, S., Kacelnik, A., Krebs, J. R., and Couzin, I. D. (2012b). Visual attention and the acquisition of information in human crowds. *Proc. Natl. Acad. Sci. U.S.A.* doi: 10.1073/pnas.1116141109. [Epub ahead of print].

German, T., Niehaus, J., Roarty, M., Giesbrecht, M., and Miller, B. (2004). Neural correlates of detecting pretense: automatic engagement of the intentional stance under covert conditions. *J. Cogn. Neurosci.* 16, 1805–1817.

Gobbini, M. I., Gentili, C., Ricciardi, E., Bellucci, C., Salvini, P., Laschi, C., Guazzelli, M., and Pietrini, P. (2011). Distinct neural systems involved in agency and animacy detection. *J. Cogn. Neurosci.* 23, 1911–1920.

Godijn, R., and Theeuwes, J. (2002). Oculomotor capture and inhibition of return: evidence for an oculomotor suppression account of IOR. *Psychol. Res.* 66, 234–246.

Greene, D. J., Mooshagian, E., Kaplan, J. T., Zaidel, E., and Iacoboni, M. (2009). The neural correlates of social attention: automatic orienting to social and nonsocial cues. *Psychol. Res.* 73, 499–511.

Gregory, C., Lough, S., Stone, V., Baron-Cohen, S., and Hodges, J. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. *Brain* 125, 752–764.

Guionnet, S., Nadel, J., Bertasi, E., Sperduti, M., Delaveau, P., and Fossati, P. (2012). Reciprocal imitation: toward a neural basis of social interaction. *Cereb. Cortex* 22, 971–978.

Harrington, L., Siegert, R. J., and McClure, J. (2005). Theory of mind in schizophrenia: a critical review. *Cogn. Neuropsychiatry* 10, 249–286.

Hayhoe, M., and Ballard, D. (2005). Eye movements in natural behavior. *Trends Cogn. Sci.* 9, 188–194.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends Cogn. Sci.* 7, 498–504.

Henderson, J. M., Williams, C. C., and Falk, R. J. (2005). Eye movements are functional during face learning. *Mem. Cognit.* 33, 98–106.

Hermens, F., and Walker, R. (2010). Gaze and arrow distractors influence saccade trajectories similarly.

Q. J. Exp. Psychol. (Hove) 63, 2120–2140.

Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychol. Res.* 73, 512–526.

Jeannerod, M. (1999). The 25th Bartlett Lecture: to act or not to act: perspectives on the representation of actions. *Q. J. Exp. Psychol. A* 52, 1–29.

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

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

Itti, L., and Koch, C. (2001). Computational modelling of visual attention. *Nat. Rev. Neurosci.* 2, 194–203.

Kingstone, A. (2009). Taking a real look at social attention. *Curr. Opin. Neurobiol.* 19, 52–56.

Kingstone, A., and Pratt, J. (1999). Inhibition of return is composed of attentional and oculomotor processes. *Percept. Psychophys.* 61, 1046–1054.

Kingstone, A., Smilek, D., and Eastwood, J. D. (2008). Cognitive ethology: a new approach for studying human cognition. *Br. J. Psychol.* 99, 317–340.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., and Eastwood, J. D. (2003). Attention, researchers! it is time to take a look at the real world. *Curr. Dir. Psychol. Sci.* 12, 176–180.

Klein, R. M., and MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychol. Sci.* 10, 346–352.

Kuhn, G., Amlani, A. A., and Rensink, R. A. (2008). Towards a science of magic. *Trends Cogn. Sci.* 12, 349–354.

Kuhn, G., and Land, M. F. (2006). There's more to magic than meets the eye. *Curr. Biol.* 16, R950–R951.

Kuhn, G., and Martinez, L. M. (2012). Misdirection - Past, present and the future. *Front. Hum. Neurosci.* 5:1–7. doi: 10.3389/fnhum.2011.00172

Kuhn, G., and Tatler, B. W. (2005). Magic and fixation: now you don't see it, now you do. *Perception* 34, 1155–1161.

Kuhn, G., Tatler, B. W., and Cole, G. G. (2009). You look where I look! effect of gaze cues on overt and covert attention in misdirection. *Vis. Cogn.* 17, 925–944.

