Human consciousness and its relationship to social neuroscience: A novel hypothesis

Michael S. A. Graziano & Sabine Kastner

Ψ Psychology Press
Taylor & Francis Group

Discussion Paper

Human consciousness and its relationship to social neuroscience: A novel hypothesis

Michael S. A. Graziano and Sabine Kastner

Department of Psychology, Princeton University, Princeton, NJ, USA

A common modern view of consciousness is that it is an emergent property of the brain, perhaps caused by neuronal complexity, and perhaps with no adaptive value. Exactly what emerges, how it emerges, and from what specific neuronal process, are in debate. One possible explanation of consciousness, proposed here, is that it is a construct of the social perceptual machinery. Humans have specialized neuronal machinery that allows them to be socially intelligent. The primary role of this machinery is to construct models of other people's minds thereby gaining some ability to predict the behavior of other individuals. In the present hypothesis, awareness is a perceptual reconstruction of attentional state; and the machinery that computes information about other people's awareness is the same machinery that computes information about our own awareness. The present paper brings together a variety of lines of evidence, including experiments on the neural basis of social perception, on hemispatial neglect, on the out-of-body experience, on mirror neurons, and on the mechanisms of decision-making, to explore the possibility that awareness is a construct of the social machinery in the brain.

Men ought to know that from the brain, and from the brain only, arise our pleasures, joys, laughter and jests, as well as our sorrows, pains, griefs and tears. Through it, in particular, we think, see, hear, and distinguish the ugly from the beautiful, the bad from the good, the pleasant from the unpleasant.

Hippocrates, fifth century BC.

A common neuroscientific assumption about human consciousness is that it is an emergent property of information processing in the brain. Information is passed through neuronal networks, and by an unknown process consciousness of that information ensues. In such a view, a distinction is drawn between the information represented in the brain, which can be studied physiologically, and the still unexplained property of being conscious of that information. In the present paper, a novel hypothesis is proposed that differs from these common intuitive notions. The hypothesis is summarized in the following five points.

First, when a person asserts, "I am conscious of X," whatever X may be, whether a color, a tactile sensation, a thought, or an emotion, the assertion depends on some system in the brain that must have computed the information; otherwise, the information would be unavailable for report. Not only the information represented by X, visual information or auditory information, for example, but also the essence of consciousness itself, the inner feeling attached to X, must be information, or we

would be unable to say that we have it. In this hypothesis, consciousness is not an emergent property, or a metaphysical emanation, but is itself information computed by an expert system. This first point raises the question of why the brain would contain an expert system that computes consciousness. The question is addressed in the following points.

Second, people routinely compute the state of awareness of other people. A fundamental part of social intelligence is the ability to compute information of the type, "Bill is aware of X." In the present proposal, the awareness we attribute to another person is our reconstruction of that person's attention. This social capability to reconstruct other people's attentional state is probably dependent on a specific network of brain areas that evolved to process social information, though the exact neural instantiation of social intelligence is still in debate.

Third, in the present hypothesis, the same machinery that computes socially relevant information of the type, "Bill is aware of X," also computes information of the type, "I am aware of X." When we introspect about our own awareness, or make decisions about the presence or absence of our own awareness of this or that item, we rely on the same circuitry whose expertise is to compute information about other people's awareness.

Fourth, awareness is best described as a perceptual model. It is not merely a cognitive or semantic proposition about ourselves that we can verbalize. Instead, it is a rich informational model that includes, among other computed properties, a spatial structure. A commonly overlooked or entirely ignored component of social perception is spatial localization. Social perception is not merely about constructing a model of the thoughts and emotions of another person, but also about binding those mental attributes to a location. We do not merely reconstruct that Bill believes this, feels that, and is aware of the other, but we perceive those mental attributes as localized within and emanating from Bill. In the present hypothesis, through the use of the social perceptual machinery, we assign the property of awareness to a location within ourselves.

Fifth, because we have more complete and more continuous data on ourselves, the perceptual model of our own awareness is more detailed and closer to detection threshold than our perceptual models of other people's awareness.

The purpose of the present paper is to elaborate on the hypothesis summarized above and to review some existing evidence that is consistent with the hypothesis. None of the evidence discussed in this article is conclusive. Arguably, little conclusive evidence yet exists in the study of consciousness. Yet the evidence suggests some plausibility to the present hypothesis that consciousness is a perception and that the perceptual model is constructed by social circuitry.

The paper is organized in the following manner. First the hypothesis is outlined in greater detail (*Awareness as a product of social perception*). Second, a summary of recent work on the neuronal basis of social perception is provided (*Machinery for social perception and cognition*). A series of sections then describes results from a variety of areas of study, including hemispatial neglect, cortical attentional processing, aspects of self-perception including the out-of-body illusion, mirror neurons as a possible mechanism of social perception, and decision-making as a means of answering questions about one's own awareness. In each case, the evidence is interpreted in light of the present hypothesis. One possible advantage of the present hypothesis is that it may provide a general theoretical basis on which to understand and fit together a great range of otherwise disparate and incompatible data sets.

AWARENESS AS A PRODUCT OF SOCIAL PERCEPTION

The hypothesis that consciousness is closely related to social ability has been suggested previously in many forms (e.g., Baumeister & Masicampo, 2010; Carruthers, 2009; Frith, 1995; Gazzaniga, 1970; Humphrey, 1983; Nisbett & Wilson, 1977). Humans have neuronal machinery that apparently contributes to constructing models of other people's minds (e.g., Brunet, Sarfati, Hardy-Baylé, & Decety, 2000; Ciaramidaro et al., 2007; Gallagher et al., 2000; Samson, Apperly, Chiavarino, & Humphreys, 2004; Saxe & Kanwisher, 2003; Saxe & Wexler, 2005). This circuitry may also contribute to building a model of one's own mind (e.g., Frith, 2002; Ochsner et al., 2004; Saxe et al., 2006; Vogeley et al., 2001, 2004). The ability to compute explicit, reportable information about our own emotions, thoughts, goals, and beliefs by applying the machinery of social cognition to ourselves can potentially explain self-knowledge.

It has been pointed out, however, that self-knowledge does not easily explain consciousness (Crick & Koch, 1990). Granted that we have self-knowledge, and that we construct a narrative to explain our own behavior, how exactly do we become conscious of that information, and how does consciousness extend to other information domains such as colors, sounds, and tactile sensations? Constructing models of one's own mental processes could be categorized as "access consciousness" as opposed to "phenomenal consciousness"

(Block, 1996). It could be considered a part of the "easy problem" of consciousness, determining the information of which we are aware, rather than the "hard problem" of determining *how* we become aware of it (Chalmers, 1995).

Social approaches to consciousness are not alone in these difficulties. Other theories of consciousness suffer from similar limitations. One major area of thought on consciousness focuses on the massive, brain-wide integration of information. For example, in his global workspace theory, Baars was one of the first to posit a unified, brain-wide pool of information that forms the contents of consciousness (Baars, 1983; Newman & Baars, 1993). A possible mechanism for binding information across brain regions, through the synchronized activity of neurons, was proposed by Singer and colleagues (Engel, König, Gray, & Singer, 1990; Engel & Singer, 2001). Shortly after the first report from Singer and colleagues, Crick and Koch (1990) suggested that when information is bound together across regions of the cortex through the synchronized activity of neurons, it enters consciousness. Many others have since proposed theories of consciousness that include or elaborate on the basic hypothesis that consciousness depends on the binding of information (e.g., Grossberg, 1999; Lamme, 2006; Tononi, 2008; Tononi & Edelman, 1998). All of these approaches recognize that the content of consciousness includes a great complexity of interlinked information. But none of the approaches explain how it is that we become aware of that information. What exactly is the inner essence, the feeling of consciousness, that seems to be attached to the information?

