Title

Social Eye Cue:
How Knowledge Of Another Person‚Äôs Attention Changes Your Own

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors

Tufft, Miles R.A.
Gobel, Matthias S
Richardson, Daniel C

Publication Date

2015

Peer reviewed

Social Eye Cue:
How Knowledge Of Another Person's Attention Changes Your Own

Miles R. A. Tufft (m.tufft.12@alumni.ucl.ac.uk)
Matthias S. Gobel (matt@eyethink.org)
Daniel C. Richardson (dcr@eyethink.org)
Department of Experimental Psychology, University College London
London, WC1E 6BT, UK

Abstract

We are highly tuned to each other's visual attention. Perceiving the eye or hand movements of another person can influence the timing of a saccade or a reach of our own. However, it is not clear whether the effect of social cues is due to the *appearance* of the cue – a hand or an eye - or the *belief* that the cues are connected to another person. In two experiments we investigated this question using a spatial cueing paradigm and measuring the inhibition of return of visual attention. When participants believed that a cue stimulus – a red dot – reflected the attentional focus of another person via an eye tracker, they responded differently to when they believed its location was determined by a computer. Despite previous claims that they are 'blind' to such factors, when a cue was imbued with a social context it exerted a stronger influence over low-level visual attention.

Keywords: attention; vision; social context; joint action

Introduction

Our eyes are in constant demand: turn signals, pointed glances, flashing banner ads, they all clamour for our visual attention. Some of these cues – glances, head turns, pointing fingers – are generated by other people; others – warning lights, traffic signals and signposts – are put there intentionally by other people as a signal. And others still – bright plumage on a bird, flashes of lighting, claps of thunder – are oblivious to our presence.

Does the visual attention system respond equally to all these cues? Or do we give special weight to locations or objects in the world that are cued by other people since they might indicate their mental states or communicative intentions? Further, if social cues do interact with basic attention mechanisms, is this interaction a response to the *social appearance* of the cues (e.g. facial features or a finger) or a response to the *belief* that the cues are connected to other intentional beings with their own states of attention?

Here, we address these questions by spatially cueing participants' attention while manipulating their beliefs about the social or non-social origins of those cues. In this way, we are able to investigate how changes in beliefs about social context interact with low-level mechanisms of visual attention (Richardson & Gobel, 2015; von Zimmermann & Richardson, 2014). We made use of a robust feature of spatial attention, and placed it in a social context that could be experimentally manipulated.

Inhibition Of Return (IOR)

People are slower to return their attention back to a location that it has previously occupied. This inhibition of return (IOR) was first demonstrated with a spatial cueing paradigm (Posner, Rafal, Choate, & Vaughan, 1985). Visual attention was cued to one location on screen, and participants responded to a second stimulus that appeared either in the same location, or a different part of the screen. Response times were slower when the stimulus appeared in the same location. IOR might therefore play an adaptive role promoting efficient visual search by biasing attention away from previously attended objects or locations (Klein & MacInnes, 1999).

Our central question is whether IOR can be influenced by beliefs about social context. Staudte & Crocker (2011) proposed that another attentional phenomenon, gaze cueing, could be explained either by a purely reflexive *visual* account, or by an *intentional* account, which makes reference to beliefs about the goals and intentions of other people. By extension, we review arguments below that IOR coud be best explained by a visual or an intentional account.

Purely Visual Accounts Of IOR

In gaze-cuing paradigms, it has been shown that the visual attention of others can reflexively direct our own attention. In such experiments, participants are typically shown a centrally presented drawing or picture of a face gazing in a certain direction, and instructed to respond as quickly as possible to targets appearing either congruent with the gaze direction or incongruent (e.g. Friesen & Kingstone, 1998). Despite gaze being non-predictive of targets it has repeatedly been shown that reactions times to congruently cued targets are facilitated, leading many to conclude that gaze cues induce rapid and involuntary shifts in attention (e.g. Friesen & Kingstone, 1998; Friesen, Ristic, & Kingstone, 2004). If the gaze of other people can act as a cue for an individual's attention, then it seems plausible that social cues would also generate the same IOR effects.