Laidlaw, K. E., Foulsham, T., Kuhn, G., and Kingstone, A. (2011). Potential social interactions are important to

social attention. *Proc. Natl. Acad. Sci. U.S.A.* 108, 5548–5553.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Prog. Retin. Eye Res.* 25, 296–324.

Land, M. F., Mennie, N., and Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception* 28, 1311–1329.

Langton, S. R. H. (2009). Gaze perception: is seeing influenced by believing? *Curr. Biol.* 19, R651–R652.

Leibenluft, E., Gobbini, M. I., Harrison, T., and Haxby, J. V. (2004). Mothers' neural activation in response to pictures of their children and other children. *Biol. Psychiatry* 56, 225–232.

Macknik, S. L., King, M., Randi, J., Robbins, A., Teller, Thompson, J., and Martinez-Conde, S. (2008). Attention and awareness in stage magic: turning tricks into research. *Nat. Rev. Neurosci.* 9, 871–879.

Mar, R. A., Kelley, W. M., Heatherton, T. F., and Macrae, C. N. (2007). Detecting agency from the biological motion of veridical vs animated agents. *Soc. Cogn. Affect. Neurosci.* 2, 199–205.

Materna, S., Dicke, P. W., and Thier, P. (2008). Dissociable roles of the superior temporal sulcus and the interparietal sulcus in joint attention: a functional magnetic resonance imaging study. *J. Cogn. Neurosci.* 20, 108–119.

McKee, D., Christie, J., and Klein, R. (2007). On the uniqueness of attentional capture by uninformative gaze cues: facilitation interacts with the Simon effect and is rarely followed by IOR. *Can. J. Exp. Psychol.* 61, 293–303.

Milner, A. D., and Goodale, M. A. (1995). *The Visual Brain in Action.* Oxford, UK: Oxford Press.

Mitchell, J. P., Banaji, M. R., and Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *Neuroimage* 28, 757–762.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., and Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* 383, 812–815.

Nation, K., and Penny, S. (2008). Sensitivity to eye gaze in autism: is it normal? Is it automatic? Is it social? *Dev. Psychopathol.* 20, 79–97.

Neisser, U. (1967). *Cognitive Psychology.* New York, NY: Appleton-Century-Crofts.

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

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., and Piven, J. (2002). Visual scanning of faces in autism. *J. Autism Dev. Disord.* 32, 249–261.

Perani, D., Fazio, F., Borghese, N. A., Tettamanti, M., Ferrari, S., Decety, J., and Gilardi, M. C. (2001). Different brain correlates for watching real and virtual hand actions. *Neuroimage* 14, 749–758.

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25.

Posner, M. I., and Cohen, Y. (1984). "Components of visual orienting," in *Attention and Performance X,* eds H. Bouma and D. G. Bouwhuis (Hillsdale, NJ: Erlbaum), 531–556.

Posner, M. I., Snyder, C. R. R., and Davidson, B. J. (1980). Attention and the detection of signals. *J. Exp. Psychol.* 109, 160–174.

Povinelli, D. J., and Preuss, T. M. (1995). Theory of mind: evolutionary history of a cognitive specialization. *Trends Neurosci.* 18, 418–424.

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526.

Prinz, W. (1997). Perception and action planning. *Eur. J. Cogn. Psychol.* 9, 129–154.

Quadflieg, S., Mason, M. F., and Macrae, C. N. (2004). The owl and the pussycat: gaze cues and visuospatial orienting. *Psychon. Bull. Rev.* 11, 826–831.

Rafal, R. D., Calabresi, P. A., Brennan, C. W., and Sciolto, T. K. (1989). Saccade preparation inhibits reorienting to recently attended locations. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 673–685.

Raymond, J. E., Shapiro, K. L., and Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *J. Exp. Psychol. Hum. Percept. Perform.* 18, 849–860.

Redcay, E., Dodell-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D. E., and Saxe, R. (2010). Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. *Neuroimage* 50, 1639–1647.

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage* 22, 1694–1703.

Ristic, J., Friesen, C. K., and Kingstone, A. (2002). Are the eyes special? It depends how on you look at it. *Psychon. Bull. Rev.* 9, 507–513.