Here we propose that the machinery for social perception provides that feeling of consciousness (Graziano, 2010). The proposal does not necessarily contradict previous accounts. It could be viewed as a way of linking social theories of consciousness with theories in which consciousness depends on informational binding. If consciousness is associated with a global workspace, or a bound set of information that spans many cortical areas, as so many others have suggested, then in the present proposal the awareness ingredient added to that global information set is provided by the machinery for social perception. In particular, awareness is proposed to be a rich descriptive model of the process of attention.

The proposal begins with the relationship between awareness and attention. The distinction between awareness and attention has been studied before (e.g., Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Jiang, Costello, Fang, Huang, & He, 2006; Kentridge, Heywood, & Weiskrantz, 2004; Koch

& Tsuchiya, 2007; Lamme, 2004; Naccache, Blandin, & Dehaene, 2002). The two almost always covary, but under some circumstances it is possible to attend to a stimulus and at the same time be unaware of the stimulus (Jiang et al., 2006; Kentridge et al., 2004; Naccache et al., 2002). Awareness, therefore, is not the same thing as attention, but, puzzlingly, the two seem redundant much of the time. Here we propose an explanation for the puzzling relationship between the two: Awareness is a perceptual model of attention. Like most informational models in the brain, it is not a literal transcription of the thing it represents. It is a caricature. It exaggerates useful, need-to-know information. Its purpose is not to provide the brain with a scientifically accurate account of attention, but to provide useful information that can help guide behavior. In the following paragraphs, we discuss first the social perception of someone else's attentional state, and then the perception of one's own attentional state.

Arguably, one of the most basic tasks in social perception is to perceive the focus of somebody else's attention. The behavior of an individual is driven mainly by the items currently in that individual's focus of attention. Hence, computing that someone is attending to this visual stimulus, that sound, this idea, and that emotion, provides critical information for behavioral prediction. The importance of computing someone else's state of attention has been emphasized by others, and forms the basis for a body of work on what is sometimes called "social attention" (e.g., Birmingham & Kingstone, 2009; Friesen & Kingstone, 1998; Frischen, Bayliss, & Tipper, 2007; Nummenmaa & Calder, 2008; Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010). One of the visual cues used to perceive someone else's attentional state is the direction of gaze. Neurons that represent the direction of someone else's gaze have been reported in cortical regions thought to contribute to social perception including, in particular, the superior temporal sulcus (STS) area of monkeys and humans (Calder et al., 2002; Hoffman & Haxby, 2000; Perrett et al., 1985; Puce, Allison, Bentin, Gore, & McCarthy, 1998; Wicker, Michel, Henaff, & Decety, 1998). Gaze is, of course, not the only cue. A variety of other cues, such as facial expression, body posture, and vocalization, presumably also contribute to perceiving the focus of somebody else's attention.

In the hypothesis proposed here, when we construct a perceptual model of someone else's focus of attention, that informational model describes awareness originating in that person and directed at a particular item. In this hypothesis, the brain explicitly computes an awareness construct, and awareness is the perceptual reconstruction of attention. Figure 1 provides an
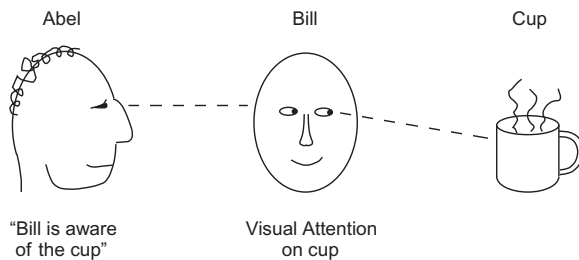
Figure 1.   Awareness as a social perceptual model of attention. Bill has his visual attention on the cup. Abel, observing Bill, constructs a model of Bill's mental state, using specialized neuronal machinery for social perception. Part of that model is the proposition that Bill is aware of the cup. In this formulation, awareness is a perceptual property that is constructed to represent the attentional state of a brain. We perceive awareness in other people. We can use the same neuronal machinery of social perception to perceive awareness in ourselves.

example to better explain this proposed relationship between awareness and attention.

In Figure 1, Abel looks at Bill and Bill looks at the coffee cup. First consider Bill, whose visual attention is focused on the cup. It is now possible to provide a fairly detailed account of visual attention, which has been described as a process by which one stimulus representation wins a neuronal competition among other representations (for review, see Beck & Kastner, 2009; Desimone & Duncan, 1995). The competition can be influenced by a variety of signals. For example, bottom-up signals, such as the brightness or the sudden onset of a stimulus, may cause its representation to win the competition and gain signal strength at the expense of other representations. Likewise, top-down signals that emphasize regions of space or that emphasize certain shapes or colors may be able to bias the competition in favor of one or another stimulus representation. Once a stimulus representation has won the competition, and its signal strength is boosted, that stimulus is more likely to drive the behavior of the animal. This self-organizing process is constantly shifting as one or another representation temporarily wins the competition. In Figure 1, Bill's visual system builds a perceptual model of the coffee cup that wins the attentional competition.

Now consider Abel, whose machinery for social perception constructs a model of Bill's mind. This model includes, among other properties, the following three pieces of information. First, awareness is present. Second, the awareness emanates from Bill. Third, the awareness is directed in a spatially specific manner at the location of the cup. These properties—the property of awareness and the two spatial locations to which it is referred—are perceptual constructs in Abel's brain.

In this formulation, Bill's visual attention is an event to be perceived, and awareness is the perceptual counterpart to it constructed by Abel's social machinery. Note the distinction between the reality (Bill's attentional process) and the perceptual representation of the reality (Abel's perception that Bill is aware). The reality is quite complex. It includes the physics of light entering the eye, the body orientation and gaze direction of Bill, and a large set of unseen neuronal processes in Bill's brain. The perceptual representation of that reality is much simpler, containing an amorphous, somewhat ethereal property of awareness that can be spatially localized at least vaguely to Bill and that, in violation of the physics of optics, emanates from Bill toward the object of his awareness. (For a discussion of the widespread human perception that vision involves something coming out of the eyes, see Cottrell & Winer, 1994; Gross, 1999.) The perceptual model is simple, easy, implausible from the point of view of physics, but useful for keeping track of Bill's state and therefore for helping to predict Bill's behavior. As in all perception, the perception of awareness is useful rather than accurate.

Consider now the modified situation in which Abel and Bill are the same person. A person is never outside a social context because he is always with himself and can always use his considerable social machinery to perceive, analyze, and answer questions about himself. Abel/Bill focuses visual attention on the coffee cup. Abel/Bill also constructs a model of the attentional process. The model includes the following information: awareness is present; the awareness emanates from me; the awareness is directed at the cup. If asked, "Are you aware of the cup?" Abel/Bill can cognitively scan the contents of this model and on that basis answer, "Yes."

If asked, "What exactly do you mean by awareness of the cup?" Abel/Bill can again scan the informational model, abstract properties from it, and report something like, "My awareness is a feeling, a vividness, a mental seizing of the stimulus. My awareness feels like it is located inside me. In a sense it *is* me. It is my mind apprehending something." These summaries reflect the brain's model of the process of attention.

