Graduate School of
Systemic Neurosciences
LMU Munich

Making eyes with others: How context information modulates attentional orienting to gaze direction

Eva Wiese

Dissertation

at the Graduate School of Systemic Neurosciences

LMU Munich

Munich, March 2013

Date of Oral Examination: June 24th, 2013

Supervisors: Prof. Dr. Hermann J. Müller

Dr. Agnieszka Wykowska

Acknowledgments

This dissertation would not have been possible without the contribution of many people and I am very grateful for their support. I thank everyone who contributed directly to my PhD project, but also those who provided logistic help, support and encouragement.

I am extremely grateful to my supervisor Prof. Dr. Hermann Müller for supporting me during my PhD studies. He accepted me as a doctoral student and gave me the opportunity to work within the framework of the Excellence Cluster of CoTeSys, with the purpose of creating robots which can understand humans and interact with them. With his insightful comments on all experiments and manuscripts originating from this PhD project and his constructive advice, he paved the way to my PhD. I cordially thank Hermann for three great years of scientific training and experience. My deepest gratitude is to my supervisor Dr. Agnieszka Wykowska who supported me during the whole time of my PhD. Whenever I needed her advice she was there for me - not only as a supervisor, but also as a friend. I have been fortunate to have an advisor who gave me the freedom and encouragement to explore things on my own and at the same time provided her competent guidance wherever it was necessary. Her patience, experience and support helped me to overcome many difficult situations and to finish this dissertation. I also want to thank my supervisor Prof. Dr. Stefan Glasauer with whom I share the interest for research on social robotics and theory of mind. As a member of my thesis advisory committee (TAC), Stefan Glasauer accompanied my research in a supportive and encouraging way and gave me great and invaluable support for my PhD project.

I am grateful to my friends Ahu, Alexander, Ellen, Jairo, Johanna, Petra, Tobias and Wolfgang from the Department of Experimental Psychology for their company, innumerable enriching discussions and relieving conversations. Lunch breaks would have been nothing without you! Furthermore I would like to thank the members of the Graduate School of Systemic Neuroscience for their help and support, especially the coordinators Dr. Alexandra Stein and Lena Bittl who have been very patient and helpful in answering my questions. I would also like to thank Dr. Jan Zwickel, who was my supervisor during the first year of my PhD and contributed interesting ideas, constructive advice and helpful comments on the manuscripts. I would like to acknowledge my collaborators, Stefan Kohlbecher, Stefan Sosnowski and Erich Schneider for programming the robots and for help with analyzing data. I am also very grateful to my student assistants Christoph Voelter, Ulrike Franke, Natalia Rabel, Verena Zierl, Regina Lanz, Nona Gergova, Jasmin Kajopoulos and Ebru Baykara for help with recruiting and testing participants.

Most importantly, none of this would have been possible without the patience of my family and friends, who have been a source of permanent support during the last three years. I’d like to thank my dearest friend Christoph who always believed in me and supported my scientific career from the beginning. I also thank my beloved parents Thea and Heinz for their life-long encouragement, financial and emotional support. My deepest gratitude and love goes to my boyfriend Mark who taught me the power of patience, positive thoughts and with whom I share the enthusiasm for science. Words cannot describe how grateful I am to my twin-sister, best friend, life-consultant, zen-master, voodoo-warrior and karma-king Lisa for her infinite support throughout all the good and bad times in my life.

Finally, I highly appreciate the financial support from the Graduate School of Systemic Neuroscience and the Excellence Cluster of CoTeSys.

Table of Contents

APPENDIX

ABSTRACT

Observing the actions of other people has a strong impact on how we process further information. For instance, when we watch other people gazing at a distinct location in space, we rapidly attend to the same location by following their gaze direction (Emery, 2000; Friesen & Kingstone, 1998; Frischen, Bayliss & Tipper, 2007). Gaze direction is an important social signal as it indicates the current object of interest that might be the target of subsequent actions. The ability to share attention based on gaze following develops early in life (Moore & Corkum, 1995; Striano & Reid, 2005) and seems to be influenced by context factors (Frith & Frith, 2006; for review), such as information about emotional (Graham & LaBar, 2007; cf. Hietanen & Leppänen, 2003) and mental states (Teufel et al., 2009, 2010).

Behavioral studies have shown that watching another person gaze at a distinct location causes shifts of the observer's attention to the same location (Driver et al., 1999; Friesen & Kingstone, 1998) – a mechanism that is commonly referred to as *gaze cueing*. Attending to gaze direction has traditionally been assumed to be a bottom-up process used by the brain to make inferences about intentions, beliefs or action goals of other people (e.g., Nummenmaa & Calder, 2009). However, this assumption has been questioned recently by proposals that even the most basic perceptual processing of social stimuli is top-down modulated by the context in which the social scene takes place (Frith & Frith, 2006; Teufel et al., 2010). The authors argued that in order to cope with perceptually ambiguous scenes, the human brain relies on a prioritization mechanism that structures perceptual information according to its social relevance based on context information.

This PhD-work systematically investigated the influence of context information on social attention mechanisms. In four studies, the interaction of bottom-up and top-down mechanisms during the process of attentional orienting to gaze direction was examined. Top-down processes were triggered by providing context information about a) the presence of objects in the periphery to which gaze direction could refer (**Study 1**), b) the probability with which gaze direction indicates a target position (**Study 2**), and c) the likelihood that mental states can be attributed to the gazer (**Study 3 & 4**). In all experiments, gaze-cueing effects were modulated by context information: **Study 1** showed general cueing effects for the whole gaze-cued hemifield when no context was provided, while the presence of reference objects in the periphery induced an additional gaze-cueing effect specific to the exact gazed-at position. **Study 2** provided evidence that specific cueing effects were observed when participants believed that nonpredictive gaze cues were predictive, while specific cueing effects were attenuated when predictive cues were believed to be nonpredictive. **Study 3** and **4** showed that the size of the cueing effects is remarkably modulated by beliefs about the mental states of the gazer. In particular, in **Study 3**, cueing effects were larger in conditions in which participants were made to believe that they were observing human behavior relative to mechanistic behavior. Most importantly, this effect was independent of the physical characteristics of the stimuli, but depended only on whether intentionality could be attributed to the gazer. **Study 4** showed that the observed modulation of cueing effects by mental state attribution takes place at early stages of perceptual processing, as ERP-components reflecting sensory gain mechanisms (i.e., P1) were affected by beliefs about the intentionality of gaze behavior.

In sum, the findings reported in this PhD-work consistently show that the allocation of attention following changes in gaze direction is top-down modulated by context information which allows gaze-triggered attention mechanisms to be sensitive to the social relevance of objects/events in the environment: while the bottom-up mechanism assures a general preparedness to social signals transmitted by other people, the top-down mechanism allows a flexible adaptation to the context of the scene. These findings led to the postulation of a theoretical model that integrates the effects of bottom-up and top-down components on gaze cueing: the bottom-up component causes a general directional bias to the cued hemifield, while the top-down component specifically guides attention to objects of high social relevance.

INTRODUCTION

1.1. The Social Brain – How we manage to understand others

For engaging in social interactions with other people we need to know who we are interacting with, and what others are going to do next (Frith & Frith, 2006). Based on this knowledge, which can be acquired directly through interacting with other people or indirectly by either observing others or by being told about them, we draw inferences about their internal states (i.e., intentions, beliefs, feelings). This mentalizing process (Frith & Frith, 2003) allows us to make predictions about others' behavior by combining information we deduce from the current state of the world with assumptions we have about possible internal states. Accordingly, the interpretation of social interactions is thought to involve a bottom-up mechanism that is activated by perceptual information in the environment, and a top-down mechanism that is based on background knowledge we have about others or infer from perceived information (e.g., Frith & Frith, 2006; Teufel, Alexis, Todd, Lawrence-Owen, Clayton et al., 2009; Teufel, Alexis, Clayton & Davis, 2010a; Teufel, Fletcher & Davis, 2010b). The top-down mechanism helps to determine the social relevance of observed content and regulates reflexive bottom-up responses accordingly.

It has traditionally been assumed that general cognitive processes involved in perception, attention, and memory are sufficient for interacting with others and for understanding behavior (Blakemore, Winston & Frith, 2004; for a discussion). In contrast to this assumption, however, recent neuroscientific evidence suggests that in order to support adequate behavior in social interactions, the human brain is equipped with a system specialized for the processing of social information (Adolphs, 1999; Brothers, 1990). The so-called *Social Brain* (Adolphs, 1999; Dunbar, 1998) acts over and above the general cognitive processes, has been shown to be unique to humans and some non-human primates (Emery, 2000) and has most probably evolved from environmental changes that made it necessary to live together in groups (Adolphs, 1999; Emery, 2000). The following paragraph discusses the important milestones in evolution that have led to the development of a social brain in primates and describes the neural structures in the brain that are exclusively dealing with the processing of social information.

1.1.1. Evolution of the Social Brain

Social living is associated with plenty of benefits, as groups provide greater opportunities for mating, together with an increased genetic diversity and a reduced risk for potential inbreeding (Emery, 2000). Social groups assure protection from predators, and provide opportunities for learning from conspecifics about how to cope with environmental challenges. Living in groups, however, can also be challenging when conspecifics constitute potential competitors for limited resources like food or mates (Emery, 2000): To deal with these challenges, it is important to have access to information about the other conspecifics in order to predict their behavior and emotional dispositions. This is particularly important for animals living in hierarchical systems, as they need to know about their own position and the position of others within the hierarchy in order to predict the outcome of social interactions and to determine when to challenge others in order to climb to a more powerful position.

Consequently, to coordinate social behavior within a group, information about both the external environment and the internal states of others needs to be communicated among group members. Group communication is thereby monitored through *social cues*, that is: signals that have acquired a special social meaning in the course of evolution and can be interpreted by all group members. Information transmitted by these cues can be spatial, that is: inform about the location of a potential danger in the environment (e.g., predator), but also about sources for food and water (Emery, 2000; Nummenmaa & Calder, 2009, for reviews). At the same time, social cues are used to transmit information about emotional states such as anger, fear or love (Adams & Kleck, 2005; Blakemore, Winston & Frith, 2004). In more evolved species, like humans and non-human primates, social cues can also be used to make inferences about the mental states of others, such as action goals, beliefs or intentions (Baron-Cohen, 1995).

The specificity of social cues is thereby determined by the sensitivity of the sensory channel that is used for communicating information (Emery, 2000): Olfactory social signals (e.g., pheromones, sweating), for instance, are spatially ambiguous and nonspecific – a characteristic making it impossible to use them to communicate the location of potential threat or objects of interest to others. However, these cues are particularly important for indicating the sexual status or the readiness to mate (Emery, 2000). Visual social signals (gaze direction, head orientation, gestures), on the other hand, are spatially more informative, less ambiguous and suitable to convey more complex information. They can be directed towards others to signal emotional states (e.g., anger, happiness) or towards the periphery to indicate the location of interesting objects (e.g., food) or events (e.g., presence of a predator). As visual social cues are superior to other cues with regard to their communicative power, it is not surprising that not only humans, but also a large number of other vertebrates (e.g., non-human primates, reptiles, birds) are capable of processing social information conveyed by visual signals (Emery, 2000, Table 2).

In the human species, the face is the most important source for social information, as it provides a wealth of information about gender, age, familiarity, emotional expression, intentions and mental states (Emery, 2000). Of particular interest for conveying information about another’s internal states is eye gaze. Information from the eye region is used for evaluating the others’ interest in specific objects/events in the environment, for inferring their emotional states and for communicating hierarchical information within a group. The use of gaze direction as social cue evolved from morphological, environmental and behavioral changes throughout the primate evolution and supplemented and/or replaced more ambiguous auditory and olfactory signals (e.g., Emery, 2000; Kobayashi & Kohshima, 1997). To avoid that information deduced from gaze direction interferes with directional information derived from other social cues like head orientation or pointing gestures, social information is weighted according to its social relevance, where eye gaze provides the most important social information, followed by head/body orientation and pointing gestures (Perrett, Hietanen, Oram & Benson, 1992).

Although information conveyed by the eyes is an important signal in many species, humans are unique in their ability to use gaze direction as an indicator for the locus of interesting objects/events

(Emery, 2000). The most important change throughout primate evolution that allowed gaze direction to be used as a social signal in humans stems from the high contrast between white sclera and black pupil in the eye: while a high iris-to-sclera ratio is reported in non-human primates (Kobayashi & Kohshima, 1997), the iris-to-sclera ratio is low in humans: that is, human eyes contain a large extent of white sclera on either side of the dark iris (Perrett & Mistlin, 1991), making it easy to determine where gaze is directed by comparing the position of the dark iris relative to the white sclera. If the dark region is in the middle of the eye, someone is looking at you, whereas when the dark part is positioned in the left/right half of the eye, someone is looking away from you to the left or to the right side. Being able to apply this simple algorithm is considered to be one of the crucial factors allowing humans to use directional information derived from the eyes to learn about objects/events by associating verbal information provided by others with visual information in the environment (Kobayashi & Kohshima, 1997).

1.1.2. Neural correlates of the Social Brain

The challenges accompanying group living, however, necessitated both neural and structural changes in the brain to account for the increased cognitive requirements for communication and social interaction. For that purpose, specific neural networks have evolved in the primate brain, which are specialized on understanding the actions, intentions and emotional states of others.

Understanding the actions of others

For understanding observed actions (i.e., hand movement), it is essential to infer the potential goals (i.e., grasping a bottle) associated with these actions (Blakemore et al., 2004; Rizzolatti & Craighero, 2002). To accomplish this, we need to simulate observed actions ourselves, as a basis for anticipating possible outcomes and underlying action goals (Rizzolatti & Craighero, 2002). For that purpose, the primate brain is equipped with a certain type of neurons, which simulate observed actions as if the observer was performing them him-/herself: the *mirror-neurons* (Di Pellegrino, Fadiga, Fogassi, Gallese & Rizzolatti, 1992; Gallese, Fadiga, Fogassi & Rizzolatti, 1996; Rizzolatti, Fadiga, Fogassi & Gallese, 1996). Studies on macaque monkeys have shown that mirror-neurons are located in the premotor cortex (i.e., the F5-area) and are activated both when an action is observed and when it is executed (e.g., Gallese, Fadiga, Fogassi & Rizzolatti, 1996): that is, every time an action performed by others is observed, the motor representation of this action is activated in the observer, which helps understanding the intentions behind this action together with its potential outcomes (Rizzolatti & Craighero, 2002). That way, the mirror-neuron system transforms observed actions into internal representations of these actions (Gazzola, Rizzolatti, Wicker & Keysers, 2007) and allows inferring intentions and action goals of others based on observation (Rizzolatti & Craighero, 2002).

Comparable to what has been found in macaque monkeys, the premotor cortex is also activated in humans when observing the actions of others (e.g., Buccino, Binkoffski, Fink, Fadiga, Fogassi et al., 2001; Rizzolatti, Fadiga, Matelli, Bettinardi, Paulesu et al., 1996). The mirror-neuron system seems to be a fundamental prerequisite for various social-cognitive skills, as it interprets the actions of others in relation to one’s own actions (Blakemore et al., 2004). Mirror-neurons have been thought to exclusively respond to biological (i.e., pointing) compared to non-biological movements (i.e., actions performed by machines/robots), based on the assumption that they are specialized on the interpretation of actions of conspecifics (Rizzolatti, Fogassi & Gallese, 2001). However, more recent studies have found mirror-neuron activity also for mechanistic robot movements (Gazzola et al., 2007; Oberman, McCleery, Ramachandran, & Pineda, 2007), pointing to the conclusion that the goal-directedness of observed actions seems to be more important for triggering mirror-neuron activity than the exact way of how observed movements are performed. In any case, the mirror-neuron system seems to prepare the observer to perform adequate action-responses to the observed behavior of others. In support of this notion, Fadiga and colleagues (1995) have shown that even when only observing actions, the same muscles in the peripheral motor system are activated that would be necessary to perform the observed action.

Importantly, the link between observed actions and our own actions is not only reflected in neuronal activation, but also translates to behavioral responses (Sebanz, Knoblich & Prinz, 2003): for instance, when a participant is performing a task together with a partner (with subtask A being performed by the participant and subtask B being performed by the partner), the other’s presence alters the participants’ actions in a comparable way, as if the participant was simultaneously performing task A and B without the partner. Furthermore, the observation of the others’ actions interferes with the execution of our own actions, which leads to a reduced performance in the executed task (Brass, Bekkering & Prinz, 2001; Craighero, Bello, Fadiga & Rizzolatti, 2002; Kilner, Paulignan & Blakemore, 2003). The presence of interference effects has been shown to be specific to the observation of human behavior, whereas the observation of robot movements seemed not to interfere with executed movements (Kilner et al., 2003). However, more recent studies provide evidence that interference effects occur even when observed actions are performed by an anthropomorphic robot (Oztop, Franklin, Chaminade & Cheng, 2005).

Understanding the minds of others

In addition to the mirror neuron system, humans are also equipped with a unique cognitive system that is specialized on *mindreading* and allows for interpreting observed behavior in mentalistic terms (Baron-Cohen, 1995; Frith & Frith, 2006; for review). On a neuronal level, when we try to understand other people, a network consisting of medial prefrontal cortex (mPFC), superior temporal sulcus (STS), fusiform gyrus (FG), amygdala and intraparietal sulcus (IPS) is activated (e.g., Adolphs, 1999; Brothers, 1990; Frith & Frith, 2006). In this network (see **Figure I1**), the STS codes for biological movements, such as head direction, pointing gestures and in particular gaze direction, with eye gaze being the most important source for inferring the mental states of others (Baron-Cohen, 1995; Frith & Frith, 2006, for review). Accordingly, the STS is more responsive to meaningful compared to random eye movements, resulting in stronger activation of the STS when gaze is directed towards an object rather than an empty space (Pelphrey, Singerman, Allison & McCarthy, 2003). The STS is also involved in face perception via its connections with the FG (Haxby, Hoffman & Gobbini, 2000, 2002; Hoffman & Haxby, 2000): while the FG is responsible for coding invariant aspects of the face (i.e., face identity), the STS detects changeable properties of the face (i.e., gaze direction). Reciprocal projections have also been found between the STS and the amygdala, a structure of the limbic system involved in the processing of facial expressions and emotional content of perceived information (Aggleton, Burton, & Passingham, 1980; Aggleton, 1993; Thomas, Drevets, Whalen, Eccard, Dahl et al., 2001).

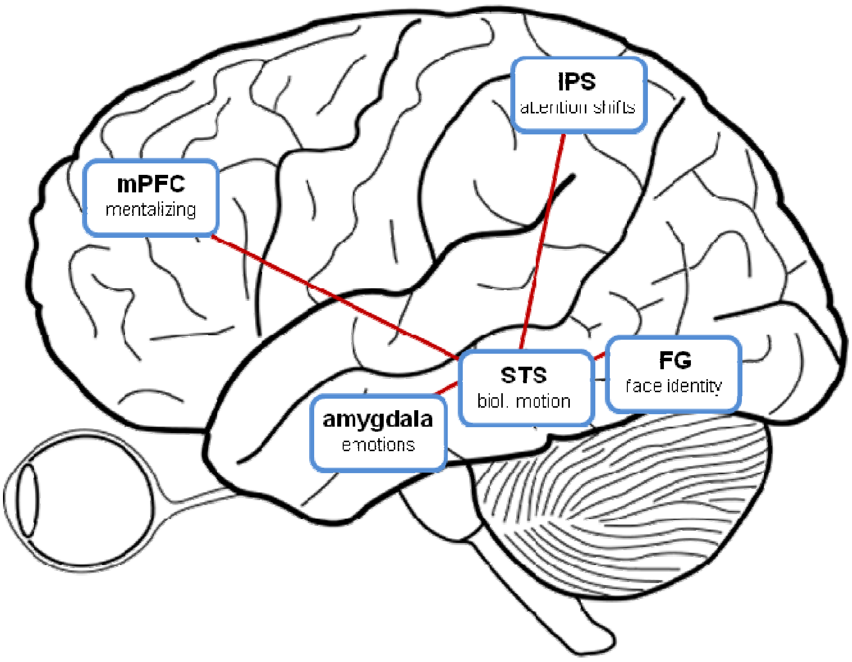

Figure I1. The mindreading system is represented by a network consisting of medial prefrontal cortex (mPFC), amygdala, superior temporal sulcus (STS), fusiform gyrus (FG) and intraparietal sulcus (IPS). While the STS is coding for biological movements (such as changes in gaze direction), the FG is sensitive to facial identity. Shifts of attention that are performed by the IPS can be top-down modulated by information about internal states of others, such as intentions (mPFC, mentalizing), or emotional states (amygdala, facial expressions).

Furthermore, the STS is involved in the orientation of attentional resources via its connections with the parietal cortex, and in particular with the IPS (Corbetta, Miezin, Shulman, & Petersen, 1993; Nobre, Sebestyen, Gitelman, Mesulam, Frackowiak et al., 1997). The perceived direction of biological movements (processed by the STS) can therefore be linked to the spatial attention system (processed by the IPS), in order to initiate reorienting of attention in direction of observed biological movements (e.g., *gaze direction*; George, Driver, & Dolan, 2001; Pelphrey, Singerman, Allison & McCarthy, 2003). The attribution of intentionality to observed behavior is exerted by the mPFC via a top-down modulation of the signals from the STS (Grezes, Frith & Passingham, 2004 a,b; Saxe & Wexler, 2005): while the STS is sensitive to the perceived movements of others, the mPFC is more responsive to judgments about the intentions behind the observed movements. Furthermore, activity in the mPFC seems to be positively correlated with the degree of background knowledge one has about an observed person (Saxe & Wechsler, 2005) and is thought to be responsible for the adaptation of perceived information to the social meaning of the observed scene.

Understanding the emotional states of others

Information about the emotional states of others is usually transmitted by facial expressions which are processed by the amygdala (Breiter, Etcoff, Whalen, Kennedy, Rauch et al., 1996), read off the face very quickly (Adolphs, 2003; Eimer, Holmes & McGlone, 2003; Sugase, Yamane, Ueno & Kawano, 1999), and processed without awareness or attentional engagement (Vuilleumier, Armony, Driver & Dolan, 2001; Whalen, Rauch, Etcoff, McIlerney, Lee et al., 1998). However, this is true only for basic emotions, such as anger, fear or happiness. In contrast, more complex emotions, such as jealousy, embarrassment or guilt, cannot simply be deduced from facial expressions, but require awareness about one’s own relation to the other person and are therefore assumed to involve the mentalizing system (e.g., Berthoz, Armony, Blair & Dolan, 2002; Farrow, Zheng, Wilkinson, Spence, Deakin et al., 2001). In addition to the processing of emotional content, humans are also able to anticipate emotional states of others and empathize with them: that is, expecting an emotional reaction from others activates the same brain areas (i.e., anterior insula, cingulate) as would be activated when the respective emotion is experienced by the person him-/herself (Singer, Seymour, O’Doherty, Kaube, Dolan et al., 2004).

However, facial expressions do not only tell us about the internal states of a person, but can also provide information about the attitude of a person towards objects in the environment (Becchio, Bertone & Castiello, 2008, for review): that is, while gaze direction indicates the location of potentially interesting objects (via the STS and IPS), facial expression (via the amygdala) is informing about the emotional reaction a gazer has towards the objects (Bayliss, Paul, Cannon & Tipper, 2006a). This mechanism is of particular importance for understanding other people, as it indicates not only how others are feeling, but also which object/event has caused that specific emotional state.

1.2. The role of gaze direction in social interactions

Gaze direction is the most important means in human development to communicate social information between interaction partners: it is used to learn about the environment, and about what others consider to be interesting. Humans employ gaze direction to signal internal states and to guide the others' attention (Frischen, Bayliss & Tipper, 2007; for a review). The following chapter describes the mechanisms that underlie the processing of gaze direction: First, it is discussed how gaze following develops in human infants. Second, it is reported how attentional orienting mechanisms are triggered by changes in observed gaze direction and what distinguishes gaze direction from other, non-social directional cues (i.e., arrows). Third, it is discussed whether or not gaze direction can be modulated by social context information.

1.2.1. Development of gaze following in human infants

Sensitivity to eye gaze can be observed from birth onwards (Striano & Reid, 2006) and even neonates prefer to look at faces with direct relative to averted gaze (Farroni, Csibra, Simion & Johnson, 2002). In the first months of life, infants preferably engage in dyadic (i.e., face-to-face) interactions in order to reciprocally share affect and emotions with others (see Striano & Reid, 2006; for review). With three months, infants start engaging in social interactions (Lavelli & Fogel, 2005), and begin to focus more on the others' eyes (Jaffe, Beebe, Feldstein, Crow, Jasnow et al., 2001; Johnson & Morton, 1991). They show signs of stress when others are reducing their engagement in social interactions with them and try to gain interest back by reducing smiling and gazing behavior (Adamson & Frick, 2003; Striano & Bertin, 2005a). Infants are also able to distinguish between relevant and irrelevant signals in social interactions by determining whether feedback (i.e., smiling) is contingent with their actions or not (Striano & Reid, 2006; for review). The ability to participate in dyadic interactions is an important pre-requisite for mentalizing about the internal states of others (Baron-Cohen, 1995; Iacoboni, Molnar-Szakacs, Gallese, Buccino, Mazziotta et al., 2005) and has been shown to be dysfunctional in infants with autism spectrum disorder (Osterling & Dawson, 1994).

Major transitions in social-cognitive development take place when infants start engaging in triadic interactions involving two people and an object that is conjointly attended by both partners (Striano & Reid, 2006). Triadic interactions are essential for language acquisition and imitation as they allow the infant to determine when important information is communicated and what it refers to (Baldwin & Moses, 2001; Brooks & Meltzoff, 2005): mutual gaze (i.e., gaze directed at one another) signals the fact that an important issue will be shared, whereas averted gaze (i.e., gaze directed to the side) indicates the spatial location of the object to which the information refers. With three months, infants start following the gaze direction of others to peripheral locations, but are not yet attending to gazed-at objects in the periphery (Hood, Willen, & Driver, 1998). At the age of nine months, they are able to conjointly attend to the same object as their social partners (Carpenter, Nagell, Tomasello, Butterworth & Moore, 1998; Tomasello, Carpenter, Call, Behne & Moll, 2005) and use gaze direction for coordination of attentional processes between people and objects (Flom & Pick, 2005; Striano & Bertin, 2005b).

While attending to others' gaze direction is based on rather low-level factors in the first year of life (i.e., triggered by moving eyes), infants at the age of 12 to 18 months begin to pay more attention to the context in which eye movements are observed and engage attentional resources more flexibly when following gaze (e.g., Brooks & Meltzhoff, 2002, 2005; Flom & Pick, 2005). Around the same age, other important social skills, such as the recognition of facial identity, face processing and identification of emotional expressions become more sophisticated (Frischen, Bayliss & Tipper, 2007; Nelson, 2001, for reviews). Based on this increased understanding of the social environment, young infants start allocating their attentional resources more efficiently depending on the social relevance of gazed-at events or objects (Baron-Cohen, 1995). They understand that people usually attend to their own actions and can infer motor intentions of others by monitoring their gaze (Castiello, 2003; Frischen et al., 2007; Jellema, Baker, Wicker & Perrett, 2000; Pierno, Becchio, Wall, Smith, Turella et al., 2006).

Around the age of 3-4 years, children start using the gaze behavior of others to make inferences about their internal states and thus acquire a *theory-of-mind* (Baron-Cohen, 1995). With a functioning theory-of-mind, children can mentalize about the intentions of others and determine that someone is looking at a particular object because he/she intends to perform an action on it or has certain beliefs about that object.

1.2.2. Gaze direction as a cue for attentional orienting

Gaze direction is not only important in childhood to communicate with others, but has also been shown to influence attentional orienting in adults by causing shifts of attention to gazed-at locations (Driver, Davis, Ricciardelli, Kidd, Maxwell et al., 1999; Friesen & Kingstone, 1998; see Frischen et al., 2007, for review). In experimental psychology, shifts of attention to particular locations have traditionally been investigated using a cueing paradigm (Posner, 1980), in which the observer’s attention is cued to peripheral locations either by abrupt onsets of a visual stimulus in the periphery (transient luminance increments: Posner, 1980; Müller & Rabbitt, 1989) or by the presentation of central symbolic cues that can carry non-social (arrows: Posner, 1980; Müller & Rabbitt, 1989; Ristic & Kingstone, 2006; words: Hommel, Pratt, Colzato & Godijn, 2001; digits: Rizzolatti, Riggio, Dascola & Umilta, 1987; Sheliga, Riggio & Rizzolatti, 1994) or social information (gaze: Friesen & Kingstone, 1998; pointing gestures: Langton & Bruce, 2000; tongues: Downing, Dodds & Bray, 2004). In response to spatial cues, attention can be allocated overtly by performing saccades to the cued locations (Hoffman & Subramaniam, 1995; Posner, 1980; Shepherd, Findlay, & Hockey, 1986) or covertly by shifting attention without moving the eyes (Posner, 1980). These shifts of attention can be triggered either in a *bottom-up manner* by stimulus properties, resulting in involuntary shifts of attention to the indicated location that are difficult to suppress (Posner, 1980), and/or in a *top-down manner*, resulting in voluntary shifts of attention to the indicated location (e.g., Jonides, 1981; Müller & Rabbitt, 1989). Top-down controlled shifts of attention are usually evoked by symbolic cues (e.g., arrows) that require interpretation, are presented centrally on the screen, and indicate target position with a high likelihood (Jonides, 1981; Müller & Rabbitt, 1989). In contrast, cues triggering shifts of attention in a bottom-up fashion are usually more salient than the surrounding (e.g., flashes), appear at peripheral locations, and induce shifts of attention even when being nonpredictive for the location of subsequent targets (Posner, 1980).

In gaze-cueing paradigms (e.g., Friesen & Kingstone, 1998; Frischen et al., 2007, for review), a face is presented centrally on the screen that gazes either straight ahead or to the left or right side of the screen. Gaze direction has been shown to shift the attentional focus of the observer towards the gazed-at location, resulting in faster responses to targets that appear in the gaze-cued hemifield compared to targets in the uncued hemifield. Gaze-cueing effects represent basic social attention mechanisms, which emerge already 150 ms after cue onset, and decay relatively quickly (about 1000 ms after cue onset). The effect of gaze direction on spatial attention is not limited to realistic faces (Hietanen & Leppänen, 2003; Langton & Bruce, 1999), but can also be induced by schematic faces (e.g., Friesen & Kingstone, 1998), as well as by various kinds of objects that carry eye-like information (Quadflieg, Mason & Macrae, 2004).

1.2.3. Is gaze direction special in cueing the attention of others?

Gaze direction has traditionally been thought to be special with regard to guiding attention, for several reasons: In contrast to other central cues (Müller & Rabbitt, 1989; Posner, 1980; Ristic & Kingstone, 2006), gaze triggers attention shifts to peripheral locations when being nonpredictive (Friesen & Kingstone, 1998; Langton & Bruce, 1999; cf. Vecera & Rizzo, 2006) or even counterpredictive with respect to the target location (Friesen, Ristic & Kingstone, 2004) – a pattern that is consistent with a reflexive, bottom-up mechanism. This reflexive orientation mechanism was thought to be exclusive to gaze direction, given that adults can voluntarily orient away from other central cues, such as counterpredictive arrows (Friesen et al., 2004) or extended tongues (Downing et al., 2004). However, the view that gaze direction provides a particularly powerful orientation signal (due to its social relevance) has been challenged recently by evidence showing that not only gaze direction, but also other overlearned symbolic cues (e.g., arrows, words) are capable of inducing shifts of attention when being nonpredictive for the target location (e.g., Hommel et al., 2001; Ristic, Friesen, & Kingstone, 2002; Tipples, 2002; cf. Friesen et al., 2004).

However, although behavioral effects caused by gaze and arrow cues partially overlap, there are still substantial differences between these cues with regard to their underlying mechanisms (Nummenmaa & Hietanen, 2009): First, while both nonpredictive and counterpredictive gaze cues cause reflexive shifts of attention to the gazed-at location, reflexive shifts of attention are only induced by nonpredictive, but not counterpredictive arrow cues (Friesen et al., 2004; cf. Hommel et al., 2001). Based on these findings, Friesen et al., (2004) concluded that orienting attention to gaze is *more* reflexive (i.e., less vulnerable to top-down control) than orienting to arrows. Second, gaze cues have been shown to induce both facilitation and inhibition effects (after 2000 ms; Frischen, Smilek, Eastwood, & Tipper, 2007), whereas cueing effects induced by arrows seem to trigger facilitatory effects only (Posner, Nissen, & Odgen, 1978). Third, there is evidence from neuroimaging studies that gaze and arrow cueing are subserved by different neural networks (Hietanen, Nummenmaa, Nyman, Parkkola, & Hämälainen, 2006): that is, arrow cues cause stronger activation in the voluntary attention network than gaze direction emphasizing the fact that attentional orienting to arrow cues might be less reflexive, but more voluntarily controlled than gaze cueing.

1.2.4. The influence of context information on gaze-cueing effects

Although gaze cues have been shown to trigger reflexive shifts of attention (Frischen et al., for review), there is also evidence that orienting to gaze direction is not purely reflexive but can rather be top-down controlled by higher-level cognitive processes (Kawai, 2011; Ristic & Kingstone, 2005; Teufel et al., 2009, 2010 a,b). The operation of top-down processes, however, has been shown to critically depend on the availability of context information (Frith & Frith, 2003; Teufel et al., 2009, 2010 a,b). In particular, preexisting assumptions concerning the observed person have been shown to influence the readiness to process and/or attend to gaze direction (Deaner, Shepherd, & Platt, 2007; Kawai, 2011; Liuzza, Cazzato, Vecchione, Crostella et al., 2011; Ristic & Kingstone, 2005; Teufel et al., 2009, 2010 a,b; cf. Bayliss & Tipper, 2006b): Teufel and colleagues (2009), for instance, reported that adaptation to gaze direction depended on whether participants believed that an observed person could actually see through a pair of goggles. Furthermore, in a study by Ristic and Kingstone (2005), participants were presented with an ambiguous stimulus that could be perceived either as a car with wheels or a face with a hat. The stimulus cued participants' attention only in conditions when they believed that they were looking at a face rather than a car. Similarly, Kawai (2011) found schematic faces to cue the participants' attention only when they believed that the target was visible to the gazer, but not when it was believed to be occluded. Furthermore, it has been shown that when the gazer is representing a stereotypical member of a group that the observer belongs to (e.g., leader of a political party), the observer is more likely to follow his/her gaze direction, resulting in increased cueing effects (Liuzza et al., 2011).

Interestingly, however, not every type of context information has been proven to influence gaze cueing. In particular, facial *identity* does not interact with orienting to gaze direction (Bayliss et al., 2006a,b; Quadflieg et al., 2004), whereas facial *expression* does (Peccinenda, Pes, Ferlazzo & Zoccolotti, 2008). With regard to facial identity, Bayliss and Tipper (2006b) showed that the identity of a gazing stimulus does not influence gaze cueing when the different faces indicate target location with different likelihoods; and Quadflieg and colleagues (2004) showed that gaze-cueing effects are not exclusive to realistic human faces, but can also be induced by all sorts of inanimate objects containing eye-like stimuli. With regard to facial expression, on the other hand, Graham and colleagues (2010) found that the size of gaze-cueing effects is modulated when gaze direction is changed with a fearful compared to a happy or neutral face. One likely explanation for the differential effects of facial expression and facial identity on gaze cueing is that the integration of information from different channels takes place at different time points in the processing stream: that is, information from brain areas that process facial identity (via FG) and signals originating from STS and IPS are integrated later in time, whereas information about facial expression (via the amygdala) is directly fed forward to the STS and therefore integrated earlier in order to exert a combined influence on gaze cueing (Hoffman & Haxby, 2000).

1.2.5. Acuity of gaze direction estimation

Utilizing gaze direction to learn about intentions and action goals requires us to identify where others are looking and to shift our own attentional focus to the corresponding object/location (for a review: Frischen et al., 2007). While it is easy to determine where an arrow cue is pointing to by following its head, it appears to be more difficult to interpret gaze direction, where directional information has to be derived from geometrical information displayed by the gazer’s eyes. However, previous studies have shown that humans are extremely precise in estimating gaze direction. This high level of precision has been found for dyadic interactions, where people have to discern direct and averted gaze of an interaction partner (e.g., Ando, 2002; Anstis, Mayhew & Morley, 1969; Cline, 1967; Gamer & Hecht, 2007; Gibson & Pick, 1963), as well as for triadic interactions, where observers have to judge gaze direction of an interaction partner towards objects in the periphery (Bock, Dicke & Thier, 2008; Gale & Monk, 2000; Schwaninger, Lobmaier & Fischer, 2005; Symons, Lee, Cedrone & Nishimura, 2004). The estimation errors ranged from 0.5° to 4° visual angle, depending on specific factors that have been varied experimentally: Besides factors directly associated with the physiological composition of the eyes, such as the iris-sclera ratio (Ando, 2002; Anstis et al., 1969; Langton, Watt & Bruce, 2000) or the luminance contrast between iris and sclera (Ando, 2002), there are also contextual factors such as head and body orientation (Gamer & Hecht, 2007; Todorović, 2006), number of visible eyes (Gamer & Hecht, 2007; Symons et al., 2004), looker-observer distance (Gamer & Hecht, 2007; Symons et al., 2004), or the presence of objects in the attended space (Lobmaier, Fischer & Schwaninger, 2006) that influence the accuracy of gaze direction estimation.

1.2.6. Spatial specificity of attentional orienting to spatial cues

Given the number of studies that have examined the precision with which gaze direction of others can be determined, it is quite surprising that the investigation of how precisely attention is shifted to gazed-at locations has so far been neglected in the literature. However, this is an important question in particular with regard to complex environments where it is difficult to determine the attentional focus of others when multiple objects are located close to each other that are of interest for the observed person.

Specific-location theories of attention, such as the *zoom-lens model* (Eriksen & St. James, 1986; Eriksen & Webb, 1989; Eriksen & Yeh, 1985) or the *gradient model* (Downing, 1988; Downing & Pinker, 1985; Henderson, 1991; LaBerge & Brown, 1989; Shulman, Wilson & Sheehy, 1985), conceive visual attention as a limited resource that can only be directed to restricted regions of the visual field. However, the zoom-lens and gradient models differ in their assumptions as to how attention levels off around cued locations. According to the zoom-lens model, attention initially covers a large area of the visual field (stimulus onset asynchrony, SOA < 100 ms), though with a low resolution. Following cue presentation, the attentional focus is gradually narrowed and the attended region becomes spatially more and more specific to the cued location. Behaviorally, this effect is indicated by fast responses to targets at the cued location and equally slowed responses to targets at uncued locations. According to the gradient model, attentional resolution diminishes in a graded fashion with increasing distance of the target from the cued location. By contrast, general-region theories of attention, such as the *meridian-boundary model* (Hughes & Zimba, 1985, 1987), propose that attention cannot be allocated to specific locations, but only to larger regions of the visual field, such as hemifields (Hughes & Zimba, 1985) or field quadrants (Hughes & Zimba, 1987).

While specific-location predictions have been shown to hold for attention shifts produced by the endogenous, voluntary attention system (e.g., Eriksen & Yeh, 1985; Shepherd & Müller, 1989, for long SOAs), there are only a few studies that have investigated how attention is allocated when the exogenous, reflexive attention system is involved (Henderson & Macquistan, 1993; Shepherd & Müller, 1989). Exogenous attention shifts, which cannot be readily controlled voluntarily, are attracted by abrupt visual onsets in the periphery (Jonides, 1981; Müller & Rabbit, 1989; Yantis & Jonides, 1984), as well as by centrally presented changes in gaze direction (Driver et al., 1999; Friesen & Kingstone, 1998; Hietanen, 1999; Hietanen & Leppänen, 2003; Langton & Bruce, 1999). However, although gaze direction has been shown to trigger quasi-automatic shifts of attention to the cued position, this mechanism is not purely reflexive, but at least partially controllable by the voluntary system (Driver et al., 1999; Downing et al., 2004; Friesen et al., 2004), leaving the question open of how attentional resources are deployed following this sort of cues.