Indeed, a 'social IOR' effect has been demonstrated using reaching behaviour. Welsh et al. (2005) asked pairs of participants to sit at opposite ends of a table and respond with a left or right button press to targets. Each participant responded twice in a row. There were longer reaction times to spatial locations that had previously been attended to, independent of whether attention was cued by the participants' own action or their partners'.

Taylor and Therrien (2005) performed an IOR experiment in which the cue stimulus was either a high-pass filtered image of a human face or a scrambled version of the face. They found the standard IOR effect across both conditions, and the magnitude of the IOR effect did not differ significantly between them. This led the authors to describe IOR as a "blind mechanism". Related studies using faces with differing emotional valance (e.g. Lange, Heuer, Reinecke, Becker, & Rinck, 2008; Stoyanova, Pratt & Anderson, 2007) found the same result that social and non-social cues were equivalent. This echoes findings from the gaze cueing literature. Ristic, Friesen and Kingstone (2002) showed that under certain conditions non-social cues could indeed reflexively cue attention in a way that was behaviourally indistinguishable from social gaze cues (see also Hommel, Pratt, Colzato, & Godijn, 2001; Pratt & Hommel, 2003; Tipples, 2002).

Intentional Accounts Of IOR

The conclusion from the studies described above is that IOR effects can be triggered by social cues, but that the magnitude of the effect is indistinguishable from non-social cues. However, in other contexts, it has been shown that when participants interpret a cue as connected to another person, they respond to it quite differently. For example, if infants (Meltzoff, Brooks, Shon, & Rao, 2010) and adults (Staudte & Crocker, 2011) have experience with a robot looking in time with linguistic information, they start to follow its gaze as they would another human. Hegel, Krach, Kircher, Wrede, and Sagerer (2008) showed that participants in a prisoners dilemma respond differently to computer, robot or human opponents to the degree to which they anthropomorphize them. These results show that beliefs about the social context of a cue can change its effects on visual attention and cognition.

There are good reasons to think that it would be adaptive for IOR mechanisms to be responsive to the intentions of others. If it is true that IOR is a mechanism to make search more efficient for an individual, then it could also make search more efficient for people working together. Individuals who are engaged in joint action (Sebanz, Bekkering & Knoblich, 2006) or joint perception (Richardson et al., 2012) are tuned to the cognitive representations and locus of their partner's attention. Brennan, Chen, Dickinson, Neider, and Zelinsky (2008) for example, found that two individuals performing a visual search task were highly efficient when they could see each other's gaze location.

However, the current literature cannot say whether IOR is best explained by a purely visual or intentional account. Past experiments all operationalize a 'social cue' as something that is social in appearance – a face onscreen or a person sat opposite – and so conflate the two explanations. Therefore our central question remains unanswered – is the effect of social cues due to the *appearance* of the cues or the *belief* that the cues are connected to other intentional beings with their own states of attention?

Experiment 1

In our experiment, we adapted the standard IOR paradigm (Posner et al., 1985) and manipulated whether or not participants believed that a cue was generated by a computer, or reflected the gaze position of another person in the laboratory. Participants were run in pairs with a confederate, sat back to back in the lab, each looking at a screen and monitored by an eye tracker. In each trial, their attention was cued by the brief appearance of a red dot in one quadrant of a screen. Following the cue, a target appeared in a congruent or incongruent location.

In one condition, participants were told that the position of the red cue was generated randomly, while in the other, they were told that it reflected the gaze position of the confederate. In reality the position of red dot was always random and computer generated. Our hypothesis was that the belief that the cue was connected to a real person would modulate the magnitude of the IOR effect.

Methods

Participants Thirty-one participants from the University College London subject pool volunteered to participate in exchange for a £6 payment. Five participants were excluded, one for guessing that the other participant was a confederate, two for falling asleep during the task and two for adverse emotional reactions to background pictures presented during the task. A total of twenty-six participants were analysed (mean age 23.0 years; 10 males, 16 females).