Saxe, R. (2006). Uniquely human social cognition. *Curr. Opin. Neurobiol.* 16, 235–239.

Scaife, M., and Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature* 253, 265–266.

Schippers, M. B., Roebroek, A., Renken, R., Nanetti, L., and Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9388–9393.

Simons, D. J., and Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends Cogn. Sci.* 9, 6–20.

Skarratt, P. A., Cole, G. G., and Gellatly, A. R. H. (2009). Prioritization for looming and receding objects: equal slopes, different intercepts. *Atten. Percept. Psychophys.* 71, 964–970.

Skarratt, P. A., Cole, G. G., and Kingstone, A. (2010). Social inhibition of return. *Acta Psychol. (Amst.)* 134, 48–54.

Spence, C., and Driver, J. (1998). Auditory and audiovisual inhibition of return. *Percept. Psychophys.* 60, 125–139.

Stuss, D. T., Gallup, G. G. Jr., and Alexander, M. P. (2001). The frontal lobes are necessary for 'theory of mind'. *Brain* 124, 279–286.

Tatler, B. W., and Kuhn, G. (2007). "Don't look now: the magic of misdirection," in *Eye Movements: A Window on Mind and Brain,* eds R. P. G. van Gompel, M. H. Fischer, W. S. Murray, and R. L. Hill (Oxford, UK: Elsevier), 697–714.

Taylor, T. L., and Klein, R. M. (2000). Visual and motor effects in inhibition of return. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1639–1656.

Teufel, C., Alexis, D. M., Clayton, N. S., and Davis, G. (2010a). Mental-state attribution drives rapid, reflexive gaze following. *Atten. Percept. Psychophys.* 72, 695–705.

Teufel, C., Fletcher, P. C., and Davis, G. (2010b). Seeing other minds: attributed mental states influence perception. *Trends Cogn. Sci.* 14, 376–382.

Teufel, C., Alexis, D. M., Todd, H., Lawrance-Owen, A. J., Clayton, N. S., and Davis, G. (2009). Social cognition modulates the sensory coding of observed gaze direction. *Curr. Biol.* 19, 1274–1277.

Tipper, S. P. (1985). The negative priming effect: inhibitory priming by ignored objects. *Q. J. Exp. Psychol. A* 37, 571–590.

Tipples, J. (2002). Eye gaze is not unique: automatic orienting in response to uninformative arrows. *Psychon. Bull. Rev.* 9, 314–318.

Tomasello, M., Kruger, A., and Ratner, H. (1993). Cultural learning. *Behav. Brain Sci.* 16, 495–552.

Torralva, T., Kipps, C. M., Hodges, J. R., Clark, L., Bekinschtein, T., Rocca, M., Calcagno, M. L., and Manes, F. (2007). The relationship between affective decision-making and theory of mind in the frontal variant of frontotemporal dementia. *Neuropsychologia* 45, 342–349.

Walker-Smith, G. J., Gale, A. G., and Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception* 6, 313–326.

Walter, H., Abler, B., Ciaramidaro, A., and Erk, S. (2005). Motivating forces of human actions. Neuroimaging reward and social interaction. *Brain Res. Bull.* 67, 368–381.

Watson, D. G., and Humphreys, G. W. (1997). Visual marking: prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychon. Bull. Rev.* 104, 90–122.

Welsh, T. N., Elliot, D., Anson, J. G., Dhillon, V., Weeks, D. J., Lyons, J. L., and Chua, R. (2005). Does Joe influence Fred's actions? Inhibition of return across different nervous systems. *Neurosci. Lett.* 385, 99–104.

Welsh, T. N., and Pratt, J. (2006). Inhibition of return in cue-target and target-target tasks. *Exp. Brain Res.* 174, 167–175.

Williams, J. H. G., Waiter, G. D., Perra, O., Perrett, D. I., and Whiten, A. (2005). An fMRI study of joint attention experience. *Neuroimage* 25, 133–140.

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128.

Yarbus, A. L. (1967). *Eye Movements and Vision*. New York, NY: Plenum.

Yantis, S., and Jonides, J. (1984). Abrupt visual onsets and selective attention: evidence from visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 601–621.

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