So far, only one study has examined in detail the spatial distribution of attention following gaze cues (Vuilleumier, 2002). The main objective of this study was to examine whether gaze induced shifts of attention would still be observed in patients with right parietal damage and left visual field neglect whose ability to attend to contralesional space was impaired. In patient samples, the author could show that in the intact, ipsilesional hemifield, attention was selectively directed to the specifically cued location; in the contralesional hemifield, however, attentional orienting was specific only to the quadrant that was gaze-cued. In healthy control samples, gaze direction led to a shift of attention to the specifically cued position whereas the other position in the cued hemifield did not receive facilitation.

1.3. Rationale of the project

Research presented in this PhD-work systematically investigates the influence of context information on the spatial distribution of attention following gaze cues. With regard to context influences on social attention mechanisms, previous studies have shown that pre-assumptions about the gazing person modulate the size of gaze-induced cueing effects (Frith & Frith, 2006, for review). For instance, there is evidence that humans are more willing to engage attentional resources in social interactions when they believe that observed behavior is originating from a social agent (i.e., a face with hat) compared to non-social stimuli (i.e., a car with wheels, Ristic & Kingstone, 2005). Stronger gaze-cueing effects are also found, when participants believe that the observed agent can actually see an aspired target object compared to when it is not visible (Kawai, 2011; Teufel et al., 2009). Furthermore, it has been shown that information about emotional states of others (transmitted via facial expression) modulates the size of gaze-cueing effects in a way that fearful faces induce stronger gaze-cueing effects than happy or neutral faces (Peccinenda et al., 2008, c.f. Hietanen et al., 2003). These findings provide strong evidence that gaze following is not a pure bottom-up process that is executed on all kinds of eye-like stimuli, but can rather be top-down controlled by context information about the social content of the scene.

However, context information about the social relevance of changes in gaze direction does not necessarily modulate attentional orienting to gaze. Rather, as has been pointed out above, it has been shown that even socially important context information, such as the identity of the gazer does not have a modulatory effect on gaze cueing (Bayliss & Tipper, 2006b). Hence, it seems that spatial information gained from the eye region and context information about the environment and/or the gazing individual have a combined influence on gaze cueing only under certain circumstances, namely, when processing of gaze direction and context information is executed by the same neural network or at least in the same time window (Hoffman & Haxby, 2000).

Thus, based on the current literature, the exact mechanisms underlying top-down modulation of gaze-cueing effects cannot be specified, leading to the following empirical questions: 1) What types of social context information (e.g., visual, verbal, empirical) have a modulatory influence on social attention mechanisms (as measured in gaze-cueing effects)? 2) How is the modulation-effect of context information on gaze-cueing effects actually caused (i.e., via which mechanism)? In order to answer these questions, the aim of the research outlined in this PhD-thesis is to systematically examine the influence of different context variables on the size and spatial distribution of gaze-cueing effects. Furthermore, it is investigated how conflicting context information is integrated when attending to gaze cues. Given that attentional orienting to gaze cues is an early process (i.e., present already 150 ms after cue presentation) and sustained for only 1000 ms (Friesen & Kingstone, 1998), we hypothesized that context information that could potentially influence orienting to gaze direction must be pre-existing or processed very rapidly. We further assumed that the modulation of gaze-cueing effects through context information is actually exerted by a change in the allocation of attentional resources *within* the cued hemifield (i.e., facilitation for the exactly cued position and/or suppression for the uncued positions).

In order to investigate these hypotheses, three different sources of context information are used as potential candidates for modulating the spatial distribution of gaze-cueing effects: *visual* information in the observed scene (Study 1), *empirical* information that can be deduced based on experience with the observed scene (Study 2) and *verbal* information provided by instruction (Study 2-4). In particular, context information was either present in form of reference objects in the periphery (Study 1) or could refer to the reliability with which gaze behavior indicated target position (either provided by instruction or inferred from experience with the observed gaze behavior, Study 2). Context information could also be given as background knowledge about the likelihood that an observed interaction partner is capable of having mental states (Study 3-4).

Hypotheses

For investigating the influence of context information on gaze cueing, we used two different paradigms: With the first paradigm, we investigated whether context information does modulate the spatial allocation of attention (Study 1-2); with the second paradigm, we examined whether beliefs about the internal states of others influence the size of gaze-cueing effects (Study 3-4).

Study 1 and 2

Study 1 and 2 investigated whether the modulation of gaze cueing by context information was caused by changes in the spatial distribution of attentional resources to gazed-at compared to other positions in the cued hemifield. The idea was to measure the spatial distribution of attention as psychometric function of the distance between cued and target position. In order to investigate this question, targets could appear at different target positions while only one position was directly cued; the other positions were not directly cued but located in different angular distances from the cued position in the cued hemifield. In conditions where context information was provided, stronger gaze-cueing effects for the gazed-at position compared to other locations were expected (in line with the *specific-location* theories of attention), whereas for conditions without context information, we expected equally strong gaze-cueing effects for all target positions in the cued hemifield (in line with the *general-region* theories of attention).

In *Study 1*, context information was provided visually, that is, in form of reference objects in the periphery. We assumed that gaze direction would cause gaze-cueing effects specific to the exact gazed-at position only when it could refer to possible reference objects in the periphery; in contrast, equally strong gaze-cueing effects for the whole cued hemifield were expected when no reference objects were presented to which gaze direction could be linked. In *Study 2*, context was provided in form of information about the reliability of observed gaze behavior (i.e., the likelihood with which gaze direction indicated target position). We hypothesized that attention would be directed to the exact gazed-at position only when gaze cues indicated target position with a high likelihood compared to with a low likelihood. Information about the reliability of gaze cues was either given by instruction or could be deduced from experience with the observed gaze behavior.

Study 3 and 4

The aim of Study 3 and 4 was to examine whether context information about the likelihood that an observed agent has a mind modulates the readiness to engage in social interactions with that agent (i.e., as is reflected in the size of the gaze-cueing effects). In particular, we used a human and a robot stimulus and manipulated participants' beliefs about the intentionality of the displayed behavior by instructions. We expected gaze-cueing effects to be larger when participants believed to observe human-like behavior compared to when they believed that they were observing mechanistic, pre-programmed behavior. Most importantly, we expected this effect to be independent of the physical appearance of the gazer, but only dependent on the degree to which intentionality could be attributed to the displayed behavior.

Behavioral effects of the instruction manipulation were investigated in *Study 3*. However, behavioral measures alone are not suited to address the issue of whether or not instruction manipulation affected attentional mechanisms (compared to other, later mechanisms related to response preparation), as performance data represent only the end effects of the processing chain. *Study 4* therefore combined behavioral measures with the EEG/ERP-methodology in order to identify the neural correlates that are underlying the top-down influences on gaze-cueing effects. We expected to find a modulation of early sensory gain control mechanisms (reflected in the ERP-components P1 and/or N1), if instruction manipulation had an effect on attentional orienting mechanisms rather than on later processes related to response preparation and execution.

1.4. References (Abstract + Introduction)

Adams, R.B., & Kleck, R.E. (2005). Effects of Direct and Averted Gaze on the Perception of Facially Communicated Emotion. *Emotion, 1,* 3-11.

Adamson, L.B., & Frick, J.E. (2003). The still-face: a history of a shared experimental paradigm. *Infancy, 4,* 451-473.

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences, 3(12),* 469-479.

Adolphs, R. (2003). Cognitive neuroscience of human social behavior. *Nature Reviews Neuroscience, 4*, 165–178.

Aggleton, J.P. (1993). The contribution of the amygdala to normal and abnormal emotional states. *Trends in Neurosciences, 16,* 328-333.

Aggleton, J.P., Burton, M.J., & Passingham, R.E. (1980). Cortical and subcortical afferents to the amygdala of the rhesus monkey (macaca mulatta). *Brain Research, 190(1980),* 347-368.

Ando, S. (2002). Luminance-induced shift in the apparent direction of gaze. *Perception, 31,* 657-674.

Anstis, S.M., Mayhew, J.W., & Morley, T. (1969). The perception of where a face or television 'portrait' is looking. *American Journal of Psychology, 82,* 474–489.

Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind.* Boston: MIT Press/ Bradford Books.

Baldwin, D.A., & Moses, L.J. (2001). Links between social understanding and early word learning: challenges to current accounts. *Social Development, 10,* 309-329.

Bayliss, A.P., Paul, M.A., Cannon, P.R., & Tipper, S.P. (2006a). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review, 6,* 1061-1066.

Bayliss, A.P., & Tipper, S.P. (2006b). Predictive gaze cues and personality judgments: should eyes trust you?. *Psychological Science, 17(6),* 514-520.

Becchio, C., Bertone, C., & Castiello, U. (2008). How the gaze of others influences object processing. *Trends in Cognitive Sciences, 12(7),* 254-258.

Berthoz, S., Armony, J.L., Blair, R.J.R., & Dolan, R.J. (2002). An fMRI study of intentional and unintentional violations of social norms. *Brain, 125(8),* 1696–1708.

Blakemore, S.-J., Winston, J., & Frith, U. (2004). Social cognitive neuroscience: where are we heading?. *Trends in Cognitive Sciences, 8(5),* 216-222.

Bock, S.W., Dicke, P., & Thier, P. (2008). How precise is gaze following in humans?, *Vision Research, 48(7),* 946-957.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in simple response tasks. *Acta Psychologica, 106,* 3-22.

Breiter, H.C., Etcoff, N.L., Whalen, P.J., Kennedy, W.A., Rauch, S.L. et al. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron, 17(5),* 875–887.

Brooks, R., & Meltzoff, A.N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology, 38(6),* 958-966.

Brooks, R., & Meltzoff, A.N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8,* 535-543.

Brothers, L. (1990). The social brain: s project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience, 1*, 27-51.

Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L. et al. (2001). Action observation activates premotor and parietal areas in somatotopic manner: an fMRI study. *European Journal of Neuroscience, 13*, 400-404.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, Serial No. 255, 63(4),* 1-143.

Castiello, U. (2003). Understanding other people's actions: Intention and attention. *Journal of Experimental Psychology: Human Perception and Performance, 29(2),* 416-430.

Cline, M.G. (1967). The perception of where a person is looking. *American Journal of Psychology, 80,* 41-50.

Corbetta, M., Miezin, F.M., Shulman, G.L., & Petersen, S.E. (1993). A PET study of visuospatial attention. *Journal of Neuroscience, 13,* 1202-1226.

Craighero, L., Bello, A., Fadiga, L. & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia, 40*, 492-502.

Deaner, R.O., Shepherd, S.V., & Platt, M.L. (2007). Familiarity accentuates gaze cuing in women but not in men. *Biology Letters, 3*, 64-67.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research, 91,* 176-180.

Downing, C.J. (1988). Expectancy and visuo-spatial attention. Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception & Performance, 14*, 188-202.

Downing, C.J., & Pinker, S. (1985). The spatial structure of visual attention. In M.I. Posner & O.S.M. Marin (Eds.), *Attention and performance XI*, pp.171-188. Hillsdale, NJ: Erlbaum.

Downing P., Dodds, C.M., & Bray, D. (2004). Why does the gaze of others direct visual attention?. *Visual Cognition, 11*, 71-79.

Dunbar, R. (1998). The social brain hypothesis. *Evolutionary Anthropology*, 178-190.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6,* 509-540.

Eimer, M., Holmes, A., & McGlone, F.P. (2003). The role of spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions. *Cognitive, Affective, & Behavioral Neuroscience, 3(2),* 97–110.

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

Eriksen, C.W., & St.James, J.D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40,* 225-240.

Eriksen, C.W, & Webb, J.M. (1989). Shifting of attentional focus within and about a visual display. *Perception & Psychophysics, 45,* 175-183.

Eriksen, C. W., & Yeh, Y.-Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception & Performance, 11,* 583-597.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology, 73,* 2608-2611.

Farroni, T., Csibra, G., Simion, F., & Johnson, M.H. (2002). Eye contact detection at birth. *PNAS, 99,* 9602-9605.

Farrow, T.F., Zheng, Y., Wilkinson, I.D., Spence, S.A., Deakin, J.F. et al. (2001). Investigating the functional anatomy of empathy and forgiveness. *Neuroreport, 12(11),* 2433–2438.

Flom, R., & Pick, A.D. (2005). Experimenter affective expression and gaze following in 7-month-olds. *Infancy, 7,* 207-218.

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

Friesen, C.K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception & Performance, 30,* 319-329.

Frischen, A., Bayliss, A.P., & Tipper, S.P. (2007). Gaze Cuing of Attention: Visual Attention, Social Cognition, and Individual Differences. *Psychological Bulletin, 133(4)*, 694-724.

Frischen, A., Smilek, D., Eastwood, J.D., & Tipper, S.D. (2007). Inhibition of return in response to gaze cues: The roles of time course and fixation cue. *Visual Cognition, 15(8),* 881-895.

Frith, U., & Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society. Series B: Biological Sciences, 358(1431),* 459-473.

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

Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119,* 593-609.

Gale, C., & Monk, A.F. (2000). Where am I looking? The accuracy of video-mediated gaze awareness. *Perception & Psychophysics, 62,* 586-595.

Gamer, M., & Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. *Journal of Experimental Psychology: Human Perception & Performance, 33(3),* 705-715.

Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *Neuroimage, 35,* 1674-1684.

George, N., Driver, J., & Raymond, J.D. (2001). Seen Gaze-Direction Modulates Fusiform Activity and Its Coupling with Other Brain Areas during Face Processing. *Neuroimage, 13,* 1102-1112.

Gibson, J.J., & Pick, A. (1963). Perception of another person's looking behavior. *American Journal of Psychology, 76,* 386-394.

Graham, R, & LaBar, K.S. (2007). Garner Interference Reveals Dependencies Between Emotional Expression and Gaze in Face Perception. *Emotion, 2,* 296-313.

Graham, R., Friesen, C.K., Fichtenholtz, H.M., & LaBar, K.S. (2010). Modulation of reflexive orienting to gaze direction by facial expressions. *Visual Cognition, 18(3)*, 331-368.

Grezes, J., Frith, C.D., & Passingham, R.E. (2004a). Brain mechanisms for inferring deceit in the actions of others. *Journal of Neuroscience, 24(24)*, 5500–5505.

Grezes, J., Frith, C.D., & Passingham, R.E. (2004b). Inferring false beliefs from the actions of oneself and the others: an fMRI study. *Neuroimage, 21(2)*, 744–750.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-233.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human Neural Systems for Face Recognition and Social Communication. *Biological Psychiatry, 51*, 59-67.

Henderson, J.M. (1991). Stimulus discrimination following covert attentional orienting to an exogenous cue. *Journal of Experimental Psychology: Human Perception & Performance, 17*, 91-106.

Henderson, J.M., & Macquistan, A.D. (1993). The spatial distribution of attention following an exogenous cue. *Perception & Psychophysics, 53(2)*, 221-230.

Hietanen, J. K. (1999). Does your gaze direction and head orientation shift my visual attention? *Neuro Report, 10*, 3443-3447.

Hietanen, J.K., & Leppänen, J.M. (2003). Does facial expression affect attention orienting by gaze direction cues? *Journal of Experimental Psychology: Human Perception and Performance, 29(6)*, 1228-1243.

Hietanen, J.K., Nummenmaa, L., Nyman, M.J., Parkkola, R., & Hämäläinen, H. (2006). Automatic attention orienting by social and symbolic cues activates different neural networks: an fMRI study. *Neuroimage, 33(1)*, 406-13.

Hietanen, J.K., Leppänen, J.M., Nummenmaa, L., & Astikainen, P. (2008). Visuospatial attention shifts by gaze and arrow cues: an ERP study. *Brain Research, 1215*, 123-136.

Hood, B.M., Willen, J.D., & Driver, J. (1998). Adult’s eyes trigger shifts of visual attention in human infants. *Psychological Science, 9*, 53-56.

Hoffman, E.A., & Haxby, J.V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience, 3(1),* 80-84.

Hoffman, J.E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57(6)*, 787-795.

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

Hughes, H.C., & Zimba, L.D. (1985). Spatial maps of directed visual attention. *Journal of Experimental Psychology: Human Perception & Performance, 11*, 409-430.

Hughes, H.C., & Zimba, L.D. (1987). Natural boundaries for spread of directed visual attention. *Neuropsychologia*, *2,* 5-18.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C. et al., (2005). Grasping the intentions of others with one’s own mirror neuron system. *PLoS Biology*, *3,* e79.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C.L., Jasnow, M.D. et al., (2001). Rhythms of dialogue in infancy: coordinated timing and social development. *Monograph of the Society for Research in Child Development, 66(2)*. Washington DC: Wiley-Blackwell.

Jellema, T., Baker, C. I., Wicker, B., & Perrett, D.I. (2000). Neural Representation for the Perception of the Intentionality of Actions. *Brain & Cognition*, *44*, 280-302.

Johnson, M.H., & Morton, J. (1991). *Biology and Cognitive Development: The Case of Face Recognition*. Washington DC: Wiley-Blackwell.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye movement. In J. Long & A. Baddeley (Eds.), *Attention and performance IX*, pp. 187-203. Hillsdale, NJ: Erlbaum.

Kawai, N. (2011). Attentional shift by eye gaze requires joint attention: Eye gaze cues are unique to shift attention. *Japanese Psychological Research, 53,* 292–301.

Kilner, J.M., Paulignan, Y., & Blakemore, S.-J. (2003). An interference effect of observed biological movement on action. *Current Biology, 13,* 522-525.

Kobayashi, H., & Kohshima, S. (1997). Morphological uniqueness of human eyes and its adaptive meaning. *Nature, 387(767),* 8.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96,* 101-124.

Langton, S.R.H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition, 6,* 541-567.

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

Lavelli, M., & Fogel, A. (2005). Developmental changes in mother-infant face-to-face communication: birth to 3months. *Developmental Psychology, 38,* 288-305.

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 in-group voters. *Plos One, 6(9),* 1-7.

Lobmaier, J.S., Fischer, M.H., & Schwaninger, A. (2006). Objects capture perceived gaze direction. *Experimental Psychology, 53,* 117-122.

Moore, C., & Corkum, V. (1995). The origins of joint visual attention. In C. Moore & P. Dunham (Eds.) *Joint attention: Its origins and role in development,* pp. 61-85. Hillsdale, NJ: Erlbaum.

Müller, H. J., & Rabbitt, P. M. A. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 315-330.

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Réserbat-Plantey, D. (1999). Expectancies for social contingency in 2-months-olds. *Developmental Science, 2*, 164-173.

Nelson, C.A. (2001). The development and neural bases of face recognition. *Infant and Child Development*, *10*, 3-18.

Nobre, A.C., Sebestyen, G.N., Gitelman, D.R., Mesulam, M.-M., Frackowiak, R.S.J. et al. (1997). Functional localization of the system for visuospatial attention using positron emission tomography. *Brain*, *120*, 515-533.

Nummenmaa, L., & Calder, A. (2009). Neural mechanisms in social attention. *Trends in Cognitive Sciences 13 (3):*135-143.

Nummenmaa, L., & Hietanen, J.K. (2009). How attentional systems process conflicting cues. The superiority of social over symbolic orienting revisited. *Journal of Experimental Psychology: Human Perception & Performance*, *35(6)*, 1738-1754.

Oberman, L.M., McCleery, J.P., Ramachandran, V.S., & Pineda, J.A. (2007). EEG evidence for mirror neuron activity during the observation of human and robot actions: Towards an analysis of the human qualities of interactive robots. *Neurocomputing*, *70*, 2194-2203.

Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: a study of first birthday home video tapes. *Journal of Autism and Developmental Disorders*, *24*, 247-257.

Oztop, E., Franklin, D.W., Chaminade, T., & Cheng, G. (2005). Human-Humanoid Interaction: Is a humanoid robot perceived as a human? *International Journal of Humanoid Robotics, 2(4)*, 537-559.

Pecchinenda, A., Pes, M., Ferlazzo, F., & Zoccolotti, P. (2008). The combined effect of gaze direction and facial expression on cueing spatial attention. *Emotion*, *8(5)*, 628-634.

Pelphrey, K.A., Singerman, J.D., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia*, *41*, 156-170.

Perrett, D.I., & Mistlin, A.J. (1991). Perception of facial characteristics by monkeys. In: Stebbins, W.C., Berkley, M.A. (eds.). *Comparative Perception*, pp. 187-215, New York: John Wiley.

Perrett, D.I., Hietanen, J.K., Oram, M.W., & Benson, P.J. (1992). Organization and function of cells responsive to faces in the temporal cortex. *Phil.Trans.Roy.Soc.Lond.:Biol.Sci., 335*, 23-30.

Pierno, A.C., Becchio, C., Wall, M.B., Smith, A.T., Turella, L., et al. (2006). When gaze turns into grasp. *Journal of Cognitive Neuroscience, 18(12)*, 2130-2137.

Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3-25.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In Pick, H. L., & Saltzman, I. J. (Eds.) *Modes of Perceiving and Processing Information*, pp. 137-157. Hillsdale, N. J.: Erlbaum.

Quadflieg, S., Mason, M.F., & Macrae, C.N. (2004). The owl and the pussycat: Gaze cues and visuospatial orienting. *Psychonomic Bulletin & Review*, *11*, 826-831.

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

Ristic, J., & Kingstone, A. (2005). Taking control of reflexive social attention, *Cognition 94(3):* B55- B65.

Ristic, J., & Kingstone, A. (2006). Attention to arrows: Pointing to a new direction. *The Quarterly Journal of Experimental Psychology*, *59(11)*, 1921-1930.

Rizzolatti, G., Riggio, L., Dascola, I., & Umilta, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia, 25,* 31-40.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., et al. (1996). Localization of grasp representations in humans in PET: Observation versus execution. *Experimental Brain Research, 111*, 246-252.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3(2)*, 131-141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2,* 661-670.

Rizzolatti, G. & Craighero, L. (2002). The mirror-neuron system. *Annual Reviews of Neuroscience, 27,* 167-192.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia, 43(10),* 1391–1399.

Schwaninger, A., Lobmaier, J.S., & Fischer, M.H. (2005). The inversion effect of gaze perception reflects processing of component information. *Experimental Brain Research, 167,* 49-55.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own?. *Cognition, 88,* B11-B21.

Sheliga, B.M., Riggio, L., & Rizzolatti, G. (1994). Orienting of attention and eye movements. *Experimental Brain Research, 98(3),* 507-522.

Shepherd, M., Findlay, J.M., & Hockey, R.J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 38(3),* 475-491.

Shepherd, M., & Müller, H.J. (1989). Movement versus focusing of visual attention. *Perception & Psychophysics, 46(2),* 146-154.

Shulman, G.L., Wilson, J., & Sheehy, J.B. (1985). Spatial determinants of the distribution of attention. *Perception & Psychophysics, 37,* 59-65.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J. et al. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science, 303,* 1157–1162.

Striano, T., & Bertin, E. (2005a). Coordinated affect with mothers and strangers: a longitudinal study of joint attention between 5 and 9 months of age. *Cognition & Emotion, 19,* 781-790.

Striano, T., & Bertin, E. (2005b). Relation among joint attention skills in 5- to 10-month-old infants. *British Journal of Developmental Psychology, 23,* 1-11.

Striano, T. & Reid, V.M. (2006). Social cognition in the first year. *Trends in Cognitive Sciences, 10(10)*, 471-476.

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, *400*, 869–873.

Symons, L.A., Lee, K., Cedrone, C.C., & Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. *Journal of General Psychology*, *131*, 451-469.

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

Teufel, C., Alexis, D. M., Clayton, N. S., & Davis, G.. (2010a). Mental attribution drives rapid, reflexive gaze following in human adults. *Attention, Perception, & Psychophysics*, *72*, 695–705.

Teufel, C., Fletcher, P. C., & Davis, G. (2010b). Seeing other minds: attributed mental states influence perception. *Trends in Cognitive Sciences*, *14*, 376–382.

Thomas, K.M., Drevets, W.C., Whalen, P.J., Eccard, C.H., Dahl, R.E., et al. (2001). Amygdala response to facial expressions in children and adults. *Biological Psychiatry*, *49*, 301-316.

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

Todorovic, D. (2006). Geometrical basis of perception of gaze direction, *Vision Research*, *46*, 3549-3562.

Tomasello, M. Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behavioral Brain Sciences*, *28*, 675-691.

Vecera, S.P., & Rizzo, M. (2006). Eye gaze does not produce reflexive shifts of attention: Evidence from frontal-lobe damage. *Neuropsychologia*, *44*, 150–159.

Vuilleumier, P., Armony, J.L., Driver, J., & Dolan, R.J.. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*, 30, 829–841

Vuilleumier, P. (2002). Perceived gaze direction in faces and spatial attention: A study in patients with parietal damage and unilateral neglect. *Neuropsychologia*, *40*, 1013-1026.

Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B. et al., (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, *18*, 411–418.

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. doi:10.1371/ journal.pone. 0045391.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception & Performance*, *10*, 601-621.

STUDIES ON INDIVIDUAL RESEARCH QUESTIONS

The influence of social context information on the *magnitude* and *spatial distribution* of gaze-cueing effects

2.1. Study 1

The importance of context information for the spatial specificity of gaze cueing

Eva Wiese[1], Jan Zwickel[1] & Hermann J. Müller[1,2]

[1]Ludwig-Maximilian University of Munich, Germany

[2]Birkbeck College, University of London, UK

Author contributions: EW, HJM conceived and designed research. EW performed research. EW analyzed data. EW, JZ, HJM wrote the paper.

Wiese, E., Zwickel, J., & Müller, H.J. (in press). The importance of context information for the spatial specificity of gaze cueing. *Attention, Perception & Psychophysics*. [Epub ahead of print], doi: 10.3758/s13414-013-0444-y

2.1.1. Abstract

In three experiments, we investigated the spatial allocation of attention in response to central gaze cues. In particular, we examined whether the allocation of attentional resources is influenced by context information, that is, the presence or absence of reference objects (i.e., placeholders) in the periphery. On each trial, gaze cues were followed by a target stimulus to which participants had to respond by key press or by performing a target-directed saccade. Targets were presented either in an empty visual field (Experiments 1 and 2) or in previewed location placeholders (Experiment 3) and appeared at one of 18 (Experiment 1) or, respectively, six (Experiments 2 and 3) possible positions. The spatial distribution of attention was determined by comparing reaction times (RTs) as a function of the distance between the cued and the target position. Gaze cueing was not specific to the exact cued position, but instead generalized equally to all positions in the cued hemifield, when no context information was provided. However, gaze direction induced an additional facilitation effect specific to the exact gazed-at position when reference objects were presented. We conclude that the presence of possible objects in the periphery to which gaze cues could refer is a prerequisite for attention shifts specific to the gazed-at position.

Keywords

Gaze cueing, top-down control, eye movements, spatial allocation of attention, psychophysics

2.1.2. Introduction

Interacting with other people in a crowded environment is a difficult task that requires complex cognitive skills. Even simple interactions, like shaking hands or passing an object from one person to another, are challenging in that two individuals must coordinate their actions. A prerequisite for successful interaction is the anticipation of action goals (Sebanz, Bekkering, & Knoblich, 2006) – which can be inferred from the other’s focus of attention, as locations or objects attended by the other are likely targets for upcoming actions. For identifying where or what others are paying attention to, we rely on directional information provided by social cues, including gestures, body and head orientation, and, in particular, gaze direction (Emery, 2000; Langton, Watt, & Bruce, 2000; Kingstone, Smilek, Ristic, Friesen & Eastwood, 2003).

Utilizing gaze direction to infer action goals requires us to identify where the other is looking at and to shift our own attentional focus to the corresponding location/object (for a review, see Frischen, Bayliss, & Tipper, 2007). Previous studies have shown that humans are quite precise in estimating gaze direction, with estimation errors ranging from 0.5° to 4° of visual angle. This high level of precision has been found for *dyadic* interactions, where observers had to discern direct and averted gaze of an interaction partner (Ando, 2002; Anstis, Mayhew & Morley, 1969; Cline, 1967; Gamer & Hecht, 2007; Gibson & Pick, 1963), as well as for *triadic* interactions, where observers had to judge to which object in space an interaction partner’s gaze was directed (Bock, Dicke & Their, 2008; Gale & Monk, 2000; Schwaninger, Lobmaier & Fischer, 2005; Symons, Lee, Cedrone, & Nishimura, 2004; Wiese, Kohlbecher, & Müller, 2012). Besides factors that are directly associated with the physical appearance of the eyes, such as the iris-sclera ratio (Ando, 2002; Anstis et al., 1969; Langton, Watt & Bruce, 2000) or the luminance contrast between the iris and the sclera (Ando, 2002), various context factors such as body and head orientation (e.g., Gamer & Hecht, 2007; Todorović, 2006), number of visible eyes (Gamer & Hecht, 2007; Symons et al., 2004), looker-observer distance (e.g., Gamer & Hecht, 2007; Symons et al., 2004), and presence of objects in the attended space (Lobmaier, Fischer, & Schwaninger, 2006) have been shown to influence the accuracy of gaze direction estimation.

Interestingly, when watching another person gazing at a distinct location in space, the attentional focus of the observer is quasi-reflexively shifted to the gazed-at location (Driver et al., 1999; Friesen & Kingstone, 1998; Hietanen, 1999; Hietanen & Leppänen, 2003; Langton & Bruce, 1999; cf. Downing, Dodds, & Bray, 2004; Driver, Davis, Ricciardelli, Kidd, Maxwell et al., 1999; Friesen, Ristic, &

Kingstone, 2004), so as to permit flexible responding to upcoming actions of the gazer. Gaze-induced shifts of attention have traditionally been investigated by using a cueing paradigm (Friesen & Kingstone, 1998) in which a schematic face is presented centrally on the screen, gazing either straight ahead or to the left or the right side of the screen. Targets appearing in the cued direction are detected, located, and discriminated faster than targets in the other, non-cued direction (for a review: Frischen, Bayliss & Tipper, 2007). Cueing effects emerge already 150 ms after cue onset, decay relatively quickly (within 1000 ms after cue presentation), and are obtained in cases when gaze direction is nonpredictive (Friesen & Kingstone, 1998) or even counterpredictive with regard to upcoming target positions (Friesen, Ristic & Kingstone, 2004). Moreover, gaze cueing effects can be induced by realistic faces (Hietanen & Leppänen, 2003; Langton & Bruce, 1999), as well as by various kinds of non-face objects that convey eye-like information (Quadflieg, Mason & Macrae, 2004).

Given the number of studies that have examined how precise observers are in determining the gaze direction of others, it is surprising that the question of how accurately they deploy their own attention to the gazed-at locations has been neglected in the literature. Arguably though, this is an important question, in particular with regard to complex environments in which multiple objects of potential interest to the observed person may be located close to each other, making it difficult for the observer to determine the other’s attentional focus precisely. For this situation, specific-location theories of spatial attention, such as the *zoom-lens model* (e.g., Eriksen & Yeh, 1985) or the *gradient model* (e.g., Downing & Pinker, 1985; Henderson, 1991; LaBerge & Brown, 1989; Shulman, Wilson, & Sheehy, 1985) would predict that attention is directed to the exact gaze-cued location, selectively enhancing attentional processing for a narrow region of the visual field. In contrast, general-region theories, such as the *meridian-boundary model* (Hughes & Zimba, 1985, 1987), would predict that gaze-cueing effects generalize to the whole cued hemifield/quadrant, as attentional allocation is constrained by neural machinery that can demarcate only broader regions of the visual field.

To our knowledge, only one, neuropsychological study (Vuilleumier, 2002, Experiments 4-5) has looked at the spatial specificity of attentional orienting in response to gaze cues (and this more as a side effect). The main objective of this study was to examine whether gaze-induced shifts of attention would still be observed in patients with right parietal damage and left visual field neglect, whose ability to attend to contralesional space was impaired. Targets appeared at one of four predefined positions inside a placeholder, while a centrally presented schematic face looked either towards the target location, another location on the same side, a location on the opposite side, or straight ahead. Gaze direction was neither relevant for the task nor predictive of the target location. In patients, Vuilleumier (2002) could show that in the intact, ipsilesional hemifield, attention was selectively directed to the exact cued location; in the contralesional hemifield, by contrast, attentional orienting was specific only to the cued quadrant. In healthy controls, gaze direction always led to a shift of attention to the exact cued position, whereas the other position in the cued hemifield did not receive any facilitation.

Although these findings provide an important first attempt at assessing the spatial specificity of gaze-cueing effects, it is difficult to generalize them for several reasons. First, while Vuilleumier (2002) showed that gaze direction can produce specific cueing effects under certain conditions, his findings do not specify how cueing effects change as a function of the distance of the target from the cued position. Second, the study is not conclusive as to how the modulation of spatial-attentional allocation by context information actually works, in particular: were specific cueing effects triggered by the gaze cues per se, or rather by the combination of directional information provided by the cue and the presence of context information in the periphery (i.e., predefined placeholders)?

These are important issues, given that context information may impact the spatial specificity of gaze cueing via two different mechanisms: On the one hand, presenting placeholders might cause specific gaze-cueing effects by restricting processing to only a few regions of the visual field where task-relevant events may happen. In this context, cueing effects caused by non-social central cues (e.g., arrows) have been shown to be widely distributed when targets were presented in an uncluttered visual field, in which case transitions in performance were evident only for the horizontal and vertical field meridians (Hughes & Zimba, 1985, 1987). Interestingly, however, when targets were presented in predefined placeholders, cueing effects were specific to the exact cued position (e.g., Eriksen & Yeh, 1985; Henderson & Macquistan, 1993; Shepherd & Müller, 1989). On the other hand, humans might expect that changes in gaze direction are related to the appearance of objects or events at the gazed-at location and would, accordingly, shift their attentional focus only under these conditions to the exact gazed-at position. Evidence for this assumption comes from developmental studies which have shown that sharing attention between two partners is only facilitated when objects are placed in the visual field, thereby providing information about what the other's gaze is referring to (Striano & Stahl, 2005).

The present study, with adult participants, was designed to investigate whether the allocation of spatial attention in response to social gaze cues is modulated by visual context information, in particular, whether spatially specific cueing effects are inherently induced by gaze cues per se, rather than being due to the interplay of gaze direction *and* context information (i.e., placeholder). Spatial-attentional allocation was determined by comparing RTs as a function of the distance between cued and target position. Based on the findings for non-social cues (Eriksen & Yeh, 1985; Hughes & Zimba, 1985, 1987; Shepherd & Müller, 1989), we hypothesized that the cueing effects for social gaze cues would be spatially specific only when objects are presented as reference frames in the periphery with which the observer's, as well as the gazer's, attentional focus could be aligned. If gaze cueing was spatially specific, facilitation would be the strongest for the exact cued position, whereas other positions in the cued hemifield would exhibit only weaker facilitation, if any. By contrast, if gaze-cueing effects were not position-specific but hemifield-specific, all positions in the cued hemifield would show equal facilitation.

2.1.3. Experiments

Experiment 1

The main goal of Experiment 1 was to determine how spatial attention is allocated in response to centrally presented gaze cues when no context information is provided in the periphery, that is: would attention, in this situation, be allocated to the exact cued location or to a more global region of the field? An answer to this question would be provided by analyzing the cueing effects as a function of the distance of the target from the cued location.

A secondary goal was to examine whether the allocation of attention following gaze cues would be modulated by the type of task to be performed, as previous studies have shown the size of the attended region to vary considerably depending on perceptual task demands (Downing, 1988; LaBerge, 1983; Müller & Findlay, 1987). To this end, two different types of tasks were introduced in Experiment 1, which differed in their demands on perceptual resolution: *target localization* (Experiment 1a), which makes low demands, and *target discrimination* (Experiment 1b), which requires greater resolution. According to Downing (1988) tasks that are more difficult to solve require narrower attentional foci, whereas less demanding tasks may still be performed effectively with a broader distribution of attention. Thus, for roughly localizing a salient target in the visual field (left vs. right hemifield decision) it may not be necessary to precisely focus attention on the exact target position. Consequently, spatially specific cueing effects might be best, or only, demonstrable in tasks in which higher perceptual resolution is required for successful performance.

Another reason why we used a target discrimination, in addition to the – ubiquitously used (e.g., Friesen & Kingstone 1998) – left-right target localization task in Experiment 1 is that localization performance is prone to spatial stimulus-response (S-R) compatibility effects. With left-right target localization (as required in Experiment 1a), such effects may be induced, in the valid-cue condition, by the central gaze cue pointing to the side on which the target-congruent manual response has to be given.

That is, the cue might directly facilitate (or pre-activate) the response, rather than solely attentionally enhancing the processing of the target. In other words, RT facilitation might not solely arise from gaze-induced orienting of attention, but also from Simon-like modulations (Simon, 1969) acting at the stages of response preparation and execution (Nummenmaa & Hietanen, 2009). Given that the cue might induce a global response bias towards the cued side this bias might obscure any spatially specific *attentional* effect deriving from the gaze cueing. Given this, on this account too, the discrimination task used in Experiment 1b (in which the cue cannot logically bias the response) would hold a greater potential to reveal the *true* distribution of attention in response to gaze cues.

The questions outlined above were investigated by using a modified version of the gaze-cueing paradigm (Friesen & Kingstone, 1998). Instead of one target position on each side, we used nine semi-circularly arranged positions in each hemifield, with gaze cues always indicating only either the left or the right position on the horizontal meridian (=*central position*, Figure 1). Our design permitted recording of RTs to targets positioned at different distances from the exact cued location, providing a measure of gaze-cueing effects as a function of the distance between cued position and target position.

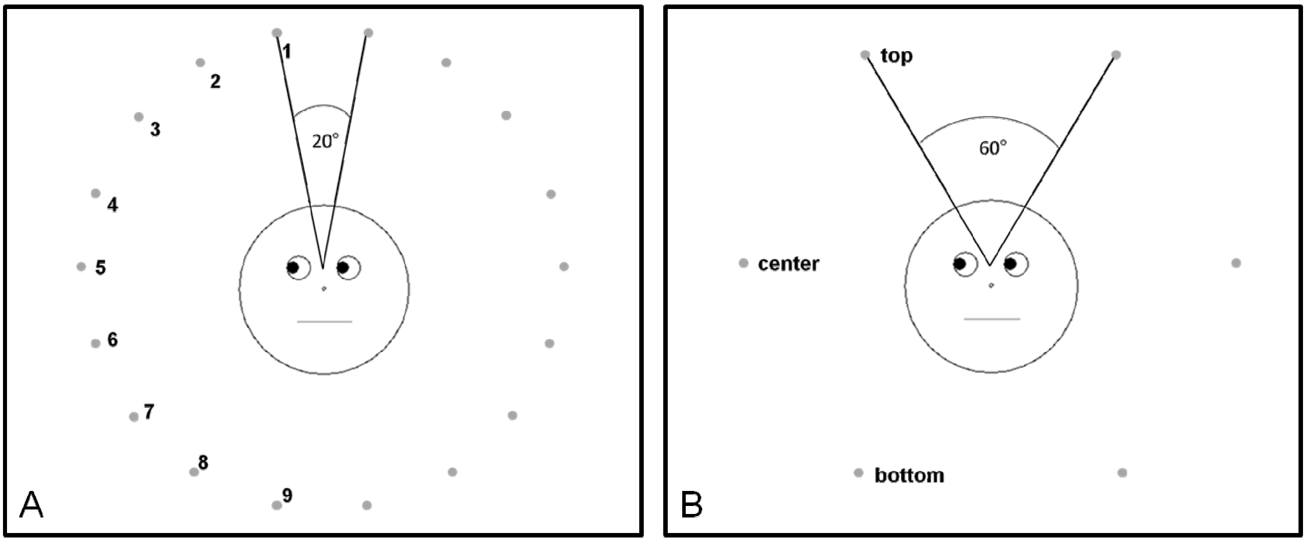

Figure 1. Possible target positions in (A) Experiment 1, and in (B) Experiments 2 and 3

Method

Participants

36 volunteers participated in Experiment 1, either for course credit or payment (8 €/hour): Eighteen participants were tested on the localization task (*Experiment 1a*; 13 women; age: *M*= 25.17, *SD*= 3.65, range: 20-32 years), the other eighteen on the discrimination task (*Experiment 1b*; 11 women; age: *M*= 22.94, *SD*= 3.21, range: 19-32 years). All participants were right-handed and reported normal or corrected-to-normal visual acuity. Testing time was about 1 h for the localization task and about 2 h for the discrimination task (split into two testing sessions).