Awareness, in this account, is one's social intelligence perceiving one's focus of attention. It is a second-order representation of attention. In that sense, the hypothesis may seem similar to proposals involving metacognition (e.g., Carruthers, 2009; Pasquali, Timmermans, & Cleeremans, 2010; Rosenthal, 2000). Metacognition generally refers to semantic knowledge about one's mental processes, or so-called "thinking about thinking." The proposal here, however, is

different. While people clearly have semantic knowledge about their attentional state, what is proposed here is specifically the presence of a rich, descriptive, perceptual model of attentional state that, like most perception, is computed involuntarily and is continuously updated. When we gain cognitive access to that perceptual model and summarize it in words, we report it as awareness. Block (1996) distinguished between phenomenal consciousness (the property of consciousness itself) and access consciousness (cognitive access to the property of consciousness). In the present theory, the perceptual representation of attentional state is akin to phenomenal consciousness. The cognitive access to that representation, which allows us abstract semantic knowledge and to report on it, is akin to access consciousness.

In the present hypothesis, we propose a similarity between perceiving someone else's awareness and perceiving one's own awareness. Both are proposed to be social perceptions dependent on the same neuronal mechanisms. Yet, do we really perceive someone else's awareness in the same sense that we perceive our own, or do we merely acknowledge in an abstract or cognitive sense that the other person is likely to be aware? In a face-to-face conversation with another person, so many perceptions and cognitive models are present regarding tone of voice, facial expression, gesture, and the semantic meaning of the other person's words, that it is difficult to isolate the specific perceptual experience of the other person's awareness. Yet, there is one circumstance in which extraneous perceptions are minimized and the perception of someone else's awareness is relatively isolated and therefore more obvious. This circumstance is illustrated in Figure 2. Everyone is familiar with the spooky sensation that someone is staring at you from behind (Coover, 1913; Titchner, 1898). Presumably built on lower level sensory cues such as subtle shadows or sounds, the perception of a mind that is located behind you and that is aware of you is a type of social perception and a particularly pure case of the perception of awareness. Other aspects of social perception are
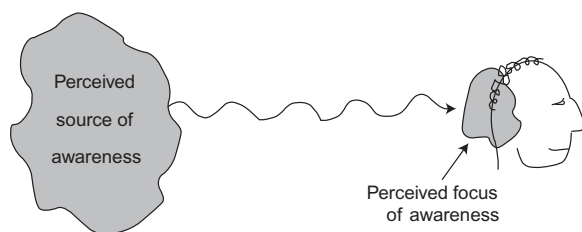
stripped away. The perceptual illusion includes a blend of three components: the perception that awareness is present, the perception that the awareness emanates from a place roughly localized behind you, and the perception that the awareness is directed at a specific object (you). This illusion helps to demonstrate that awareness is not only something that a brain perceives to be originating from itself—I am aware of this or that—but something that a brain can perceive as originating from another source. In the present argument, awareness is a perceptual property that can be attributed to someone else's mind or to one's own mind.

The present hypothesis emerges from the realization that social perception is not merely about reconstructing someone else's thoughts, beliefs, or emotions, but also about determining the state of someone else's attention. Information about someone else's attention is useful in predicting the likely moment-by-moment behavior of the person. The social machinery computes that Bill is aware of this, that, and the other. Therefore social perception, when applied to oneself, provides not only a description of one's own inner thoughts, beliefs, and feelings, but also a description of one's awareness of items in the outside environment. It is for this reason that awareness, awareness of anything, awareness of a color, or a sound, or a smell, not just self-awareness, can be understood as a social computation.

The examples given above focus on visual attention and visual awareness. The concept, however, is general. In the example in Figure 1, Bill could just as well attend to a coffee cup, a sound, a feeling, a thought, a movement intention, or many other cognitive, emotional, and sensory events. In the present hypothesis, awareness is the perceptual reconstruction of attention, and therefore anything that can be the subject of attention can also be the subject of awareness.

MACHINERY FOR SOCIAL PERCEPTION AND COGNITION

Arguably, social neuroscience began with the discovery by Gross and colleagues of hand and face cells in the inferior temporal cortex of monkeys (Desimone, Albright, Gross, & Bruce, 1984; Gross, Bender, & Rocha-Miranda, 1969). Further work indicated that a neighboring cortical area, the superior temporal polysensory (STP) area, contains a high percentage of neurons tuned to socially relevant visual stimuli including faces, biological motion of bodies and limbs, and gaze direction (Barraclough, Xiao, Oram, & Perrett, 2006; Bruce, Desimone, & Gross, 1981; Jellema & Perrett,

Figure 2. The perceptual illusion that somebody behind you is staring at you.

2003, 2006; Perrett et al., 1985). In humans, through the use of functional magnetic resonance imaging (fMRI), a region that responds more strongly to the sight of faces than to other objects was identified in the fusiform gyrus (Kanwisher, McDermott, & Chun, 1997). Areas in the human STS were found to become active during the perception of gaze direction and of biological motion such as facial movements and reaching (Grossman et al., 2000; Pelphrey, Morris, Michelich, Allison, & McCarthy, 2005; Puce et al., 1998; Thompson, Hardee, Panayiotou, Crewther, & Puce, 2007; Vaina, Solomon, Chowdhury, Sinha, & Belliveau; Wicker et al., 1988). This evidence from monkeys and humans suggests that the primate visual system contains a cluster of cortical areas that specializes in processing the sensory cues related to faces and gestures that are relevant to social intelligence.

Other studies in social neuroscience have investigated a more cognitive aspect of social intelligence sometimes termed "theory of mind" (Frith & Frith, 2003; Premack & Woodruff, 1978). Tasks that require the construction of models of the contents of other people's minds have been reported to engage a range of cortical areas, including the STS, the temporoparietal junction (TPJ), and the medial prefrontal cortex (MPFC), with a greater but not exclusive activation in the right hemisphere (Brunet et al., 2000; Ciaramidaro et al., 2007; Fletcher et al., 1995; Gallagher et al., 2000; Goel, Grafman, Sadato, & Hallett, 1995; Saxe & Kanwisher, 2003; Saxe & Wexler, 2005; Vogeley et al., 2001).

The relative roles of these areas in social perception and cognition are still in debate. It has been suggested that the TPJ is selectively recruited during theory-of-mind tasks, especially during tasks that require constructing a model of someone else's beliefs (Saxe & Kanwisher, 2003; Saxe & Wexler, 2005). Damage to the TPJ is associated with impairment in theory-of-mind reasoning (Apperly, Samson, Chiavarino, & Humphreys, 2004; Samson et al., 2004; Weed, McGregor, Feldbaek Nielsen, Roepstorff, & Frith, 2010).

The STS, adjacent to the TPJ, has been argued to play a role in perceiving someone else's movement intentions (Blakemore et al., 2003; Pelphrey, Morris, & McCarthy, 2004; Wyk, Hudac, Carter, Sobel, & Pelphrey, 2009). Not only is the STS active during the passive viewing of biological motion, as noted above, but the activity distinguishes between goal-directed actions, such as reaching to grasp an object, and non-goal-directed actions, such as arm movements that do not terminate in a grasp (Pelphrey et al., 2004). Even when a subject views simple geometric shapes that move on a computer screen, movements that are perceived as intentional activate the STS whereas movements that appear mechanical do not (Blakemore et al., 2003).