Apparatus Participant and confederate were seated in reclining chairs at opposite corners of a $25m^2$ room such that they had their backs to each other and could not see each other's screens or their partner's actions. Participants faced an arm mounted 19" LCD screen positioned approximately 60cm away on which stimuli were presented. A custom-built remote eye tracker was positioned at the base of the screen. Participants wore headsets throughout the experiment through which they could hear a tone signifying the beginning of each trial. They were also provided with a wireless computer mouse that was used to register the button press target detection response. Though in reality it was not used, all equipment was replicated for the confederate as well as calibration and set-up procedures in order to maintain deception. An iMac computer presented stimuli and recorded RTs

Design This was a 2x2 within subjects design with two factors, congruence and cue condition. Congruence had two levels (validly cued, invalidly cued) and cue condition had two levels (social, non-social).

Validly cued trials were trials on which the blue box was presented in the same location as the red dot; invalidly cued trials were when the blue box was presented in one of three alternate positions. As per previous studies, to ensure the cue was not predictive of target the probability of a validly cued trial was at chance (1 in 4). Social trials were trials on

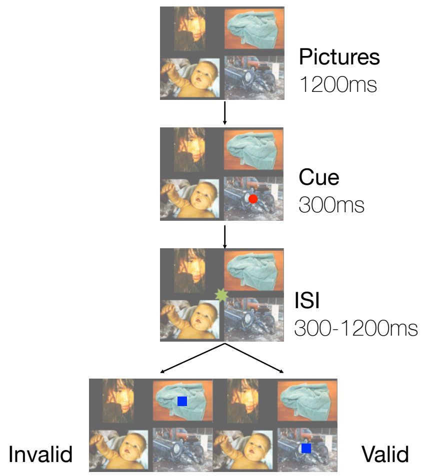

Figure 1. Trial design Experiment 1.

which participants believed the red dot's position represented the gaze of the confederate; non-social trials were trials on which participants believed the red dot's position was random.

Trials employed a classic cue-target spatial cuing paradigm with four possible cue-target locations under overt orientating conditions (eye movements allowed). Each trial began with a 1200ms presentation of a set of four pictures. These images were set at 40% transparency and taken from a normed database (Lang, Bradley, & Cuthbert, 2005). Each set comprised of one negative photo (e.g. crying child), one positive photo (e.g. puppy) and two neutral photos with no strong valence (e.g. spoon). The purpose of the pictures was to provide a legitimate reason for participants to scan a scene adding to the believability that gaze direction information was real and being exchanged. Picture valence was not analysed as a factor in this study. After 1200ms, a red dot cue appeared for 300ms (1cm in diameter; salience was increased by a transient size increase from 1cm to 3cm diameter across the 300ms presentation) in the exact centre of one of the four pictures. This was followed immediately by a centrally (centre of screen) presented green star (1cm diameter; salience was increased by giving the star a rotating motion) presented for one of 300, 600, 900 or 1200ms (randomized within block) before a blue box target ($1cm^2$) was presented in the exact centre of one of the four pictures. Participants were required to press the left mouse button with the thumb of their dominant (writing) hand as quickly and as accurately as possible on seeing the blue box (target detection task). If no response was given after 3000ms the trail ended and the next trial began.

Participants completed a total of 288 trials during a single session comprising of four blocks (two social, two non-social) of 72 trials. Social and non-social blocks were presented alternately with the first block type counterbalanced across sessions. Of the 72 trials per block on average 8 were catch trials during which no blue target was presented as means to maintain vigilance. The other trials were then split between the four different cue locations with an equal probability, such that on average 4 trials would be validly cued and 12 invalidly cued.

Procedure A briefing was provided beforehand during which the paradigm was explained to the participant (and confederate) in some detail. During the briefing, participants were told that they would be performing a simple reaction time task. It was explained that each trial would begin with the presentation of a set of four images and that they were free to inspect these images. Following this, a series of shapes would appear on top of the pictures, these would be a red dot followed by a green star and then a blue box. They were instructed to fixate the red dot and green box and then respond only to the blue box with a single button press as quickly and accurately as possible.