Apparatus

The experiment was controlled by a Dell Precision 390 computer, with stimuli presented on a 17-inch Graphics Series CRT G90fB monitor, with the refresh rate set at 85 Hz. RT measures were based on keyboard responses. Participants viewed the monitor from a distance of 57 cm, with their head position centered with respect to the screen and keyboard. The experiment was set up using *Experiment Builder* (SR Research Ltd., Ontario, Canada).

Stimuli

Schematic faces, constructed following Friesen and Kingstone (1998), were used as gaze-cue providers. Faces were drawn in black against a white background. The round face covered an area of 6.8° of visual angle and contained two circles representing the eyes, a smaller circle symbolizing the nose, and a straight line representing the mouth. The eyes subtended 1.0°, were located 1.0° from the central vertical axis and centered on the central horizontal axis of the display. The nose subtended 0.2°, was located 0.9° below the eyes and served as fixation point. The mouth was 2.2° in length and centered 1.3° below the nose. Black filled circles appeared within the eyes representing the pupils. Pupils subtended 0.5°, were centered vertically and horizontally in the eyes (straight gaze) or shifted either leftwards (gaze to the left) or rightwards (gaze to the right) until they touched the outline eye circles.

In the localization task (Experiment 1a), the target stimulus was a gray dot of a diameter of 0.5°. In the discrimination task (Experiment 1b), targets were black capital letters, F or T, measuring 0.8° wide and 1.3° high. Targets could appear at eighteen positions (not marked by placeholders) that were equally distributed around an imaginary circle of a radius of 6.0° centered at the fixation point (Figure 1A). The resulting radial distance between adjacent target positions was 20°. Gaze direction was manipulated orthogonal to target position, that is: in one third of the trials, gaze was directed to the side on which the target appeared (*cued hemifield*), and in another third to the other side (*uncued hemifield*); in the remaining third of the trials, the face was gazing straight ahead. Importantly, on trials with changes in gaze direction, only the central position on each side was directly gazed-at. Consequently, targets in the cued hemifield could appear at radial distances of 0°, 20°, 40°, 60°, or 80° from the specifically cued position.

Design

Experiment 1a (localization task) consisted of 884 trials; and Experiment 1b (discrimination task) of 1768 trials split into two test sessions of 884 trials each. Experimental sessions started with a block of 20 practice trials preceding eight experimental blocks of 108 trials each. In the localization task, participants were asked to decide whether a target dot was shown on the left or the right side of the screen, and to press “D” with the left index finger for left and “K” with the right index finger for right. In the discrimination task, participants were asked to determine whether an F or a T was presented in the target display. To minimize any target-response associations, half of the participants responded to F by pressing “D” with the left index finger and to T by pressing “K” with the right index finger, and vice versa for the other half of the participants.

Gaze direction (straight, left, right), target side (left, right), target position (1 to 9), and, in Experiment 1b, target identity (F, T) levels appeared pseudo-randomly and with equal frequency within each block. In the localization task, SOA (300 ms, 600 ms) was randomized throughout the experiment; in the discrimination task, SOA was blocked and counterbalanced across the two test sessions (i.e., half the participants started with the short and half with the long SOA).

Cue validity was defined in terms of the combination of gaze direction and target side. Trials with straight-ahead gaze served as *neutral* condition. On *valid* trials, gaze direction and target side matched, whereas on *invalid* trials targets appeared opposite to the gaze-cued hemifield.

Procedure

Figure 2A illustrates the sequence of events on a trial. Trial start was signaled by the onset of a fixation cross at the center of the screen. 400 ms later, a face with blank eyes appeared on the screen. After a random time interval of 700 to 1000 ms, pupils appeared within the eyes, looking left, right, or straight ahead. Following the gaze cue, a target dot appeared at one of eighteen target positions with an SOA of either 300 ms or 600 ms, measured from the onset of the pupils (i.e., the cue) to the onset of the target. Schematic face, pupils, and target remained on the screen until a response was given or 1200 ms had elapsed. The inter-trial-interval (ITI) was 680 ms.

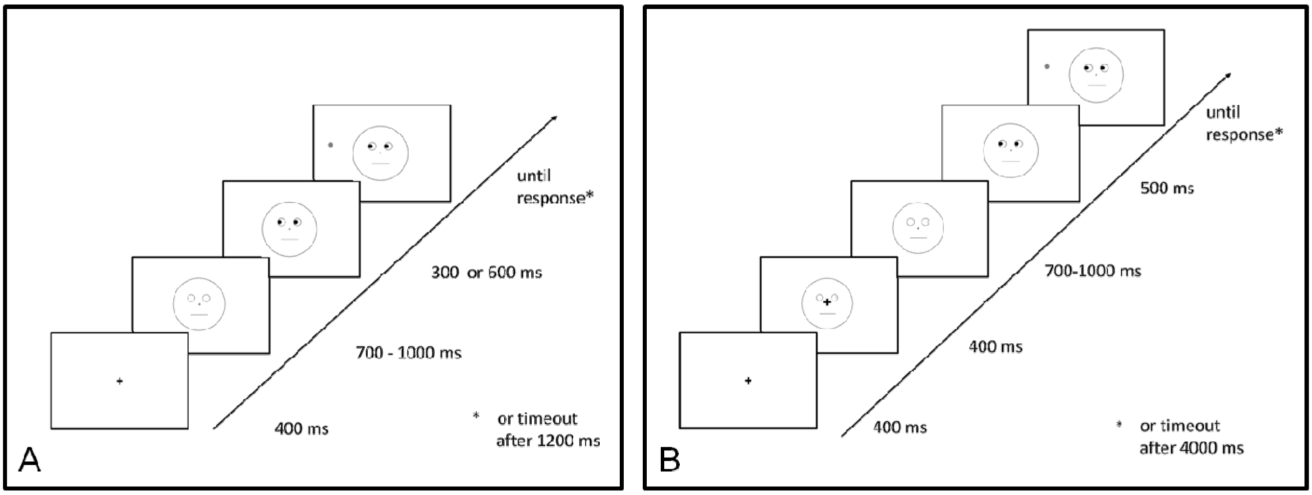

Figure 2. Sequence of trials in (A) localization and discrimination task, and (B) saccade task

Prior to the test sessions, participants were instructed to fixate the central cross as long as it was shown on the screen. They were told that following the fixation cross, a drawing of a face with blank eyes would appear at the center of the screen and that they now had to fixate their eyes on the nose of the face. Further, participants were advised that after the presentation of the face, pupils would appear in the eyes (looking left, right, or straight ahead), followed by a gray target dot/capital letter that could appear anywhere in the field. Participants were expressly informed that the direction in which the eyes looked was not predictive for the location of the target, and they were asked to respond as quickly and accurately as possible to detecting the target.

Analysis

Statistical analyzes focused on the comparison of valid and invalid trials as a function of target position. Cueing effects were examined in terms of *costs-plus-benefits* (invalid-valid), rather than *benefits* (neutral-valid) and *costs* (invalid-neutral) with respect to the neutral condition – because neutral trials may not provide an adequate baseline for the separate assessment of cueing effects (Jonides & Mack, 1984). In fact, neutral trials were found to elicit longer RTs than valid or invalid trials, likely owing to straight-ahead gaze having a *holding effect* on attention (George & Conty, 2008), making it difficult for the target onset to summon an orienting response (Senju & Hasegawa, 2011).

The spatial specificity of gaze cueing was assessed in a three-way repeated-measures analysis of variance (ANOVA) of the RTs, with the factors cue validity (valid, invalid), target position (1 to 9), and SOA (300 ms, 600 ms). Spatially specific gaze-cueing effects would be evidenced by a significant interaction between position and validity (over and above a main effect of validity), with enhanced cueing effects for the exact cued position compared to other positions in the same hemifield. By contrast, spatially nonspecific gaze cueing would manifest in terms of a main effect of validity (not accompanied by a position x validity interaction), with equal facilitation for all positions in the cued hemifield.

To more precisely determine whether and how gaze-cueing effects vary as function of the distance of the target from the cued position, a one-way repeated-measures ANOVA was conducted on the gaze-cueing effects ($\Delta_{RT^{\text{invalid-valid}}}$) with the factor cue-target distance (0°, 20°, 40°, 60°, 80°). Cueing effects were calculated as the RT difference for a given position (e.g., the position 40° in the upper left quadrant) between trials on which this position was validly cued (i.e., gaze directed to the left) compared to when this position was invalidly cued (i.e., gaze directed to the right), with cueing effects collapsed across the two hemifields.

Results

Misses (localization: 0.44%; discrimination: 0.29%) and incorrect responses (localization: 4.20%; discrimination: 3.31%) as well as outliers (±2.5 SD from individual participants' condition means) were excluded from analysis. Mean RTs and Standard Errors for valid, neutral, and invalid trials are presented in Table 1 as a function of target position and SOA. As SOA did not have an influence on the spatial specificity of cueing [localization: $F(16,272)= 1.054$, $p= .401$; discrimination: $F < 1$], the data were collapsed across this factor and only validity and position effects were considered further. Results of statistical analyzes for all trial types (neutral, valid, invalid) are summarized in Table 2.

Experiment 1a: Localization task

The ANOVA of the RTs, with the factors validity and position, revealed the following effects: Mean RTs were shorter on valid compared to invalid trials [*validity*: $F(1,17)= 29.543$, $p< .001$, $\eta^2= .635$]. Furthermore, responses were slower to targets that appeared closer to, compared to targets located further away from, the vertical midline of the display [*position*: $F(8,136)= 64.030$, $p< .001$, $\eta^2= .790$]. Importantly, general RT benefits for valid compared to invalid trials appeared to be of equal magnitude for all target positions, indicating that gaze cueing was equally strong for all positions in the cued hemifield [*validity x position*: *F*(8,136)= 1.497, *p*= .164, η^2= .081, 1-β= .406]. In line with this, the gaze-cueing effects were not modulated by the distance between the cue and the target [main effect of *distance* in ANOVA of cueing effects: *F*(4,68)= .282, *p*= .889, η^2= .016, 1-β= .677]; see Figure 3A.

Table 1: Mean Response Times (in ms), and Standard Errors (in ms) for localization and discrimination task

	Localization			Discrimination		
Condition	valid	neutral	invalid	valid	neutral	invalid
300-ms SOA						
Pos 1	404 (12)	412 (16)	433 (15)	428 (11)	432 (11)	428 (12)
Pos 2	360 (11)	369 (11)	377 (11)	431 (11)	430 (11)	430 (12)
Pos 3	348 (9)	367 (11)	375 (11)	429 (11)	434 (12)	435 (12)
Pos 4	345 (9)	358 (8)	362 (9)	422 (11)	429 (11)	428 (10)
Pos 5	340 (8)	360 (9)	363 (9)	424 (11)	430 (11)	432 (12)
Pos 6	344 (8)	359 (9)	363 (9)	425 (12)	430 (10)	427 (11)
Pos 7	346 (8)	361 (9)	365 (8)	428 (11)	435 (11)	427 (11)
Pos 8	355 (9)	370 (11)	375 (11)	428 (11)	430 (13)	437 (12)
Pos 9	395 (13)	399 (12)	413 (15)	430 (11)	439 (12)	436 (13)
600-ms SOA						
Pos 1	376 (11)	387 (12)	399 (14)	413 (13)	417 (14)	419 (14)
Pos 2	343 (11)	351 (10)	358 (9)	416 (13)	411 (12)	415 (13)
Pos 3	333 (9)	344 (8)	341 (9)	411 (15)	413 (15)	420 (13)
Pos 4	317 (8)	327 (8)	340 (9)	411 (13)	418 (14)	414 (13)
Pos 5	318 (8)	334 (9)	335 (8)	412 (13)	414 (13)	411 (15)
Pos 6	322 (9)	329 (8)	338 (10)	410 (13)	418 (15)	419 (14)
Pos 7	327 (8)	331 (9)	341 (8)	416 (14)	417 (14)	416 (13)
Pos 8	339 (9)	342 (9)	352 (9)	413 (12)	419 (12)	422 (14)
Pos 9	381 (12)	387 (12)	379 (12)	415 (12)	425 (14)	416 (15)

Experiment 1b: Discrimination task

Similar to Experiment 1a, RTs were shorter on valid relative to invalid trials [*validity*: *F*(1,17)= 11.957, *p*= .003, η^2= .413], indicating that gaze cues facilitated the discrimination of targets in the cued hemifield. In contrast to Experiment 1a, RTs did not differ among the nine target positions in the cued hemifield [*position*: *F*(8,136)= 1.418, *p*= .194, η^2= .077, 1-β= .406]; that is, in contrast to target localization, target discrimination was equally difficult however close a target was to the vertical midline. Importantly, gaze-cueing effects were equivalent for all target positions [*distance*: *F*(4,68)= .088, *p*= .986, η^2= .005, 1-β= .677], providing evidence that gaze facilitated target discrimination generally for the whole cued hemifield, rather than specifically for the exact gazed-at location [*validity x position*: *F*(8,136)= 1.166, *p*= .324, η^2= .064, 1-β= .406], see Figure 3A.

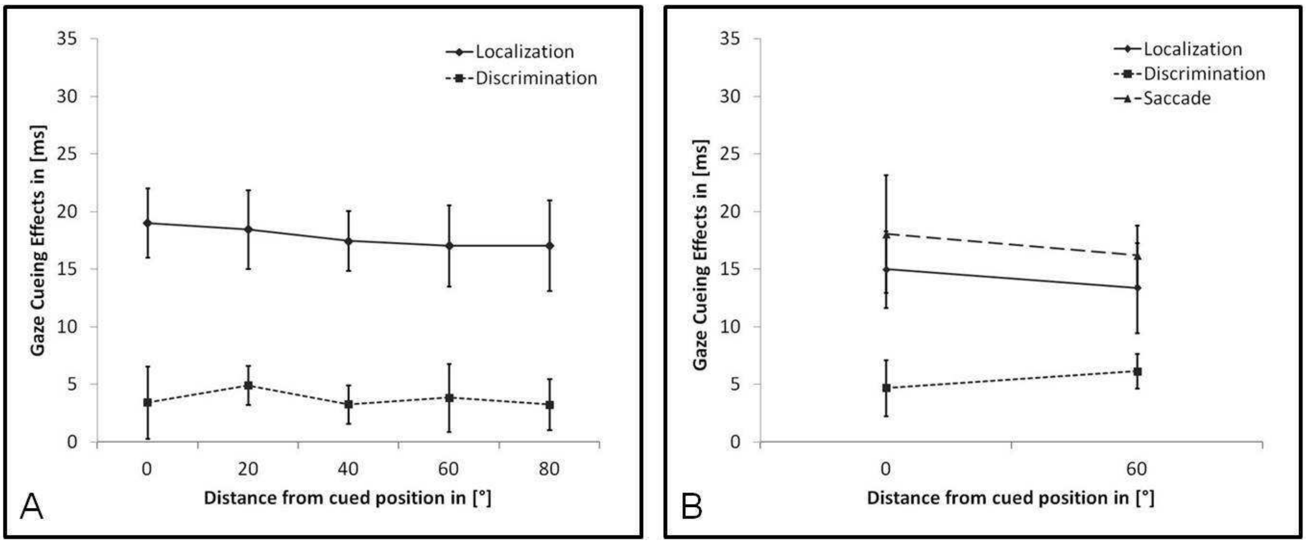

Figure 3. Gaze-cueing effects as function of the distance between cued position and target position (uncluttered visual field). (A) shows results of Experiments 1a and b (18 target positions), and (B) results of Experiments 2a and b (6 target positions). The solid line depicts results for the target localization, the dotted line for the target discrimination, and the dashed line for the saccade task.

Comparison between Experiments 1a and 1b

In a follow-up ANOVA, the gaze-cueing effects were compared between the target localization (Experiment 1a) and target discrimination (Experiment 1b) tasks. This ANOVA, with task type (localization, discrimination) as between-subjects factor and validity (valid, invalid) and position (1 to 9) as within-subjects factors, revealed the essential pattern of cueing effects to be similar for both types of task: While the RTs were overall shorter, and the cueing effects overall smaller for the discrimination versus the localization task [*task*: $F(1,34)= 18.900$, $p< .001$, $\eta^2= .357$; *task x validity*: $F(1,34)= 16.324$, $p< .001$, $\eta^2= .324$], the spatial distribution of the cueing effects [*validity x position*: $F(8,272)= 1.354$, $p= .238$, $\eta^2= .038$, $1-\beta= .903$] was not influenced by the task to be performed [*task x validity x position*: $F(8,272)= 1.389$, $p= .224$, $\eta^2= .039$, $1-\beta= .903$]: the effects were of equivalent magnitude for all positions in the cued hemifield for both task types. The only other effect involving task was the *task x position* interaction [$F(8,272)= 43.793$, $p< .001$, $\eta^2= .563$], due to RTs varying as a function of the distance of the target from the vertical midline in the localization task, but not in the discrimination task.

Discussion

The main goal of Experiment 1 was to investigate the distribution of attention following a central gaze cue (when no context information is available in the periphery). Our findings clearly showed equal RT benefits for all targets that appeared (at one of the nine positions) in the cued hemifield (relative to the corresponding positions in the uncued hemifield), indicating that changes in gaze direction caused cueing effects that were not specific to the exact gazed-at location. In the localization task, RTs were the slower the closer the target was located to the vertical midline, likely owing to the fact that it is harder to make a left/right localization decision for targets appearing closer to, rather than further away from, the vertical axis. Consistent with this, a main effect of position was not found in the discrimination task; rather, the difficulty of discriminating between an F and a T was equivalent for all nine target positions, however close they were to the vertical midline. Importantly, the position effect in the localization task was not modulated by validity, that is: gaze cueing was not enhanced for the exact cued position relative to the other positions within the cued hemifield. Furthermore, at variance with the expectation that specific cueing effects would more likely be observed in tasks requiring greater perceptual resolution, the cueing effects were equally distributed across the cued hemifield even in the discrimination task. Thus, taken together, the cueing effects were spatially nonspecific, with all positions in the cued hemifield receiving equal facilitation, in both the localization and the discrimination task. Changes in RTs occurred only when the vertical meridian was crossed from the cued to the uncued hemifield (Δ_{RT}= 17 ms). Furthermore, the cueing effects were global for short (300 ms) as well as for long (600 ms) cue-target SOAs.

Table 2: Results of statistical analyzes for all trial types (neutral, valid, invalid), for Experiments 1-3.

	validity	position	task	val x pos	task x pos	val x task	val x pos x task
Exp 1a	$F(2,34)= 25.7$	$F(8,136)= 74.4$		$F(16,272)= 1.3$			
	$p< .001$	$p< .001$		$p= .230$			
	$\eta^2= .602$	$\eta^2= .814$		$\eta^2= .069$			
Exp 1b	$F(2,34)= 10.1$	$F(8,136)= 2.2$		$F(16,272)= 1.2$			
	$p< .001$	$p= .064$		$p= .317$			
	$\eta^2= .372$	$\eta^2= .113$		$\eta^2= .065$			
Exp 2a	$F(2,34)= 11.2$	$F(2,34)= 3.6$	$F(1,17)= 132.4$	$F(4,68)= .6$	$F(2,34)= 5.7$	$F(2,34)= 3.6$	$F(4,68)= 1.3$
	$p< .001$	$p= .067$	$p< .001$	$p= .700$	$p= .009$	$p= .037$	$p= .291$
	$\eta^2= .397$	$\eta^2= .174$	$\eta^2= .886$	$\eta^2= .031$	$\eta^2= .252$	$\eta^2= .177$	$\eta^2= .070$
Exp 2b	$F(2,34)= 41.8$	$F(2,34)= 7.4$		$F(4,68)= .4$			
	$p< .001$	$p= .002$		$p= .842$			
	$\eta^2= .723$	$\eta^2= .315$		$\eta^2= .022$			
Exp 3	$F(2,46)= 6.0$	$F(2,46)= 13.4$	$F(1,23)= 129.8$	$F(4,92)= 3.9$	$F(2,46)= 14.7$	$F(2,46)= 3.7$	$F(4,92)= 2.1$
	$p= .005$	$p< .001$	$p< .001$	$p= .006$	$p< .001$	$p= .032$	$p= .095$
	$\eta^2= .208$	$\eta^2= .369$	$\eta^2= .850$	$\eta^2= .145$	$\eta^2= .390$	$\eta^2= .139$	$\eta^2= .081$

Although gaze-cueing effects were spatially nonspecific with both types of task, changing the task had a considerable influence on the overall size of cueing benefits: they were significantly reduced in the discrimination compared to the localization task. Given that the effect of gaze cueing decays with time (e.g., Friesen & Kingstone, 1998), this reduction might, in part, be attributable to the extended processing time required to solve the discrimination task. Another explanation is that gaze cueing not only induces attentional orienting towards the cued hemifield, but also interacts with spatial S-R compatibility effects along the lines sketched above. Compared to the localization task in which the cue might induce a general bias to produce a directionally corresponding left/right response (in addition to inducing attentional orienting), the discrimination task measures cueing effects independent of spatial S-R compatibility. This additional response bias component might explain why the cueing effect was larger in the localization task. Nevertheless, given that the cue-response compatibility effect is additive to that of the attentional effect of cueing, the localization task remains a valid alternative to the discrimination task, for three reasons: (i) processing of the target and response preparation after cue presentation can be accomplished faster for localization (than for discrimination) tasks, thus making them more sensitive for examining cue-induced effects on attentional orienting at different SOAs; (ii) target localization is easier for participants to perform and (iii) requires only half as many trials as the discrimination task.

Experiment 2

Although the essential results were consistent in Experiments 1a and 1b, they are not directly comparable with those of Vuilleumier (2002), who used a smaller number of possible target positions (two per hemifield). Importantly, differences in set size could account for differences in the spatial allocation of attention, as has been shown for the inhibition-of-return (IOR) effect (Birmingham, Visser, Snyder, & Kingstone, 2007). This is because increasing the number of target positions increases the proximity between cued and uncued locations, while at the same time decreasing the spatial predictivity of the cues. Thus, to reveal possible effects arising from the number of target positions, Experiment 2 examined the spatial specificity of gaze cueing using a reduced number of positions, that is: would spatially nonspecific gaze-cueing effects still be obtained when the number of potential target positions within a hemifield is reduced from nine (Experiment 1) to three?

There were two further objectives: The first was to test more directly for differences between target localization and target discrimination by manipulating task type within participants (Experiment 2a), in contrast to the between-subjects manipulation in Experiment 1a and b. The other objective was to examine whether a pattern of spatially nonspecific gaze-cueing effects would also be observed with *overt* shifts of attention (involving eye movements to the target), rather than the *covert* shifts required in Experiment 1. In the context of gaze-cueing effects, it has been shown that the mere observation and the actual execution of eye movements activate similar cortical regions (Grosbras et al., 2005), and that even covert shifts of attention activate the motor system, triggering the preparation of an eye movement to the cued position (Friesen & Kingstone, 2003; Rizzolatti et al., 1987). However, based on the results of Experiment 1, it would appear that the mere programming of a saccade is not sufficient to induce specific cueing effects for the exact gaze-cued position. Rather, it might be that the requirement to actually execute eye movements to the target, together with the preparation of a saccade in response to gaze cues, is essential for inducing spatially specific cueing effects. If this were true, participants would be faster to make saccades to targets appearing at gazed-at positions, compared to targets at uncued locations in cued hemifield. This prediction was tested in Experiment 2b, in which participants were required to make a speeded saccadic response to the target.

As cue-target SOA had no influence on the spatial specificity of gaze cueing in Experiment 1, it was kept constant at 500 ms in Experiment 2. In all other respects, Experiments 2a and 2b were similar to Experiment 1.

Method

Participants

36 volunteers participated in Experiment 2: Eighteen participants (13 women, 5 men) performed localization and discrimination tasks (*Experiment 2a*: age: *M*= 25.11, *SD*= 4.03, range 19-34 years; one left-

handed); the other eighteen participants (14 women, 4 men) performed a saccade task (*Experiment 2b*: age: *M*= 24.94, *SD*= 5.02, range 21-38; two left-handed). In Experiment 2b, the data of one participant had to be excluded from analysis because of eye tracking problems. Testing time was about 30 minutes for the localization and the saccade task, and 1 h for the discrimination task. None of the participants had taken part in Experiment 1.

Stimuli

In Experiment 2, targets could appear at only six positions, three within each hemifield – resulting in a radial distance of 60° between adjacent target positions (Figure 1B). As before, targets could appear at only one of these positions on a trial, and only the position on the horizontal meridian was cued.

Experiment 2a: Apparatus, Design, Procedure, and Analysis

The apparatus in Experiment 2a was the same as in Experiment 1. Experiment 2a was split into two sessions: participants performed the localization task in one session and the discrimination task in the other session. Half of the participants started with the localization task, and the other half with the discrimination task. The localization task consisted of 380 trials, with a block of 20 practice trials preceding the 20 test blocks of 18 trials each. The discrimination task consisted of 740 trials, with a block of 20 practice trials and 20 test blocks of 36 trials each. Gaze direction (straight, left, right), target side (left, right), and target position (top, center, bottom) were selected pseudo-randomly and appeared with equal frequency within each block (with SOA between cue and target onset fixed at 500 ms). In the discrimination task, target identity (F, T) was also varied.

Data were analyzed by a three-way repeated-measures ANOVA with the factors task-type (localization, discrimination), validity (valid, invalid), and position (top, center, bottom).

Experiment 2b: Apparatus, Design, Procedure, and Analysis

In Experiment 2b, stimuli were presented on a Dell CRT color monitor, with the refresh rate set at 100 Hz. Monocular eye tracking of the left eye was performed using an Eyelink 1000 system in combination with a Tower Mount chin-and-head rest (SR Research Ltd., Ontario, Canada). Viewing distance was fixed at 55 cm, and eye data were sampled at 1000 Hz with a spatial resolution of 0.01°.

Experiment 2b consisted of 380 trials, with a block of 20 practice trials preceding 5 test blocks of 72 trials each. Figure 2B provides an illustration of the sequence of events in the saccade task, in which participants were instructed to make, as quickly and accurately as possible, a saccade from the nose of the schematic face (starting position) to the location of the target.

Prior to the experiment, the eye tracker was calibrated. Participants had to fixate the central fixation cross for 400 ms to initiate a trial. Technically, this required the participant's eye position to remain for a 400-ms period within a tolerance region of 1° in diameter around the cross. Achieving successful fixation triggered the appearance of a schematic face with blank eyes, while the fixation cross remained visible on the screen. After a further 400 ms, the fixation cross was replaced by the nose of the schematic face, which participants were instructed to fixate. After a random time interval (700 to 1000 ms), pupils appeared within the eyes, looking left, right, or straight ahead. Participants were told that the direction in which the eyes were looking was not predictive of target location. After a successful fixation for a further 500 ms, the target appeared at one of six possible positions. If fixation was not maintained successfully, the instruction "*Please fixate the nose to proceed*" was displayed at fixation. Schematic face, pupils, and target remained on the screen until a successful saccade was made or 4000 ms had elapsed, whichever came first. During the experiment, the Saccade Reaction Time (SRT) was measured, which was defined as the time interval between target onset and successful landing of the saccade in the region of interest (1° in diameter) around the target. The ITI between consecutive trials was 680 ms. Participants were instructed to take a short break after each test block. Before starting the next block, the eye tracker was recalibrated.

Results

In the saccade task, the only error that could occur was a miss. Misses (localization: 0.15%, discrimination: 0.75%, saccade: 3.42 %), incorrect responses (localization: 1.85%, discrimination: 5.37%), as well as outliers (±2.5 SD from individual participants' means) were excluded from analysis. Mean RTs (or SRTs) for valid, neutral, and invalid trials are presented in Table 3 as a function of target position. Results of statistical analyzes for all trial types (neutral, valid, invalid) are given in Table 2.

Table 3: Mean Response Times and Standard Errors (in ms) for manual localization, discrimination, and saccade tasks

	Localization			Discrimination			Saccade task		
Condition	valid	neutral	invalid	valid	neutral	invalid	valid	neutral	invalid
500-ms SOA									
Pos 1	327 (9)	342 (10)	343 (10)	436 (12)	438 (12)	441 (12)	218 (9)	234 (10)	231 (8)
Pos 2	310 (9)	325 (9)	331 (13)	433 (12)	435 (12)	440 (12)	211 (8)	227 (6)	230 (10)
Pos 3	339 (13)	341 (15)	349 (13)	444 (12)	446 (12)	449 (12)	253 (14)	270 (13)	271 (14)

Experiment 2a

RTs were shorter for valid compared to invalid trials [*validity*: $F(1,17)= 26.257$, $p< .001$, $\eta^2= .607$], but did not vary as a function of position [*position*: $F(2,34)= 2.970$, $p= .096$, $\eta^2= .149$, $1-\beta= .627$]. Overall, participants were faster in localizing than in discriminating targets [*task*: $F(1,17)= 157.788$, $p< .001$, $\eta^2= .903$], and gaze-cueing effects were more marked in the former than in the latter [*task x validity*: $F(1,17)= 4.996$, $p= .039$, $\eta^2= .227$]. Most importantly, the distance between cued position and target position did not modulate the size of gaze-cueing effects [*distance*: $F(1,17)= .002$, $p= .967$, $\eta^2< .001$, $1-\beta= .717$], indicating that all positions in the cued hemifield were equally facilitated [*validity x position*: $F(2,34)= .563$, $p= .575$, $\eta^2= .032$, $1-\beta= .627$]. This pattern of spatially nonspecific cueing effects was evident for the localization and the discrimination task [*task x validity x position*: *F*(2,34)= .627, *p*= .490, η^2= .051, 1-β= .627; *task x distance*: *F*(1,17)= .562, *p*= .464, η^2= .032, 1-β= .717], see Figure 3B.

Comparison between Experiments 1 and 2a

To examine whether the number of target positions influenced the size of gaze-cueing effects, we compared the results of Experiment 2a (three positions per hemifield) with Experiment 1 (nine positions per hemifield) in a meta-analysis of the effect sizes (Rosenthal & DiMatteo, 2001). As the size of the cueing effect is reflected in the main effect of validity, two separate meta-analyses were carried out on the estimated effect sizes for validity: one for the localization and one for the discrimination task. Estimated effect sizes *r* were as follows: Experiment 1a: r_1 = .797 (Loc, 18 pos), Experiment 1b: r_2 = .643 (Discr, 18 pos), Experiment 2a: r_3 = .678 (Loc, 6 pos), and r_4 = .543 (Discr, 6 pos). There was no significant difference in estimated effect sizes for either target localization (r_1 vs. r_3, *z*= .725, *p*= .154) or target discrimination (r_2 vs. r_4, *z*= .421, *p*= .183), indicating that the size of cueing effects was comparable between experiments and not systematically influenced by the number of possible target positions.

Experiment 2b

Comparable to Experiments 1 and 2a (covert attention), SRTs in the saccade task were shorter for valid relative to invalid trials [*validity*: *F*(1,16)= 63.562, *p*< .001, η^2= .799], indicating that gaze cues expedited saccadic reactions when the gaze was directed towards the hemifield in which the target was subsequently presented. Furthermore, SRTs were generally shorter for targets presented on the horizontal meridian (center position), compared to targets in the upper and lower quadrants (top and bottom position) [*position*: *F*(2,32)= 7.161, *p*= .003, η^2= .309]. However, position did not modulate the size of the cueing effects [*validity x position*: *F*(2,32)= .323, *p*= .726, η^2= .020, 1-β= .599], which were equivalent for all positions in the cued hemifield, independent of the distance of the target from the cued position [*distance*: $F(1,16)= .198$, $p= .662$, $\eta^2= .012$, $1-\beta= .819$], see Figure 3B.

Discussion

Experiment 2a

Experiment 2a demonstrated that (i) spatially nonspecific cueing effects are also found when the number of target positions within a hemifield is reduced, while (ii) the size of the gaze-cueing effect itself is unaffected by this reduction. Thus, the spatially nonspecific cueing effects are more likely attributable to attention being distributed equally over the cued (but unstructured) hemifield, rather than to set-size effects.

Note that the finding of spatially nonspecific cueing effects, even with the reduced number of target positions, is unlikely attributable to the use of schematic, instead of realistic, gaze stimuli. In a prior study with a human and a robot gazer, we found participants to be very precise in indicating the gaze direction of the robot's *schematic* eyes; indeed, acuity did not differ between realistic human and *schematic* robot eyes (Wiese, Kohlbecher, & Müller, 2012). This strongly suggests that the *schematic* cues used in the present study were precise enough for participants to perceive them as pointing into a specific direction, rather than just globally indicating a whole hemifield.

Furthermore, the spatially nonspecific cueing effects are unlikely due to the fact that the schematic face was consistently looking at the central position on the horizontal midline (and never at any of the other potential target positions) in a hemifield, which might, for some reason, induce a global bias towards the cued side. This is effectively ruled out by a control experiment in which the (spatially non-predictive) gaze cue was equally likely directed to the position in the lower and, respectively, the upper hemifield quadrant, as well as the central position on the horizontal midline. In all other respects, the experiment was comparable to the localization task of Experiment 2a: The results obtained from ten participants (6 female; age: M=25 years, range 20-30 years; all right-handed) mirrored the findings of Experiment 2a, as evidenced by a repeated-measures ANOVA with the factors validity (valid, invalid), gaze direction (top, center, bottom), and target position (top, center, bottom): RTs were shorter on valid compared to invalid trials [*validity*: $F(1,9)=13.831$, $p= .005$, $\eta^2= .606$], whether the lower, the central, or the upper target position was cued. The RTs were also longer for positions located closer to the vertical midline [*target position*: $F(2,18)=25.590$, $p< .001$, $\eta^2= .740$], as is typical for left-right localization. Importantly, the cueing effects did not differ between the exact gazed-at position and the respective other positions in the cued hemifield, whichever position has been cued [*validity x gaze direction x target position*: $F(4,36)= .816$, $p= .523$, $\eta^2= .083$, $1-\beta= .494$]. Thus, general gaze-cueing effects for the whole gaze cued hemifield are still found under conditions in which a central gaze cue is directed to one of all possible target positions (rather than just the central position) on one or the other side.

Experiment 2b

The goal of Experiment 2b was to examine the hypothesis that gaze-cueing effects would be spatially specific if the task requires participants to actually execute a saccade to the target (together with the implicit, covert preparation of an eye movement in response to the cue). However, at variance with this hypothesis, SRTs mirrored the previous findings of nonspecific cueing effects: facilitation was equivalent for all target positions within the cued hemifield, with no extra enhancement for the exact gazed-at position. Most importantly, as shown in a meta-analysis of the effect sizes for the main effect of validity, overt orienting in response to gaze cues was not more robust compared to mere covert orienting ($r_1 = .779$ (Experiment 2a); $r_2 = .894$ (Experiment 2b); r_1 vs. r_2, $z= 1.07$, $p= .113$), which is in line with Friesen and Kingstone (2003).

Thus, going beyond previous findings, the results of Experiment 2b show that even for overt shifts of attention, gaze-cueing effects are not confined to the exact gazed-at position; rather, gaze cues equally facilitate the programming of an eye movement to any position in the cued hemifield. This pattern is consistent with Rizzolatti et al. (1987), who showed that the eye movement program induced by symbolic cues specifies only the left/right direction parameter of a saccade, resulting in a global benefit for the cued hemifield, but not for the exact cued position. This finding is very plausible given that oblique saccades are composed of a vertical and a horizontal component which are implemented by separate neuronal channels in the brain stem and are executed by distinct groups of motor neurons and eye muscles (Bahill & Stark, 1977; King, Lisberger, & Fuchs, 1986). Consequently, given that vertical and horizontal components of saccadic eye movements are planned and generated independently of each other and that oblique saccades are more difficult to generate in general (as they involve the coordination of two separate components), it is possible that observing a lateral gaze shift also triggers the preparation of an eye movement in only one, dominant (i.e., horizontal), but not the other (i.e., vertical), channel. Thus, when being presented with horizontal gaze shifts, only the horizontal (but not the vertical) component of an eye movement is prepared, resulting in SRT benefits for left/right, but not up/down saccadic eye movements.

In sum, Experiments 1 and 2 indicate that when no visual context information is provided to which gaze cues could refer, (i) spatial attention is not allocated to the exact cued position, but facilitates target localization and discrimination at all positions in the cued hemifield; (ii) nonspecific gaze-cueing effects are independent of task demands; (iii) the size of gaze cueing is not influenced by the number of target positions; and (iv) nonspecific cueing effects are also found for overt attention shifts.

Experiment 3

Experiment 3 was designed to investigate whether the allocation of attention induced by gaze cues is sensitive to context information presented in the visual field. Assuming that gaze direction provides important information about other persons’ internal states, we hypothesized that gaze cues are only interpreted specifically in relation to external reference objects that represent the targets of the others' attentional focus. If so, presenting targets inside predefined reference objects would be expected to turn spatially nonspecific gaze-cueing effects into specific effects, as a link between gaze direction and potential objects of interest can already be established before the target appears.

Method

The methodological details were generally the same as in Experiment 2a, with the following exceptions: Twenty-four volunteers (17 women, 7 men) participated in Experiment 3 (age: M= 24.54, SD= 2.63, range 21-30 years; all right-handed); none of them had taken part in the previous experiments. Targets were presented in predefined placeholder objects that consisted of a white rectangle surrounded by a black line (1.0° wide; 1.5° high); the peripheral placeholders appeared simultaneously with the blank-eyed face in the display center. Next, pupils appeared inside the eyes, creating the impression of the face looking at one of the placeholders. 500 ms later the target – either a dot (localization task) or a letter (discrimination task) – was presented centered inside one of the placeholders. The spatial specificity of gaze cueing was assessed by a two-way repeated-measures ANOVA of the RTs, with the factors validity (valid, invalid) and position (top, center, bottom), as well as by an ANOVA of the cueing effects (i.e., $\Delta RT_{\text{invalid-valid}}$), with the factor cue-target distance (0°, 60°).

Results

Misses (localization: 1.89%; discrimination: 0.39%), incorrect responses (localization: 1.54%; discrimination: 4.74%) and outliers (±2.5 SD from individual participants' means) were excluded from analysis. Mean RTs for valid, neutral, and invalid trials are presented in Table 4 as a function of target position. Results of statistical analyzes for all trial types (neutral, valid, invalid) are presented in Table 2.

As in the first two experiments, RTs were shorter for valid compared to invalid trials [*validity*: $F(1,23)= 16.926$, $p< .001$, $\eta^2= .424$], and participants were always faster in localizing than in discriminating targets [*task*: $F(1,23)= 130.607$, $p< .001$, $\eta^2= .850$]. Again, gaze cueing effects were larger in the localization than in the discrimination task [*task x validity*: $F(1,23)= 5,920$, $p= .023$, $\eta^2= .205$], and performance in the localization task, but not in the discrimination task, depended on target position [*task x position*: $F(2,46)= 9.508$, $p< .001$, $\eta^2= .292$]. Most importantly, this time, cueing effects were modulated by target position, with larger effects for the exact gazed-at position compared to the other two positions in the cued hemifield [*validity x position*: $F(2,46)= 9.113$, $p< .001$, $\eta^2= .284$]. In line with this, the gaze-cueing effects within the cued hemifield varied as a function of the distance between cue and target [*distance*: $F(1,23)= 13.133$, $p= .001$, $\eta^2= .363$]. Note that although gaze cueing was strongest for the exact cued position, there was still facilitation for the other two positions in the cued hemifield. This pattern indicates that, under the conditions of Experiment 3, gaze cues caused a significantly stronger facilitation effect for the exact gazed-at position compared to the other two positions in the cued hemifield.