The role of the MPFC is not yet clear. It is consistently recruited in social perception tasks and theory-of-mind tasks (Brunet et al., 2000; Fletcher et al., 1995; Frith, 2002; Gallagher et al., 2000; Goel et al., 1995; Passingham, Bengtsson, & Lau, 2010; Vogeley et al., 2001), but lesions to it do not cause a clear deficit in theory-of-mind reasoning (Bird, Castelli, Malik, Frith, & Husain, 2004). Some speculations about the role of the MPFC in social cognition are discussed in subsequent sections.

Taken together, these studies suggest that a network of cortical areas, mainly but not exclusively in the right hemisphere, collectively build models of other minds. Different areas within this cluster may emphasize different aspects of the model, though it seems likely that the areas interact in a cooperative fashion.

The view that social perception and cognition are emphasized in a set of cortical areas dedicated to social processing is not universally accepted. At least two main rival views exist. One view is that the right TPJ and STS play a more general role related to attentional processing rather than a specific role related to social cognition (e.g., Astafiev, Shulman, & Corbetta, 2006; Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Mitchell, 2008; Shulman et al., 2010). A second alternative view is that social perception is mediated at least partly by mirror neurons in the motor system that compute one's own actions and also simulate the observed actions of others (e.g., Rizzolatti & Sinigaglia, 2010). Both of these alternative views are discussed in subsequent sections. Much of the discussion below, however, is based on the hypothesized role of the right TPJ and STS in social perception and social cognition.

PREDICTION 1: DAMAGE TO THE MACHINERY FOR SOCIAL PERCEPTION SHOULD CAUSE A DEFICIT IN AWARENESS

If the present proposal is correct, if awareness is a construct of the machinery for social perception, then damage to the right TPJ and STS, the brain areas most associated with constructing perceptual models of other minds, should sometimes cause a deficit in one's own awareness. These cortical regions are heterogeneous. Even if we assume their role in social cognition, different sub-areas probably emphasize different functions. As discussed in the next section, the sub-regions of TPJ involved in attention may be partially distinct from sub-regions involved in theory of

mind (Scholz, Triantafyllou, Whitfield-Gabrieli, Brown, & Saxe, 2009). Therefore, even in the present hypothesis, a lesion to the TPJ and STS should not always affect all aspects of social cognition equally. A range of symptoms might result. The present hypothesis does, however, make a clear prediction: Damage to the right TPJ and STS should often be associated with a deficit in consciousness.

The clinical syndrome that comes closest to an awareness deficit is hemispatial neglect, the loss of processing of stimuli usually on the left side of space after damage to the right hemisphere of the brain (Brain, 1941; Critchley, 1953). Patients classically fail to report, react to, or notice anything on the left half of space, whether visual, auditory, tactile, or memory. The left half of space, and any concept that it ever existed, are erased from the patient's awareness.

It is now generally accepted that there is no single neglect syndrome. A range of lesion sites can result in neglect, and different neglect patients can have somewhat different mixtures of symptoms (Halligan, Fink, Marshall, & Vallar, 2003; Halligan & Marshall, 1992; Vallar, 2001). It is therefore not correct to attribute neglect to a single brain area or mechanism. It has been reported, however, that a strong form of neglect, the almost total loss of conscious acknowledgment of the left side of space or anything in it, occurs most often after damage to the right TPJ (Valler & Perani, 1986). In at least one subset of neglect patients who lacked any accompanying low-level blindness, and in this sense were more "pure" in their neglect symptoms, the most common lesion site was ventral and anterior to the TPJ, in the right STS (Karnath, Ferber, & Himmelbach, 2001). Temporary interference with the right TPJ by transcranial magnetic stimulation (TMS) has also been reported to induce symptoms of left hemispatial neglect (Meister et al., 2006).

Some have argued that neglect is more commonly associated with damage to the posterior parietal lobe rather than to the more ventral TPJ or STS (Mort et al., 2003). A parietal locus for neglect is certainly a more traditional view (Brain, 1941; Critchley, 1953; Gross & Graziano, 1995). Neglect symptoms can also be observed after frontal lesions, though they tend to be less severe (Heilman & Valenstein, 1972a; Mesulam, 1999; Ptak & Schnider, 2010). These parietal and frontal sites for neglect are consistent with the proposal of a parietofrontal network for the top-down control of attention (Ptak & Schnider, 2010; Szczepanski, Konen, & Kastner, 2010). How can parietofrontal sites for neglect be reconciled with the observation of severe neglect from lesions in the TPJ and STS? One of the primary reasons for these differences among studies may be a disagreement over the definition of neglect. Different tests for neglect may result in a selection of different patient populations and therefore different observed lesion sites (Ferber & Karnath, 2001; Rorden, Fruhmann, Berger, & Karnath, 2006). In particular, the two most common clinical tests for neglect may measure different deficits. In a line cancellation test, the patient crosses out short line segments scattered over a visual display. This test measures awareness because the patient cancels only the line segments that reach awareness and fails to cancel the line segments that do not reach awareness. In contrast, in a line-bisection task, the patient attempts to mark the center of a long horizontal line. This test does not measure awareness or the lack thereof, since the patient is always aware of the horizontal line. Instead the test measures a relative response bias toward one side, perhaps caused by an underlying attentional bias. Different variants of neglect, revealed by these different tests, might be associated with different lesion sites. It is of interest, given the present hypothesis, that neglect defined as an awareness deficit through the use of line cancellation was associated with more ventral lesions in the STS (Ferber & Karnath, 2001; Karnath et al., 2001; Rorden et al., 2006).

The point of the present section is not that neglect is "really" an awareness deficit instead of an attentional bias, and not that it is "really" caused by lesions to the right TPJ or STS instead of to the right inferior parietal lobe. Undoubtedly, different symptoms and lesions sites can fall under the more general label of hemispatial neglect. The present hypothesis predicts that there should be at least two different kinds of neglect associated with two different brain systems, one associated with the process of *controlling* attention (neglect caused by damage to parietofrontal attentional mechanisms) and the other associated with the process of perceptually *representing* attention (neglect caused by TPJ and STS damage).

Does the unilateral nature of neglect argue against the present interpretation? Shouldn't a consciousness area, if damaged, lead to a total loss of conscious experience and not merely a unilateral loss? We do not believe the laterality of neglect argues against the present interpretation. If social perception and social cognition were found to activate a right hemisphere region only, then damage to that region, according to the present theory, might eliminate all consciousness. The studies reviewed above on the neuronal basis of social perception and social cognition, however, suggest that these functions are represented in a bilateral manner with a strong emphasis on the right side. If consciousness is a construct of the social machinery, as suggested here, then it is portioned in some unequal manner between the hemispheres. One would expect,

therefore, that damage to the system for social perception would, depending on the hemisphere, have an asymmetric effect on consciousness.

Some controversy surrounds the exact explanation of the unilateral nature of neglect. One view is a representational hypothesis (Heilman & Valenstein, 1972b; Mesulam, 1981, 1999). In that hypothesis, the right hemisphere contains some critical type of representation needed for awareness, which covers both sides of space, whereas the left hemisphere contains a representation only of the right side of space. Damage to the right hemisphere therefore leaves the patient without awareness of the left side of space, whereas damage to the left hemisphere leaves the patient mainly behaviorally intact. An alternative explanation is based on the concept of interhemispheric competition among controllers of attention (Kinsbourne, 1977; Szczepanski et al., 2010). In that view, damage to the parietal lobe on one side reduces the ability of the attentional control system that serves the opposite side of space. An imbalance is created and attention is unavoidably drawn to the ipsilesional side of space. The present theory of consciousness may be able to accommodate both of these seemingly conflicting views of the mechanisms of neglect. In the present theory, the brain contains at least two general processes related to attention. One is the control of attention, perhaps more emphasized in a parietofrontal network, lesions to which may result in a more competition-style imbalance of attention; and the other is the perceptual representation of attention, emphasized in more ventral areas involved in social perception, including the TPJ and the STS, damage to which may result in a more representational-style neglect.