It was also explained that during the task eye trackers would monitor their eye movements and that on certain designated blocks of trials their eye tracker and that of the confederate would be 'linked' together so that they would both be able to see which picture their partner had just looked at. During these 'social blocks' this information would be conveyed by the location of the red dot that would act to highlight the picture their partner had just looked at. In this way participants believed they were engaged in a two-way exchange of information. During 'non-social' blocks participants were told the eye trackers were 'unlinked' and the red dot's location would be chosen at random by the computer and no other aspect of the design would change. However, at no point was any gaze direction information exchanged between confederate and participant and at all points the cue location was randomized and non-predictive of the target.

Following their initial briefing participants were taken through an eye tracker calibration sequence in order to demonstrate that the eye tracker worked and was capable of determining where they were looking on their screen. On completing calibration they were presented with an instruction screen that reiterated instructions given during the briefing and reminded them that for certain blocks their eye tracker would be 'linked' to the confederate's. In addition, at the beginning of each block participants viewed an information screen, which informed them whether the eye trackers were 'linked' or 'not linked', a point the experimenter would also verbally reiterate.

Results & Discussion

Mean RTs (ms) were calculated for all trials, excluding catch trials and any RTs less than 100ms or greater than 3 SD from the participants' mean RT by block, in line with other studies (Taylor & Therrien, 2005). In total 1.8% of trials were excluded.

A two-way repeated measures ANOVA was performed on the mean RTs by participant. The analysis revealed a significant main effect of congruence, $F(1, 25) = 32.69$, $p < .01$, with validly cued trials ($M = 466.7$, $SD = 80.2$) slower than invalidly cued trials ($M = 446.8$, $SD = 77.7$) consistent with an overall IOR effect. There was also a main effect of condition, $F(1, 25) = 5.17$, $p < .05$, with participants significantly slower overall on social ($M = 466.9$, $SD = 92.7$) compared to non-social trials ($M = 446.6$, $SD = 69.0$). Importantly, in combination with a main effect of congruence there was a significant interaction effect between congruence and condition, $F(1, 25) = 4.84$, $p < .05$, with IOR magnitude on social trials ($M = 26.1$, $SD = 20.9$) significantly greater than on non-social trials ($M = 13.7$, $SD = 24.7$).

To our knowledge, these results demonstrate for the first time that the manipulation of *beliefs* about the social context of non-social cues can modulate IOR magnitude. Since no sensory characteristics of the stimuli were changed across conditions, IOR differences cannot be attributed to any low level sensory or perceptual effects. Instead differences must derive from high level interpretations made by the participants.

One concern is that the use of background pictures in our design could have led to potential confounds, with some pictures attracting more attention than others by increased visual salience or emotional arousal. There are two main reasons though why we believe this is unlikely to be the case. Firstly, the picture locations were fully randomized across conditions so all permutations of cue, target, and picture location will have been sampled at each SOA. Secondly, a post-hoc analysis of the three-way interaction between cue picture type (positive, negative or neutral), condition (social, non-social) and congruence (valid, invalid) was non-significant, $F(2, 52) = 0.40$, $p > .05$ (i.e. the magnitude difference seen was the same irrespective of the type of picture behind the presentation of the cue). This provides confidence that the presence of the pictures was not associated with the change in IOR between conditions.

Despite these reasons, the background pictures could still represent a mediating factor. Since a small number of the pictures contain social information such as faces, these could have triggered the effects we found.

Experiment 2

We sought to replicate our first experiment without the presence of background pictures. In addition, we decided to run participants in pairs rather than with a confederate. This served the obvious benefit of improving the rate of data collection (data from two participants collected in a single session) but also controlled for any potential confounds arising from the actions of one particular confederate.

Method

The experimental methods were identical to experiment 1, except for the details listed below.

Participants Thirty-two volunteers recruited from the UCL psychology participant pool volunteered to participate in exchange for a £5 payment. No confederate was used therefore they all performed the experiment in randomly assigned pairs (acting, as it were, as each other's confederate). Following initial introductions, subsequent interactions between participants were minimal. Four participants were excluded, three for falling asleep during the task and one for revealing prior knowledge of the task purpose. A total of twenty-eight participants were analysed (mean age 22.8 years; 12 males; 16 females). Each session lasted no more than fifty minutes.