Table 4: Mean Response Times and Standard Errors (in ms) for localization and discrimination (with placeholders)

	Localization			Discrimination		
Condition	valid	neutral	invalid	valid	neutral	invalid
500-ms SOA						
Pos 1	342 (13)	352 (9)	361 (8)	428 (20)	431 (20)	429 (19)
Pos 2	320 (17)	343 (8)	348 (9)	428 (16)	424 (19)	433 (16)
Pos 3	355 (17)	367 (13)	363 (10)	439 (17)	444 (19)	440 (17)

As can be seen from Figure 4, spatially specific gaze cueing was evident for both types of task (dashed lines). Statistically, the *task x validity x position* interaction [$F(2,46)= 1,301$, $p= .282$, $\eta^2= .054$, $1-\beta= .770$] and the *task x distance* interaction [$F(1,23)= 2.049$, $p= .166$, $\eta^2= .082$, $1-\beta= .936$] were non-significant. And when analyzing the localization and discrimination tasks separately, the critical *validity x position* interaction [localization: $F(2,46)= 6.054$, $p= .005$, $\eta^2= .208$; discrimination: $F(2,46)= 3.632$, *p*= .034, η^2= .136] as well as the effect of cue-target *distance* [localization: *F*(1,23)= 8.091, *p*= .009, η^2= .260; discrimination: *F*(1,23)= 8.675, *p*= .007, η^2= .274] were significant for both types of task.

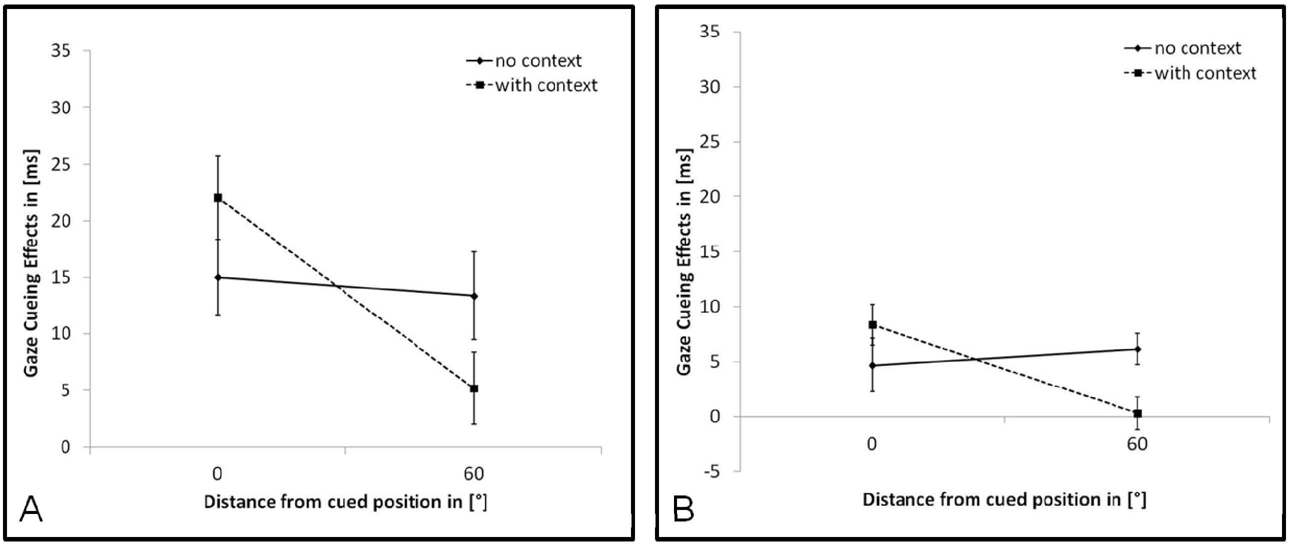

Figure 4. Comparison between Experiment 2a (without context information, solid line) and Experiment 3 (with context information, dashed line). (A) shows results of the localization task, (B) shows results of the discrimination task. With both types of task, adding context information in the form of predefined position placeholders changed a general cueing effect for the whole cued hemifield into a more specific cueing effect for the exact gazed-at position.

Importantly, as depicted in Figure 4, a comparison between Experiment 3 (with placeholders – dashed lines) and Experiment 2a (without placeholders – solid lines) revealed the spatial specificity effect to be significantly enhanced when context information was presented in the periphery [*distance x experiment*: *F*(1,40)= 8.262, *p*= .006, η^2= .171], with both the localization and the discrimination task [*distance x experiment x task: F*(1,40)= .515, *p*= .477, η^2= .013, 1-β= .997; *distance x experiment* (localization): *F*(1,40)= 4.206, *p*= .046, η^2= .095; *distance x experiment* (discrimination): *F*(1,40)= 6.610, *p*= .014, η^2= .142]. This pattern indicates that visual context information changes the general cueing effect that is observable under conditions without context information into a specific cueing effect for the exact cued position, with significantly stronger facilitation for the exact gazed-at position compared to the other positions in the cued hemifield. Figure 4 suggests that this change in the spatial specificity of gaze-cueing effects is due to increased facilitation for the exact gazed-at position, coupled with reduced facilitation for the uncued positions in the cued hemifield. However, when comparing cueing effects for the same positions (exact, other) between conditions with and without context information (across experiments), separately for the localization and discrimination tasks, only the reduced facilitation for the other positions in the discrimination task turned out to be just significant ($t(40)= 2.753$, $p=.036$; all other ts < 1.506, ps > .140).

Discussion

Experiment 3 examined whether the specificity of gaze cueing is modulated by visual context information in the periphery, in particular, whether spatially specific cueing effects are dependent on context information, such as objects of interest being available in the field to which the gazer's attentional focus could be directed. The assumption that cueing effects are modulated by context information is consistent with previous findings on the spatial specificity of non-social symbolic cues: While Hughes and Zimba (1985, 1987) found general cueing effects for larger regions of the visual field (i.e., hemifields or quadrants) when no context was provided, spatially specific cueing effects were observed when targets were presented at predefined placeholder locations (e.g., Eriksen & Yeh, 1985; Henderson & Macquistan, 1993; Shepherd & Müller, 1989). Based on these observations, we hypothesized that the manifestation of specific gaze-cueing effects critically depends on the availability of context information (i.e., placeholders) in the periphery, which would restrict processing to only a few regions of the visual field where task-relevant events are likely to occur. Experiment 3 clearly supported this hypothesis. In contrast to the first two experiments, it revealed an interaction between validity and position: gaze-cueing effects were significantly enhanced for the exact gazed-at position compared to the other two positions within the cued hemifield.

This pattern points to an interplay of two components which determine the spatial specificity of gaze-cueing effects: (i) a general orienting component that is activated bottom-up by gaze cues, whether or not further context information is provided; and (ii) a top-down component that comes into play only if contextual information is available, permitting the gaze direction to be linked to specific reference objects. While the bottom-up component results in a general gaze-cueing effect for the whole gaze-cued hemifield, the top-down component appears to induce a facilitation effect which is spatially specific to the exact gazed-at position.

One possible challenge to this interpretation of the data is that when placeholder structures are available in the periphery, participants are more likely to make a saccade in response to the gaze cues, which could potentially generate a pattern of spatially specific cueing effects. In order to validate that participants did heed the instructions to maintain fixation on the central fixation cross throughout a trial, we tested six additional participants (age: M=27, range 23-30 years; 5 female; right-handed) in an eye-tracking experiment, using the same displays, with potential target positions in the periphery marked by placeholders, as in Experiment 3. Only the discrimination task was tested, based on the assumption that the likelihood of making eye movements to the cued (or the target) location would be highest when task performance requires greater perceptual resolution. However, even under these conditions, fixations remained on the central cross (i.e., within a tolerance area of ±1.5°) in 94.6% of all trials. Of the remaining 5.4% of the trials, only 0.2% of the eye movements went to the target (with a tolerance area of ±1.5°), indicating that participants followed the instructions and did not execute saccades towards the current target location. For those trials on which participants maintained central fixation during the whole trial, the results were comparable to Experiment 3 (first 12 participants): RTs were overall shorter for targets appearing in the cued hemifield rather than the uncued hemifield (442 vs. 447 ms) [*validity*: $F(1,16)= 4.485$, $p= .05$, $\eta^2= .219$], as well as for targets presented on the horizontal meridian compared to targets in the upper and lower field quadrants (437 ms vs. 448 ms) [*position*: $F(2,32)=19.260$, $p< .001$, $\eta^2= .546$]. Importantly, gaze-cueing effects were larger for the exact gazed-at position compared to the other positions in the cued hemifield (12 ms vs. 1 ms) [*validity x position*: $F(2,32)= 8.780$,

p= .001, η^2= .354]. Most importantly, the pattern of results did not differ between the control experiment and Experiment 3 (discrimination task): the interactions *experiment x validity* [F(1,16)= .143, p= .710, η^2= .009, 1-β= .712], *experiment x position* [F(2,32)=2.038, p= .147, η^2= .105, 1-β= .625], and *experiment x validity x position* [F(2,32)=2.326, p= .114, η^2= .127, 1-β= .625] were all non-significant. This makes it unlikely that the spatially specific gaze-cueing effects observed in Experiment 3 can be attributed to systematic saccades to potential target positions marked by placeholders.

2.1.4. General Discussion

The goal of the current study was to investigate the spatial allocation of attention induced by nonpredictive gaze cues. Experiments 1 and 2 examined how gaze-cueing effects are distributed in an uncluttered visual field, dependent on the type of task, the number of possible target positions (set size), and on whether attention is oriented covertly or overtly. Experiment 3 then examined the influence of context information, in the form of peripheral position placeholders, on visuo-spatial orienting in response to gaze cues. The results revealed the spatial specificity of gaze cueing to be critically dependent on the availability of context information: Nonspecific gaze-cueing effects were found consistently when there were no placeholder objects in the periphery to which the gaze cues could refer; and this general cueing effect was independent of the type of task to be performed (localization, discrimination), the set size, and whether attention had to be shifted covertly or overtly. In contrast, when reference objects were provided in the periphery, gaze cues induced a cueing effect specific to the exact gazed-at position.

The current results help integrate different theories of spatial attention with regard to their validity for gaze cueing. On the one hand, they confirm the assumptions of general-region theories of attention, such as the *meridian-boundary model* (Hughes & Zimba, 1985, 1987) – by showing that gaze cues give rise to spatially nonspecific, hemifield cueing effects when no further information about potential reference objects in the periphery is provided. On the other hand, they are also consistent with specific-location theories of attention, such as the *zoom-lens model* (Eriksen & Yeh, 1985) or the *gradient model* (Downing & Pinker, 1985) – by showing that attentional resources can be allocated to specific, narrow regions of the visual field when objects are presented in the periphery that can serve as reference points for the other’s gaze direction.

The present findings argue in favor of the idea that gaze direction cues can initiate both a general and a specific component of attentional orienting, dependent on the availability of context information in the visual scene: If no further information is provided, only a general directional orienting component is activated (*bottom-up*), yielding a global cueing effect for the whole gaze-cued hemifield. By contrast, if context information is available, a spatially specific component (*top-down*) comes into play, inducing a facilitation effect for the exact gazed-at position. However, it is not entirely clear from the presented data whether the modulation of the spatial specificity of gaze-cueing effects is due to a trade-off relationship between the top-down and bottom-up components, or whether the top-down component is simply additive to the bottom-up component. Whatever the precise relationship between the two orienting components, the specific component seems to critically depend on the availability of context information and is, thus, likely to result from the combination of two separate sources of information: the linking of context information provided in the visual field with directional information from the gazer’s eyes.

The proposal that visual information from the eye region and the periphery are integrated when processing gaze direction is in line with Lobmaier et al. (2006). They showed that objects in the visual field can capture perceived gaze direction of an interaction partner, causing systematic biases in estimating where the other is looking. Arguably, rather than being achieved directly (in *bottom-up* manner), this linkage may involve complex, higher-level processes of cross-referencing central gaze direction with peripheral structures, which may then exert a *top-down-like* effect on the allocation of attention. However, this does not rule out that the computations involved are highly automatized and that, once it is established which peripheral object is referred to by the gaze cue, this object becomes an effective attractor for the allocation of attention.

The idea that gaze direction can induce both a general and a specific cueing effect, the latter dependent on the availability of peripheral context information, is also consistent with Gibson, Thompson, Davis, and Biggs (2011). They showed that the extent to which attention is oriented to specific locations in response to symbolic cues depends on the expectancies observers have about the direction and distance of the upcoming target: while directional information is provided by the cue itself, distance information is either derived from experience (e.g., learning that targets always appear at the same eccentricity) or provided by position placeholders (see also Shepherd & Müller, 1989). However, in the present study, simply presenting targets consistently at the same distance from fixation did not give rise to spatially specific cueing effects (Experiments 1 and 2); rather, specific effects were dependent on the presence of position placeholders (Experiment 3), suggesting that reference objects are critical for the exact interpretation of gaze direction.

This notion would also be consistent with joint-attention theories, according to which attending to the attentional focus of others is an important prerequisite for sharing information in social interactions (Butterworth & Jarrett, 1991; Scaife & Bruner, 1975). Of importance with respect to the present findings, it has been shown that joint attention can only be established when objects are placed in the visual field, with which the other’s gaze can be aligned (Butterworth & Jarrett, 1991; Moore & Corkum, 1995). In line with this, our findings show that the presence of objects in the visual field is also essential for inducing spatially specific gaze-cueing effects in adults.

The evidence available to date suggests that what has been established in the present study with regard to the spatial specificity of gaze cueing may also apply to non-social symbolic (e.g., arrow) cues in general: symbolic cueing towards an uncluttered visual field produces spatially non-specific cueing effects (Hughes & Zimba, 1985, 1987), whereas the cueing effects are spatially specific when possible target locations are predefined by placeholders (Eriksen & Yeh, 1985; Henderson & Macquistan, 1993; Shepherd & Müller, 1989). Nevertheless, given that these studies were quite different with regard to the stimulus arrangements and tasks used (the only commonality being the use of central arrow cues), the generality issue would need to be examined in further, systematic experiments adapting the designs of the present study. However, even if gaze cueing turned out to be a specific instance of symbolic cueing as concerns the role of visual context for the spatial specificity of the cueing effects, it would remain that gaze direction induces cueing effects faster and in a more reflexive manner compared to arrow cues (Friesen et al., 2004). Also, gaze-cueing effects may well be particularly reliant on the availability of external objects. Further work would be required to elucidate the exact mechanisms that relate the central information from the cues with peripheral context information, and whether their dynamics is the same with social as with non-social symbolic cues.

In summary, the present results reveal a degree of flexibility in the gaze-cueing system that allows for the integration of multiple sources of information to guide attention: when information about possible reference objects in the visual field is lacking, attention is allocated to a broader area, whereas attention shifts are specific to the cued location when a relation between gaze direction and objects of interest can be established. This context-dependent flexibility is adaptive, in that it allows for rapid detection of relevant objects in a constantly changing social environment.

Acknowledgments

We thank Verena Zierl, Christoph Voelter, Natalia Rabel, and Regina Lanz for help with data collection and Mark D. Schroeder for help with programming. This work has been supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) within the Excellence Cluster Cognition for Technical Systems (CoTeSys). We also thank the Graduate School of Systemic Neurosciences (GSN) Munich for their support.

2.1.5. References

Ando, S. (2002). Luminance-induced shift in the apparent direction of gaze, *Perception, 31*, 657-674.

Anstis, S.M., Mayhew, J.W., & Morley, T. (1969). The perception of where a face or television 'portrait' is looking. *American Journal of Psychology, 82*, 474–489.

Bahill, A.T., & Stark, L. (1977). Oblique saccadic eye movements – Independence of horizontal and vertical channels. *Archives of Ophthalmology, 95*, 1258-1261.

Birmingham, E., Visser, T.A.W., Snyder, J.J., & Kingstone, A. (2007). Inhibition of return: Unraveling a paradox. *Psychonomic Bulletin & Review*, *14(5)*, 957-963.

Bock, S.W., Dicke, P., & Thier, P. (2008). How precise is gaze following in humans?, *Vision Research*, *48(7)*, 946-957.

Butterworth, G., & Jarrett, N.L.M. (1991). What minds have in common space: spatial mechanisms for perspective taking in infancy. *British Journal of Developmental Psychology*, *9*, 55-72.

Cline, M.G. (1967). The perception of where a person is looking. *American Journal of Psychology, 80*, 41-50.

Downing, C.J. (1988). Expectancy and visuo-spatial attention. Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception & Performance*, *14*, 188-202.

Downing, C.J., & Pinker, S. (1985). The spatial structure of visual attention. In M.I. Posner & O.S.M. Marin (Eds.), *Attention and performance XI*, pp.171-188. Hillsdale, NJ: Erlbaum.

Downing, P., Dodds, C.M., & Bray, D. (2004). Why does the gaze of others direct visual attention? *Visual Cognition*, *11*, 71-79.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, *6*, 509-540.

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

Eriksen, C. W., & Yeh, Y.-Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception & Performance, 11*, 583-597.

Frischen, A., Bayliss, A.P., & Tipper, S.P. (2007). Gaze Cuing of Attention: Visual Attention, Social Cognition, and Individual Differences. *Psychological Bulletin*, *133(4)*, 694-724.

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

Friesen, C.K., & Kingstone, A. (2003). Abrupt onsets and gaze direction cues trigger independent reflexive attentional effects. *Cognition, 87*, B1-B10.

Friesen, C.K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception & Performance*, *30*, 319-329.

Gale, C., & Monk, A.F. (2000). Where am I looking? The accuracy of video-mediated gaze awareness. *Perception & Psychophysics*, *62*, 586-595.

Gamer, M., & Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. *Journal of Experimental Psychology: Human Perception & Performance*, *33(3)*, 705-715.

George, H., & Conty, L. (2008). Facing the gaze of others. *Clinical Neurophysiology*, *38(3)*, 197-207.

Gibson, J.J., & Pick, A. (1963). Perception of another person's looking behavior. *American Journal of Psychology*, *76*, 386-394.

Gibson, B. S., Thompson, A. N., Davis, G. J., & Biggs, A. T. (2011). Going the distance: Extrasymbolic contributions to the symbolic control of spatial attention. *Visual Cognition*, *19*, 1237–1261.

Grosbras, M.H., Laird, A.R., & Paus, T. (2005). Cortical regions involved in eye movements, shifts of attention and gaze perception. *Human Brain Mapping*, *25*, 140-154.

Henderson, J.M. (1991). Stimulus discrimination following covert attentional orienting to an exogenous cue. *Journal of Experimental Psychology: Human Perception & Performance*, *17*, 91-106.

Henderson, J.M., & Macquistan, A.D. (1993). The spatial distribution of attention following an exogenous cue. *Perception & Psychophysics*, *53(2)*, 221-230.

Hietanen, J.K. (1999). Does your gaze direction and head orientation shift my visual attention? *NeuroReport*, *10*, 3443-3447.

Hietanen, J.K., & Leppänen, J.M. (2003). Does facial expression affect attention orienting by gaze direction cues? *Journal of Experimental Psychology: Human Perception & Performance*, *29*, 1228-1243.

Hughes, H.C., & Zimba, L.D. (1985). Spatial maps of directed visual attention. *Journal of Experimental Psychology: Human Perception & Performance*, *11*, 409-430.

Hughes, H.C., & Zimba, L.D. (1987). Natural boundaries for spread of directed visual attention. *Neuropsychologia*, *2*, 5-18.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin, 96(1)*, 29-44.

King, W.M., Lisberger, S.G., & Fuchs, A.F. (1986). Oblique saccadic eye movements of primates. *Journal of Neurophysiology*, *56(5)*, 769-784.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C.K., & Eastwood, J.D. (2003). Attention researchers! It is time to take a look at the real world. *Current Directions in Psychological Science*, *12*, 176-180.

LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception & Performance*, *9*, 371-379.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, *96*, 101-124.

Langton, S.R.H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, *6*, 541-567.

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

Lobmaier, J.S., Fischer, M.H., & Schwaninger, A. (2006). Objects capture perceived gaze direction. *Experimental Psychology, 53*, 117-122.

Moore, C., & Corkum, V. (1995). The origins of joint visual attention. In C. Moore & P. Dunham (Eds.) *Joint attention: Its origins and role in development*, pp. 61-85. Hillsdale, NJ: Erlbaum.

Müller, H.J., & Findlay, J.M. (1987). Sensitivity and criterion effects in the spatial cuing of visual attention. *Perception & Psychophysics, 42,* 383-399.

Nummenmaa, L., & Hietanen, J.K. (2009). How attentional systems process conflicting cues. The superiority of social over symbolic orienting revisited. *Journal of Experimental Psychology: Human Perception & Performance, 35(6),* 1738-1754.

Quadflieg, S., Mason, M.F., & Macrae, C.N. (2004). The owl and the pussycat: Gaze cues and visuospatial orienting. *Psychonomic Bulletin & Review, 11(5),* 826-831.

Rizzolatti, G., Riggio, L., Dascola, I., & Umilta, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favour of a premotor theory of attention. *Neuropsychologia, 25,* 31-40.

Rosenthal, R., & DiMatteo M.R. (2001). Meta-analysis: recent developments in quantitative methods for literature reviews, *Annual Review of Psychology, 52,* 59-82.

Scaife, M., & Bruner, J. (1975). The capacity for joint visual attention in the infant. *Nature, 253(5489),* 265-266.

Schwaninger, A., Lobmaier, J.S., Fischer, M.H. (2005). The inversion effect of gaze perception reflects processing of component information. *Experimental Brain Research, 167,* 49-55.

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

Senju, A., & Hasegawa, T. (2011). Direct gaze captures visuospatial attention. *Visual Cognition, 12(1),* 127-144.

Shepherd, M., & Müller, H.J. (1989). Movement versus focusing of visual attention. *Perception & Psychophysics*, *46(2)*, 146-154.

Shulman, G.L., Wilson, J., & Sheehy, J.B. (1985). Spatial determinants of the distribution of attention. *Perception & Psychophysics, 37*, 59-65.

Simon, J.R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81(1)*, 174-176.

Striano, T., & Stahl, D. (2005). Sensitivity to triadic attention in early infancy. *Developmental Science, 8(4),* 333-343.

Symons, L.A., Lee, K., Cedrone, C.C., & Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. *Journal of General Psychology*, *131*, 451-469.

Todorovic, D. (2006). Geometrical basis of perception of gaze direction, *Vision Research*, *46*, 3549-3562.

Vuilleumier, P. (2002). Perceived gaze direction in faces and spatial attention: A study in patients with parietal damage and unilateral neglect. *Neuropsychologia*, *40*, 1013-1026.

Wiese, E., Kohlbecher, S., & Müller, H.J. (2012). Acuity in estimating gaze direction in Human-Human and Human-Robot Interaction. In: *Proceedings of the 7th ACM/IEEE HRI 2012*. March 5, 2012, Boston, MA, USA.

2.2. Study 2

What we observe is biased by what other people tell us – Beliefs about the reliability of gaze behavior modulate attentional orienting to gaze cues

Eva Wiese[1], Agnieszka Wykowska[1] & Hermann J. Müller[1,2]

[1]Ludwig-Maximilian University of Munich, Germany

[2]Birkbeck College, University of London, UK

Author contributions: EW, HJM conceived and designed research. EW performed research. EW analyzed data. EW, AW, HJM wrote the paper.

Wiese, E., Wykowska, A., & Müller, H.J. (under review). What we observe is biased by what other people tell us: Beliefs about the reliability of gaze behavior modulate attentional orienting to gaze cues. *Journal of Experimental Psychology: General*

2.2.1. Abstract

For interacting with other people, information about the physical surrounding has to be integrated with information about the social interaction partner. Processing social information is usually guided by a *bottom-up* mechanism that is reflexively triggered by stimulus-related information in the social scene, combined with a *top-down* mechanism activated by task-related context information. The present study investigated whether these components interact during attentional orienting to changes in gaze direction. In particular, we examined whether the spatial specificity of gaze cueing is modulated by expectations about the reliability of gaze behavior that were induced either by experience with the gaze behavior or by instruction. Spatially specific cueing effects were observed for highly predictive cueing conditions, and importantly, also when participants believed that nonpredictive gaze cues were highly predictive. Analogously, cueing effects for the whole gazed-at hemifield were observed for nonpredictive cues and when predictive cues were believed to be nonpredictive, specific cueing effects were attenuated. This pattern indicates that instruction-induced expectations about gaze behavior modulate social attention, even when they contradict information available in the social scene.

Keywords

Gaze cueing, top-down control, spatial allocation of attention, cue predictivity

2.2.2. Introduction

For engaging in interactions with other people we need to know *who* we are interacting with, and *what* others are going to do next (Frith & Frith, 2006). Based on this knowledge, which can be acquired *directly* through interacting with people or *indirectly* by either observing someone interacting with another person or by being told about a person, we make inferences about the other's internal states, including intentions, beliefs, and feelings. At the core of this *mentalizing* process (Frith & Frith, 2003) is that our predictions about others are based not simply on the actual state of the world, but also on our assumptions about the others' internal states. Accordingly, the interpretation of social scenes is thought to involve (i) a *bottom-up* mechanism that is activated by perceptual information in the social scene, and (ii) a *top-down* mechanism that is based on background knowledge we have about others, or inferences we draw from perceived information, and that interplays with the bottom-up component. To allow adequate behavior in social interactions, the human brain developed a system specialized for processing social information consisting of *medial prefrontal cortex* (mPFC), *superior temporal sulcus* (STS), *orbitofrontal cortex*, *amygdala* and *anterior insula* (Adolphs, 1999; Brothers, 1990). While bottom-up responses to social signals are thought to be evoked in the STS, top-down modulation of these responses is assumed to originate from the mPFC (Grezes et al., 2004 a,b; Saxe & Wexler, 2005).

One fundamental mechanism employed in the processing of social information is gaze following. Gaze direction is very informative, as it indicates another's focus of interest and encourages the observer to shift attention to the same location (for a review: Frischen, Bayliss, & Tipper, 2007). Gaze-triggered attention shifts have been investigated using a cueing paradigm (Friesen & Kingstone, 1998) in which a face is presented centrally that gazes either straight ahead, to the left, or to the right. Reactions to targets appearing in the gazed-at hemifield are typically faster than those to targets in the opposite hemifield (e.g., Friesen & Kingstone, 1998; Langton & Bruce, 1999; Wiese, Zwickel & Müller, in press).

Gaze direction has traditionally been thought to be special with regard to guiding attention. In contrast to other central cues (Müller & Rabbitt, 1989; Posner, 1980; Ristic & Kingstone, 2006), gaze triggers shifts of attention to peripheral locations when being nonpredictive (Friesen & Kingstone, 1998; Langton & Bruce, 1999; cf. Vecera & Rizzo, 2006) or even counterpredictive with respect to the target location (Friesen, Ristic & Kingstone, 2004) – a pattern that is consistent with a reflexive mechanism.

However, the view that gaze cues provide particularly powerful attentional orienting signals (reflecting their social relevance) has recently been challenged by evidence showing that not only gaze, but also other overlearned symbolic (e.g., arrow) cues are capable of inducing shifts of attention when being nonpredictive (Hommel et al., 2001; Ristic, Friesen, & Kingstone, 2002; Tipples, 2002; cf. Friesen, Ristic & Kingstone, 2004). Furthermore, orienting of attention in response to gaze direction can be top-down controlled if appropriate context information is available (Nummenmaa & Hietanen, 2009; Teufel et al., 2009; Wiese et al., in press). In particular, pre-existing assumptions concerning the observed stimulus have been shown to influence gaze cueing (Deaner, Shepherd, & Platt, 2007; Ristic & Kingstone, 2005; Teufel et al., 2010; Wiese, Wykowska et al., 2012; cf. Bayliss & Tipper, 2006). For example, when humans believe that the observed gaze behavior is intentional, cueing effects are larger compared to when the gazer is believed to display only mechanistic behavior (Wiese, Wykowska et al., 2012). Similarly, when the gazer represents a stereotypical member of a group that the observer belongs to (e.g., political party), the observer is more likely to follow his/her gaze direction (Liuzza et al., 2011).

Taken together, these findings suggest that gaze cues, in addition to a bottom-up component of attentional orienting, evoke a top-down mechanism, which is dependent on whether or not task-relevant context information is available. In support of this dual-component model, Wiese et al. (in press) have shown that when targets were presented in an unstructured visual field, cueing was not specific to the exact gazed-at position, but facilitated all positions within the cued hemifield to an equal degree. However, when additional context information was provided in the form of peripheral position placeholders, cueing effects were strongest for the exact gazed-at location. The authors took this pattern to indicate that bottom-up and top-down mechanisms are co-active in gaze cueing: while the bottom-up (*reflexive*) component causes a general directional bias for the whole cued hemifield, the top-down component triggers facilitation specific to the particular gazed-at position.

The present study was designed to investigate whether gaze-induced attentional orienting can be top-down modulated by participants’ expectations about observed gaze behavior. Expectations were induced by either *actual predictivity* of gaze behavior (i.e., likelihood with which targets appeared at gazed-at locations) or *instructed predictivity* (independent of the actual predictivity). In Experiment 1, actual (i.e., experienced) predictivity tallied with instructed (i.e., believed) predictivity, whereas in Experiment 2 it did not. Based on the two-component model of Wiese et al. (in press), we expected that in Experiment 1, the predictivity with which gaze direction indicates target position would influence the specificity of gaze cueing as follows: nonpredictive gaze behavior would only activate the bottom-up component, resulting in equal-size gaze-cueing effects for the whole hemifield; predictive gaze behavior, by contrast, would additionally invoke the top-down component giving rise to facilitation for the exact gazed-at position. In Experiment 2, we examined whether believed predictivity would interact with experienced predictivity, that is: would nonspecific gaze cueing triggered by nonpredictive cues be spatially more specific when the cue was believed to be predictive (Experiment 2), relative to when it was believed to be nonpredictive (Experiment 1); analogously, would specific gaze cueing induced by predictive cues be less specific when the gaze-cue was believed to be nonpredictive (Experiment 2) compared to when it was believed to be predictive (Experiment 1)?

2.2.3. Experiments

Experiment 1

In Experiment 1, gaze cues either predicted the target location with a high likelihood (80%), or they were nonpredictive (≈ 17%). Participants were explicitly informed about these probabilities. There were three semi-circularly arranged target positions in each hemifield (Figure 1A). Importantly, these positions were *not* marked by placeholders (see Wiese et al., in press, for the effects of nonpredictive gaze cues without versus with placeholder information). Participants had to make a speeded localization (left vs. right hemifield) response to the target. We expected predictive gaze cues to produce the strongest gaze-cueing effect for the exact gazed-at position, whereas nonpredictive cues would generate equal effects for all target positions within the cued hemifield.

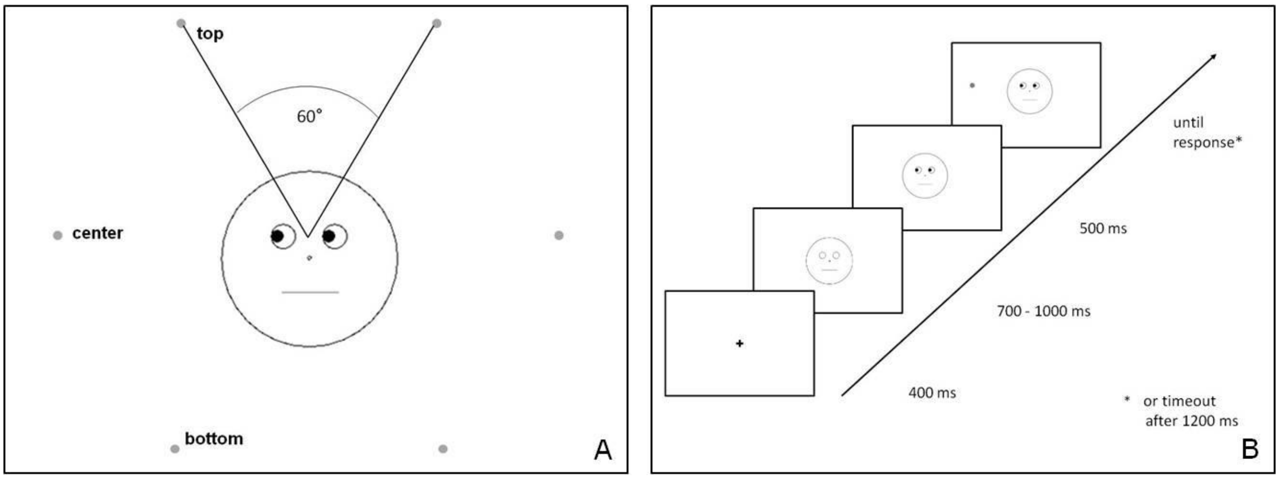

Figure 1. Stimulus and target positions (A) and sequence of events within a trial (B).

Methods and Materials

Participants

Twelve volunteers (8 women; mean age: 25 years, range: 20–30 years; all right-handed, with normal or corrected-to-normal visual acuity; all having given written informed consent) participated in the experiment either for course credit or payment (8€/h). Testing time was two hours, split into two sessions.

Apparatus

Stimuli were presented on a 17” Graphics Series G90fB CRT monitor with the refresh rate of 85 Hz. Reaction time (RT) measures were based on standard keyboard responses. The experiment was controlled by *Experiment Builder* (SR Research Ltd., Ontario, Canada). Participants were seated 57 cm away from the monitor, centered with respect to display and keyboard.

Stimuli

Schematic faces, constructed in line with Friesen and Kingstone (1998), were presented in the center of the display as black drawings against a white background. The round face outline circumscribed an area of 6.8° of visual angle and contained two circles representing the eyes, a smaller circle symbolizing the nose, and a straight line representing the mouth. The eyes subtended 1.0° and were positioned on the horizontal midline, at a distance of ±1.0° from the vertical midline. The nose subtended 0.2°, was located 0.9° below the eyes, and served as fixation point. The mouth was 2.2° in length and centered 1.3° below the nose. Black filled circles, subtending 0.5°, appeared within the eyes, representing the pupils. Gaze cues were implemented by moving the pupils sideways into one of six different directions: pupils were either shifted left- or rightwards on the central horizontal axis or rotated up- or downwards relative to the midline by an angle of 60°, until they touched the outline eye circles. The target stimulus was a gray dot 0.5° in diameter. Targets could appear at one of six positions equally distributed on an imaginary circle with a radius of 6.0° around the fixation point within the central face (Figure 1A). The angular distance between adjacent targets was 60°.

Design

Each session of the experiment consisted of 740 trials, with a block of 20 practice trials preceding 20 experimental blocks of 36 trials each. Gaze direction (left, right), gaze position (top, center, bottom), target side (left, right), and target position (top, center, bottom) were presented pseudo-randomly. Cue predictivity was blocked: one testing session was devoted to nonpredictive and the other to predictive cues, with session order counterbalanced across participants. In the ***nonpredictive condition***, targets appeared at each of the six target positions with the same likelihood (≈17%); by contrast, in the ***predictive condition***, targets appeared with a likelihood of 80% at the exact gazed-at position and a likelihood of 4% each at one of the other five positions.

Procedure

Figure 1B illustrates the sequence of events on a trial. Trials started with the onset of a central fixation cross. 400 ms later, a face with blank eyes was presented. After a random interval of 700–1000 ms, pupils appeared within the eyes looking at one of the six target positions (Figure 1A). Following the cue, a target dot appeared at one of the six target positions at a stimulus onset asynchrony (SOA) of 500 ms. Schematic face, pupils, and target remained on the screen until a response was given or 1200 ms had elapsed. Participants were asked to determine, as fast and accurately as possible, whether targets were presented on the left or right side of the screen, pressing the “D”- or “K”-key with their left or right index finger for a target on the left or right side, respectively. The inter-trial-interval (ITI) was 680 ms.

Participants were veridically informed about the predictivity of the gaze cues: ***Instruction 1*** informed that gaze direction was not predictive of the location of the upcoming target; ***Instruction 2*** informed that the target would appear with a high likelihood at the gazed-at position.

Analysis

The specificity of gaze cueing was assessed in a repeated-measures analysis of variance (ANOVA) on the gaze-cueing effects, with the factors *gaze position* (top, center, bottom), *target position* (top, center, bottom), and *predictivity* (low, high).[1] Cueing effects were calculated as the RT-difference between a validly cued position (gaze direction and target side matched) and the respective invalidly cued position (gaze direction and target side did not match) on the same horizontal axis. For instance, cueing effects for the top-position (60° in the upper quadrant) on the left side were calculated as the RT difference between trials on which this position was validly cued (i.e., gaze directed to the left) compared to when this position was invalidly cued (i.e., gaze directed to the right). For the ANOVA, cueing effects were collapsed across the two hemifields. Specific cueing effects would manifest as a significant interaction between gaze position and target position, with stronger cueing effects for the gazed-at position than for the other positions in the same hemifield. By contrast, nonspecific gaze cueing would yield equal facilitation for all positions in the cued hemifield (i.e., main effect of validity[1], in the absence of a gaze position x target position interaction on the gaze-cueing effects). If predictivity influenced the specificity of gaze cueing, the interaction among predictivity, gaze position, and target position should be significant, with the interaction between gaze position and target position being significant only for predictive cues.

Results

Anticipations (defined as responses with latency < 100 ms, 1.29%), misses (defined as responses with latency > 1200 ms, 3.69%), and incorrect responses (1.49%) were excluded from analysis. Figure 2 presents the cueing effects for predictive and nonpredictive trials as a function of gaze position and target position. Mean RTs and Standard Errors for valid and invalid trials are reported in the SI-section (see Table S1); and results of the ANOVA with the factors validity, gaze position, target position, and predictivity are summarized in Tables S2 and S3. The results of the ANOVA on the gaze-cueing effects are reported below.

The ANOVA of the cueing effects revealed the gaze-cueing effects to be overall larger with predictive cues (Δ_{RT} = 61 ms) than with nonpredictive cues (Δ_{RT} = 11 ms) [predictivity: $F(1,11)= 44.716$, $p< .001$, $\eta^2= .803$]. Moreover, the spatial distribution of the gaze-cueing effects was dependent on the relation of the gazed position to the actual target position in the cued hemifield [gaze position x target position: $F(1,11)= 44.716$, $p< .001$, $\eta^2= .803$]. Importantly, however, the spatial distribution of cueing effects differed significantly between predictive and nonpredictive cues [predictivity x gaze position x target position: $F(4,44)= 15.265$, $p< .001$, $\eta^2= .581$]. All other effects were non-significant (all *F*s < 2.543, all *p*s > .101).