PREDICTION 2: THE MACHINERY FOR SOCIAL PERCEPTION SHOULD CARRY SIGNALS THAT CORRELATE WITH ATTENTION

In the present hypothesis, a basic task of social perception is to reconstruct the focus of someone else's attention. In the same manner, in perceiving oneself, the social machinery reconstructs one's own constantly changing focus of attention. A prediction that follows from this hypothesis is that tasks that involve focusing or shifting attention should evoke brain activity not only in areas that participate in attentional control (such as the parietofrontal attention network; for review, see Beck & Kastner, 2009) but also in the social circuitry that generates a reconstruction of one's focus

of attention. The prediction is therefore that the right TPJ and STS should be active in association with changes in attention.

Area TPJ and the adjacent, caudal STS show elevated activity in response to stimuli that are unexpected, that appear at unexpected locations, or that change in an unexpected manner (Astafiev et al., 2006; Corbetta et al., 2000; Shulman et al., 2010). This response to salient events is strongest in the right hemisphere. It has therefore been suggested that the TPJ and the STS are part of a right-lateralized system, a "ventral network" that is involved in some aspects of spatial attentional processing (Astafiev et al., 2006; Corbetta et al., 2000; Shulman et al., 2010).

The attention-related responses found in the STS and the TPJ could be viewed as contradictory to the social functions proposed for the same areas. It could be argued that the attention findings challenge the view that the right STS and TPJ are in any way specialized for social processing. Perhaps they serve a more general function related to attention. In the present hypothesis, however, the two proposed functions are compatible. There is no contradiction. The social machinery is constantly computing and updating a model of one's own mind, and central to that model is a representation of one's focus of attention. Tracking an individual's attention is a fundamental task in social perception.

An alternate explanation of why attentional functions and social functions are both represented in the right TPJ and STS is that the two functions are emphasized in different cortical areas that happen to be so near each other that fMRI cannot easily resolve them. One functional imaging study indicated that the two types of tasks recruited overlapping areas of activity in the TPJ with spatially separated peaks (Scholz et al., 2009). This attempt to separate the "attention" sub-area of TPJ from the "theory-of-mind" sub-area, however, does not entirely address the underlying issue of proximity. The cortex is generally organized by functional proximity. Similar or related functions tend to be processed near each other (Aflalo & Grazianio, 2011; Graziano & Aflalo, 2007). This trend toward functional clustering is certainly true in the social domain. Even if some spatial separation exists between attentional functions and theory-of-mind functions, why are the two functions clustered so closely within the right TPJ and STS? Rather than view this proximity of different functions as a contradiction, or suggest that one function must be correct and the other a mistake, or suggest that the functions must be separate from each other and co-localized merely as an accident of brain organization, it is suggested here that the functions share a

common cause and therefore are not mutually contradictory. In the present hypothesis, reconstructing someone's beliefs (such as in a false belief task) is only one component of social perception. Another component is tracking someone's attention; a third is constructing a model of someone's movement intentions. Presumably, many other components of social perception exist, and these components may be represented cortically in adjacent, partially separable regions.

PREDICTION 3: THE MACHINERY FOR SOCIAL PERCEPTION SHOULD CONSTRUCT A SELF-MODEL

A direct prediction of the present hypothesis is that the cortical machinery that constructs models of other minds should also construct a perceptual model of one's own mind.

Some of the most compelling evidence on the self-model involves the out-of-body illusion, in which the perceived location of the mind no longer matches the actual location of the body. It is essentially an error in constructing a perceptual model of one's own mind. Perceptual mislocalizations of the self can be induced in normal, healthy people by cleverly manipulating visual and somatosensory information in a virtual-reality setup (Ehrsson, 2007; Lenggenhager, Tadi, Metzinger, & Blanke, 2007). A complete out-of-body illusion can also be induced with electrical stimulation of the cortex. In one experiment, Blanke, Ortigue, Landis, and Seeck (2002) electrically stimulated the cortical surface in an awake human subject. When electrical stimulation was applied to the TPJ in the right hemisphere, an out-of-body experience was induced. The stimulation evidently interfered with the machinery that normally assigns a location to the self. Subsequent experiments showed that tasks that involve mentally manipulating one's spatial perspective evoke activity in the right TPJ (Blanke et al., 2005; Zacks, Vettel, & Michelon, 2003) and are disrupted by TMS of the right TPJ (Blanke et al., 2005). The evidence suggests that a specific mechanism in the brain is responsible for building a spatial model of one's own mind, and that the right TPJ plays a central role in the process. Given that the right TPJ has been implicated so strongly in the social perception of others, this evidence appears to support the hypothesis that the machinery for social perception also builds a perceptual model of one's own mental experience.

An entire subfield of neuroscience and psychology is devoted to the subject of the body schema (for review, see Graziano & Botvinick, 2002). For example,

experiments have examined the neuronal mechanisms by which we know the locations of our limbs, incorporate tools into the body schema, or experience body-shape illusions. However, among the many studies on the body schema, the particular phenomenon of the out-of-body experience is uniquely relevant to the discussion of awareness. The reason is that the out-of-body experience is specifically a spatial mislocalization of the source of one's awareness.

The very existence of an out-of-body illusion could be interpreted as evidence in support of the present hypothesis, for the following reasons. Suppose that awareness is not explicitly computed, as hypothesized here, but instead is an emergent property produced as a byproduct of neuronal activity. Neurons compute and transmit information, and, as a result, somehow awareness of that information occurs. Why would we feel the awareness as located in, and emanating from, our own selves? By what mechanism does the awareness feel as though it is anchored here or anywhere else? The fact that awareness comes with a perceived location suggests that it is a perceptual model that, like all perception, is associated with a computed spatial location. Usually, one does not notice that awareness has a perceived location. Its spatial structure, emanating from inside one's own body, is so obvious and so ordinary that we take it for granted. The out-of-body illusion, in which the spatial computation has gone awry, allows one to realize that awareness comes with a computed spatial arrangement. Evidently, perceptual machinery in the brain computes one's awareness and assigns it a perceived source inside one's body, and interference with the relevant circuitry results in an error in the computation. The out-of-body experience demonstrates that awareness is a computation subject to error.

In a second general approach to the self-model, subjects in a series of experiments performed self-perception tasks such as answering questions about their own feelings or beliefs, or monitoring their own behavior while brain activity was measured through functional imaging (Frith, 2002; Ochsner et al., 2004; Passingham et al., 2010; Saxe et al., 2006; Vogeley et al., 2001). The network of areas recruited during self-perception included the right TPJ, STS, and MPFC. The activity in the MPFC was the largest and most consistent. Again, the findings appear to support the general hypothesis that the machinery for social perception is recruited to build a perceptual model of the self. These functional imaging studies on the self-model, however, are less pertinent to the immediate question of awareness. The reason is that the studies examined self-report of emotions or thoughts, not the

report of awareness itself. The focus of the present proposal is not self-awareness in particular, but awareness in general. It is extremely difficult to pinpoint the source of awareness in a functional imaging study of this type, because regardless of the task, whether one is asked to be aware of one's own mind state or of someone else's, one is nonetheless aware of something, and therefore the computation of awareness is present in both experimental and control tasks.