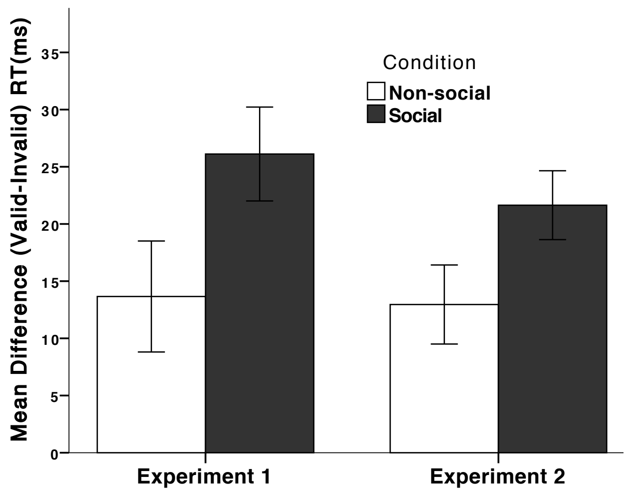

Figure 2. IOR magnitude in Experiments 1 and 2.

Apparatus, Design & Procedure Set up of the experimental apparatus was identical for participant 1 as for the single participant in experiment 1. Participant 2 sat in the same chair originally used by the confederate however data was now collected using exactly the same set up as participant 1. A second iMac was introduced to control stimuli and collect RT data for participant 2.

The design was identical to experiment 1, except the background pictures were not displayed. The cue and target stimuli appeared on white backgrounds in a 2x2 grid.

Results & Discussion

Following experiment 1, mean RTs (ms) were calculated for non-catch trials; and any RTs that were less than 100ms or greater than 3 standard deviations from the participants' mean RT excluded. In total 1.6% of trials were excluded.

A two-way repeated measures ANOVA was performed on the mean RTs by participant. The analysis revealed a significant main effect of congruence, $F(1, 27) = 44.28$, $p < .01$, with validly cued trials ($M = 404.2$, $SD = 68.5$) slower than invalidly cued trials ($M = 386.9$, $SD = 67.0$) consistent with an overall IOR effect. There was also a main effect of condition, $F(1, 27) = 6.89$, $p < .05$, with participants significantly slower overall on social ($M = 403.8$, $SD = 72.1$) compared to non-social trials ($M = 387.4$, $SD = 66.7$). Crucially, experiment 2 replicated experiment 1 and showed a significant interaction effect between congruence and condition, $F(1, 27) = 4.91$, $p < .05$ with IOR magnitude on social trials ($M = 21.6$, $SD = 15.9$) significantly greater than on non-social trials ($M = 13.0$, $SD = 18.3$).

This replication rejects the hypothesis that the original result was due to an interaction from background pictures or the use of a confederate. We therefore conclude that together these results demonstrate that IOR magnitude is influenced by the beliefs about the social context of non-social cue events.

General Discussion

The attentional focus of another person can act as strong cue for our own attention, as elegantly demonstrated in the gaze cueing literature (e.g. Friesen & Kingstone, 1998; Friesen, et al., 2004). Someone else's eye and hand movements can even trigger our own inhibition of return mechanisms (e.g. Skarratt, Cole, & Kingstone, 2010; Welsh et al., 2005). But in all of these experimental demonstrations, the social cue is *visibly* social: a face, a hand or a head turn. And in each case, researchers have shown that the attentional effect of a social cue is roughly equivalent to a non-social cue such as an arrow.

Our results make two contributions to our understanding of attention and social context. They suggest that social cues can have a stronger effect upon IOR attention mechanisms than non-social cues, in the right circumstances. And they show that these differences in IOR effects do not depend on a social appearance, but can rest purely on a belief about the *social context* of a cue. In this way, a red dot that looks like any other can have a more significant influence on low-level visual attention when a participant believes that it is connected to someone else – an *intentional account of IOR*.