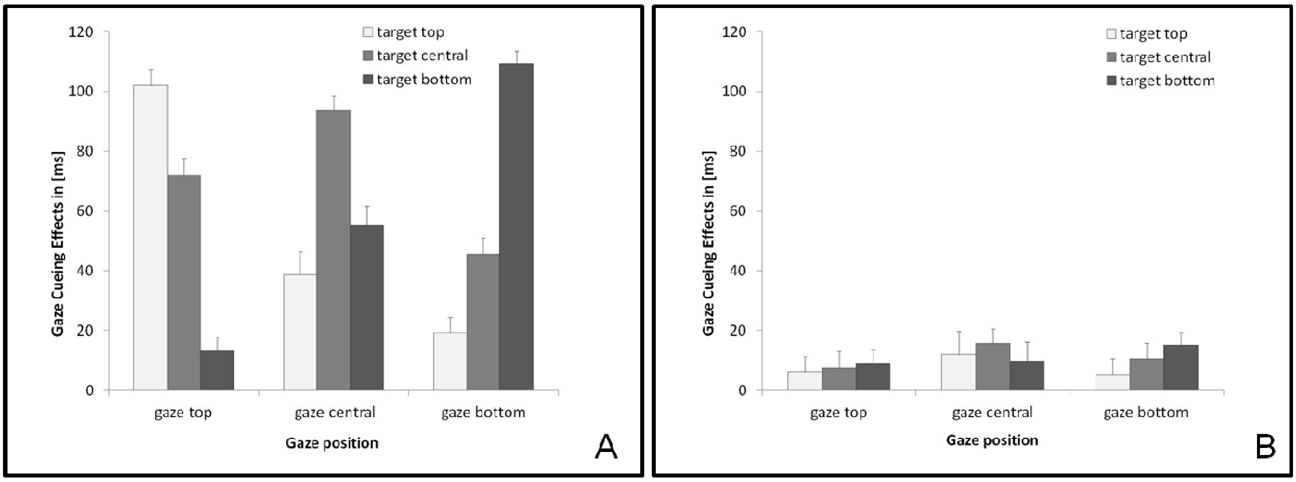

Figure 2. Gaze-cueing effects as function of gaze and target position for (A) **high actual** and **instructed** predictivity; for (B) **low actual** and **instructed** predictivity. Depicted error bars represent standard errors of the mean adjusted to the within-participants design (Cousineau, 2005).

To statistically test whether the spatially specific component manifested only with predictive, but not with nonpredictive, cues, the cueing effects were examined in follow-up ANOVAs with only the factors *gaze position* (top, center, bottom) and *target position* (top, center, bottom), conducted separately for each of the predictivity conditions. With nonpredictive cues, the cueing effects were of comparable size for all target positions in the cued hemifield [gaze position x target position: $F(4,44)= 1.078$, $p= .379$, $\eta^2= .088$]; see Table S3 for the main effect of validity. By contrast, with predictive cues, the size of the gaze-cueing effect depended on the congruency of the gazed-at and the target position [gaze position x target position: $F(4,44)= 18.309$, $p< .001$, $\eta^2= .625$]. All other effects were non-significant (all *F*s < 1.973, all *p*s > .163).

To examine the deployment of attention with respect to particular positions in the gaze-cued hemifield for the predictive condition, cueing effects were compared between the exact gazed-at position and the other two locations (averaged together) in the cued hemifield; this comparison confirmed the gaze-cueing effects to be significantly larger for the exact cued position than for the other two locations ($\Delta_{\text{GCexact-other}}= 61$ ms, $t(11)= 6.111$, $p< .001$, *two-tailed*).[2]

Discussion

Experiment 1 investigated whether attentional orienting is influenced by explicit information about the predictivity of gaze behavior. The results showed that for *predictive cues*, gaze cueing was significantly stronger for targets that appeared at the exact gazed-at position, relative to targets at one of the other two positions in the cued hemifield. *Nonpredictive cues*, by contrast, generated significant gaze-cueing effects (see Table S3) that were equally strong for all target positions within the cued hemifield.

The finding that predictivity influences both the size and spatial distribution of gaze-cueing effects raises an interesting question, namely: is the observed pattern mediated by instruction-induced expectations, or does it emerge as a result of acquired experience with cues of various degrees of predictivity? The results of Experiment 1 cannot answer this question, as the experienced (=*actual*) and believed (=*instructed*) predictivity were always congruent. In order to disentangle the effects of experience versus belief, instructed and actual predictivity must be varied orthogonally: With highly predictive cues, participants would be told that the gaze cue is nonpredictive; with nonpredictive cues, by contrast, participants would be told that the cue is highly predictive.

Experiment 2

In Experiment 2, the effects of actual and believed predictivity were contrasted. Participants received either ***Instruction 1***: they were told that the cue was highly predictive, when it actually was nonpredictive (*actual predictivity*: 17%; *instructed predictivity*: 80%); or ***Instruction 2***: they were told that the cue was nonpredictive, when it actually was highly predictive (*actual predictivity*: 80%, *instructed predictivity*: 17%). The order of instruction was counterbalanced across participants. In order to examine the influence of experienced vs. believed predictivity on gaze cues, we compared conditions with the same actual but different instructed predictivities between experiments in a mixed four-way ANOVA on cueing effects with the within-participant factors *gaze position* (top, center, bottom), *target position* (top, center, bottom), and *actual predictivity* (high, low), and the between-participant factor *experiment* (Exp.1: experience congruent with instruction, Exp.2: experience incongruent with instruction).

Method

Methods in Experiment 2 were the same as in Experiment 1, with the only exception of congruency between actual and instructed predictivity (congruent in Experiment 1, incongruent in Experiment 2).

Participants

Twelve new volunteers (10 women; mean age: 25 years, range: 20–28 years; all right-handed, all with normal or corrected-to-normal visual acuity; all having given written informed consent) participated in Experiment 2, either for course credit or payment (8€/h).

Results and Discussion

Anticipations (0.82 %), misses (0.09 %), and incorrect responses (3.86%) were excluded from analysis.

Mean RTs and Standard Errors for valid and invalid trials are reported in the SI-section (see Table S4). ANOVA-effects of interest are reported below; complete results are summarized in Table S5.

The ANOVA of the cueing effects revealed the actual cue predictivity to influence the allocation of spatial attention induced by gaze cues (see Figure 3): cues with high actual predictivity gave rise to larger cueing effects than nonpredictive cues [actual predictivity: $F(1,22)= 64.975$, $p< .001$, $\eta^2= .803$]. Moreover, highly predictive cues generated cueing effects specific to the gazed-at position [actual predictivity x gaze position x target position: $F(4,88)= 15.130$, $p< .001$, $\eta^2= .407$], with significant differences between the exact cued versus the other positions: all *t*s> 2.295, *p*s< .031, *two-tailed*).

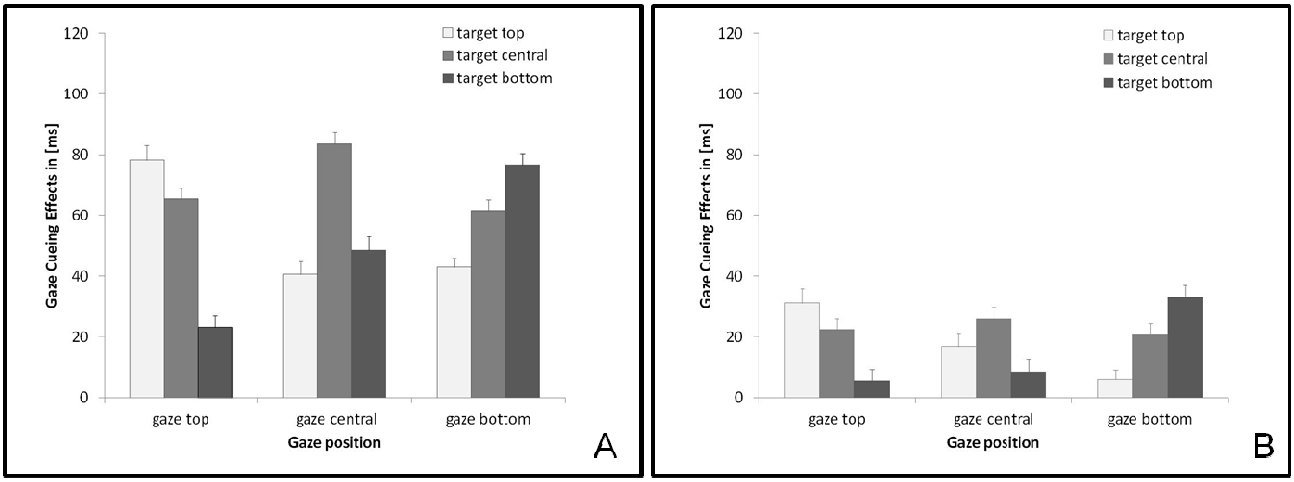

Figure 3. Gaze-cueing effects as function of gaze and target position for (A) **high actual** predictivity and **low instructed** predictivity; for (B) **low actual** predictivity and **high instructed** predictivity. Depicted error bars represent corrected standard errors of the mean adjusted to within-participants design (Cousineau, 2005).

Crucially, this pattern was modulated by believed predictivity [experiment x actual predictivity x gaze position x target position: $F(4,88) = 5.419$, $p = .001$, $\eta^2 = .198$], that is: the allocation of spatial attention in response to the experienced (i.e., actual) cue predictivity was top-down modulated by expectations based on the believed (i.e., instructed) cue predictivity – see Figure 4.

In subsequent analyses, the spatial specificity of gaze cueing and its modulation by instructed predictivity were analyzed for each actual predictivity condition separately. ***Nonpredictive cues*** generated spatially nonspecific cueing effects when participants believed that the cue was not predictive (Exp.1), whereas the same cues produced specific effects when participants believed the cues were predictive (Exp.2) [experiment x gaze position x target position: $F(4,88) = 5.649$, $p < .001$, $\eta^2 = .204$]. Planned comparisons revealed the cueing effects to be significantly larger for the exact gazed-at position than for the other positions within the cued hemifield when participants were told that the cues are predictive (Exp.2, $\Delta_{GCcued\text{-}other} = 17$ ms), compared to being told that they are nonpredictive (Exp.1, $\Delta_{GCcued\text{-}other} = 3$ ms); [$t(21) = 3.478$, $p = .002$, two-tailed]; see Figure 4A.[2] Similarly, believed predictivity modulated the spatial specificity of gaze cueing for ***predictive cues*** [experiment x gaze position x target position : *F*(4,88)= 2.583, *p*= .043, η^2= .105]: the spatially specific component was significantly stronger for cues believed to be predictive (Exp.1, $\Delta_{\text{GCcued-other}}$= 61 ms) compared to cues believed to be nonpredictive (Exp.2, $\Delta_{\text{GCcued-other}}$= 32 ms), [*t(*21)= -2.216, *p*= .037, *two-tailed*]; see Figure 4B.[2] Complete results are reported in Table S6.

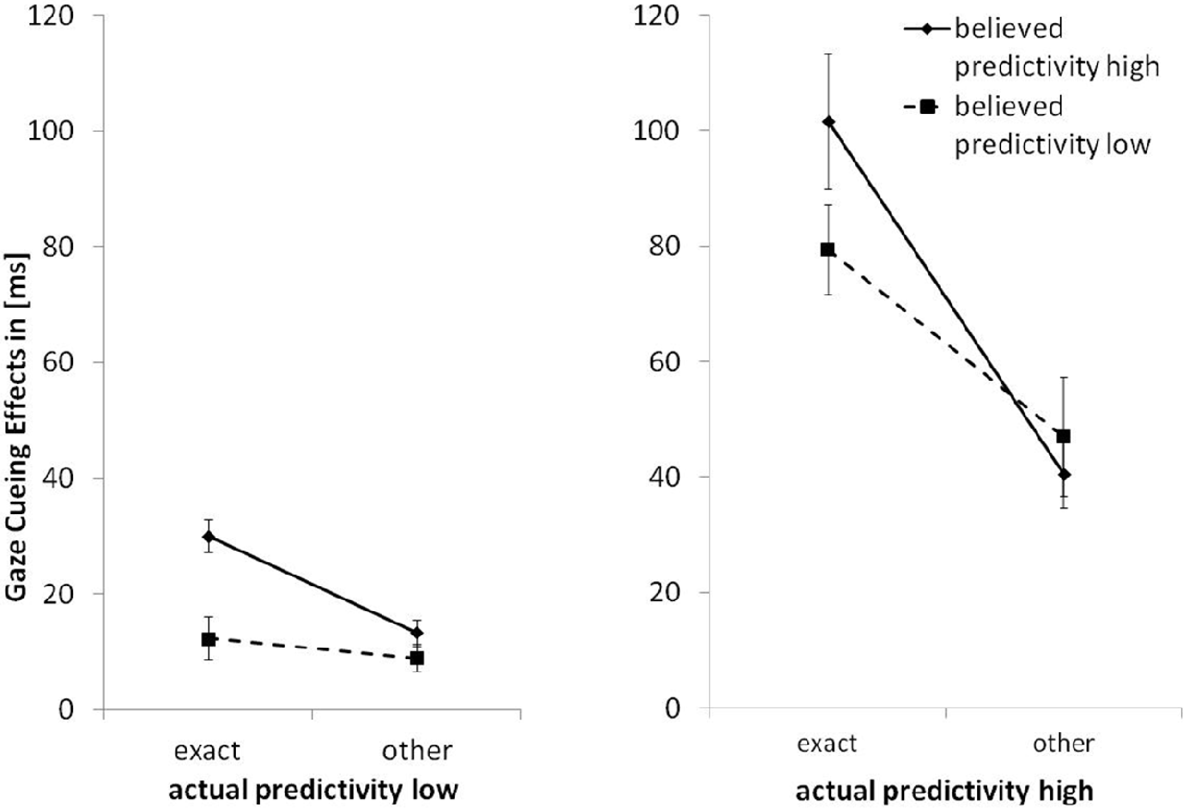

Figure 4. Comparison between Experiments. Gaze-cueing effects as function of target position (exact gazed-at position vs. other positions in cued hemifield), believed predictivity (high: ***solid line,*** low: ***dashed line***) and actual predictivity (high: ***left side,*** low: ***right side***). Note that the bigger the difference (the steeper the depicted line) between gaze-cueing effects for the exact and the other positions in the cued hemifield, the more specific the allocation of attention to the gazed-at position. Depicted error bars represent corrected standard errors adjusted to within-subject designs (see Cousineau, 2005).

2.2.4. General Discussion

The goal of the present study was to investigate whether fundamental mechanisms of social cognition such as orienting of attention in response to gaze direction are influenced by context information about the predictivity of gaze behavior. Information about predictivity was provided either explicitly by instruction, or could have been inferred from actual gaze behavior. Actual predictivity either did (Exp.1) or did not match instructed predictivity (Exp.2). When actual and instructed predictivity matched, specific cueing effects for the gazed-at location were expected in the predictive condition; and cueing effects for the whole gazed-at hemifield were expected in the nonpredictive condition. When actual and instructed predictivity did not match, believed predictivity was presumed to modulate the spatial allocation of gaze-cueing effects evoked by actual predictivity.

This pattern of results was predicated on Wiese et al. (in press), who showed that a general gaze-cueing effect for the whole gazed-at hemifield could be complemented by a cueing effect specific for the gazed-at position, though only in conditions in which peripheral position placeholders were presented that could be referred to by the central gaze. This pattern led the authors to propose a two-component model of gaze cueing, according to which spatially specific gaze-cueing effects are mediated by a context-dependent top-down component that is integrated with a bottom-up component producing a general directional bias towards the gaze-cued hemifield.

The present findings provide further support for this view, however, with context referring not to external, visual information, but to the internal (instructed / experienced) reliability of the cue information: with predictive cues, RT benefits were significantly larger for targets that appeared at the exact gazed-at position relative to targets at the other two positions within the cued hemifield; nonpredictive cues, by contrast, gave rise to facilitation effects of equivalent magnitude for all positions within the cued hemifield. Importantly, the effects of experienced predictivity were modulated by expected predictivity: that is, nonpredictive cues believed to be predictive caused cueing effects specific to the gazed-at position, compared to when nonpredictive cues were veridically instructed to be nonpredictive (Figure 4, left). In contrast, specific gaze-cueing effects caused by actually predictive cues were significantly reduced when the cue was believed to be nonpredictive (Figure 4, right).

Thus, the present results extend previous findings by showing that gaze-cueing effects may not only be up-, but also down-regulated depending on the context information that is provided about the predictivity of the cues: a specific cueing effect caused by actually predictive cues is reduced in its spatial specificity when participants believe that the cue is nonpredictive; by the same token, spatially nonspecific gaze-cueing effects induced by actually nonpredictive cues yield increased spatial specificity when participants are told that the cue is predictive (Figures 2, 3, and 4). This is consistent with the view that top-down modulation of the spatial distribution of gaze cueing effects can be induced by various types of context information: visual information provided in the scene (e.g., placeholders) or purely verbal information (provided by instruction) about observed behavior.

The observation that explicit knowledge about *who* we are interacting with does influence basic attentional processes involved in social interactions is consistent with Frith and Frith (2006) and Teufel et al. (2010), who have suggested that bottom-up orienting to gaze direction can be top-down controlled by contextual information about the gazer. Our results are also in line with empirical evidence showing that background information based on familiarity with the gazer (gender effect for women only: Deaner et al., 2007) or stereotypical information about the gazer (member of a political party: Liuzza et al., 2011) modulate the size of gaze-cueing effects. Only Bayliss and Tipper (2006) reported findings that might be interpreted as being at variance with the present results: they did not observe a modulation of gaze-cueing effects by cue predictivity. In their paradigm, different faces (randomized across trials) indicated the target location with different validities, and the association of facial identity with cue predictivity had to be learned over the course of the experiment.[3] The lack of a predictivity-related modulation of gaze cueing in their study might have been due to the fact that information about the predictivity of gaze behavior was coupled with facial identity. Critically, Hoffman and Haxby (2000) have shown that the interpretation of gaze direction and facial identity are subserved by different neural networks and that these components are integrated only at later stages of information processing. Given that spatial cueing produces relatively fast-acting effects on attentional orienting, it is likely that cueing studies fail to disclose effects of slower-acting facial identity information on the response to gaze cues.

In summary, our findings show that early operations of spatial attention are highly penetrable by cognitive processes related to social context. The involvement of a context-modulated mechanism in gaze cueing is very plausible, as gaze-triggered mechanisms of attention are specifically sensitive to the social relevance of the environment within which they operate: the bottom-up component assures a general preparedness for social signals conveyed by other people, while the top-down mechanism allows a flexible adaptation to the social context of a scene. The present findings reveal that in integrating context information within social attention mechanisms, humans tend to incorporate what they are told about others into their own experience and observation.

2.2.5. References

Adolphs, R. (1999). Social cognition and the human brain. *TICS*, *3(12)*, 469-479.

Bayliss, A.P., & Tipper, S.P. (2006). Predictive gaze cues and personality judgments: should eyes trust you? *Psychological Science*, *17(6)*, 514-520.

Brothers, L. (1990). The social brain: s project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience*, *1*, 27-51.

Deaner, R.O., Shepherd, S.V., & Platt, M.L. (2007). Familiarity accentuates gaze cuing in women but not in men. *Biology Letters*, *3*, 64-67.

Downing, C.J., & Pinker, S. (1985). The spatial structure of visual attention. In M.I. Posner & O.S.M. Marin (Eds.), *Attention and performance XI*, pp.171-188. Hillsdale, NJ: Erlbaum.

Eriksen, C.W., & St.James, J.D. (1986). Visual attention within and around the field of focal attention: A zoom lens model, *Perception & Psychophysics*, *40*, 225-240.

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

Friesen, C.K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *JEP:HPP, 30*, 319-329.

Frischen, A., Bayliss, A.P., & Tipper, S.P. (2007). Gaze Cuing of Attention: Visual Attention, Social Cognition, and Individual Differences. *Psychological Bulletin*, *133(4)*, 694-724.

Frith, U., & Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society. Series B: Biological Sciences, 358(1431)*, 459-473.

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

Grezes, J., Frith, C.D., & Passingham, R.E. (2004a). Brain mechanisms for inferring deceit in the actions of others. *Journal of Neuroscience*, *24(24)*, 5500-5505.

Grezes, J., Frith, C.D., & Passingham, R.E. (2004b). Inferring false beliefs from the actions of oneself and others: an fMRI study. *NeuroImage*, *21(2)*, 744-750.

Henderson, J.M. (1991). Stimulus discrimination following covert attentional orienting to an exogenous cue. *JEP:HPP*, *17*, 91-106.

Hoffman, E.A., & Haxby, J.V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3(1)*, 80-84.

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

Hughes, H.C., & Zimba, L.D. (1985). Spatial maps of directed visual attention. *JEP:HPP*, *11*, 409-430.

Hughes, H.C., & Zimba, L.D. (1987). Natural boundaries for spread of directed visual attention. *Neuropsychologia*, *2*, 5-18.

Langton, S.R.H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, *6*, 541-567.

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)*, 1-7.

Müller, H.J., & Rabbitt, P.M. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *JEP:HPP*, *15(2)*, 315-330.

Nummenmaa, L., & Hietanen, J.K. (2009). How attentional systems process conflicting cues. The superiority of social over symbolic orienting revisited. *JEP:HPP*, *35(6)*, 1738-1754.

Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3-25.

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

Ristic, J., & Kingstone, A. (2005). Taking control of reflexive social attention. *Cognition*, *95*, B55-B65.

Ristic, J., & Kingstone, A. (2006). Attention to arrows: Pointing to a new direction. *The Quarterly Journal of Experimental Psychology*, *59(11)*, 1921-1930.

Saxe, R. & Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*, *43(10)*, 1391-1399.

Shulman, G.L., Wilson, J., & Sheehy, J.B. (1985). Spatial determinants of the distribution of attention. *Perception & Psychophysics*, *37*, 59-65.

Teufel, C., Alexis, D.M., Todd, H., Lawrence-Owen, A.J., Clayton, N.S., & Davis, G. (2009). Social cognition modulates sensory coding of observed gaze direction, *Current Biology*, *19(15)*, 1274-1277.

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, 314-318.

Vecera, S.P., & Rizzo, M. (2006). Eye gaze does not produce reflexive shifts of attention: Evidence from frontal-lobe damage. *Neuropsychologia*, *44(1),* 150-159.

Wiese, E., Zwickel, J., & Müller, H.J. (in press). The importance of context information for the spatial specificity of gaze cuing. *Attention, Perception & Psychophysics,* doi: 10.3758/s13414-013-0444-y.

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.*

Footnotes

1. Note that we also ran an ANOVA on mean RTs with the factors *validity* (valid, invalid), *gaze position* (top, center, bottom), *target position* (top, center, bottom), and *predictivity* (low, high), in order to examine whether the basic cueing effects were significant. Mean RTs, standard errors, and ANOVA results are reported in Tables S1 and S2 in the Supplementary Materials sections. Results of post-hoc ANOVAs on RTs with the factors *validity* (valid, invalid), *gaze position* (top, center, bottom), *target position* (top, center, bottom) for each predictivity separately are reported in Table S3.
2. All t-tests were Bonferroni-corrected for multiple comparisons.
3. In our paradigm, facial identity was not varied throughout the experiment (the same schematic face was used in all conditions) and cue predictivity was blocked. Further, participants were provided with explicit information about the observed gaze behavior via instruction. Our results indicate that although knowledge acquired through experience with given gaze behavior influences the interpretation of social scenes, explicit knowledge is essential to modulate early attentional processes involved in gaze cueing.

2.2.6. Supplementary Materials

Table S1. Mean Response Times and Standard Errors (in ms) for **actual and instructed** predictivity high vs. low (*Exp. 1*).

		actual and instructed predictivity high			actual and instructed predictivity low		
		Gaze top	Gaze central	Gaze bottom	Gaze top	Gaze central	Gaze bottom
Target top	valid	279 (10)	339 (24)	378 (16)	327 (13)	328 (14)	335 (13)
	invalid	381 (10)	378 (21)	398 (10)	333 (12)	341 (12)	341 (12)
Target central	valid	306 (11)	261 (15)	320 (13)	310 (12)	304 (10)	312 (12)
	invalid	378 (12)	354 (17)	365 (16)	318 (11)	319 (10)	323 (11)
Target bottom	valid	355 (12)	325 (24)	273 (10)	324 (12)	323 (12)	314 (13)
	invalid	368 (19)	380 (21)	382 (10)	333 (12)	333 (12)	328 (12)

Table S2. F-and p-values for the four-way ANOVA on **RTs** with the factors: validity, gaze position, target position and predictivity (actual and instructed predictivity congruent, *Exp.1*).

	F-value	p-value	effect size η^2
validity	$F(1,11)= 109.437$	$p< .001$	$\eta^2= .909$
target position	$F(2,22)= 49.954$	$p< .001$	$\eta^2= .820$
gaze position	$F(2,22)= 1.090$	$p= .354$	$\eta^2= .090$
predictivity	$F(1,11)= 1.057$	$p= .326$	$\eta^2= .088$
validity x target position	$F(2,22)= 2.543$	$p= .101$	$\eta^2= .188$
validity x gaze position	$F(2,22)= .526$	$p= .598$	$\eta^2= .046$
validity x predictivity	$F(1,11)= 44.716$	$p< .001$	$\eta^2= .803$
gaze position x target position	$F(4,44)= 14.766$	$p< .001$	$\eta^2= .573$
gaze position x predictivity	$F(2,22)= .778$	$p= .472$	$\eta^2= .066$
target position x predictivity	$F(2,22)= 1.247$	$p= .307$	$\eta^2= .102$
validity x gaze pos x target pos	$F(4,44)= 18.716$	$p< .001$	$\eta^2= .630$
validity x gaze pos x predictivity	$F(2,22)= .647$	$p= .533$	$\eta^2= .056$
validity x target pos x predictivity	$F(2,22)= 1.268$	$p= .301$	$\eta^2= .103$
gaze pos x target pos x predictivity	$F(4,44)= 8.560$	$p< .001$	$\eta^2= .438$
validity x gaze pos x target pos x predictivity	$F(4,44)= 15.265$	$p< .001$	$\eta^2= .581$

Table S3. F-values and p-values for the post-hoc ANOVAs on **RTs** with the factors: validity, gaze position, and target position for each actual predictivity condition separately (*Exp.1*).

	actual and instructed predictivity high			actual and instructed predictivity low		
	F-value	*p*-value	effect size η^2	*F*-value	*p*-value	effect size η^2
validity	$F(1,11)= 79.447$	$p< .001$	$\eta^2= .878$	$F(1,11)= 19.413$	$p= .001$	$\eta^2= .638$
target position	$F(2,22)= 20.016$	$p< .001$	$\eta^2= .645$	$F(2,22)= 32.301$	$p< .001$	$\eta^2= .746$
gaze position	$F(2,22)= .948$	$p= .403$	$\eta^2= .079$	$F(2,22)= .518$	$p= .603$	$\eta^2= .045$
validity x target position	$F(2,22)= 1.973$	$p= .163$	$\eta^2= .152$	$F(2,22)= 1.126$	$p= .342$	$\eta^2= .093$
validity x gaze position	$F(2,22)= .346$	$p= .711$	$\eta^2= .031$	$F(2,22)= 1.864$	$p= .179$	$\eta^2= .145$
target position x gaze position	$F(4,44)= 12.276$	$p< .001$	$\eta^2= .527$	$F(4,44)= 5.470$	$p= .001$	$\eta^2= .332$
validity x gaze pos x target pos	$F(4,44)= 18.309$	$p< .001$	$\eta^2= .625$	$F(4,44)= 1.078$	$p= .379$	$\eta^2= .089$

Table S4. Mean Response Times and Standard Errors (in ms) for **actual** predictivity low/**believed** predictivity high vs. **actual** predictivity high/**believed** predictivity low (*Exp. 2*).

		actual pred. low/ believed pred. high			actual pred. high/ believed pred. low		
		Gaze top	Gaze central	Gaze bottom	Gaze top	Gaze central	Gaze bottom
Target top	valid	280 (13)	321 (20)	338 (19)	315 (14)	330 (14)	339 (13)
	invalid	358 (11)	361 (13)	381 (15)	345 (12)	347 (13)	345 (12)
Target central	valid	292 (15)	275 (12)	301 (18)	311 (13)	306 (14)	312 (11)
	invalid	357 (12)	358 (10)	363 (11)	333 (13)	331 (10)	333 (10)
Target bottom	valid	337 (22)	312 (19)	273 (13)	337 (13)	332 (14)	309 (12)
	invalid	360 (12)	361 (11)	349 (10)	342 (12)	340 (11)	342 (10)

Table S5. F-values and p-values for the four-way ANOVA on **gaze-cueing effects** with the factors: gaze position, target position, **actual** predictivity and experiment (**instructed** predictivity, *Exp. 2*).

	F-value	*p*-value	effect size η^2
actual predictivity	*F*(1,22)= 64.975	$p < .001$	$\eta^2 = .747$
target position	*F*(2,44)= 7.581	$p = .001$	$\eta^2 = .256$
gaze position	*F*(2,44)= .092	$p = .913$	$\eta^2 = .004$
experiment	*F*(1,22)= .222	$p = .642$	$\eta^2 = .010$
actual pred x experiment	*F*(1,22)= 1.146	$p = .296$	$\eta^2 = .049$
actual pred x gaze position	*F*(2,44)= .047	$p = .954$	$\eta^2 = .002$
actual pred x target position	*F*(2,44)= 3.715	$p = .032$	$\eta^2 = .144$
experiment x gaze position	*F*(2,44)= .612	$p = .547$	$\eta^2 = .027$
experiment x target position	*F*(2,44)= 1.278	$p = .289$	$\eta^2 = .055$
gaze position x target position	*F*(4,88)= 29.955	$p < .001$	$\eta^2 = .577$
actual pred x gaze pos x target pos	*F*(4,88)= 15.130	$p < .001$	$\eta^2 = .407$
experiment x gaze pos x target pos	*F*(4,88)= .634	$p = .639$	$\eta^2 = .028$
actual pred x experiment x gaze pos	*F*(2,44)= .766	$p = .471$	$\eta^2 = .034$
actual pred x experiment x target pos	*F*(2,44)= .097	$p = .908$	$\eta^2 = .004$
actual pred x experiment x target pos x gaze pos	*F*(4,88)= 5.419	$p = .001$	$\eta^2 = .198$

Table S6. F-values and p-values for the post-hoc ANOVAs on **gaze-cueing effects** with the factors: gaze position, target position, and experiment (instructed predictivity) for each actual predictivity condition separately (*Exp.2*).

	actual predictivity high			actual predictivity low		
	F-value	*p*-value	effect size η^2	*F*-value	*p*-value	effect size η^2
experiment	*F*(1,22)= .080	*p*= .780	η^2= .004	*F*(1,22)= 7.051	*p*= .014	η^2= .243
target position	*F*(2,44)= 6.291	*p*= .004	η^2= .222	*F*(2,44)= 2.968	*p*= .062	η^2= .119
gaze position	*F*(2,44)= .024	*p*= .976	η^2= .001	*F*(2,44)= .401	*p*= .672	η^2= .018
experiment x target position	*F*(2,44)= .595	*p*= .556	η^2= .026	*F*(2,44)= 2.153	*p*= .128	η^2= .089
experiment x gaze position	*F*(2,44)= .469	*p*= .629	η^2= .021	*F*(2,44)= 2.152	*p*= .128	η^2= .089
target position x gaze position	*F*(4,88)= 24.200	*p*< .001	η^2= .524	*F*(4,88)= 11.728	*p*< .001	η^2= .348
experiment x gaze pos x target pos	*F*(4,88)= 2.583	*p*= .043	η^2= .105	*F*(4,88)= 5.649	*p*< .001	η^2= .204

2.3. Study 3

I see what you mean - How attentional selection is shaped by ascribing intentions to others

Eva Wiese[1*], Agnieszka Wykowska[1*], Jan Zwickel[1], & Hermann J. Müller [1,2]

[1]LMU, Munich Germany

[2]Birkbeck College, University of London, UK

Author contributions: EW, JZ, HJM designed research. EW performed experiments. EW and AW analyzed the data. EW, AW, HJM wrote the paper. AW theoretical conceptualization.

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. doi: 10.1371/journal.pone.0045391.

*equal contribution

2.3.1. Abstract

The ability to understand and predict others’ behavior is essential for successful interactions. When making predictions about what other humans will do, we treat them as intentional systems and adopt the *intentional stance,* i.e., refer to their mental states such as desires and intentions. In the present experiments we investigated whether the *mere belief* that the observed agent is an intentional system influences basic social attention mechanisms. We presented pictures of a human and a robot face in a gaze cueing paradigm, and manipulated the likelihood of adopting the intentional stance by instruction: in some conditions, participants were told that they were observing a human or a robot, in others, that they were observing a human-like mannequin or a robot whose eyes were controlled by a human. In conditions in which participants were made to believe they were observing human behavior (*intentional stance likely*) gaze cueing effects were significantly larger as compared to conditions when adopting the intentional stance was less likely. This effect was independent of whether a human or a robot face was presented. Therefore, we conclude that adopting the intentional stance when observing others’ behavior fundamentally influences basic mechanisms of social attention. The present results provide striking evidence that high-level cognitive processes, such as beliefs, modulate bottom-up mechanisms of attentional selection in a top-down manner.

Keywords

Gaze Cueing, Intentional Stance, Social Attention, Theory of Mind, Human-Robot-Interaction

2.3.2. Introduction

Can we design machines that think? This old question has not yet been settled, despite the progress in the fields of artificial intelligence and cognitive science. For us humans, however, equally important questions are: would we be inclined to treat thinking artificial systems equally to other humans, and would we be ready to engage in social interactions with non-human agents that have minds? Alan Turing postulated that observed behavior is the only source of information based on which we ascribe minds to others (Turing, 1950). Accordingly, for us to engage in social interactions, it would be critical that we *perceive* other agents as thinking, whether or not they *actually* have a mind.

In this paper, we ask a fundamental question, namely: would the *mere belief* that an agent has a mind be sufficient to engage in interactions with him/her in a social way, compared to if he/she was believed to be just a machine. The belief that an agent has a mind might lead us to adopt the *intentional stance* (Dennett, 2003) towards him/her, which involves "treating the object whose behavior you want to predict as a rational *agent* with beliefs and desires and other mental states" (Dennett, 2003, p. 372). Dennett argues that the intentional stance is the best predictive strategy, given that the system whose behavior we want to predict is truly intentional. Throughout our lifelong experience with other humans, we have learned that humans are truly intentional systems in this sense – and, therefore, adopting the *intentional stance* towards human agents should be more successful in predicting their behavior as compared to adopting other strategies (e.g., the design or the physical stance).

Importantly, adopting the intentional stance towards other agents might not only be a successful predictive strategy, but also play a decisive role for one's own readiness to engage in social interactions. For instance, if I believe that you are pointing to a location with the intention of showing me something, I will be likely to direct my attention there; but I will be unlikely to attend to where a lever of a broken machine is pointing, as I will not interpret the lever's behavior as conveying communicative content (Grice, 1975; Tomasello, 2010). The present study was designed to investigate the fundamental issue of whether a belief concerning the minds of others has an impact on *basic* mechanisms underlying social cognition.

Understanding others' behavior in social interactions

On the neuronal level, interpreting others' intentions is supported by a system consisting of *medial prefrontal cortex* (mPFC), *superior temporal sulcus* (STS), *orbitofrontal cortex* (OFC), *amygdala* and *anterior insula* (Adolphs, 1999; Brothers, 1990; Frith & Frith, 2006). Furthermore, activation in the *anterior paracingulate cortex* has been reported to be related specifically to the adoption of the intentional stance (Gallagher, Jack, Roepstorff & Frith, 2002). At the cognitive level of description, Baron-Cohen (Baron-Cohen, 1995) has postulated that higher-level processes of mental-state attribution (*Theory of Mind*, ToM) are informed by low-level perceptual mechanisms: the *Intentionality Detector* (ID) and the *Eye-Direction Detector* (EDD).

Although the interpretation of social scenes involves low-level perceptual processes, it is unlikely that information is fed forward uni-directionally to higher-level processes of mental-state attribution, as everyday interactions require mechanisms that modulate perceptual information according to its social relevance. For these reasons, Teufel et al. (Teufel, Alexis, Todd, Lawrence-Owen, Clayton et al., 2009; Teufel, Fletcher & Davis, 2010a; Teufel, Alexis, Clayton & Davis, 2010b) proposed that low-level processes not only inform higher-level processes but are themselves modulated by the latter, that is: specifying particular mental states requires the integration of bottom-up information provided by the stimulus and top-down information reflecting various context variables. In support of this view, Teufel and colleagues (2009) reported that adaptation to gaze direction depended on whether participants believed that the observed person could actually or could not see through a pair of goggles.

Reading out others' mental states based on gaze direction

Gaze direction provides the basis for making inferences about the other's focus of interest, and encourages the observer to shift attention to the corresponding location (for review: Frischen, Bayliss & Tipper, 2007). Attention shifts triggered by gaze direction are typically investigated using a gaze-cueing paradigm (Friesen & Kingstone, 1998), in which a schematic face is presented centrally on the screen that gazes either straight ahead or to the left or right. Targets appearing in the gaze-cued hemifield are detected, localized, and discriminated more rapidly compared to targets in the uncued hemifield.

Orienting attention in response to perceived gaze direction has traditionally been regarded as a reflexive, bottom-up process involved in making inferences about the other's mental states (Nummenmaa & Calder, 2009). In line with this view, children as young as three months have been found to attend more quickly to peripheral objects that are gazed at by a human face compared to objects not directly gazed at (Hood, Willen & Driver, 1998). Moreover, adult participants have been found to reflexively orient attention in the direction of another's gaze even when gaze cues are counterpredictive of the target location (Driver, Davis, Ricciardelli, Kidd, Maxwell et al., 1999), whereas they voluntarily orient away from counterpredictive arrows (Friesen, Ristic & Kingstone, 2004) or extended tongues (Downing, Dodds & Bray, 2004).

However, although gaze-cueing has been shown to be triggered in a bottom-up fashion, a growing body of evidence suggests that attending to where others gaze is not purely reflexive, but can rather be modulated by higher-level cognitive processes. For instance, Ristic and Kingstone (2005) presented participants with an ambiguous stimulus that could be perceived as either a car with wheels or a face with a hat, and manipulated their beliefs by instruction; the stimulus was found to cue participants' attention only when they believed that they were looking at a face, rather than at a car. Similarly, Kawai (2011) found that schematic faces caused gaze-cueing effects only when participants believed that a potential target was visible to the gazer, but not when it was occluded.

Aim of study

The present study was designed to address a more fundamental issue, namely, whether adopting the intentional stance based on a mere belief concerning the observed agent would affect basic mechanisms of social attention. In classical studies examining the processes involved in inferring others' mental states (Leslie & Frith, 1988; Wimmer & Perner, 1983), participants are typically observing *intentional* agents and asked to provide a description of the agents' behavior making use of mentalistic vocabulary. Similarly, in the studies of Teufel and colleagues (2009, 2010a, 2010b), participants have always been observing intentional agents exhibiting particular mental states (such as being either able or unable to see through a pair of goggles). However, inferring particular mental states from observed behavior presupposes that the agent is construed as an intentional system that is *capable* of having intentions. Given this, based on the previous studies on that topic, one cannot decide whether basic mechanisms of social perception and attention were influenced by adopting the intentional stance *per se*, rather than by processes of reasoning about *particular* mental states of others. For resolving this fundamental issue, it is important to examine whether the likelihood of adopting the intentional stance *per se* (i.e., assuming that the observed agent has mental states) has an effect on one's own social cognition.

Thus, critically, we distinguish between processes of *mentalizing* about others' internal states and *adopting the intentional stance*. The former involves an active process of reasoning about mental states that underlie particular behavior, whereas the latter is, fundamentally, a result of activating a set of pre-existing representations of what it means to be "a human", which contains – amongst others – properties like *having a mind* or *being capable of having intentions*. By contrast, a representation of a mechanistic device (such as a robot) is probably devoid of mind-related characteristics. Thus, when predicting and/or explaining behavior of observed systems, humans either adopt the intentional stance or use other predictive strategies (such as the design or physical stances), dependent on the activated representations. In sum, adopting the intentional stance is based on a decision as to whether or not an observed agent is capable of having intentions; mentalizing, by contrast, is concerned with reasoning about what specific intentions are underlying behavior displayed by an agent that has already been classified as having a mind.