CHALLENGE 1: HOW CAN THE INNER "FEELING" OF AWARENESS BE EXPLAINED?

Thus far, we have presented a novel hypothesis about human consciousness, outlined three general predictions that follow from the hypothesis, and reviewed findings that are consistent with those predictions. The evidence is by no means conclusive, but does suggest that the hypothesis has some plausibility. Human consciousness may be a product of the same machinery that builds social perceptual models of other people's mental states. Arguably, the strongest evidence in favor of the hypothesis is that damage to that social machinery causes a profound deficit in consciousness. The hypothesis, however, also has some potential challenges. Several aspects of conscious experience are not, at the outset, easily explainable by the present hypothesis. We address three of these potential challenges to the hypothesis. In each case, we offer a speculation about how the hypothesis might explain, or at least be compatible with, the known phenomena. These sections are necessarily highly speculative, but the issues must be addressed for the present hypothesis to have any claim to plausibility.

The first challenge we take up concerns the difference between constructing an informational model about awareness and actually "feeling" aware. In the present hypothesis, networks that are expert at social computation analyze the behavior of other people and compute information of the type, "Bill is aware of X." The same networks, by hypothesis, compute information of the type, "I am aware of X," thereby allowing one to report on one's own awareness. Can such a machine actually "feel" aware, or does it merely compute an answer without any inner experience?

The issue is murky. If you ask yourself, "Am I merely a machine programmed to answer questions about awareness, or do I actually *feel* my own awareness," you are likely to answer yes, you actually feel it. You might even specify that you feel it inside you as a somewhat amorphous but nonetheless real thing. But in doing so, you are merely computing the answers to other questions. Distinguishing between actually having an inner experience, and merely computing, when asked, that you have an inner experience, is a difficult, perhaps impossible task.

In the present account, we are not hypothesizing that a set of semantic, symbolic, or linguistic propositions can take the place of the inner essence of awareness, but that a representation, an informational picture, comprises awareness. In the distinction between phenomenal consciousness and access consciousness suggested by Block (1996), the informational representation proposed here is similar to phenomenal consciousness, and our ability to cognitively access that representation and answer questions about it is similar to access consciousness.

Yet, even assuming a rich representational model of attention, and even assuming that the model comprises what we report to be awareness, why does it "feel" like something to us? Why the similarity between awareness and feeling? To address the issue, we begin with a set of illusions, both related to the spatial structure assigned to the model of awareness.

As mentioned in the previous section, the out-of-body experience is the illusion of floating outside your own body. The perceived source of your awareness no longer matches the actual location of your body. The illusion can be induced by direct electrical stimulation applied to the right TPJ (Blanke et al., 2002). Similar mislocalizations of the self can be induced in people by manipulating visual and somatosensory information in a virtual-reality setup (Ehrsson, 2007; Lenggenhager et al., 2007). Evidently, a specific mechanism in the brain is responsible for embodiment, or building a spatial model of one's own awareness.

Does a similar property of embodiment apply when constructing a model of someone else's awareness? The illusion illustrated in Figure 2 suggests that a model of someone else's awareness does come with an assigned location. The perception that someone is staring at you from behind (Coover, 1913; Titchner, 1898) includes a spatial structure in which awareness emanates from a place roughly localized behind you and is directed at you.

These illusions demonstrate a fundamental point about social intelligence that is often ignored. When building a model of a mind, whether one's own or someone else's, the process is not merely one of inferring disconnected mental attributes—beliefs, emotions, intentions, awareness. The model also has a spatial embodiment. The model, like any other perceptual model, consists of a set of attributes assigned to a location. Social perception is not normally compared directly to sensory perception, but in this manner they are similar. They both involve computed properties—

mental properties or sensory properties—as well as a computed location, bound together to form a model of an object.

In this hypothesis, one's own consciousness is a perceptual model in which computed properties are attributed to a location inside one's body. In this sense, consciousness is another example of somesthesis, albeit a highly specialized one. It is a perceptual representation of the workings of the inner environment. Like the perception of a stomach ache, or the perception of joint rotation, or the perception of light-headedness, or the perception of cold and hot, the perception of one's own awareness is a means of monitoring processes inside the body.

A central philosophical question of consciousness could be put this way: Why does thinking *feel* like something? When Abel solves a math problem in his head, why doesn't he merely process the numerical information and output the number without feeling it? Why does he feel as though a process is occurring inside his head? To rephrase the question more precisely: Why is it that, on introspection, when he engages a decision process to compare thinking to somatosensory processing, he consistently concludes that a similarity exists? We speculate that a part of the answer may be that his awareness of his own thinking is a model of the internal environment, assigned to a location inside his body, and as such is a type of body perception. He therefore reports that thinking shares associations with processing in the somatosensory domain.

Consider again the illusion diagrammed in Figure 2, the feeling that somebody is looking at you from behind. Subjectively, you feel something on the back of your neck or in your mind that seems to warn you. You feel the presence of awareness. But it is not your own awareness. It is a perceptual model of someone else's awareness. How can the feeling be explained? The perceptual model includes a property (awareness) and two spatial locations (the source of the awareness behind you and the focus of the awareness on you). One of these spatial locations overlaps with your self-boundary. It "feels" like something in the sense that a property is assigned to a location inside you. It literally belongs to the category of somesthesis because it is a perceptual reconstruction of the internal environment. What is key here is not merely localizing the event to the inside of the body, but the rich informational associations that come with that localization, the associations with other forms of somesthesis including touch, temperature, pressure, and so on. In this speculation, the perceptual representation of someone else's attention on the back of one's head comes with a vast

shadowy complex of information that is linked to it and is subtly activated along with it.

In addition to the circumstance of "feeling" someone else's awareness directed at you, consider the opposite circumstance in which your own awareness is directed at something else. Do you "feel" your own awareness of the item? In the present hypothesis, your machinery for social perception constructs a model of your own attentional processes. That self-model includes the property of awareness, a source of the awareness inside you, and a focus of the awareness on the object. One of these spatial locations overlaps with your self-boundary. On introspection, you can decide that you "feel" the source of your own awareness, in the sense that it is a perceptual model of your internal environment and therefore shares rich informational associations with other forms of somesthesis.

In summary, we speculate that even though conscious experience is information computed by expert systems, much like the information in a calculator or a computer, it nonetheless can "feel" like something in at least the following specific sense: A decision process, accessing the self-model, arrives at the conclusion and triggers the report that awareness shares associations with somesthesis. This account, of course, does not explain conscious experience in its entirety; but it helps to make a central point of the present perspective. 'Experienceness' itself may be a complex weaving together of information, and it is possible to tease apart at least some of that information (such as the similarity to somesthesis) and understand how it might be computed.

CHALLENGE 2: HOW CAN THE MACHINERY FOR SOCIAL PERCEPTION GAIN ACCESS TO MODALITY-SPANNING INFORMATION?