Such an account goes against a purely visual account of IOR as a socially blind, bottom-up, stimulus-driven process (e.g. Taylor and Therrien, 2005). It highlights the importance of beliefs on low-level attention mechanisms and builds on similar results from the gaze cuing and joint attention literature. For example, gaze cue effects can be modulated by changes in beliefs about the gazer. Dalmaso, Pavan, Castelli and Galfano (2012) showed that a subject's belief about the relative social status of gazers modulates their gaze-cuing effects. Participants were presented with a series of faces together with fictitious resumes that described the person as being of either high or low status. Subsequently, those faces associated with a higher social status produced greater gaze-cuing effects in participants than those with a lower social status. Furthermore, related studies have shown that in-group membership (Pavan, Dalmaso, Galfano, & Castelli, 2011) and shared political views (Liuzza et al., 2011) can all lead to increases in cuing effects, hinting that some top-down, social processes have an effect on low-level attention.

Teufel, Alexis, Clayton, and Davis (2010) tested this directly by devising an experiment that allowed manipulation of observers' beliefs about the gazer while keeping all cue stimuli the same across all conditions. They presented participants with pre-recorded video sequences of a real-life model turning his head to the left or to the right. Subjects were made to believe the video was live and that they were engaging with a real person. In the video the model wore mirrored goggles so that their eyes were occluded. Observers were told that the model would wear one of two types of goggles clearly indicated by their colour. They were informed that one pair was transparent from the perspective of the model (seeing condition) and the other pair opaque (non-seeing condition). It was found that a cuing effect was only present when the observer believed that the model could see. They concluded that the attribution of a mental state was critical for reflexive gaze following.

Building on this idea, Wiese, Wykowska, Zwickel, and Müller (2012) measured gaze-cuing effects while manipulating observers' beliefs about the capacity of the gazer to hold mental states, termed taking an "intentional stance". In experiment 1 participants viewed gaze cues made by pictures of a real human face and that of a robot. In experiment 2 they viewed cues from the same human face or robot but were either informed that the robot was now controlled by a human being or that the human face was that of a realistic mannequin. They found that gaze cuing effects were only present for gaze cues originating from stimuli where an intentional stance was likely, human face or robot face controlled by a human. When an intentional stance was unlikely, mannequin face or robot face, then gaze cuing effects were significantly reduced.

Even when people have no information about each other's attentional state, research suggests that they try to follow (what they imagine to be) each other's gaze. In a series of studies (Richardson et al., 2012), pairs of participants were instructed to look at sets of pictures, some with positive valence and some with negative valence. Half of the time, they believed that they were looking at the same images, and half of the time that they were looking at different images. This social context changed randomly on a trial-by-trial basis, and participants reported that they mostly ignored the information about their partner's condition. Despite this reported behavior, however, simply knowing that another person was attending to the same stimuli—even though they could not see each other or have any verbal interaction—shifted participants' attention. When participants believed that they were looking at the images together with another person, they tended to look towards the more negative images.

In another experiment from that series (Richardson et al., 2012), participants were told to either (a) search a set of pictures for an "X" or (b) memorize a set of pictures. Each participant was given one of these tasks and was told which of these tasks their partner would be doing as well. In this study, we again observed the powerful effects of social context and belief can have on lower-level behavior: believing their partner was experiencing the same *stimuli* but did not share the same *task* did not result in joint perception. Joint perception only occurred when participants believed that their partner was engaged in exactly the same task (Richardson et al., 2012). One explanation is that when the stimuli were believed to be shared, participants looked towards the images that they thought their partner would also be looking at. In other words, even with this minimal social context of no interaction or visual information about each other, participants were seeking to coordinate their visual attention.

Though many studies in visual perception take place in the solitary confinement of an experimental cubicle, people often use their perceptual faculties in a rich social environment, where findings may not easily generalise. Recent work on face perception, for example, has shown that people produce very different gaze patterns when looking at a live or pre-recorded video (Laidlaw, Foulsham, Kuhn & Kingstone, 2011), or whether or not they believe the person they are watching can look back at them (Gobel, Kim & Richardson, 2015). Thus, social context may shift and structure attentional mechanisms in ways that we do not fully appreciate in the laboratory.