2.3.3. Experiments

In three experiments, we investigated whether social attention is modulated by the likelihood of adopting the intentional stance. To this end, we used two different types of gazers to orient participants' attention in a gaze-cueing paradigm: either a human face or a robot face (see *Figure 1A*). The likelihood of adopting the intentional stance was manipulated by instruction rather than by the appearance of the gazer (see Gallagher et al., 2002). In **Experiment 1**, participants were instructed (*Instruction 1*) that they would observe either a human (*intentional stance likely*) or a robot (*intentional stance unlikely*). In **Experiment 2**, *Instruction 2* informed participants that they were observing a human or a robot whose gaze behavior was controlled by a human (*intentional stance likely* in both cases), while in *Instruction 3* participants were told they would be observing a human-like mannequin or a robot (*intentional stance unlikely* in either case). In both experiments, participants had to perform a target discrimination task. To ensure that variations of gaze-cueing effects could not be attributed to physical differences between stimuli, the same stimuli were used for all experimental groups. On the intentional-stance hypothesis, we expected to find stronger gaze-cueing effects for stimuli representing *intentional agents* (Experiment. 1: *human*; Experiment 2, Instruction 2: *human* and *robot*) relative to stimuli representing agents who were less likely to be treated as intentional systems (Experiment 1: *robot*; Experiment 2, Instruction 3: *human* and *robot*). Please note that our paradigm did not involve actual social interaction. That is, participants observed photographs representing either intentional or non-intentional agents. Yet, we believe that it is still possible to adopt the intentional stance towards virtual humans. For example, when one watches a movie, one predicts behavior of characters depicted in that movie by adopting the intentional stance, even though one is not actually interacting with the characters. At the same time one might not adopt the intentional stance to a virtual robot due to pre-existing representations of what a robot is.

Experiment 3 examined whether the effects in question are generalizable to tasks with different attentional demands. The pattern of results revealed in Experiments 1 and 2 was replicated in Experiment 3 (localization task), which is reported in the supporting information (SI) section. Please see also *Figure S1* and *Table S1*.

Materials and Methods

Participants

24 participants participated in ***Experiment 1*** (15 women; mean age: 25 years (M=25.25, range: 19-32), two left-handed). In ***Experiment 2***, 48 participants were randomly assigned to two groups with different instructions: *Instruction 2* (human – human controlled robot): 16 women; mean age: 23 years (M= 23.17, range: 18-30) and *Instruction 3* (human-like mannequin, robot): 18 women; mean age: 23 years (M=22.58, range: 18-30). Participants reported normal or corrected-to-normal visual acuity. Testing time was about 40 minutes. One participant had to be excluded from Experiment 1 because of significantly increased error rates compared to other participants (M=16.7% compared to M=4.7%); one participant was excluded from Experiment 2 (Instruction 2) because the participant did not complete the experiment.

Ethics Statement

The experiments were conducted at the Department of Experimental Psychology at the LMU Munich, where all experimental procedures with purely behavioral data collection (e.g., RTs and error rates) of healthy adult participants, that do not include invasive or potentially dangerous methods are approved by the ethics committee of the Department of Psychology, LMU Munich, in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Data were stored and analyzed anonymously. Participants gave their informed consent and were either paid or received course credit for participating.

Apparatus

Stimuli were presented on a 17-in. Graphics Series G90fB monitor with the refresh rate set at 85 Hz. RT measures were based on standard keyboard responses. Participants were seated approximately 57 cm from the monitor, and the experimenter ensured that participants were centered with respect to the monitor. The experiment was controlled by *Experiment Builder* (SR Research Ltd., Ontario, Canada).

Stimuli

In the human condition, digitized photos of a female face (F 07) were used as stimuli, chosen from the *Karolinska Directed Emotional Faces* database (Lundqvist, Flykt & Öhman, 1998). We have received written informed consent (as outlined in the PloS consent form) from Karolinska Institute (Department of Clinical Neuroscience, Section Psychology) to use the photograph (F 07) for experimental investigation and illustration of the stimuli in publications. In the robot condition, photos of a humanoid robot (*EDDIE; developed by TU Munich*) were used. Stimuli were 6.4° wide and 10.0° high, depicted on a white background and presented in full frontal orientation with eyes positioned on the central horizontal axis of the screen (*Figure 1A*). For left- and rightward gaze, irises and pupils in the eyes were shifted with Photoshop$_{TM}$ and deviated 0.4° from direct gaze. The target stimulus was a black capital letter (F or T), measuring 0.8° in width and 1.3° in height. Targets appeared on the horizontal axis, located 6.0° from the center of the screen.

Procedure

Figure 1B illustrates the sequence of events in the present experiments. The beginning of every trial was signaled by a fixation cross at the center of the screen. 1000 ms later, a face with straight gaze appeared on the screen while the fixation cross remained in its position. After a random time interval of 700 to 1000 ms, gaze either remained straight or was shifted left- or rightwards. Following the gaze cue with a SOA of 500 ms, the target letter appeared either on the left or the right side of the screen. SOA was measured as the interval between the onset of the gaze shift to the onset of the target. Face and target remained on the screen until a response was given or after 1200 ms had elapsed. The inter-trial-interval (ITI) was 680 ms.

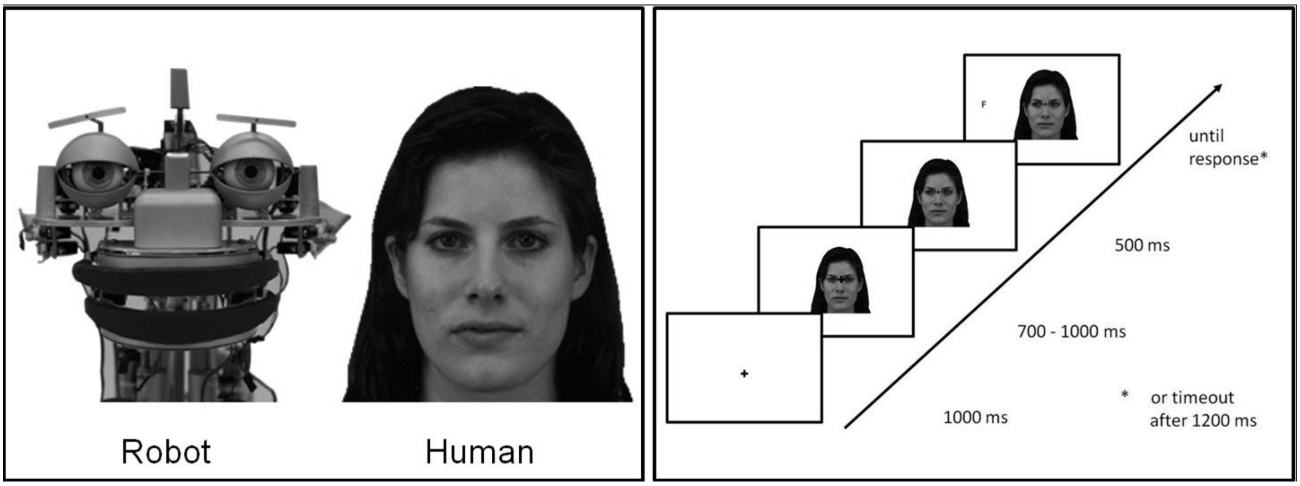

Figure 1. Stimuli and trial sequence: Pictures for the Robot and Human Gazer are shown in (A). The sequence of events within a trial is shown in (B). The human face (F 07) is taken from the Karolinska Directed Emotional Faces (KDEF) database (Lundqvist et al., 1998). We have received written informed consent (as outlined in the PLoS consent form) from Karolinska Institute (Department of Clinical Neuroscience, Section Psychology) to use the photograph for experimental investigations and illustration of the stimuli in publications. The picture of the robot face is made by LSR (TU Munich) and depicts the research robot EDDIE (made by LSR, TU Munich).

At the beginning of each session, participants were told to fix their gaze on a centrally presented cross. They were also instructed that after the fixation cross a photo of either a human or a robot would appear in the center of the screen but that they should still keep their eyes fixated on the fixation cross. Further, participants were advised that after its initial presentation the face gaze could remain straight or shifted left- or rightwards, subsequently followed by a target letter. Participants were asked to respond to target identity as quickly and as accurately as possible. For half of the participants F was assigned to the “D” key and T to the “K” key on the keyboard, for the other half of the participants stimulus-response mapping was reversed. The key labels were covered with a sticker to prevent letter interference effects. All instructions were given in written form and the experimenter was not informed about the purposes of this experiment.

Each session of the experiment was composed of 500 trials, with a block of 20 practice trials preceding 10 experimental blocks of 48 trials each. Gaze direction (straight, left, right), target side (left, right), target identity (F, T) and cue identity (human, robot) were selected pseudo-randomly and every combination appeared with equal frequency. Gaze direction was manipulated orthogonal to target position: that is, in one third of the trials, gaze was directed to the side on which the target appeared (*valid*), in another third of the trials to the other side (*invalid*) and in another third of the trials, the face was gazing straight ahead (*neutral*).

Analysis

Gaze-cueing effects were examined by comparing valid vs. invalid trials, i.e., in terms of costs-plus-benefits (invalid-valid) rather than benefits (neutral-valid) and costs (invalid-neutral) with respect to the neutral condition. This was done because the latter condition may not provide an adequate baseline measure for the separate assessment of cueing effects (Jonides & Mack, 1984). In fact, in all conditions of the present study, neutral trials were found to elicit longer RTs than valid and invalid trials – for two likely reasons: (i) straight-ahead gaze might have a holding effect on attention (George & Conty, 2008), making it difficult for the target onset to disengage attention and summon an orienting response (in line with Senju & Hasegawa, 2011); (ii) with straight-ahead gaze being maintained in the neutral condition, there was no similar temporal warning-signal effect to that induced by the gaze shift in valid and invalid conditions. In this regard, gaze cueing paradigms with naturalistic faces differ from those with schematic faces (Friesen & Kingstone, 1998): in the latter, trials typically start with face-like stimuli in which the eyes contain no pupils; pupils appear only later, so that also straight-ahead gaze involves a visual change that can serve as a temporal warning signal. This is not done with naturalistic faces (Bayliss & Tipper, 2006), as empty eyes without pupils are thought to be emotionally disturbing, potentially interfering with attentional orienting.

Results

Experiment 1

Missed (0.69%), and incorrect responses (3.69%), as well as RTs deviating by more than ±2.5 SD from individual participants' means were removed prior to analyses. The statistical analyses focused on the conditions of interest: mean RTs on valid versus invalid trials as a function of cue type (human vs. robot). Results of statistical analyses for all trial types (neutral, valid, invalid) are presented in *Table 1*, along with Mean RTs and Standard Errors. *Figure 2* depicts the corresponding gaze-cueing effects ($\Delta_{\mathrm{RT}\,\mathrm{invalid\text{-}valid}}$) for both types of cue; for the results of Experiment 1, see the top row in *Table 1* and the left-hand side of *Figure 2*.

Table 1. Mean RTs and SEM (in ms) as a function of cue validity and instruction, for human and robot cues.

	Human			Robot			Statistics
	Valid	Invalid	Neutral	Valid	Invalid	Neutral	Cue type x Validity
Instruction 1	447 (10)	461 (11)	479 (10)	451 (11)	456 (12)	481 (11)	$F(2,44)= 3.7$, $p< .05$
Instruction 2	454 (12)	469 (13)	482 (12)	453 (12)	467 (12)	484 (12)	$F(2,44)= 0.1$, $p> 0.8$
Instruction 3	451 (11)	455 (12)	474 (12)	450 (11)	454 (11)	477 (13)	$F(2,46)= 0.4$, $p> 0.6$

A two-way repeated-measures analysis of variance (ANOVA) of mean RTs with the factors cue validity (valid, invalid) and cue-type (human, robot) revealed the main effect of validity [$F(1,22)=$

22.131, $p < .001$, $\eta_p^2 = .501$] to be significant: valid trials yielded faster responses than invalid trials (449 ms vs. 458 ms). Importantly, the interaction between validity and cue-type was significant [$F(1,22) = 14.113$, $p < .002$, $\eta_p^2 = .391$]: under ***Instruction 1***, gaze-cueing effects were twice as strong for the human (Δ_{RT}= *14 ms*, [$t(22) = -5.954$, $p < .001$]) as for the robot (Δ_{RT}= *5 ms*, [$t(22) = -2.211$, $p < .04$]), though reliable in both conditions. The main effect of cue-type (454 ms for the human vs. 453 ms for the robot) was not significant [$F(1,22) = .017$ $p > .8$, $\eta_p^2 = .001$].

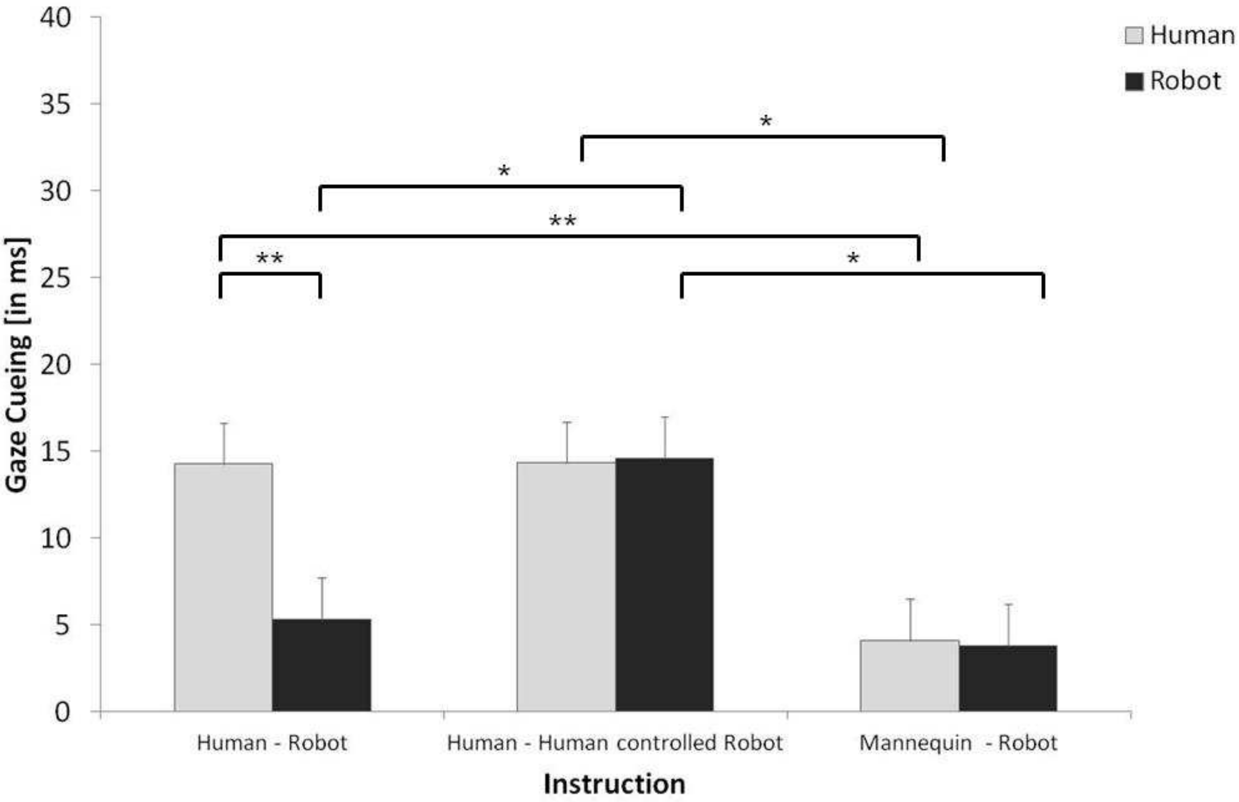

Figure 2. Size of gaze-cueing effects as function of Cue Type and Instruction. Error bars represent standard errors of the mean adjusted to within-subject designs (see Cousineau, 2005). * p < .05, ** p< .01.

Experiment 2

Experiment 2 was designed to further investigate the influence of adopting the intentional stance on gaze-cueing, independently of the physical characteristics of the stimuli. To realize this, stimuli were kept the same across conditions, while instruction was manipulated: ***Instruction 2*** – human versus robot controlled by human (*intentional stance likely*) and ***Instruction 3*** – human-like mannequin versus robot (*intentional stance unlikely*). The setup was comparable to Experiment 1, with one exception: in order to investigate the temporal dynamics of attentional orienting in response to gaze cues, the stimulus onset asynchrony (SOA) between cue and target presentation was varied: SOA was either short (250 ms) or long (600 ms).

Missed (0.44%) and incorrect responses (4.20%) as well as outliers (±2.5 SD from individual participants' means) were excluded from analysis. For the results of Experiment 2, see the middle (Instruction 2) and bottom (Instruction 3) rows in *Table 1* and the middle (Instruction 2) and right-hand sides (Instruction 3) of *Figure 2*. Mean RTs were examined in a mixed-design ANOVA with the between-subject factor instruction (Instruction 2, Instruction 3) and the within-subjects factors SOA (250 ms, 600 ms), validity (valid, invalid), and cue-type (human, robot). Results of statistical analyses for all trial types (neutral, valid, invalid) are summarized in *Table 1*. As SOA did not interact with any effects of interest, all Fs < 1.3, ps > .2, data were collapsed over this factor.

Again, there was a main effect of validity [$F(1,45)= 33.790$, $p< .001$, $\eta_p^2 = .429$], with shorter RTs for valid relative to invalid trials (452 ms vs. 461 ms), and no main effect of cue type (457 ms for the human vs. 456 ms for the robot, [$F(1,45)= .641$, $p> .4$, $\eta_p^2 = .014$]). Most importantly, the interaction between validity and instruction was significant [$F(1,45)= 11.087$, $p< .003$, $\eta_p^2 = .02$]: gaze-cueing effects were larger when adopting the intentional stance was likely (Δ_{RT} =*14 ms* for the human and Δ_{RT} =*15 ms* for the robot) compared to when this was unlikely (Δ_{RT} = *4 ms* for the human and for the robot). Note that this effect was independent of cue-type, as evidenced by a non-significant interaction between instruction, validity, and cue-type [$F(1,45)= .012$, $p> .9$, $\eta_p^2< .001$]. No other effect reached significance [$F< .02$, $p> .8$].

Comparisons across Experiments

To compare gaze-cueing effects among all three instructions and cue-types, post-hoc t-tests (Bonferroni-corrected for multiple comparisons) were conducted. Comparisons confirmed that the size of the gaze-cueing effect was not influenced by the cue-type as such (human vs. robot), but only by the likelihood of adopting the intentional stance towards the cue provider. In more detail, gaze-cueing effects did not differ between human and robot in conditions in which participants believed they observed either human behavior (Experiment 1, ***human***: $\Delta_{RT} = 14$ *ms*, vs. Experiment 2, Instruction 2, ***robot***: $\Delta_{RT} = 15$ *ms;* [$t(44) = -.094$, $p > .9$]) or non-human behavior (Experiment 1, ***robot***: $\Delta_{RT} = 5$ *ms*, vs. Experiment 2, Instruction 3, ***human***: $\Delta_{RT} = 4$ *ms;* [$t(45) = .328$, $p > .7$]). But the same cue-type elicited cueing effects of *different* sizes depending on whether or not the intentional stance was likely to be adopted towards the cue provider (Experiment 1, ***human***: $\Delta_{RT} = 14$ *ms*, vs. Experiment 2, Instruction 3, ***human:*** $\Delta_{RT} = 4$ *ms* [$t(45) = 2.727$, $p < .01$]; Experiment 2, Instruction 2, ***robot:*** $\Delta_{RT} = 15$ *ms*, vs. Experiment 1, ***robot:*** $\Delta_{RT} = 5$ *ms;* [$t(44) = 2.644$, $p < .02$]).

2.3.4. General Discussion

The present study investigated whether the *mere belief* that the observed stimulus is representing an agent with a mind influences basic social attention mechanisms, as measured by gaze-cueing effects. Rather than solely manipulating perceptual aspects of the cue provider (Admoni, Bank, Tan, Toneva & Scassellati, 2011; Bayliss & Tipper, 2006; Quadflieg, Mason & Macrae, 2004), we varied participants' beliefs about the gazer through instruction (while keeping the stimuli constant). We hypothesized that gaze-cueing would be increased when adopting the intentional stance was likely, whatever the identity of the gazer.

Our findings clearly support this hypothesis: while both human and robot induced attention shifts to gazed-at positions, cueing effects were twice as large when adopting the intentional stance towards the gazer was likely, as compared to when this was unlikely. In particular, gaze-cueing effects were significantly smaller for the robot than for the human when no explicit instruction was provided. Importantly, however, the same stimuli elicited gaze-cueing effects to varying degrees when different beliefs were induced: the human face condition yielded reduced cueing effects (comparable to the robot condition) when it was believed to represent a mannequin, while the robot face elicited enhanced cueing effects (comparable to the human condition) when it was believed to be controlled by a human. The results of Experiment 3 show that this pattern is robust, generalizing to other tasks with very different attentional demands, such as target localization (Posner & Cohen, 1984).

This pattern of results shows that basic social attention mechanisms are modulated by the observers' beliefs, induced solely by instruction, about whether or not the cue provider represents an intentional system. That is, social attention mechanisms are modulated fundamentally by the observer adopting the intentional stance towards others, rather than simply by attributing particular mental states (Kawai, 2011; Teufel et al., 2009, 2010a, 2010b).

Reflexive behavior under the control of beliefs

The present results provide evidence that attentional orienting in response to gaze direction is not purely reflexive, but prone to top-down modulation induced by higher-level cognitive processes. That is, gaze direction triggers social attention mechanisms based on a combination of two components: a bottom-up component that reflexively directs attention to where others are looking, and a top-down component that incorporates social context information relating to the observed scene into attentional guidance. In the present study, the bottom-up component produced a weak attentional bias towards stimuli at the gazed-at location, whether or not participants construed the cue provider as an intentional agent (in line with: Admoni et al., 2011; Quadflieg et al., 2004). The top-down component came into play in conditions in which adopting the intentional stance towards the cue provider was likely. Thus, attentional mechanisms involved in low-level processes of social perception not only influence, but *are* themselves influenced by beliefs humans hold about social stimuli they observe (in line with: Teufel et al., 2009, 2010a, 2010b). Interactions between lower- and higher-level processes are also supported by neuroimaging evidence: while the STS appears to trigger bottom-up responses to social signals, top-down control of these responses is thought to originate from the mPFC, adapting the system to the social context of the scene (Grezes, Frith & Passingham, 2004a, 2004b; Saxe & Wexler, 2005).

In this context, it is important to distinguish between i) reflexive vs. top-down modulated shifts of attention on the side of the observer; and ii) reflexive vs. intentional shifts of gaze on the side of the gazer. As discussed, the present data show that attentional shifts on the side of the observer are due to a reflexive mechanism that can be modulated by top-down component. At the same time, one needs to note that the gaze shifts on the side of the gazer can also be either reflexive (due to attentional orienting to a salient event in the periphery) or intentional, i.e., carrying social communicative content (the gazer shifts his/her gaze to the periphery in order to communicate a certain intention to the observer). The manipulation in the present study was concerned rather with the latter, as there was no salient event in the periphery that the gazer could reflexively shift gaze to. Therefore, we discuss the results in the context of intentional gaze behavior on the side of the gazer.

Humans do not engage in social interactions with just any agent

The finding that social attention processes are modulated by adopting the intentional stance when observing others’ behavior raises a fundamental question: why does the belief that another agent is an intentional system influence the way we allocate attentional resources? Clearly, an attentional system that is sensitive to social context information is highly advantageous from an evolutionary perspective: it permits adaptation to the social relevance of the scenario in which an interaction takes place. Allocating attention to where another person is attending serves the purpose of establishing *shared intentional-*

ity (Tomasello, 2010), which enables us to engage in collaborative activities by sharing goals, intentions, knowledge, and beliefs with others. The present results suggest that humans opt to engage in shared intentionality only with those who are believed to have intentions and are expected to display predictable, goal-oriented behavior. Given this, humans might be reluctant to adopt the intentional stance when observing a robot as compared to other humans. Importantly, what is crucial for adopting the intentional stance and, as a result, for readiness to engage in social interactions is not whether the observed agent *actually* has mental states, but whether the agent is *believed* to have mental states.

Interestingly in this context, there have been several reports (Abell, Happé & Frith, 2000; Castelli, Happé, Frith & Frith, 2000; Heider & Simmel, 1944; Klein, Zwickel, Prinz & Frith, 2009; Zwickel, 2009) that humans tend to provide mentalistic descriptions of the behavior of simple geometrical figures in dynamic motion scenarios. However, participants in those studies may not have actually adopted the intentional stance towards the observed stimuli, but only described behavior using mentalistic vocabulary – similarly to when one says “my computer did not want to start”. Alternatively, participants may have adopted the intentional stance by treating the geometric figures as *representations* of intentional agents – in a similar way to the present study where the robot was instructed to be controlled by a human. Hence, the novelty of our study is that through instruction manipulation, we triggered the activation of preexisting representations of the observed agents: the representation of a human as being an intentional agent, versus that of a robot being a mechanistic, non-intentional object. These representations in turn modulated the degree to which social attention mechanisms were employed.

Consequently, if humans tend not to adopt the intentional stance towards robots, they would ascribe less *social* relevance to its behavior compared to that displayed by humans. Hence, the present findings are not only of theoretical interest, but are also of significance to applied domains in which artificial systems are to be involved in interactions with humans (e.g., social robotics). If attribution of mental states is a crucial factor for enabling efficient social interactions, social robotics might need to address the issue of humans being hesitant to adopt the intentional stance towards a robot.

Concluding remarks

The present findings indicate that a *mere belief* that the observed agent represents a human triggers the concept of an *intentional agent*, and encourages adopting the intentional stance – in contrast to when the observed agent is believed to represent a mechanistic system (e.g., a robot). Consequently, social attention mechanisms are more readily employed when the intentional stance is adopted. This seems plausible, especially given that two types of intentions are communicated through gaze behavior that leads to directing others' attention (Tomasello, 2010): *referential* – what is the object of attention; and *social* – why do I direct your attention to this object? If an observer believes that the latter component is missing and is not convinced that the observed agent is capable of communicating social intentions, he/she might allocate attention to a lesser degree to the gazed-at object. On this basis, we propose that adopting the intentional stance plays a pivotal role in basic attention mechanisms involved in social interactions. For us humans to recruit these mechanisms, it seems not to matter whether the observed agents can *actually* think – but rather whether we *believe* they do!

Acknowledgments

We thank Verena Zierl and Regina Lanz for help with data collection. This work has been supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) within the Excellence Cluster Cognition for Technical Systems (CoTeSys). We also thank the Graduate School of Systemic Neurosciences (GSN) Munich for their support.

2.3.5. References

Abell, F., Happé, F., & Frith U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development, 15*(1), 1-16.

Admoni, H., Bank, C., Tan, J., Toneva, M., & Scassellati, B. (2011) Robot gaze does not reflexively cue human attention. *Proceedings of the 33rd CogSci*, 1983-1988.

Adolphs, R. (1999). Social cognition and the human brain. *TICS, 3(12)*, 469-479.

Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind.* Boston: MIT Press/ Bradford Books.

Bayliss, A.P., & Tipper, S.P. (2006). Predictive gaze cues and personality judgments: should eyes trust you?. *Psychological Science, 17(6)*, 514-520.

Brothers, L. (1990). The social brain: a project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neurosciences, 1*, 27-51.

Castelli, F., Happé, F., Frith, U., & Frith, C.D. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage, 12*(3), 314-325.

Cousineau, D. (2005). Confidence intervals in within-subjects designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology, 1(1)*, 42-45.

Dennett, D.C. (2003). True believers: the intentional strategy and why it works. In O'Connor, T., & Robb, D. (eds.). *Philosophy of Mind: Contemporary Readings*, pp. 370-390. London: Routledge.

Downing, P., Dodds, C.M., & Bray, D. (2004). Why does the gaze of others direct visual attention? *Visual Cognition, 11*, 71-79.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E. et al. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6*, 509-540.

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

Friesen, C.K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *JEP:HPP, 30,* 319-329.

Frischen, A., Bayliss, A.P., & Tipper. S.P. (2007). Gaze Cuing of Attention: Visual Attention, Social Cognition, and Individual Differences. *Psychological Bulletin, 133(4),* 694-724.

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

Gallagher, H.L., Jack, A.I., Roepstorff, A., & Frith, C.D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage, 16(3*), 814-821.

George, N., & Conty, L. (2008). Facing the gaze of others. *Neurophysiology, 38*, 197-207.

Grezes, J., Frith, C.D., & Passingham, R.E. (2004a). Brain mechanisms for inferring deceit in the actions of others. *Journal of Neurosciences, 24(24),* 5500-5505.

Grezes, J., Frith, C.D., & Passingham, R.E. (2004b). Inferring false beliefs from the actions of oneself and the others: an fMRI study. *Neuroimage, 21(2),* 744-750.

Grice, H.P. (1975). Logic and conversation. In Cole, P., Morgan, J.L. (eds.) *Syntax and Semantics. Speech acts*, pp.41-58. New York: Academic Press.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*, 246-259.

Hood, B.M., Willen, J.D., &Driver J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science, 9(2),* 131-134.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin, 96,* 29-44.

Kawai, N. (2011). Attentional shift by eye gaze requires joint attention: Eye gaze cues are unique to shift attention. *Japanese Psychological Research, 53,* 292–301.

Klein, A., Zwickel, J., Prinz, W., & Frith, U. (2009). Animated triangles: An eye tracking investigation. *Quarterly Journal of Experimental Psychology, 62(6),* 1189-1197.

Leslie, A.M., & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology, 6*, 315–324.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces (KDEF).* Stockholm: Department of Neurosciences Karolinska Hospital.

Nummenmaa, L., & Calder, A. (2009). Neural mechanisms in social attention. *TICS, 13 (3),* 135-143.

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

Quadflieg, S., Mason, M.F., & Macrae, C.N. (2004). The owl and the pussycat: Gaze cues and visuospatial orienting. *Psychonomic Bulletin & Review, 11(5),* 826-831.

Ristic, J., & Kingstone, A. (2005). Taking control of reflexive social attention, *Cognition 94(3):* B55-B65.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia, 43(10),* 1391-1399.

Senju, A., & Hasegawa, T. (2011). Direct gaze captures visuospatial attention. *Visual Cognition, 12(1),* 127-144.

Teufel, C., Alexis, D.M., Todd, H., Lawrence-Owen, A.J., Clayton, N.S. et al. (2009). Social cognition modulates sensory coding of observed gaze direction, *Current Biology, 19(15),* 1274- 1277.

Teufel, C., Fletcher, P.C., & Davis, G. (2010a). Seeing other minds: attributed mental states influence perception, *TICS, 14(8),* 376-382.

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

Tomasello, M. (2010). *Origins of Human Communication*. MIT Press: Cambridge, MA.

Turing, A.M. (1950). Computing machinery and intelligence. *Mind, 59,* 433-460.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13(1),* 103–128.

Zwickel, J. (2009). Agency attribution and visuospatial perspective taking. *Psychonomic Bulletin & Review, 16(6),* 1089-1093.

2.3.6. Supplementary Materials

Experiment 3

Experiment 3 was conducted to examine whether the pattern of results revealed in Experiments 1 and 2 (discrimination task) would generalize to other types of task that make different demands on attentional processing resources. To this end, we selected a speeded target localization task, as this type of task is typically used in gaze cueing studies (Friesen & Kingstone, 1998) and known to be sensitive to picking up small cueing effects.

Method and Materials

Experiment 3 was methodologically the same as Experiments 1 and 2 (instruction was varied between participants, while SOA, cue type, and cue validity were varied within participants), except for the task participants had to perform: participants were asked to localize the target as fast as possible by pressing the "D" key for targets presented on the left side and the "K" key for targets on the right side.

Participants, Apparatus, Stimuli, Design and Procedure

72 volunteers participated in Experiment 3, they were randomly assigned to three experimental groups

with 24 participants each: ***Group 1*** (20 women; mean age: 25 years (M=25.00, range: 19-31); two left-handed) received *Instruction 1* (human, robot); ***Group 2*** (19 women; mean age: 25 years (M=24.79, range: 21-36); two left-handed) received *Instruction 2* (human, human-controlled robot); and ***Group 3*** (16 women; mean age: 25 years (M=24.71, range: 20-32); two left-handed) received *Instruction 3* (human-like mannequin, robot). None of the participants has taken part in either Experiment 1 or Experiment 2. One participant from Group 1 (Instruction 1) was excluded from analysis because of significantly increased error rates compared to other participants (M= 11.46% compared to M=1.38%).

Ethics Statement

The experiments were conducted at the Department of Experimental Psychology at the LMU Munich, where all experimental procedures with purely behavioral data collection (e.g., RTs and error rates) of healthy adult participants, that do not include invasive or potentially dangerous methods are approved by the ethics committee of the Department of Psychology, LMU Munich, in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Data were stored and analyzed anonymously. Participants gave their informed consent and were either paid or received course credit for participating.

Table S1. Mean RTs and SEM (in ms) as a function of cue validity and instruction, for human and robot cues

	Human			Robot			Statistics
	Valid	Invalid	Neutral	Valid	Invalid	Neutral	Cue type x Validity
Instruction 1	322 (13)	349 (13)	359 (13)	324 (13)	343 (13)	357 (13)	$F(2,44)= 3.4, p< 0.05$
Instruction 2	306 (10)	335 (11)	355 (10)	305 (10)	335 (10)	353 (11)	$F(2,46)= 0.3, p> 0.75$
Instruction 3	333 (10)	340 (11)	369 (10)	333 (10)	342 (10)	363 (9)	$F(2,46)= 2.1, p> 0.13$

Results

Missed responses (0.34%), incorrect responses (1.04%), and outliers (±2.5 SD from individual participants' means) were excluded from analysis. Mean RTs and Standard Errors for neutral, valid, and invalid trials are presented in *Table S1* as a function of validity, cue type, and instruction. *Figure S1* depicts gaze-cueing effects as a function of cue type and instruction. As SOA did not interact with validity [$F(1,68) = .197$, $p > .6$, $\eta_p^2 = .003$], and did not have an influence on the interaction between instruction and validity [$F(2,68) = .058$, $p > .9$, $\eta_p^2 = .002$] or the interaction between validity, cue type and instruction [$F(2,68) = .304$, $p > .7$, $\eta_p^2 = .009$], data were collapsed over this factor for subsequent analyses.

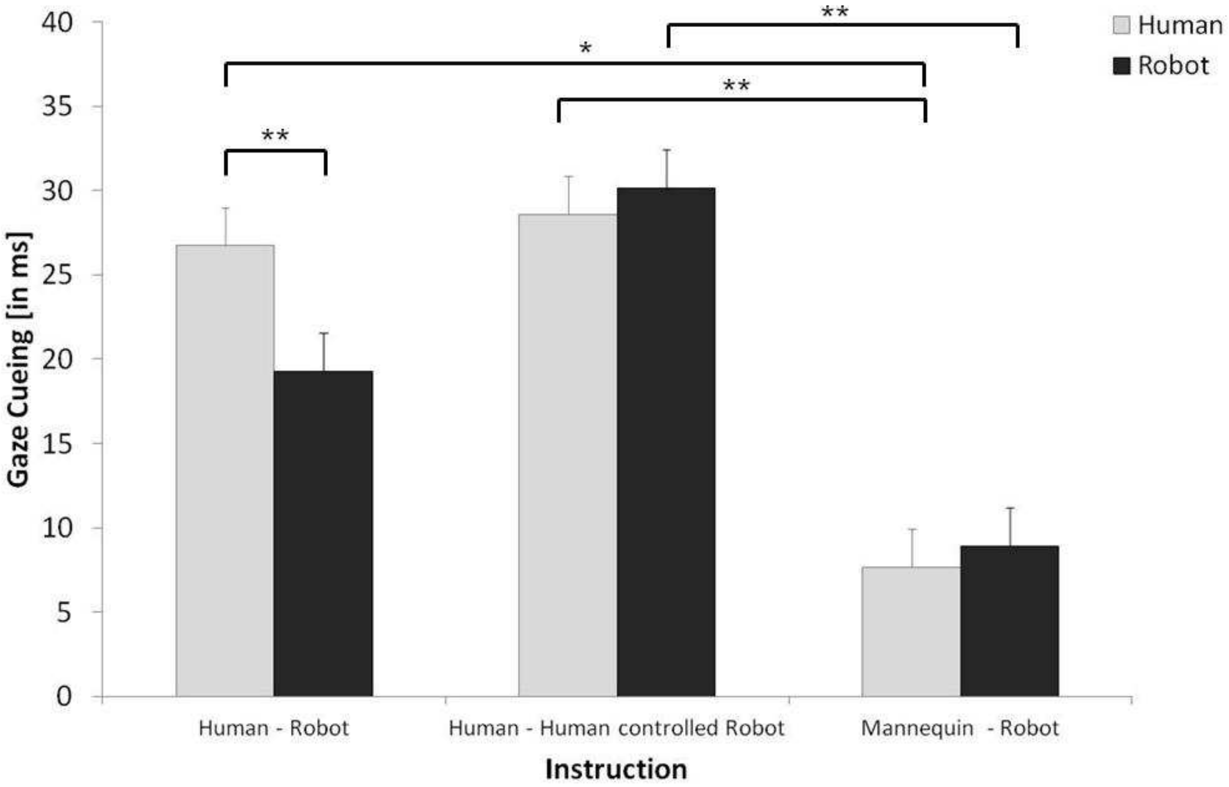

Figure S1. Size of gaze-cueing effects as function of Cue Type and Instruction. Error bars represent standard errors of the mean adjusted to within-subject designs (see Cousineau, 2005). * $p < .05$, ** $p < .01$.

As can be seen from *Figure S1* (and *Table S1*), Experiment 3 closely replicated the results of

Experiments 1 and 2. In more detail, valid cues yielded overall faster mean RTs than invalid cues [$F(1,68) = 67.471$, $p < .001$, $\eta_p^2 = .498$]. When no further instruction was given (***Instruction 1***), the robot condition elicited smaller cueing effects than the human condition [$F(1,22) = 9.821$, $p < .006$, $\eta_p^2 = .309$]. However, when both were believed to represent human behavior (***Instruction 2***), gaze-cueing effects for the human and robot condition were of similar magnitude [$F(1,23) = .296$, $p > .5$, $\eta_p^2 = .013$] and comparable to that in the human condition with no further instruction [$t(45) = -.478$, $p = .635$, *two-tailed*]. Similarly, when both agents were believed to represent non-human behavior (***Instruction 3***), the gaze cueing effects were also comparable in size [$F(1,23) = .076$, $p > .7$, $\eta_p^2 = .003$] and even smaller compared to the robot condition with no further instruction [$t(45) = 2.271$, $p < .03$, *two-tailed*]. This overall pattern of results mirrors the findings of Experiments 1 and 2, showing that the essential effects are independent of the type of task.

2.4. Study 4

Beliefs about the minds of others influence our own mind! An EEG study on the modulation of sensory gain control by higher-order cognition.

Agnieszka Wykowska[1*], Eva Wiese[1*], Aaron Prosser[1], & Hermann J. Müller[1,2]

[1]Ludwig Maximilian University, Munich Germany

[2]Birkbeck College, University of London, UK

Author contributions: AW, EW, HJM designed research; AW programmed experiment; EW, AP performed experiments. AW analyzed data. AW, EW, AP, HJM wrote paper. AW theoretically conceptualized the study. AP formalized study within the predictive coding framework.