Social perception presents a challenge to neuronal representation. In color perception, one can study the areas of the brain that receive color information. In auditory perception, one can study auditory cortex. But in social perception, a vast range of information must converge. Is the other person aware of that coffee cup? Is he aware of the chill in the room? Is he aware of the abstract idea that I am trying to get across, or is he distracted by his own thoughts? What emotion is he feeling? Social perception requires an extraordinary, multimodal linkage of information. All things sensory, emotional, or cognitive that might affect another person's behavior must be brought together and considered, in order to build a useful predictive model of the

other person's mind. Hence, the neuronal machinery for social perception must be an information nexus. It is implausible to talk about a specific region of the brain that performs all of social perception. If the TPJ and STS play a central role, which seems likely on the evidence reviewed above, they must serve as nodes in a brain-wide network.

In the present hypothesis, awareness is part of a social perceptual model of one's own mind. Consider the case in which you report, "I am aware of the green apple." Which part of this linked bundle of information is encoded in the social machinery, perhaps including the TPJ or STS, and which is encoded elsewhere in the brain? Can the TPJ or STS contain a complete, unified representation of the awareness of green apples? Or, instead, does the social circuitry compute the "awareness" construct, and the visual cortex represent the "green apple," while a binding process links the two neuronal representations? On sheer speculation, the second possibility seems more likely. Certainly, the importance to consciousness of binding information across wide regions of the brain has been discussed before (e.g., Baars, 1983; Crick & Koch, 1990; Damasio, 1990; Engel & Singer, 2001; Grossberg, 1999; Lamme, 2006; Llinas & Ribary, 1994; Newman & Baars, 1993; Treisman, 1988; von der Malsburg, 1997). From the perspective of information theory, it has been suggested that consciousness is massively linked information (Tononi, 2008; Tononi & Edelman, 1998). The contribution of the present hypothesis is to suggest that binding of information across cortical areas is not, by itself, the raw material of awareness, but instead awareness is specific information, a construct about the nature of experience, that is computed by and represented in specific circuitry and that can be bound to larger, network-spanning representations.

In Treisman's feature integration theory (Treisman & Gelade, 1980), when we perceive an object, different informational features, such as shape, color, and motion, are bound together to form a unified representation. That binding requires attention. Without attention to the object, binding of disparate information about the object is possible but incomplete and often in error. If awareness acts like a feature that can be bound to an object representation, then attention to an object should be necessary for consistent or robust awareness to be attached to the object. In contrast, awareness of an object should not be necessary for attention to the object. In other words, it should be possible to attend to an object without awareness of it, but difficult to be aware of an object without attention to it. This pattern broadly matches the literature on the relationship between attention and awareness (e.g., Dehaene et al., 2006; Jiang

et al., 2006; Kentridge et al., 2004; Koch & Tsuchiya, 2007; Lamme, 2004; Naccache et al., 2002).

Koch and Tsuchiya (2007) argued that, qualitatively, it seems possible to be aware of stimuli at the periphery of attention, and therefore awareness must be possible with minimal attention. But note that, in this circumstance, speaking qualitatively again, one tends to feel aware of something without knowing exactly what the something is. Consider the proverbial intuition that something's wrong, or something's present, or something is intruding into awareness at the edge of vision, without being able to put your finger on what that item is exactly. This feeling could be described as awareness that is not fully or reliably bound to a specific item. Only on redirecting attention does one become reliably aware of the item. In this sense, awareness acts precisely like the unreliable or false conjunctions occurring outside the focus of awareness in feature integration theory. It acts like computed information about an object, such as a feature that, without attention, is not reliably bound to the object representation.

In the present hypothesis, therefore, the relationship between attention and awareness is rather complex. Not only is awareness a perceptual reconstruction of attention, but also to bind awareness to a stimulus representation requires attention, because attention participates in the mechanism of binding.

CHALLENGE 3: IS THE PRESENT HYPOTHESIS COMPATIBLE WITH SIMULATION THEORY?

In the present hypothesis, the human brain evolved mechanisms for social perception, a type of perception that allows for predictive modeling of the behavior of complex, brain-controlled agents. There is no assumption here about whether perception of others or perception of oneself emerged first. Presumably they evolved at the same time. Whether social perception is applied to oneself or to someone else, it serves the adaptive function of a prediction engine for human behavior.

An alternative view is that awareness emerged in the human brain for reasons unknown, perhaps as an epiphenomenon of brain complexity. Social perception was then possible by the use of empathy. We understand other people's minds by comparison to our own inner experience. This second view has the disadvantage that it does not provide any explanation of what exactly awareness is or why it might have evolved. It merely postulates that it exists. These two potential views of consciousness have their counterparts in the literature on social perception.

There are presently two main rival views of social perception. The first view, reviewed in the sections above, is that social perception depends on expert systems probably focused on the right TPJ and STS that evolved to compute useful, predictive models of minds. The second view is simulation theory. In simulation theory, social perception is the result of empathy. We understand other people's minds by reference to our own internal experience. The hypothesis about consciousness proposed here was discussed above almost entirely in the framework of the "expert systems" view of social perception. Can the hypothesis be brought into some compatibility with simulation theory? The present section argues that the expert systems view and simulation theory are not mutually exclusive and could, at least in principle, operate cooperatively to allow for social perception. The hybrid of the two mechanisms is consistent with the present hypothesis about consciousness.

The experimental heart of simulation theory is the phenomenon of mirror neurons. Rizzolatti and colleagues first described mirror neurons in the premotor cortex of macaque monkeys, in a region thought to be involved in the control of the hand for grasping (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Each mirror neuron became active during a particular, complex type of grasp such as a precision grip or a power grip. The neuron also became active when the monkey viewed an experimenter performing the same type of grip. Mirror neurons were therefore both motor and sensory. They responded whether the monkey performed or saw a particular action.

Mirror-like properties were reported in the human cortex in fMRI experiments in which the same area of cortex became active whether the subject performed or saw a particular action (Buccino et al., 2001; Filimon, Nelson, Hagler, & Sereno, 2007; Iacoboni et al., 2005). The hypothesized mirror-neuron network includes sensory-motor areas of the parietal lobe, anatomically connected regions of the premotor cortex, and perhaps regions of the STS (Rizzolatti & Sinigaglia, 2010).

The hypothesized role of mirror neurons is to aid in understanding the actions of others. In that hypothesis, we understand someone's hand actions by activating our own motor machinery and covertly simulating the actions. The mirror-neuron hypothesis is in some ways an elaboration of Liberman's original hypothesis about speech comprehension (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), in which we categorize speech sounds by using our own motor machinery to covertly mimic the same sounds.

The concept of mirroring can in principle be generalized beyond the perception of other people's hand actions to all social perception. We may understand someone else's joy, complete with nuances and psychological implications, by using our own emotional machinery to simulate that joy. We may understand someone else's intellectual point of view by activating a version of that point of view in our own brains. We may understand other minds in general by simulating them with the same machinery in our own brains. The extent to which mirror neurons directly cause social perception, however, or are a product of more general associative or predictive mechanisms, has been the subject of some discussion (e.g., Heyes, 2010; Kilner, Friston, & Frith, 2007).

One difficulty facing simulation theory is that it does not provide any obvious way to distinguish one's own thoughts, intentions, and emotions from someone else's. Both are run on the same hardware. If simulation theory is strictly true, then one should be unable to tell the difference between one's own inner experience and one's perception of someone else's.

A second difficulty facing simulation theory is that it contains some circularity. Before brain A can mirror the state of brain B, it needs to know what state to mirror. brain A needs a mechanism that generates hypotheses about the state of brain B.