References

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, *106*(3), 1465-1477.

Dalmaso, M., Pavan, G., Castelli, L., & Galfano, G. (2012). Social status gates social attention in humans. *Biology letters*, *8*(3), 450-452.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic bulletin & review*, 5(3), 490-495.

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(2), 319.

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

Hegel, F., Krach, S., Kircher, T., Wrede, B., & Sagerer, G. (2008, August). Understanding social robots: A user study on anthropomorphism. In *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on* (pp. 574-579). IEEE.

Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological Science*, *12*(5), 360-365.

Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological science*, *10*(4), 346-352.

Laidlaw, K. E., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. *Proceedings of the National Academy of Sciences*, *108*(14), 5548-5553.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). International affective picture system (IAPS). *Gainesville, FL: University of Florida*.

Lange, W. G., Heuer, K., Reinecke, A., Becker, E. S., & Rinck, M. (2008). Inhibition of return is unimpressed by emotional cues. *Cognition and Emotion*, *22*(8), 1433-1456.

Liuzza, M. T., Cazzato, V., Vecchione, M., Crostella, F., Caprara, G. V., & Aglioti, S. M. (2011). Follow my eyes: the gaze of politicians reflexively captures the gaze of ingroup voters. *PloS one*, *6*(9), e25117.

Meltzoff, A. N., Brooks, R., Shon, A. P., & Rao, R. P. (2010). "Social" robots are psychological agents for infants: A test of gaze following. *Neural Networks*, *23*(8), 966-972.

Pavan, G., Dalmaso, M., Galfano, G., & Castelli, L. (2011). Racial group membership is associated to gaze-mediated orienting in Italy. *PLoS One*, *6*(10), e25608.

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive neuropsychology*, *2*(3), 211-228.

Pratt, J., & Hommel, B. (2003). Symbolic control of visual attention: The role of working memory and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(5), 835.

Richardson, D. C, & Gobel, M. S. (2015). *Social Attention.* In J. Fawcett, E. Risko & A. Kingstone (Eds). *Handbook of Attention*. Cambridge MA: MIT Press.

Richardson, D. C., Street, C. N., Tan, J. Y., Kirkham, N. Z., Hoover, M. A., & Cavanaugh, A. G. (2012). Joint perception: gaze and social context. *Frontiers in human neuroscience*, *6*.

Ristic, J., Friesen, C. K., & Kingstone, A. (2002). Are eyes special? It depends on how you look at it. *Psychonomic Bulletin & Review*, *9*(3), 507-513.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in cognitive sciences*, *10*(2), 70-76.

Skarratt, P. A., Cole, G. G., & Kingstone, A. (2010). Social inhibition of return. *Acta psychologica*, *134*(1), 48-54.

Staudte, M., & Crocker, M. W. (2011). Investigating joint attention mechanisms through spoken human–robot interaction. *Cognition*, *120*(2), 268-291.

Stoyanova, R. S., Pratt, J., & Anderson, A. K. (2007). Inhibition of return to social signals of fear. *Emotion*, *7*(1), 49.

Taylor, T. L., & Therrien, M. E. (2005). Inhibition of return for faces. *Perception & psychophysics*, *67*(8), 1414-1422.

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

Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, *9*(2), 314-318.

Welsh, T. N., Elliott, D., Anson, J. G., Dhillon, V., Weeks, D. J., Lyons, J. L., & Chua, R. (2005). Does Joe influence Fred's action?: Inhibition of return across different nervous systems. *Neuroscience Letters*, *385*(2), 99-104.

Wiese, E., Wykowska, A., Zwickel, J., & Müller, H. J. (2012). I see what you mean: how attentional selection is shaped by ascribing intentions to others. *PloS one*, *7*(9), e45391

Zimmermann, von, J., & Richardson, D. C. (2014). Joint Perception. In S. S. Obhi & E. Cross, *Shared Representations: Sensorimotor Foundations of Social Life* (pp. 1–12). Cambridge University Press.