Wykowska, A.*, **Wiese, E.***, Prosser, A. & Müller, H.J. (under review). Beliefs about the minds of others influence our own mind! An EEG study on the modulation of sensory gain control by higher-order cognition. *Journal of Neuroscience.*

*equal contribution

2.4.1. Abstract

Sensory gain control allows for improved processing of attended stimuli relative to ignored stimuli and has been examined by manipulating spatial attention (Hillyard et al., 1998). Furthermore, amplification of sensory processing has also been found with feature-based attention (Zhang & Luck, 2009). This is the first study, however, to examine the impact of *higher-order social cognition* on sensory gain control. With the use of a paradigm in which attention was guided to a location by the gaze of a centrally presented face, we manipulated participants' beliefs about the gazer: gaze behavior was believed to result either from operations of a mind or from a machine. In Experiment 1, beliefs were manipulated by cue identity (human or robot), whereas in Experiment 2, cue identity (robot) remained identical across conditions, and beliefs were manipulated solely via instruction, which was irrelevant to the task. ERP results and behavior showed that participants' attention was guided by gaze only when gaze was believed to be controlled by a human. Specifically, the P1 was more enhanced for validly cued versus invalidly cued trials only when participants believed that the gaze behavior was the result of a mind, rather than of a machine. This shows that sensory gain control can be influenced by higher-order (task-irrelevant) beliefs about the observed scene. We propose an interdisciplinary model of social attention: the Intentional Stance Model (ISM), which conceptualizes social cognition's influence on the sensory gain control within the predictive-coding framework (Friston, 2005).

2.4.2. Introduction

Directing attention to a location improves the processing of stimuli presented at that location. A neural mechanism underlying this phenomenon has been identified as *sensory gain control* (Hillyard et al., 1998; Mangun et al., 1993), which increases the signal-to-noise ratio (SNR) for stimuli at attended, relative to other, locations (Hawkins et al., 1990; Müller & Findlay, 1987). Sensory gain control has been examined using a variety of methods: single-unit neurophysiology (Luck et al. 1997), neuroimaging (Corbetta et al., 1990), and psychophysics (Lu & Dosher, 1998), with all studies providing converging evidence for the notion that attention influences sensory processing by amplifying stimulus-related neuronal signals. Based on the event-related potential (ERP) technique of scalp-recorded EEG, the P1-N1 complex at posterior-occipital electrode sites has been identified as the ERP index of the sensory gain control. For example, Mangun and colleagues (1993) observed that when spatial attention was deployed to a location, stimuli subsequently presented there elicited more enhanced P1 and N1 components relative to stimuli at other locations. The sensory gain mechanism has been studied extensively with a variety of procedures designed to modulate spatial attention: exogenous cues (Luck et al., 1994), central cueing (Eimer, 1994; Mangun & Hillyard, 1991); sustained attention (Mangun et al., 1993), or with the use of more naturalistic cues, such as faces with directional gaze (Schuller & Rossion, 2001). Sensory amplification has also been observed with feature-based attention (Corbetta, 1990; Zhang & Luck, 2009; though see Hillyard et al., 1998, for an interpretation of feature-based attention effects in terms of sustained bias). However, the actual sources of attentional control over sensory processing and the question of whether sensory gain is sensitive to task-irrelevant higher-order cognitive processes remain to be examined.

The aim of the present study was to investigate whether or not task-irrelevant beliefs about the observed scene can modulate sensory gain control. In our paradigm, attention was guided to a location by the gaze direction of a centrally presented face, and we manipulated beliefs regarding whether the face’s gaze behavior resulted from the operations of a mind or of a machine. Crucially, these beliefs were irrelevant to the task, and were manipulated either by cue identity (Experiment 1: presenting a human or robot face) or solely by instruction, with cue identity remaining identical across conditions (Experiment 2: presenting only a robot face but informing participants that its gaze behavior was either human-controlled or pre-programmed). We reasoned that attentional control over sensory gain would be enhanced when the gaze behavior was believed to result from the operations of a mind, rather than a machine, as attending to locations gazed-at by an intentional agent is adaptive from the social and evolutionary perspective (Tomasello, 2010).

Given this, the present study is the first to examine the social source of attentional control over the sensory gain mechanism, and whether sensory gain can be governed not just by intrinsically visual factors (e.g., spatial or feature-based selection), but also by higher-order cognitive processes, such as beliefs about the observed scene.

2.4.3. Experiments

Materials and Methods

Stimuli and apparatus

Stimuli were presented on a 17-inch computer screen with a 100-Hz refresh rate, placed at a distance of 80 cm from the observer. In the human-face condition of Experiment 1, a digitized photo, 5.7° × 5.7° of visual angle in size, of the face of the same female individual, chosen from the Karolinska Directed Emotional Faces (KDEF, Lundqvist et al., 1998) database, was used in the human condition. In the robot-face condition of Experiments 1 and 2, a photo of an anthropomorphic robot (EDDIE, LSR, TU München), of the same size as the human face, was presented. See Figure 1 for example stimuli.

Both human and robot faces were presented frontally without changes in head orientation. To produce gaze direction cues, irises and pupils within the eyes were shifted (using Photoshop$_{TM}$) left- or rightwards to deviate by 0.2° from straight-ahead gaze, in both the human and the robot condition. Stimuli were presented centrally on a white background, with eyes positioned on the central horizontal axis of the screen. The midpoints of the human and robot faces were positioned 0.2° and, respectively, 1.1° below the central horizontal axis; this slight difference in positioning with respect to the y-axis ensured that the peripheral target letters were always presented at the same level as the eyes of the human or robot face on the central horizontal axis.

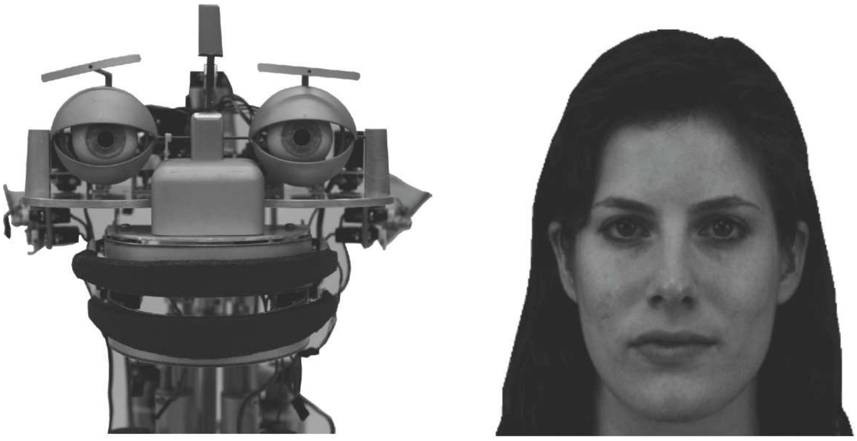

Figure 1. The human face (left) and the robot face (right) used in Experiment 1. In Experiment 2, only the robot face (left) was presented.

The target stimulus was a black capital letter (F or T), 0.2° × 0.2° in size, which was presented on the central horizontal axis at an eccentricity of 5.7° with respect to the screen center (Figure 2). Target positions (left or right) were determined pseudo-randomly such that targets appeared with equal frequency at either of the two positions.

Gaze direction was not predictive of target position, that is: in Experiment 1, on target-present trials (80% in total), gaze was directed either to the side on which the target appeared (valid trials, 33% of target present trials) or to the other side (invalid trials, 33% of target-present trials), or it remained gazing straight-ahead, with targets equally likely appearing on either side (neutral trials, 33% of target-present trials). The neutral condition was introduced in order to examine for possible differential effects related to physical dissimilarities between the human and robot conditions. In Experiment 2, neutral trials were not included in the design. That is, the face could gaze to only the left or the right (50% trials with each direction, in target-present trials), with the target presented either on the right or on the left side of the screen. In both experiments, 20% of all trials were catch trials (no target presented). These target-absent trials were introduced to permit subtraction of the EEG signal on target-absent trials from that on target-present trials, so as to eliminate ERP potentials elicited by the cue, which overlapped with potentials related to the target.

Participants

Sixteen volunteers took part in the Experiment 1 (5 women; mean age: 24 years; age range 20 to 30 years; all right-handed; and all with normal or corrected-to-normal vision; none of the observers had taken part in an experiment with such a paradigm before), and twenty-eight volunteers in the Experiment 2 (19 women; mean age: 24.4 years; age range 19 to 34 years; 7 left-handed; and all with normal or corrected-to-normal vision; none of the observers had taken part in Experiment 1 or any other experiment with such a paradigm); they received an honorarium for their participation. The experiments were conducted with the full understanding and written consent of each participant. In Experiment 1, the data of two participants had to be discarded due to technical problems during recording of the EEG data.

Procedure

Participants were seated in a dimly lit chamber with a keyboard under their hands. A trial started with a fixation point (2 pixels) presented for 850 ms. Subsequently, a face with gaze directed straight-ahead (in the direction of the observer) appeared on the screen for 850 ms, while the fixation dot remained visible (in-between the eyebrows of the face). The straight-ahead gazing face was followed by a gaze shift (cue) to the left or the right (valid and invalid trials), or the gaze remained straight (neutral trials) for another 600 ms. Next, the target letter was presented on either the left or the right side of the screen, at a fixed stimulus onset asynchrony (SOA) of 600 ms relative to the onset of the gaze cue. Following this event, the face and target remained on the screen for another 30 ms only, in order to minimize eye movements in the critical time window. Participants were then asked to respond as quickly and accurately as possible to the identity of the target letter (F or T) using the ‘o’ or, respectively, ‘i’ key on a standard keyboard (all other letters were removed and the o/i letters were covered with green/blue stickers), with response assignment (o=F/i=T vs. o=T/i=F) counterbalanced across participants. The keys were to be pressed with the index finger of the left and the right hand, respectively. The display was blank for the duration of response. Upon response, another trial started with the presentation of the fixation dot in the screen center. On target-absent trials, no response was required and the trial continued (blank screen) for another 800 ms. For an illustration of the trial sequence, see Figure 2.

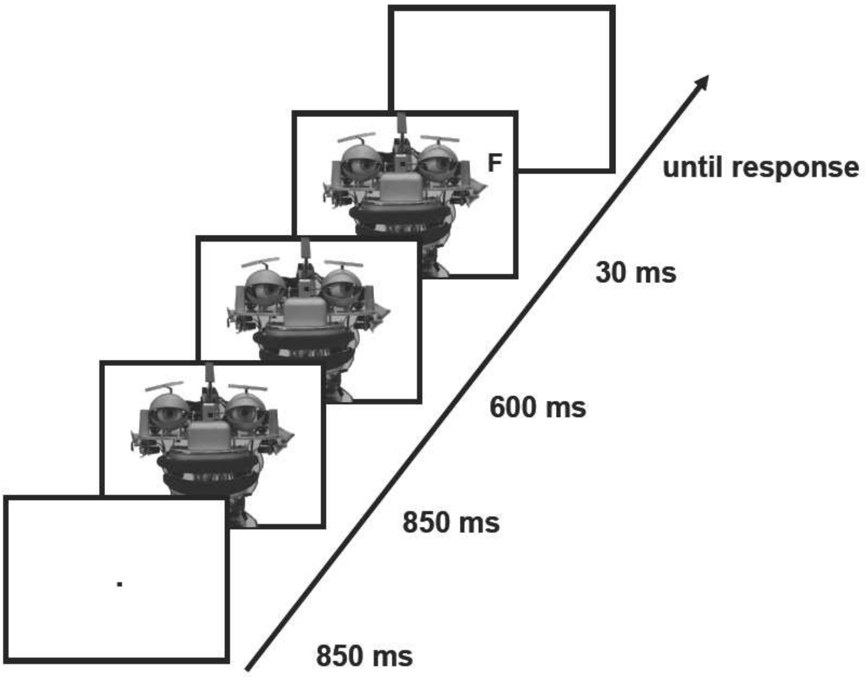

Figure 2. A trial sequence. Participants first fixated on a fixation dot for 850 ms. Subsequently, a robot or human face (Experiment 1) or always a robot face (Experiment 2) gazing straight-ahead was presented for another 850 ms. Next, the gaze direction changed to either the left or the right for another 600 ms, which was then followed by target presentation (30 ms) either at the gazed-at location (valid-cue trial) or the opposite location (invalid-cue trials). Participants were then asked to respond to target identity, with a blank screen presented until the response. On catch trials, the display with a face gazing to the left/right was presented for another 30 ms.

Experiment 1 consisted of 900 trials and all conditions were randomly mixed within 10 blocks of 90 trials each. In Experiment 2, there were altogether 960 trials, split into 2 sessions with two different instructions (on the same day, with a 15-30 min break in between). Each of the participants received both instructions (Instruction 1: human-controlled, Instruction 2: pre-programmed), with order counterbalanced across participants.

EEG recording

EEG was recorded with Ag-AgCl electrodes from 64 electrodes of an active electrode system (Acti-Cap, Brain Products, GmbH, Munich, Germany). Horizontal and vertical EOG were recorded bipolar from the outer canthi of the eyes and from above and below the observer's left eye, respectively. All electrodes were referenced to Cz and re-referenced offline to the average of all electrodes. Electrode impedances were kept below 5 kΩ. Sampling rate was 500 Hz, with an online high cut-off filter of 125 Hz.

Data analysis

We hypothesized that the directional gaze shift would guide attention to the gazed-at location. Hence, we expected validity effects (superior performance, and enhanced amplitudes of the P1-N1 ERP complex, for valid- vs. invalid-cue trials). Moreover, we expected the validity effects to be modulated by cue type (Experiment 1) and instruction (Experiment 2) – the rationale being that gaze following makes more sense if the gaze potentially conveys communicative content, relative to when it only reflects mechanistic behavior. In sum, the main factors of interest for all our analyses were: validity (valid vs. invalid) and cue-type (human vs. robot: Experiment 1) or, respectively, instruction (human-controlled vs. pre-programmed: Experiment 2). The analyses focused on valid and invalid trials, as neutral trials of Experiment 1 did not constitute a proper baseline – owing to the fact that in gaze cueing paradigms with naturalistic stimuli, neutral, straight-ahead gaze towards the observer is special due to producing an arousal effect and/or having a holding effect on attention, making it difficult to disengage attention (the central, straight-ahead gazing face) and shift it to peripheral targets (George & Conty, 2008; Senju & Hasegawa, 2005; Wiese, Wykowska et al., 2012). Neutral trials were only analysed with respect to a main effect of cue type in Experiment 1, in order to examine for differential effects related to physical dissimilarities of the cue stimuli.

EEG data

The data were averaged over a 700-ms epoch including a 200-ms pre-stimulus baseline, with epochs time-locked to target onset. Trials with eye movements and blinks on any recording channel (indicated by any absolute voltage difference in a segment exceeding 80 µV or voltage steps between two sampling points exceeding 50 µV) were excluded from analyses. Additionally, channels with other artefacts were separately excluded if amplitude exceeded ±80 µV or any voltage was lower than 0.10 µV for a 100-ms interval. Only trials with correct responses were analyzed. One participant in Experiment 1 and three participants in Experiment 2 were excluded from analyses due to extensive eye blinks, and one participant was excluded due to abnormal alpha activity (Experiment 2).

Target-locked early sensory effects (P1, N1)

Analyses were conducted on correct target-present trials with ERPs time-locked to target onset. The two types of target (F and T) as well as the side of presentation (left and right) were averaged together. Target-absent (catch) trials were subtracted from target-present trials, to eliminate overlapping potentials related to gaze cue onset and, thus, extract the potentials related to the targets. The subtraction was conducted on epoched data, separately for each type of cue (human vs. robot), each gaze direction (left vs. right), time-locked to target onset. For the ERP data, too, the analyses focused on the comparison between valid and invalid trials. The EEG signal was averaged for the two validity conditions (valid vs. invalid) and the two types of cue (human vs. robot). Mean amplitudes in the typical time window of the P1 component (100–140 ms, i.e., ±20 ms from the latency of the grand average peak amplitude (120 ms) in the 100–150-ms time window in the human valid condition, in which P1 was most pronounced; regarded as the typical P1 time window, in line with Luck, 2005) in Experiments 1 and 2; as well as in the subsequent N1 component (150–190 ms in Experiment 1 and 170–210 ms in Experiment 2; i.e., ±20 ms from the latency of the grand average peak amplitude in the 140–200-ms time window in the robot valid condition, in which N1 was most pronounced) for the lateral posterior-occipital electrodes (O1, O2, PO7, PO8) were subjected to ANOVAs with the factors *electrode* (O1, O2, PO7, PO8), *cue type* (human vs. robot), and *cue validity* (valid vs. invalid). Where appropriate, statistics were corrected according to Greenhouse-Geisser for potential nonsphericity.

Behavioral data

Prior to the reaction time (RT) analysis, trials with response errors or RTs faster than 150 ms and longer 1200 ms were excluded. Median RTs and mean error rates were computed for each participant. The statistical analyses focused on the comparison between valid and invalid trials. Individual median RTs and mean error rates were submitted to a 2×2 analysis of variance (ANOVA) with the factors *cue type* (human vs. robot, Experiment 1) *instruction* (human-controlled, pre-programmed, Experiment 2) and *cue validity* (valid, invalid). The same participants whose EEG data were analysed were subjected to behavioral analyses.

Results

Experiment 1

P1. The 2×2×4 ANOVA with the factors *cue validity* (valid vs. invalid), *cue type* (human vs. robot), and *electrode* (O1, O2, PO7, PO8) for the mean amplitudes in the P1 time window (100–140 ms) revealed the cue type x cue validity interaction to be significant, F (1, 12) = 6.596, p = .025, $\eta_p^2 = .355$, with a more positive P1 amplitude for valid trials (M = 1.13 µV, SEM = .48) than for invalid trials (M = .88 µV, SEM = .43) in the human face condition; in the robot face condition, by contrast, his pattern tended to be reversed: the P1 amplitude was slightly less positive for valid trials (M = .21 µV, SEM = .47) relative to invalid trials (M = .73 µV, SEM = .39); see Figure 3. This interaction was stable across all four electrode sites (three-way interaction with electrode: p = .31). The main effect of cue type was not significant, F (1, 12) = 1.04, p = .35. Subsequent analyses conducted separately for the human and robot faces showed that in the human-face condition, the validity effect interacted with electrode site, *F* (1.6, 19.5) = 5.598, *p* = .016, η_p^2 = .318, with a significant validity effect for the right electrodes (O2, PO8), *F* (1, 12) = 6.545, *p* = .025, η_p^2 = .35, and a non-significant effect for the left electrodes (O1 and PO7), *p* = .79. The robot face condition, by contrast, did not yield any significant effects or interactions of interest (validity effect: *p* = .08; validity × electrode interaction: *p* = .42)[1].

N1. An analogous analysis for the later time window (150–190 ms) of the N1 ERP component revealed a main effect of cue validity, *F* (1, 12) = 8.032, *p* = .015, η_p^2 = .401, with valid trials eliciting a more negative mean amplitude (*M* = -2.33 µV, *SEM* = .5) compared to invalid trials (*M* = -1.7 µV, *SEM* = .58). This effect was not influenced by the type of cue, *p* = .79, or by electrode site, *p* = .24; Figure 3.

Behavior

Error rates. The 2×2 ANOVA with the factors of *cue type* (human vs. robot face) and *validity* (valid vs. invalid) revealed a significant interaction between cue type and validity, *F* (1, 12) = 5.902, *p* = .032, η_p^2 = .33, with the validity effect being more pronounced for human faces (ΔER= 2 %) than for robot faces (ΔER= -0.4%). Planned comparisons showed that for the human-face condition, error rates were significantly lower for valid trials (*M* = 3.8 %, *SEM* = .8) relative to invalid trials (*M* = 5.7%, *SEM* = .9), *t* (12) = 2.44, *p* = .031, two-tailed; by contrast, there was no difference between these two types of trial in the robot-face condition (valid: *M* = 5.2%; invalid: 4.8%, p = .706).

RTs. An analogous analysis on the median RTs revealed no significant main effects or interactions. Numerically, RTs were slightly faster on valid trials (*M* = 404 ms, *SEM* = 11) compared to invalid tri-

---

[1] Note that visual inspection of the grand-averaged ERP waveforms suggested a differential effect in a time window preceding that of the P1 component, on the negative deflection of the waveform; see Figure 3. However, statistical analysis on this time window (60–100 ms) failed to yield any significant effects, in particular, the interaction of cue and validity was nonsignificant, *F* = .009, *p* = .926 (cue type, *F* = 1.609, *p* = .229, and validity, *F* = .031, *p* = .862).

als (*M* = 410 ms, *SEM* = 10), main effect of validity: *F* (1, 12) = 3.46, *p* = .088, $\eta_p^2 = .224$, and the difference was numerically larger for the human faces ($RT_{valid} = 405$ ms; $RT_{invalid} = 411$ ms) than for the robot faces ($RT_{valid} = 404$ ms; $RT_{invalid} = 408$ ms), though not reliable statistically (cue type × cue validity interaction: *F* (1, 12) = .421, *p* = .529).

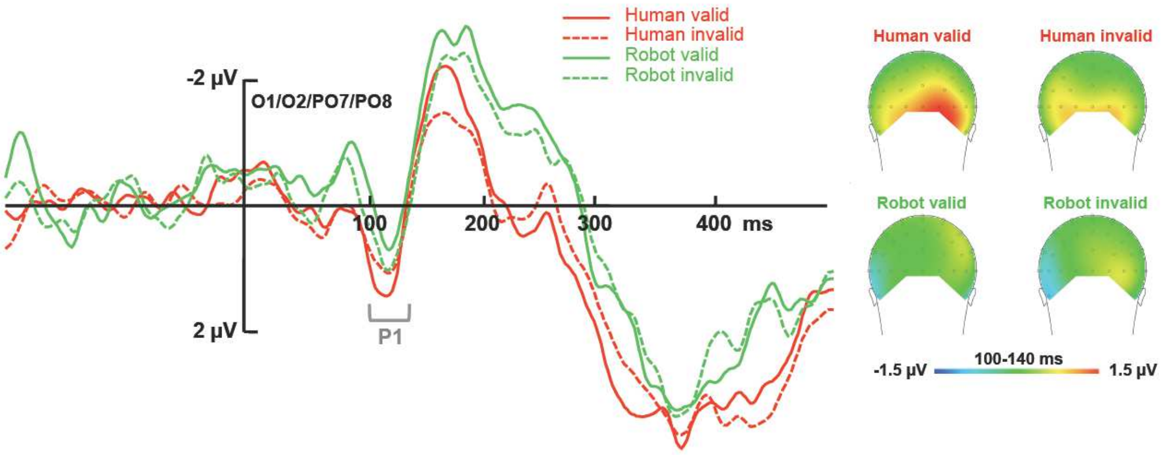

Figure 3. Left: Grand average ERP waveforms time-locked to the target onset for the pool of O1/O2/PO7/PO8 electrodes as a function of cue validity (solid lines: valid trials; dashed lines: invalid trials) and type of cue provider (red: human faces, green: robot faces), in Experiment 1. The two types of targets (F and T) as well as left/right sides of visual field were averaged together. The displayed ERPs are the subtracted waveforms (target present–target absent) and filtered with a 30-Hz high cut-off filter (Butterworth zero phase, 24 dB/Oct) for illustration purposes. Right: Topographical maps of voltage distribution for the time interval 100–140 ms, presented from posterior view; the human condition is depicted in the upper panel, the robot condition in the lower panel. Mean amplitudes are presented on the left for valid trials, and on the right for invalid trials.

Inverse Efficiency Scores (IES). When both behavioral measures (RTs and accuracy) were combined into a single dependent variable, namely: “Inverse Efficiency Scores (IES)” (see Bruyer & Brysbaert, 2011; Townsend & Ashby, 1978; Townsend & Ashby, 1983), by dividing individual median RTs by individual accuracy scores (percentages of correct responses), a 2×2 ANOVA with the factors *cue type* (human vs. robot face) and *validity* (valid vs. invalid) revealed a marginally significant interaction between cue type and validity, *F* (1, 12) = 4.536, *p* = .055, $\eta_p^2 = .274$, with the validity effect being more pronounced for human faces (ΔRT= 15 ms) than for robot faces (ΔRT= 3 ms). Planned comparisons showed that for the human-face condition, the cue validity effect was significant, *t* (12) = 2.739, *p* =

.018, two-tailed; by contrast, there was no significant effect in the robot-face condition *t* (12) = .507, *p* = .621.

Neutral trials (gaze-straight ahead). In order to examine whether physical dissimilarity between human and robot cues as such has an influence on the amplitude of the early sensory P1 component, we compared the neutral-trial P1 mean amplitude between robot and human faces. A 4×2 ANOVA with the factors *electrode* (O1, O2, PO7, PO8) and *cue type* (robot vs. human) revealed no main effect of cue type, *F* (1, 12) = .131, *p* =.723, and no interaction with electrode, *F* (2, 24) = .671, *p* = .521. The behavioral data, too, showed no indication of differential processing on neutral trials between robot and human faces: *t* (12) = .422, *p* = .68 for the error rates, and *t* (12) = .628, *p* = .542 for the median RTs; see Table 1 for the mean RTs, error rates, and P1 amplitude in the neutral condition.

Table 1. Mean RTs, error rates and mean amplitude of the P1 component (100-140 ms time window) for the neutral cue condition (gaze straight-ahead) as a function of type of cue provider (human vs. robot).

	Median RTs	Mean error rates	Mean P1 amplitude
Human	425 ms	4.10%	1.123 µV
Robot	427 ms	3.80%	1.204 µV

Experiment 2

Experiment 2 was designed to eliminate the possibility that physical differences between the two types of cue providers were responsible for the pattern of effects observed in Experiment 1, rather than differences in mind-attribution. This alternative explanation is unlikely, given that: (i) a comparison of the human versus the robot condition did not reveal any significant P1 effects; (ii) valid trials were compared with invalid trials within each of the two cue type conditions (robot, human) rather than across conditions; and (iii) a main effect of cue type was observed neither in the P1 nor in the N1 time window, indicating that the amplitudes of those components were not influenced by the type of stimulus as such. Nevertheless, it is important to clearly show that the modulation of the sensory gain control is due solely to the higher-order belief that the gaze behaviour was the result of the operations of a mind, rather than of a machine. Thus, we needed to isolate this top-down modulation to rule out the alternative explanation of the P1 effect. To do this, we used only one, physically identical cue provider across all conditions while manipulating participants' beliefs via instruction. Specifically, participants were presented with the same robot face (gazing to the left or the right; see Figure 1 right) across all conditions; crucially, in one experimental session, they were told that the robot's gaze behavior was pre-programmed (Instruction 1), and in the other session, they were told that the eyes of the robot were controlled by a human (Instruction 2).

P1. The 2×2×4 ANOVA on the mean amplitudes in the P1 time window (100–140 ms) with the factors *cue validity* (valid, invalid), *instruction* (human-controlled, pre-programmed), and *electrode* (O1, O2, PO7, PO8) revealed a significant interaction of validity and instruction, F (1, 23) = 5.566, p = .027, η_p^2 = .195, with a more positive P1 amplitude for valid trials (M = 1.45 µV, *SEM* = .28) than for invalid trials (M = 1.247 µV, *SEM* = .29) in the *human-controlled* condition, and a slightly less positive amplitude for valid trials (M = 1.45 µV, *SEM* = .4) versus invalid trials (M = 1.526 µV, *SEM* = .39) in the *pre-programmed* condition (Figure 4). This interaction was stable across all four electrode sites (three-way interaction with electrode: F(1, 23) = .434, p= .694). Subsequent analyses conducted separately for the *human-controlled* and *pre-programmed* instructions revealed the validity effect to be significant in the *human-controlled* condition, t (23) = 1.794, p =.043, one-tailed, but not in the *pre-programmed* condition t (23) = .652, p = .26, one tailed.

N1. An analogous analysis on the mean amplitudes in the N1 time window (170–210 ms) revealed no main effect of validity, F (1, 23) = .099, p = .756, and no significant interaction of validity and instruction, F (1, 23) = 1.021, p = .323.

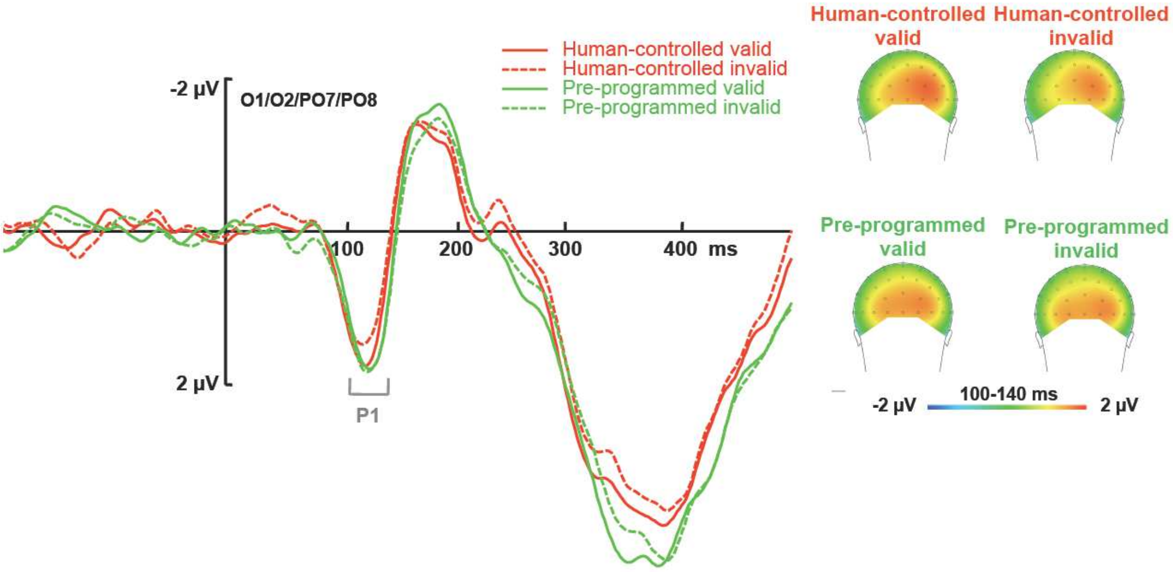

Figure 4. *Left*: Grand average ERP waveforms time-locked to target onset for the pool of O1/O2/PO7/PO8 electrodes, as a function of cue validity (solid lines: valid trials; dashed lines: invalid trials) and instruction (red: human-controlled, green: pre-programmed), in Experiment 2. The two types of targets (F and T) as well as left/right sides of visual field were averaged together. The displayed ERPs are the subtracted waveforms (target present-target absent) and filtered with a 30-Hz high-cutoff filter (Butterworth zero phase, 24 dB/Oct) for illustration purposes. *Right*: Topographical maps of voltage distribution for the time interval 100–140 ms, presented from posterior view; the human-controlled condition is depicted in the upper panel, the pre-programmed condition in the lower panel. Mean amplitudes are presented on the left for valid trials and on the right for invalid trials.

Behavior. A 2×2 ANOVA on median RTs with the factors *instruction* (human-controlled, pre-programmed) and *cue validity* (valid, invalid) yielded a significant interaction between instruction and cue validity, F (1, 23) = 4.269, p = .05, η_p^2 = .157. Planned comparisons revealed that in the *human-controlled* condition, validly cued targets elicited significantly shorter RTs (*M* = 406 ms, *SEM* = 9.5), compared to invalidly cued targets (*M* = 409 ms, *SEM* = 10.6), t(23) = 1.805, p = .042, one-tailed); in the *pre-programmed* condition, by contrast, the validity effect was not significant, t = .785, p = .22, one-tailed, with slightly longer RTs for valid trials (*M* = 410 ms, *SEM* = 10) relative to invalid trials (*M* = 409 ms, *SEM* = 10.6). An analogous analysis of the error rates revealed no significant effects or interactions, all *F*s< 2, *p*s> .2. Numerically, the error pattern reinforced the RT and ERP data: valid trials yielded slightly fewer errors (*M* = 6.1% in the *human-controlled* condition and 6.7% in the *pre-*

programmed condition) than invalid trials (M = 6.6% in the *human-controlled* condition and 7.1% in the *pre-programmed* condition).

2.4.4. General Discussion

The purpose of the present study was to examine whether sensory gain control can be modulated by higher-order cognitive processes, such as beliefs about the observed scene. In our paradigm, attention was guided by gaze, and beliefs about the observed gazer were manipulated either by the identity of the face (Experiment 1) or solely by instruction, with the gazer's identity remaining identical across conditions (Experiment 2). We hypothesized that attentional control over the sensory processing (the sensory gain control) would be enhanced when observers believed that gaze (which guided their attention to a location) was controlled by a mind, rather than a machine. Our reasoning followed Tomasello's (2010) distinction between two types of intention communicated through gaze: *referential* and *social*. The first concerns the *object* of attention, the second *why* attention is directed to this object. The idea is that gaze behavior is often only informative when it originates from a mind, because mental states not only cause gaze behavior, but they also give meaning (the "why") to it. Therefore, if an observer believes that an agent with a mind is directing gaze to a location, the observer may expect something relevant for communication at that location, and thus allocates their attention there as well. By contrast, if the observer believes that a machine is directing "eyes" to a location, the observer may not allocate their attention there because the machine's gaze behavior lacks communicative content, for it is not attributable to the operations of a mind.

Our data supported this hypothesis, showing that the attentional control over sensory processing can be modulated by higher-order beliefs about the observed scene. In two experiments, the target-locked P1 was more enhanced for the valid-cue trails relative to invalid-cue trials, but only when the gazer's behavior was believed to result from operations of a mind. This ERP effect was paralleled by the behavioral data: target-related performance was better on valid-cue, relative to invalid-cue, trials when participants believed the gazer had a mind and was not a machine. The ERP and behavioral data are particularly intriguing because participants' beliefs about the gazer were completely irrelevant to the discrimination task they performed.

Interestingly, the N1 validity effect was not modulated by cue type in Experiment 1, and no validity or instruction effect on the N1 was observed in Experiment 2. The P1 and N1 components have previously been proposed to reflect different modes of control over sensory gain: the P1 has been argued to reflect a suppression mechanism for ignored locations, whereas the N1 indexes enhanced and discriminative processing of stimuli at the attended locations (Hillyard et al., 1998; Luck & Hillyard, 1995; Mangun & Hillyard, 1991). Given this, the differential effects between the P1 and N1 component suggest that when target stimuli are presented very briefly, higher-level cognitive processes influence only the earlier, suppression-related mechanism to increase the SNR, but not the later, discriminative processes at the attended locations.

In sum, this is the first study to show that sensory gain control can have a social source and that higher-order, task-irrelevant beliefs about the observed scene can influence early sensory processing by modulating stimulus-related neuronal activity, dependent on whether the stimulus location has been signaled by a meaningful social cue (gaze direction of an agent with a mind) or not (gaze direction of a machine).

Theoretical considerations

The present findings can be interpreted in the context of the idea that humans adopt various "stances" in order to predict and understand behavior of various systems with which they interact: the physical stance, the design stance, or the *intentional stance* (Dennett, 2003). Based on experience, humans know which stance works best for which system. For example, when explaining the workings of a machine, it is best to adopt the Design Stance (DS) and understand its behavior with reference to how it is designed to behave. In contrast, when explaining other humans' behavior, the most efficient strategy is to engage in *mentalizing* (Frith & Frith, 1999): predicting and understanding behavior with reference to particular mental states (e.g., beliefs, desires, intentions).

However, we argue that before one can engage in mentalizing processes (i.e., refer to any particular mental state), one needs to assume that the entity whose behavior one is explaining is actually capable of having mental states. Restated, one needs to adopt the *Intentional Stance* (IS) towards the observed entity by assuming that the entity is an *intentional system* (a system with a mind). Our findings show that attentional control over sensory processing (sensory gain control) is exerted depending on whether or not one adopts the Intentional Stance towards an observed entity.

To account for the findings, we propose the *Intentional Stance Model* (ISM) of social attention (Figure 5). The basic idea is that the brain uses probabilistic (generative) models to make predictions about the world's causal structure, which are tested against sensory data to update the brain's beliefs about this structure (Friston, 2005; 2010). According to the ISM, when the brain adopts the IS (which most likely occurs in the anterior paracingulate cortex, cf. Gallagher et al., 2002), a generative model $M(x, \Psi)$ is activated which predicts the likelihood $p(\Psi|x)$ that the observed behavior x of an entity is caused by an intentional system (Ψ), rather than another system (e.g., Φ). Here, Ψ and Φ are drawn from a set of representations $\vartheta \supset \{\Phi, \Psi, \Gamma, ...\}$, where Φ = mechanistic system, Ψ = intentional system, etc. This likelihood is computed by combining the likelihood $p(x|\Psi)$ of receiving a sensory input x given that the entity is a Ψ, and the prior probability $p(\Psi)$ of adopting IS towards this system. The likelihood of interpreting an observed behavior as intentional is therefore given by Equation 1: $p(\Psi|x) \propto p(x|\Psi)p(\Psi)$.

If the predictions of the generative model $M(x, \Psi)$ (i.e., adopting the IS) are correct (i.e., low or no error), the brain can begin to construct specific models $m_i(x, \psi_i)$ about specific mental states (ψ_1, ψ_2,

ψ_3) in order to mentalize about particular instances of behavior displayed by the entity. However, if the model's predictions yield a high error, the brain activates a different model, such as $D(x, \Phi)$. Here, $D(x, \Phi)$ corresponds to adopting the Design stance (DS). Following the same logic, when the predictions of the generative model $D(x, \Phi)$ are correct, other specific models $d_i(x, \varphi_i)$ referring to *mechanistic states* (φ_1, φ_2, φ_3) *are constructed* in order to explain particular instances of behavior. Thus, the same behavior (e.g., *A* gazes at an apple) can be explained with reference to either specific mental states (*A* wants to eat the apple: ψ_1) if predictions of $M(x, \Psi)$ are correct (e.g., eye behavior generally socially meaningful), or to specific mechanistic states (*A*'s machinery shifts the camera lens around: φ_1) if the predictions of $M(x, \Psi)$ yield a high error (e.g., eye behavior does not carry social meaning).

Previous research has shown that mentalizing influences perceptual processing (Teufel et al., 2010a). For example, the magnitude of gaze-guided attentional effects has been found to be modulated depending on whether observers believed that the gazer could see through a pair of goggles (Teufel et al., 2010b). Teufel and colleagues (2010a) proposed the so-called *perceptual mentalizing model* (PMM) to account for these mentalizing-dependent effects. According to PMM, when observers perform a gaze cueing task, brain areas implicated in mentalizing, such as the medial prefrontal cortex (mPFC) and temporo-parietal junction (TPJ), generate signals which modulate neural activity in social perception areas, such as the superior temporal sulcus (STS). The STS in turn interacts with the parietal attention mechanisms of the intraparietal sulcus (IPS) in order to shift attention in the direction of the gaze by increasing the commitment of neural resources to the gazed-at location.

One major limitation of the PMM, however, is that it does not account for the impact of adopting the Intentional Stance on sensory processing. As described above, mentalizing logically and functionally presupposes adopting the Intentional Stance, because the brain must first assume that the observed entity is actually capable of having mental states before it can infer mental states underlying particular behaviors.

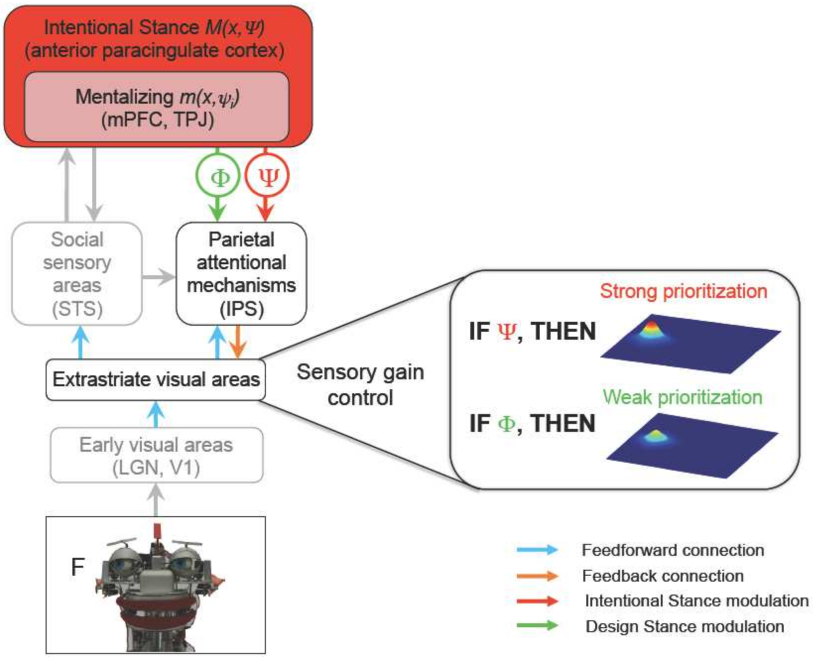

Figure 5. The Intentional Stance Model (ISM) of social attention. A visual stimulus (the robot face, bottom) is processed in the visual pathway from the lowest-level (early visual areas box) to higher-level areas (e.g., STS). The Attentional network (IPS) is involved in orienting attention to the stimulus (the F) that is cued by the gaze. One of the core claims of ISM is that mentalizing is dependent on the Intentional Stance (IS), because it logically and functionally presupposes the adoption of the IS. When the brain adopts the IS (which most likely occurs in the anterior paracingulate cortex, Gallagher et al., 2002), a generative model $M(x, \Psi)$ is activated which predicts the likelihood $p(\Psi|x)$ that the observed behavior x of an entity (e.g., robot gazing left or right) is caused by an intentional system (Ψ). Conversely, the Design Stance (DS) predicts that behavior x is caused by a mechanistic system (Φ). Mentalizing, which constructs specific models $m(x,\psi_i)$ of specific mental states, only occurs once the brain activates $M(x, \Psi)$.The IS and DS predictions feed-back into the parietal attentional mechanisms, subsequently modulating the sensory gain control in the extrastriate visual areas (right). When observing an entity's gaze behavior while adopting the IS, this higher-order belief modulates the sensory gain control in the extrastriate areas, increasing the priority of an item cued by the gaze (represented by a higher peak of neural activity on the right; the other peak depicts an invalidly cued object). This additional prioritization does not occur when the brain adopts the DS. Thus, beliefs about the mind of others influence one's own mind. LGN = lateral geniculate nucleus, V1 = primary visual cortex, STS = superior temporal sulcus, IPS = intraparietal sulcus, mPFC = medial prefrontal cortex, TPJ = temporo-parietal junction. Processes of social cognition and perception that are the focus of this paper and are essential for the core claims of the ISM are highlighted in black and color, while gray boxes represent other processes of social perception/cognition that are not in the focus of this paper.

The Intentional Stance Model of social attention presented here accounts for this limitation, because specific generative models of specific mental states (ψ_i) are constructed only when the IS is adopted. If so, it follows that the Intentional Stance is functionally independent and presupposed by mentalizing. Finally, unlike the PMM, the ISM specifies the neuro-cognitive machinery by which adopting the Intentional Stance and mentalizing exert their top-down influences on social attention.

When the IS predictions feedback to lower levels of the processing hierarchy, it modulates the *sensory gain control*. Accordingly, the IS modulations seem to reach as low as the extrastriate visual areas, where stimulus coding is modulated by the sensory gain mechanism (Hillyard et al., 1998).

To conclude, the present study showed that a general perceptual selection mechanism – sensory gain control – is governed not just by intrinsically visual factors, such as spatial or feature-based selection, but is sensitive to higher-order task-irrelevant beliefs about others. This implies that beliefs about the minds of others influence our own mind.

Author Contributions

AW, EW, and HM conceived and designed the study; AW programmed the experiment; AP and EW performed the experiment; AW analyzed data; AW, EW, AP, and HM wrote paper; AW theoretically conceptualized the study in the context of Intentional/Design Stance; AP formalized the Intentional Stance within the predictive coding framework.

2.4.5. References

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)?. *Psychologica Belgica, 51(1),* 5-13.

Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G.L., & Petersen, S.E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science, 248,*1556-1559.

Dennett, D.C. (2003). True believers: the intentional strategy and why it works. In T. O'Connor & D. Robb (eds): *Philosophy of Mind: Contemporary Readings*, pp. 370-390. London: Routledge.

Eimer, M. (1994). ´Sensory gating´ as a mechanism for visuospatial orienting: electrophysiological evidence from trial-by-trial cueing experiments. *Perception & Psychophysics*, *55*, 667-675.

Friston, K. (2005). A theory of cortical responses. Philosophical Transactions. *Biological Sciences, 360,* 815–836.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Review Neuroscience, 11*, 127–38.

Frith, C.D., & Frith, U. (1999). Interacting minds: a biological basis. *Science, 286,* 1692-1695.

Gallagher, H.L., Jack, A.I., Roepstorff, A., & Frith, C.D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage, 16,* 814-821.

George, N., & Conty, L. (2008). Facing the gaze of others. *Neurophysiology, 38,* 197-207.

Hawkins, H.L., Hillyard, S.A., Luck, S.J., Mouloua, M., Downing, C.J., & Woodward, D.P. (1990). Visual attention modulates signal detectability. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 802–811.

Hillyard, S.A., Vogel, E.K., & Luck, S.J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions. Biological Sciences, 353*, 1257-1270.

Lu, Z., & Dosher, B.A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38*, 1183–1198.

Luck, S.J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

Luck, S.J., & Hillyard, S.A. (1995). The role of attention in feature detection and conjunction discrimination: An electrophysiological analysis. *International Journal of Neuroscience, 80*, 281-297.

Luck, S.J., Chelazzi, L., Hillyard, S.A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology, 77,* 24-42.

Luck, S.J., Hillyard, S.A., Mouloua, M., Woldorff, M.G., Clark, V.P., & Hawkins, H.L. (1994). Effects of spatial cuing on luminance detectability: psychophysical and electrophysiological evidence for early selection. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 887–904.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces (KDEF).* (Department of Neurosciences Karolinska Hospital, Stockholm).

Mangun, G.R., & Hillyard, S.A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 1057-1074.

Mangun, G.R., Hillyard, S.A., & Luck, S.J. (1993). Electrocortical substrates of visual selective attention. In D. Meyer & S. Kornblum (eds): *Attention and Performance XIV*, pp. 219-243. Cambridge, MA: MIT Press.

Müller, H.J., & Findlay J.M. (1987). Sensitivity and criterion effects in the spatial cuing of visual attention. *Perception & Psychophysics, 42,* 383-399.

Schuller, A.-M., & Rossion, B. (2001). Spatial attention triggered by eye gaze increases and speeds up early visual activity. *Neuroreport, 12*, 2381-2386.

Senju, A., & Hasegawa, T. (2011). Direct gaze captures visuospatial attention. *Visual Cogniton, 12(1),* 127-144.

Teufel, C., Fletcher, P.C., & Davis, G. (2010a). Seeing other minds: attributed mental states influence perception. *Trends in Cognitive Sciences, 14,* 376–382.

Teufel, C., Alexis, D.M., Clayton, N.S., & Davis, G. (2010b). Mental attribution drives rapid, reflexive gaze following in human adults. *Attention, Perception, & Psychophysics, 72,* 695–705.

Tomasello, M. (2010). *Origins of Human Communication*. Cambridge, MA: MIT Press.

Townsend, J.T., & Ashby, F.G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory*, *Vol.3*, (pp. 200-239). Hillsdale, N.J.: Erlbaum

Townsend, J.T., & Ashby, F.G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.

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. doi:10.1371/journal.pone.0045391.

Zhang, W., & Luck, S.J. (2009). Feature-based attention modulates feed-forward visual processing. *Nature Neuroscience, 12(1),* 24-25.

GENERAL DISCUSSION

3.1. Synopsis of results

The goal of the presented PhD-project was to conduct a comprehensive series of experiments that investigate whether and how context information influences basic mechanisms of social attention. We showed that gaze-cueing effects were modulated by social context information in all experiments. In particular, we provided evidence that *visual* (placeholders), *verbal* (instructions) and *empirical* (experience) context information affected the size and the spatial distribution of gaze-cueing effects. Our findings suggest that the context-sensitive modulation of gaze cueing originates from top-down influences on the early P1 ERP-component, which is reflecting a sensory gain control mechanism suppressing the processing of stimuli at unattended locations. We take this pattern to indicate that bottom-up and top-down mechanisms interact to induce gaze-cueing effects and exert a combined influence on the allocation of spatial attention following gaze cues.

3.1.1. Context information modulates the spatial allocation of social attention

In the first two series of experiments, we investigated whether the allocation of attentional resources following changes in gaze direction is influenced by the context in which the gaze cues are presented: Social context was either provided in form of reference objects (i.e., placeholders) in the periphery to which gaze direction could be linked (**Study 1**) or by supplying information about the predictivity with which gaze direction indicated target position (**Study 2**). In both cases, we expected to find general cueing effects for the whole cued hemifield when no context information was provided and gaze-cueing effects specific to the exact gazed-at position when context information was available.

In **Study 2**, we additionally investigated how information about experienced (i.e., actual) and expected (i.e., believed) cue-predictivity is integrated in order to guide attention to the most likely target position: When experienced and expected cue-predictivity matched, we expected to observe specific cueing effects for the exact gazed-at location in the predictive condition, and nonspecific cueing effects for the whole cued hemifield in the nonpredictive condition. When experienced and expected predictivity did not match, we presumed the allocation pattern of cueing effects induced by experienced predictivity to be modulated corresponding to what would have been expected based on the believed predictivity of the gaze cues.

Our results showed that the allocation of spatial attention following gaze cues can be modulated by both sources of social context information: the presence of reference objects in the periphery and the predictivity of the cues. In **Study 1**, we found nonspecific cueing effects if there were no objects in the periphery to which gaze direction could refer. This general cueing effect was independent of set size, task type and of whether attention was shifted covertly or overtly. However, if reference objects were present in the periphery, gaze induced an additional cueing effect specific to the gazed-at position. In **Study 2**, we showed that the reliability of observed gaze behavior modulated the spatial specificity of gaze cueing: with predictive cues, cueing effects were significantly larger for targets that appeared at the exact gazed-at position relative to targets at the other two positions; nonpredictive cues, by contrast, gave rise to facilitation effects of equivalent magnitude for all positions in the cued hemifield. Most importantly, the effect of the experienced (i.e., actual) predictivity was modulated by the expected (i.e., instructed) predictivity, that is: nonpredictive cues believed to be predictive caused additional cueing effects specific to the gazed-at positions, compared to nonpredictive cues believed to be nonpredictive (which caused general cueing effects for the whole hemifield). In contrast, specific gaze-cueing effects caused by actually predictive cues were significantly reduced when the cue was believed to be nonpredictive, compared to when it was believed to be predictive.

Based on the results of the first two studies, it appears that attentional orienting to gaze always induces a general cueing effect for the whole cued hemifield in a *bottom-up* manner. When social context information is provided, a cueing effect is observed that is specific to the gazed-at position acting over and above the general cueing effect by causing a *top-down* modulation of the stimulus-triggered allocation of attentional resources. However, based on **Study 2**, it seems that a specific component may not simply be added to a general cueing effect, but can rather be up- and down-regulated depending on the beliefs participants have about the observed scene: the spatial specificity of cueing effects is reduced, if participants believe that gaze direction does not inform about the most likely target position, whereas it is enhanced in cases when participants believed that gaze direction is guiding them to the most likely target position. Hence, the first two studies extend previous findings by showing that a top-down modulation of orienting to gaze direction can be invoked by various types of context information, namely visual information provided in the social scene (placeholders), verbal information (instruction) about observed behavior and empirical information deducible from experience with observed gaze behavior.

Findings from Study 1 and 2 allow the integration of different theories of spatial attention with regard to their validity for social attention mechanisms. On the one hand, they are in line with the assumptions of the general-region theories of attention, such as the *meridian-boundary model* (Hughes & Zimba, 1985, 1987), by showing that changes in gaze direction cause spatially nonspecific cueing effects under conditions when no context information is provided. On the other hand, the results also confirm the predictions of the specific-location theories of attention, such as the *zoom-lens model* (Eriksen & Yeh, 1985) or the *gradient model* (e.g., Downing & Pinker, 1985) – by showing that attentional resources can be allocated to specific regions of the visual field when the observed scene is embedded within a social context created, for instance, by presenting reference objects in the periphery or by providing information about the gazer that allows intentionality to be attributed to observed gaze behavior.

3.1.2. Readiness to engage in social interactions depends on attributing the mind to others

In **Study 3** and **4**, we investigated whether the belief that the observed stimulus is representing an agent with a mind has an influence on the readiness to engage in social interactions with that agent. The degree of attentional engagement was measured as the size of the gaze-cueing effects that were induced by the observed agent (human or robot). Importantly, rather than manipulating perceptual aspects of the cue provider, we varied participants' beliefs through instruction (while keeping the physical appearance of the stimuli constant). We hypothesized that gaze-cueing effects would be increased when attributing the mind towards the gazer was likely and that this effect would be independent of the gazer's identity (i.e., physical appearance).

The results of **Study 3** support the intentionality hypothesis: while both the human and the robot face induced attention shifts to the gazed-at position, cueing effects were twice as large when attributing the mind to the gazer was likely (i.e., when eye movements were believed to be caused by a human) compared to when it was unlikely (i.e., when eye movements were believed to be mechanistic). Importantly, the same stimulus elicited cueing effects to varying degrees depending on which belief has been induced by instruction: the human face yielded reduced cueing effects (comparable to the robot face) when it was believed to represent a mannequin, while the robot face elicited enhanced gaze-cueing effects (comparable to the human face) when it was believed to be controlled by a human. This pattern has been shown consistently in all experiments and generalized from quite complex discrimination tasks to other tasks with different attentional demands (i.e., target localization).

In **Study 4**, the EEG/ERP-method was used to identify the neural correlates underlying the top-down modulation of gaze-cueing effects by attributing the mind towards the gazer. We showed that attribution of intentionality imposes a modulatory effect on early components of perceptual processing, such as the sensory gain control mechanism. This mechanism improves perceptual coding by increasing the signal-to-noise ratio by enhancing the processing of stimuli presented at attended locations and/or suppressing processing at other, interfering locations. In particular, we found larger gaze-cueing effects for intentional relative to mechanistic agents, indicating that attribution of the mind to observed agents modulates social attention mechanisms already at the very early stages of perceptual processing.

Taken together, these findings indicate that changes in gaze direction believed to be mechanistic are followed to a lesser degree than eye movements believed to be intentional. In line with the first two studies, the bottom-up component involved in gaze cueing produced a general attentional bias towards stimuli in the cued hemifield, whether or not participants construed the gazer as intentional agent (in line with Bayliss et al., 2006; Quadflieg et al., 2004). However, the top-down component came into play only in conditions in which attributing the mind to the gazer was likely. Thus, attentional mechanisms involved in low-level processes of social attention not only influence, but are themselves influenced by beliefs about observed social agents (in line with Frith & Frith, 2006; Teufel et al., 2009, 2010).

The assumption of gaze-cueing effects being caused by interactions between bottom-up and top-down processes is further supported by evidence from neuroimaging studies showing that while the STS triggers bottom-up responses to social signals, the top-down control of these responses originates from the mPFC, thereby adapting the system to the social context of the scene (Grezes et al., 2004 a,b; Saxe et al., 2005). A modulation of gaze cueing via attribution of intentionality to observed behavior is very plausible, as attending to where another agent attends is beneficial only if gaze direction carries informative content. However, changes in gaze direction are informative only if they reflect the intentions/goals of another agent or are generally resulting from actions of the mind, and thus help predicting and interpreting behavior in a meaningful way and preparing appropriate action responses.

Why does the belief that an observed agent is an intentional system influence how we allocate attentional resources? Being sensitive to context information when attending to the behavior of others is highly advantageous as it permits adaptation to the social relevance of the scenario in which an interaction takes place. Furthermore, allocating attention to where another person is attending serves the purpose of establishing shared intentionality (Tomasello, 2010), and enables us to engage in activities with others by sharing goals, intentions, knowledge, and beliefs. Our findings suggest that humans tend to engage in shared intentionality only with those who are believed to have intentions and are expected to display predictable behavior. Given this, it seems that what is crucial for attributing the mind and for engaging in social interactions is not whether the observed agent actually has mental states, but whether the agent is believed to have mental states.

3.2. Context effects on spatial attention – an integrative two-component model of gaze cueing

The present findings argue in favor of the idea that gaze cues can initiate both a general and a specific component of attentional orienting, dependent on the availability of context information: If no context information is provided, only a general directional orienting effect is induced (*bottom-up*) that causes a global, but small cueing effect for the whole gaze-cued hemifield. By contrast, if context information is available, a spatially specific component is evoked that induces a facilitation effect for the exact gazed-at position and/or suppression for the other (uncued) positions in the cued hemifield. However, based on the available data, it is not entirely clear whether the context-dependent modulation of gaze cueing is caused by an interaction of top-down and bottom-up components or is rather due to a top-down effect that is purely additive to the general bottom-up effect.

Whatever the exact relationship between bottom-up and top-down components is, one can clearly conclude that specific gaze-cueing effects can be observed under conditions with context information, but only a general cueing effect for the cued hemifield without context information. The specific orienting component, however, is likely to result from the combination of two separate sources of information: the linking of social context information with directional information deduced from the gazer’s eyes. Interestingly, our results show that this linking is independent of the modality in which context is provided: visual, verbal and empirical information have all been shown to influence the size and/or specificity of gaze cueing.

Based on these findings, we postulate an integrative two-component model of social attention that combines the influences of the bottom-up and top-down components (**Figure GD1**): The bottom-up component induces an orientation of attention to the gaze-cued hemifield, and is carried out by a network consisting of visual areas (i.e., V1), STS and IPS (Nummenmaa & Calder, 2009, for review). The top-down modulation of the signal that has been generated bottom-up is exerted via connections with the mPFC (e.g, Grezes et al., 2004 a,b; Saxe et al., 2005) and the anterior cingulate (Gallagher et al., 2002) which weight attentional deployment to specific locations in the cued hemifield according to the social relevance of the cue. Due to this weighting, attentional resources are allocated in a way that objects/events that are relevant for the social interaction are prioritized.

Our conclusions are in line with previous postulates in the literature. That is, the proposal that visual information from the eye region and the periphery are integrated when processing gaze direction is in line with findings of Lobmaier and colleagues (2006), who have shown that objects in the visual field can cause systematic biases in estimating where the other is looking. The observation that explicit knowledge about our interaction partners does influence basic social attention processes is consistent with Frith and Frith (2006) and Teufel and colleagues (2010), who have suggested that bottom-up orienting to gaze direction can be top-down controlled by context information about the gazer. Our results are also in line with empirical evidence showing that familiarity with the gazer (gender effect for women only: Deaner et al., 2007) and stereotypical information about the gazer (member of a political party: Liuzza et al., 2011) modulate the size of gaze-cueing effects.

Taken together, findings presented in this PhD-thesis reveal a degree of flexibility in the gaze-cueing system that allows for the integration of multiple sources of information to guide attention: whenever context information about the social relevance of the scene is missing, attention is directed to a broader area, whereas attention shifts are specific to the cued location when a relation between gaze direction and context information can be established. This context-dependent flexibility is adaptive, in that it facilitates detection of relevant social information (e.g., objects of interest) in a constantly changing environment.

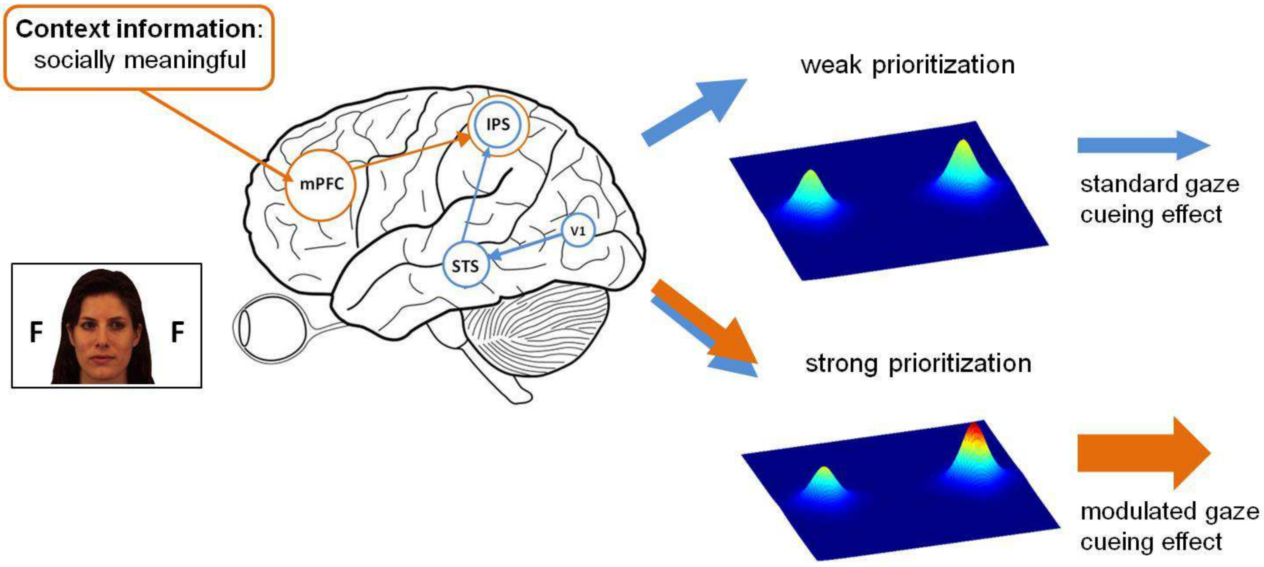

Figure GD1. Bottom-up orienting to gaze cues is carried out by a network consisting of visual areas (i.e., V1), STS and IPS (in *blue*), which causes a standard gaze-cueing effect for targets presented in the cued hemifield. The top-down modulation of bottom-up signals is exerted via connections with the mPFC (in *orange*). The top-down mechanism weights the deployment of attentional resources to specific locations in the cued hemifield according to the social relevance of the cue, resulting in stronger gaze-cueing effects in conditions when social context information is provided.

3.3. Conclusions

Over the course of evolution, social signals such as head and body orientation, pointing gestures and in particular gaze direction have acquired a special meaning in peer-to-peer communication, as they tend to indicate the location of important objects or events in the environment (Emery, 2000). In order to interpret gaze direction in a meaningful way, however, one needs more information about the context in which eye movements are performed/observed: who is the gazer, where is the gaze direction referring to, how is the gazer feeling about the attended object/event and why is the gazer interested in what he or she is looking at? The more information we have about the social context of the scene, the better we are able to make predictions about what the gazer is planning to do next, which in turn provides the basis for preparing appropriate action responses when being engaged in interactions with him/her.

In support of this notion, the present findings provide evidence that the readiness to engage in social interactions and the efficiency with which intentions and action goals of others can be inferred is significantly modified by the availability of context information. That is, the degree to which attention is allocated to the gazed-at object/event critically depends on whether or not observed behavior can be explained in a meaningful way. For instance, while one is hesitant to follow the gaze of a mechanistic system or gaze that is directed towards empty space, one is more willing to do so, if an agent capable of having intentions is gazing at objects/events in the periphery. We showed that the modulatory effect of context information on how we process social information is a very profound mechanism that exerts its influence already on early sensory processing components, such as P1. Thus, on the basis of the present findings, we propose that social attention mechanisms are not only influencing higher-order cognitive processes (i.e., mentalizing), but are also influenced by beliefs we hold about the observed gazer.

3.4. Future Directions

The findings presented in this PhD-work have several implications for further research. First, given that gaze direction is not the only cue used to communicate information about internal and external states in social interactions, it would be interesting to investigate whether the context variables discussed above also have an influence on the interpretation of other social signals that provide spatial information. If an interaction between directional information conveyed by the cue and context information about what is observed were specific to social scenarios, top-down modulation of attentional orienting should also be found for other social cues (i.e., pointing); in contrast, cues that transmit nonsocial information (i.e., arrows) should not interact with contextual information. Such dissociation between the functionality of social and nonsocial cues would support the assumption that the human brain is specifically sensitive to the processing of social information resulting in deeper processing of the context in which the observed scene takes place.

Second, future studies would need to identify further contextual factors that influence the spatial allocation of attention following gaze cues. One factor that may have an impact on the spatial specificity of gaze-cueing effects is the facial expression accompanying observed eye movements. In particular, facial expressions that indicate potential danger in the environment (fear, anger) or potential sources of joy (happiness, surprise) might induce highly specific attention shifts, whereas faces with a neutral expression might give rise to spatially nonspecific shifts. Identifying further context factors would thus help understanding the mechanisms that underlie social-attention processes, while also supporting the notion that gaze is special with regard to cueing spatial attention.

Third, the present findings are not only of theoretical interest, but are also of significance to applied domains in which artificial systems are involved in interactions with humans (e.g., social robotics). The problem is that, if humans tend not to attribute the mind towards robots, they would ascribe less *social* relevance to the robot's behavior compared to that displayed by humans. However, since attribution of mental states is a crucial factor for enabling efficient social interactions, social robotics needs to address the reluctance of humans to attribute the mind to robots by equipping robots with the capability of showing intentional behavior. A first step in that direction would be to provide robots with a model that can be used by the robot to interpret gaze direction of human interaction partners and to display meaningful (i.e., action-congruent) gaze behavior itself.

Fourth, the presented research has implications for developing training applications for clinical populations, in particular for people with autism spectrum disorder (ASD). It has been shown that these patients have reduced abilities to parse socially relevant information in the environment (Baron-Cohen, 1995; Dawson et al., 1998; Klin et al., 2002) and exhibit reduced mentalizing skills: that is, they have difficulties inferring mental states from observed behavior (e.g., Baron-Cohen, Leslie & Frith, 1985). However, as the presented research suggests that attending to gaze direction and being able to infer mental states reciprocally impact each other in healthy individuals, the question arises of how these capabilities are causally linked in patients with ASD: that is, are both abilities equally impaired or is rather the inability to mentalize causing impairments in following eye gaze? If attending to others' gaze direction was possible in ASD-patients under certain circumstances, remaining gaze processing skills could be used in order to increase the functionality of the whole mentalizing network. Based on the assumption that people with ASD do not follow human gaze because they are unable to explain the observed behavior in a meaningful (i.e., intentional) way, one would admittedly predict that they would be able to follow the gaze of a mechanistic system (i.e., robot), whose behavior does not have to be explained by attributing intentionality, but can rather be understood based on analytical reasoning using systematic rules (i.e., in a mechanistic way). Thus, given that people diagnosed with ASD have intact systemizing skills (Baron-Cohen, 2002), it is reasonable to assume that they would follow the gaze of a robot and that their orientation to robot gaze could be used to train them to link observed changes in gaze direction to making predictions about behavior.

3.5. References (General Discussion)

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition*, *21*, 37-46.

Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind.* Boston: MIT Press/ Bradford Books.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, *6(6)*, 248-254.

Bayliss, A.P., & Tipper, S.P. (2006a). Predictive gaze cues and personality judgments: should eyes trust you?. *Psychological Science, 17(6),* 514-520.

Dawson, G., Meltzoff, A.N., Osterling, J., Rinaldi, J. & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders, 28,* 479-485.

Deaner, R.O., Shepherd, S.V., & Platt, M.L. (2007). Familiarity accentuates gaze cuing in women but not in men. *Biology Letters*, *3*, 64-67.

Downing, C.J., & Pinker, S. (1985). The spatial structure of visual attention. In M.I. Posner & O.S.M.

Marin (Eds.), *Attention and performance XI*, pp.171-188. Hillsdale, NJ: Erlbaum.

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

Eriksen, C. W., & Yeh, Y.-Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception & Performance, 11*, 583-597.

Frith, C. D., & Frith, U. (1999). Interacting minds – a biological basis. *Science,* 286, 1692-1695.

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

Gallagher, H.L., Jack, A.I., Roepstorff, A., & Frith, C.D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage, 16(3*), 814-821.

Grezes, J., Frith, C.D., & Passingham, R.E. (2004a). Brain mechanisms for inferring deceit in the actions of others. *Journal of Neuroscience, 24(24),*5500-5505.

Grezes, J., Frith, C.D., & Passingham, R.E. (2004b). Inferring false beliefs from the actions of oneself and the others: an fMRI study. *Neuroimage, 21(2),*744-750.

Hughes, H.C., & Zimba, L.D. (1985). Spatial maps of directed visual attention. *Journal of Experimental Psychology: Human Perception & Performance, 11*, 409-430.

Hughes, H.C., & Zimba, L.D. (1987). Natural boundaries for spread of directed visual attention. *Neuropsychologia, 2,* 5-18.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59(9),* 809-816.

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),* 1-7.

Lobmaier, J.S., Fischer, M.H., & Schwaninger, A. (2006). Objects capture perceived gaze direction. *Experimental Psychology*, *53*, 117-122.

Nummenmaa, L., & Calder, A. (2009). Neural mechanisms in social attention. *Trends in Cognitive Sciences 13(3):*135-143.

Quadflieg, S., Mason, M.F., & Macrae, C.N. (2004). The owl and the pussycat: Gaze cues and visuospatial orienting. *Psychonomic Bulletin & Review*, *11*, 826-831.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia, 43(10)*, 1391-1399.

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

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

Tomasello, M. (2010). *Origins of Human Communication*. MIT Press, Cambridge, MA.

Appendix

CURRICULUM VITAE

Education

09/2009 - present

Neuro-Cognitive Psychology, LMU Munich
PhD in Neuroscience, Graduate School of Systemic Neurosciences

- Research on the influence of context information on social attention and perception, theory of mind, and human-robot-interaction

- supervised by Prof. Hermann Müller, Dr. Agnieszka Wykowska, Prof. Stefan Glasauer, CoTeSys Cluster of Excellence

10/2002 – 04/2008

Otto-Friedrich University of Bamberg
Diploma in Psychology (MSc-Equivalent): 1.0 (with distinction)

- **Major:** neuropsychology, cognitive psychology, empirical methodology
- **Minor:** statistics (1.0) and applied informatics (1.0)
- **Thesis:** "As much as needed" or "as much as possible": Inference and schema generalization in analogical problem-solving (1.0)
  supervised by Prof. U. Schmid, Prof. S. Weinert

10/2000 – 09/2002

University of Bayreuth, *Studies of Physics*

Professional Experience

05/2008 – 08/2009

Fraunhofer Institute for Production Systems and Design Technology, Berlin
Researcher

Courses: Human Factors, Human-Machine-Interaction, neurophysiological methodology, neuro-computation; **Research:** Embodied cognition, spatial reasoning, prospective design of HMI

09/2007 – 02/2008

Daimler AG Stuttgart, Customer Research Centre
Internship (Research)

Research on customer acceptance and attitudes towards innovation; preparation of test vehicles; collecting, interpreting and presenting data

06/2007 – 09/2007

Daimler AG Stuttgart, Executive Management Development
Internship

Operative work, data analysis and data management; conception, organization and administration of competence workshops; life-long learning

07/2006 – 10/2006

Max-Planck Institute for Human Brain and Cognitive Sciences, Leipzig
Internship (Research)

Conception and conduction of experiments with babies and infants (eye-tracking, EEG); data-preparation (Interact), analysis (SPSS, Excel) and interpretation

08/2005 – 09/2005

Neurological and Neuropsychological Rehabilitation Centre, Lenggries
Internship (Research)

Conception, planning and conduction of neuropsychological studies on attention, and memory degeneration in dementia; diagnostics and intervention; training

04/2003 – 04/2005

Department of Developmental Psychology, University of Bamberg
Student assistant

Conception, planning and conduction of behavioral studies on speech development, joint attention, and theory of mind; data preparation and analysis

Key Skills

Software
MS Office, Photoshop, GIMP, EndNote/Mendeley, SPSS, Experiment Builder

Programming and cognitive modeling
XML, Java, Scheme, Matlab, Prolog, ACT-R

Languages
German (mother tongue), English, French, Latin

Awards

Push-Price for best thesis at the Otto-Friedrich-University of Bamberg (12/2008)

Interests

Music, Volleyball, Sketching, Comics, Virtual Reality, Analogies, Literature

Academic References

- **Prof H Müller,** General and Experimental Psychology, LMU Munich, Tel. +49 (0)89-2180-5327, hmueller@psy.lmu.de
- **Prof U Schmid,** Applied Informatics, Otto-Friedrich-University of Bamberg, Tel. +49 (0)951-863-2860, ute.schmid@uni-bamberg.de
- **Prof D Manzey**, Engineering Psychology, Berlin University of Technology, Tel. +49(0)30-31421340, dietrich.manzey@tu-berlin.de

PUBLICATIONS

Abstracts

(1) Wykowska, A., **Wiese, E.**, & Müller, H.J. (2012). Treating others as intentional agents influences our own perception. An EEG study. *Perception*, *41*, p.28.

(2) **Wiese, E.**, Zwickel, J., & Müller, H.J. (2011). Social effects on spatial attention. How precise are shifts in visual attention following centrally presented gaze cues? *Perception*, *40*, p.68

(3) **Wiese, E.**, Zwickel, J., & Müller, H.J. (2011). Soziale Einflussfaktoren auf räumliche Aufmerksamkeitsprozesse: Wie spezifisch ist die Aufmerksamkeitsverschiebung beim Beobachten der Blickrichtung anderer. In: K.Bittrich, S.Blankenberger & J.Lukas (eds.). *Beiträge zur 53. Tagung experimentell arbeitender Psychologen*, Halle, Germany, p. 193.

(4) **Wiese, E.**, Zwickel, J., & Müller, H.J. (2010). Bewirkt das Beobachten von Blickbewegungen eine räumlich-spezifische Aufmerksamkeitsverschiebung? In: C.Frings, A.Mecklinger, D.Wentura, & H.Zimmer (eds.). *Beiträge zur 52. Tagung experimentell arbeitender Psychologen*, Saarbrücken, Germany, p. 304.

(5) **Wiese, E.**, & Schmid, U. (2007). As much as needed or as much as possible – Inference and schema generalization in analogical reasoning. In: *KogWis 2007*, Saarbrücken, p. 24.

Book Chapter

(1) **Wiese, E.** & Wiese, L. (2012). Design Problem-Solving with External Representations. In: J.Adenauer & J.Petruschat (eds). *Prototype! physical, virtual, hybrid, smart* (Form+Zweck Verlag, Berlin), pp. 160-185.

Journal Articles

---

under review / in preparation

(1) **Wiese, E.**, Wykowska, A., & Müller, H.J. (under review). What we observe is biased by what other people tell us: Beliefs about the reliability of gaze behavior modulate attentional orienting to gaze cues. *Journal of Experimental Psychology: General.*

(2) Wykowska, A.*, **Wiese, E.***, Prosser, A., & Müller, H.J. (under review). Beliefs about the minds of others influence our own mind! An EEG study on the modulation of sensory gain control by higher-order cognition. *Journal of Neuroscience.*

(3) **Wiese, E.**, Wykowska, A., & Müller, H.J. (in preparation). The spatial specificity of gaze-cueing effects is modulated by facial expression.

published

(1) **Wiese, E.**, Zwickel, J., & Müller, H.J. (in press). The importance of context information for the spatial specificity of gaze cueing. *Attention, Perception & Psychophysics*. [Epub ahead of print], doi: 10.3758/s13414-013-0444-y.

(2) **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. doi:10. 1371/journal.pone.0045391.

(3) **Wiese, E.**, Zwickel, J., Müller, H.J. (2009). Im Auge des Anderen – Wie uns die Anwesenheit anderer beeinflusst. In: *In-Mind-Magazine, 2*.

(4) Bachinger, S., **Wiese, E.**, & Israel, J.H. (2009). Establishing a semantic differential on product prototype aesthetics: a research approach. In: *MMI Interaktiv, Special Issue HMI,* p. 1-8.

(5) Israel, J.H., **Wiese E.**, Mateescu, M., & Stark, R. (2009). Investigating three-dimensional sketching for early conceptual design – Results from expert discussions and user studies. *Computer & Graphics, 33(4),* p.462-473.

*equal contribution

Proceedings (peer-reviewed)

(1) **Wiese, E.**, Kohlbecher, S., & Müller, H.J. (2012). Acuity in estimating gaze direction in Human-Human and Human-Robot-Interaction. In: *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction* (HRI 2012), Boston, USA.

(2) Kohlbecher, S., **Wiese, E.**, et al. (2012). Studying gaze-based Human-Robot Interaction: An experimental platform. In: *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction* (HRI 2012), Boston, USA.

(3) **Wiese, E.**, Israel, J.H., Meyer, A., & Bongartz, S. (2010). Investigating the Learnability of Immersive Free-Hand Sketching. In: M.Alexa & E. Do (Eds) *EUROGRAPHICS SBIM*, p. 135-142.

(4) **Wiese, E.**, Israel, J.H., Zöllner, C., Pohlmeyer, A.E. & Stark, R. (2009). The potential of immersive 3D-sketching environments for design problem-solving. In: Stephanidis, C. et al. (Eds). *Proceedings of the 13th International Conference on HCI 2009*. San Diego, USA, p. 485-489.

(5) **Wiese, E.**, Israel, J.H., Zöllner, C., Mateescu, M. & Stark, R. (2009). 3D-Skizzieren in virtuellen Umgebungen – Neue Wege für kreatives Design-Problemlösen. In: Lichtenstein, A. et al. (Eds). *Proceedings of the 8th BWMMS*, Berlin, Germany, p. 110-115.

(6) **Wiese, E.** (2009). Design problem-solving with external representations. In: *Proceedings of the Mensch & Computer*, Berlin, Germany, p.34-36.

(7) **Wiese, E.**, Konerding, U. & Schmid, U. (2008). Mapping and inference in analogical problem-solving: As much as needed or as much as possible? In: Love, B.C., McRae, K., Sloutsky, V.M. (Eds). *Proceedings of the 30th Annual Conference of the Cognitive Science Society,* Washington D.C, USA. Mahwah, NJ: Lawrence Erlbaum, 2008, p. 927-932.

EIDESSTATTLICHE VERSICHERUNG / AFFIDAVIT

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation **Making eyes with others: How context information modulates attentional orienting to gaze direction** selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

I hereby confirm that the dissertation **Making eyes with others: How context information modulates attentional orienting to gaze direction** is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

München, den 28.03.2013

Munich, date 28.03.2013 Eva Wiese

Author Contributions

published:

(1) **Wiese, E.**, Zwickel, J., & Müller, H.J. (in press). The importance of context information for the spatial specificity of gaze cueing. *Attention, Perception & Psychophysics*. [Epub ahead of print], doi: 10.3758/s13414-013-0444-y

Author contributions: EW, HJM conceived and designed research. EW performed research. EW analyzed data. EW, JZ, HJM wrote the paper.

(2) **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. doi:10.1371/journal.pone.0045391.

Author contributions: EW, JZ, HJM conceived and designed research. EW performed experiments. EW and AW analyzed the data. EW, AW, HJM wrote the paper. AW theoretical conceptualization.

(3) **Wiese, E.**, Zwickel, J., & Müller, H.J. (2009). Im Auge des Anderen – Wie uns die Anwesenheit anderer beeinflusst. In: *In-Mind-Magazine, 2*.

Author contributions (review): EW theoretical conceptualization. EW, JZ, HJM wrote paper.

under review:

(1) **Wiese, E.**, Wykowska, A., & Müller, H.J. (under review). What we observe is biased by what other people tell us: Beliefs about the reliability of gaze behavior modulate attentional orienting to gaze cues. *Journal of Experimental Psychology: General.*

Author contributions: EW, HJM designed research. EW performed research. EW analyzed data. EW, AW, HJM wrote the paper.

(2) Wykowska, A.*, **Wiese, E.***, Prosser, A. & Müller, H.J. (under review). Beliefs about the minds of others influence our own mind! An EEG study on the modulation of sensory gain control by higher-order cognition. *Journal of Neuroscience.*

Author contributions: AW, EW, HJM conceived and designed research; AW programmed the experiment; EW, AP performed the experiment. AW analyzed data. AW, EW, AP, HJM wrote the paper. AW theoretically conceptualized the study. AP formalized study within the predictive coding framework.