Both of these difficulties with simulation theory disappear in a hybrid system proposed here. The system diagramed in Figure 3 contains expert systems (including the TPJ and STS) that contribute to generating models of other minds, and mirror-neuron networks that simulate and thereby refine those models.

As discussed in previous sections, a perceptual model of a mind includes a spatial location assigned to the model. It is true perception in the sense that the perceived properties have a perceived location, to form a perceived object. This spatial embodiment allows one to keep track of whether it is a model of one's own mind or of this or that other person's mind, solving the first difficulty of simulation theory. You know whether the perceived mental states are yours or someone else's in the same way that you know whether a certain color belongs to one object or another—by a spatial tag.

A perceptual model of a mind can also supply the necessary information to drive the mirror-neuron simulations. The mirror-neuron system, in this proposal, knows to simulate a reach because the biological motion detectors in the STS have used visual cues to categorize the other person's action as a reach. This
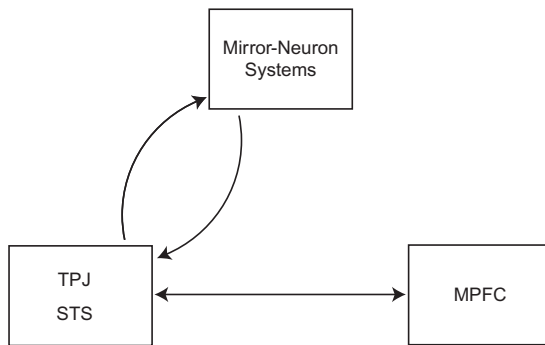
Figure 3. A proposed scheme that integrates simulation theory with the theory of dedicated cortical areas for social cognition. The box labeled "TPJ, STS" represents a cluster of cortical areas that contributes to building perceptual models of minds, including a self-model and models of other minds. The box labeled "MPFC" represents a prefrontal area that contributes to decisions in the social domain. The box labeled "Mirror-Neuron Systems" represents brain-wide networks that simulate and thereby refine the models of minds generated in TPJ and STS.
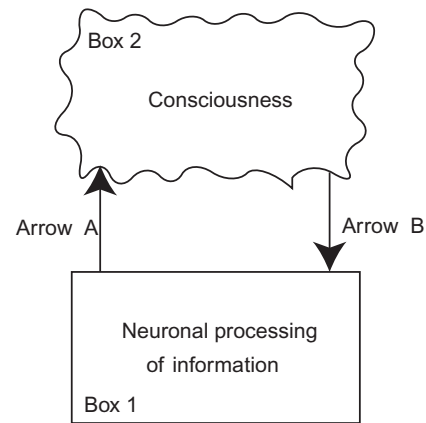
Figure 4. A traditional view in which consciousness emerges from the information processed in the brain. Consideration of Arrow A, how the brain creates consciousness, leads to much controversy and little insight. Consideration of Arrow B, how consciousness affects the brain, leads to the inference that consciousness must be information, because only information can act as grist for decision machinery, and we can decide that we have consciousness.

process solves the second difficulty of simulation theory. The likely dependence of mirror neurons on an interaction with the STS has been emphasized before (Rizzolatti & Sinigaglia, 2010).

Mirror neurons therefore should not be viewed as rivals to the theory that social perception is emphasized in specialized regions such as the TPJ and STS. Instead, the two mechanisms for social perception could in principle operate in a cooperative fashion. Areas such as the TPJ and STS could generate hypotheses about the mind states and intentions of others. Mirror neurons could then use these generated hypotheses to drive simulations. The simulations have the potential to provide a detailed, high-quality feedback, resulting in a more elaborate, more accurate model of the other mind. In this proposed scheme (Figure 3), the mirror neuron system is an extended loop adding to and enhancing the machinery that constructs models of minds.

Think how much more complicated, in a recursive, loop-the-loop way, the system becomes when the process of social perception is turned inward. Suppose your self-model includes the hypothesis that you are happy right now. To enhance that hypothesis, to enrich the details through simulation, the machinery that constructs your self-model contacts and uses your emotion-generating machinery. In that case, if you were not actually happy, the mirroring process should make you become so as a side product. If you were already happy, perhaps you become more so. Your self-model and the self that is being modeled become

intertwined in complicated ways. Perceiving your own mind changes the thing being perceived, a phenomenon long known to psychologists (Bandura, 2001; Beauregard, 2007).

CONSCIOUSNESS AS INFORMATION

Figure 4 diagrams a traditional way to conceptualize the problem of consciousness. When information is processed in the brain in some specific but as yet undetermined way (Box 1), a subjective experience emerges (Box 2).

For example, imagine that you are looking at a green apple. Your visual system computes information about the absorption spectrum of the apple. The presence of this wavelength information in the brain can be measured directly by inserting electrodes into visual areas such as cortical area V4. As a result of that information, for unknown reasons, you have a conscious experience of greenness. We could say that two items are relevant to the discussion: the computation that the apple is green (Box 1 in Figure 4), and the "experienceness" of the green (Box 2).

In a religious or spiritual view, the "experienceness," the consciousness itself, is a nonphysical substance, something like plasma. It is ectoplasm. It is spirit. In New Age thinking, it is energy or life force. In traditional Chinese medicine, it is *chi*. On introspection, people describe it as a feeling, a sense, an experience, a vividness, a private awareness hovering inside

the body. In the view of Descartes (1984 [1641]), it is *res cogitans* or "mental substance." In the view of the eighteenth-century physician Mesmer and the many practitioners who subscribed to his ideas for more than a century, it is a special force of nature called animal magnetism (Alvarado, 2009). In the view of Kant (1966 [1781]), it is fundamentally not understandable. In the view of Searle (2007), it is like the liquidity of water; it is a state of a thing, not a physical thing. According to many scientists, whatever it is, and however it is caused, it can only happen to information that has entered into a complex, bound unit (Baars, 1983; Crick & Koch, 1990; Engel & Singer, 2001; Tononi, 2008).

Arrow A in Figure 4 represents the process by which the brain generates consciousness. Arrow A is the central mystery to which scientists of consciousness have addressed themselves, with no definite answer or common agreement. It is exceedingly difficult to figure out how a physical machine could produce a nonphysical feeling. Our inability to conceive of a route from physical process to mental experience is the reason for the persistent tradition of pessimism in the study of consciousness. When Descartes (1984 [1641]) claims that *res extensa* (physical substance) can never be used to construct *res cogitans* (mental substance), when Kant (1966 [1781]) indicates that consciousness can never be understood by reason, when Creutzfeldt (1987) argues that science cannot give an explanation to consciousness, and when Chalmers (1995) euphemistically calls it the "hard problem," all of these pessimistic views derive from the sheer human inability to imagine how any Arrow A could possibly get from Box 1 to Box 2.

It is instructive, however, to focus instead on Arrow B, a process that is underemphasized both scientifically and philosophically. Arrow B represents the process by which consciousness can impact the information-processing systems of the brain, allowing people to report on the presence of consciousness. We suggest that much more can be learned about consciousness by considering Arrow B than by considering Arrow A. By asking what, specifically, consciousness can *do* in the world, what it can affect, what it can cause, one gains the leverage of objectivity.

Whatever consciousness is, it can ultimately affect speech, since we can talk about it. Here some clarification is useful. Much of the work on consciousness has focused on the information of which one becomes conscious. You can report that you are conscious of this or that. But you can also report on the consciousness itself. You can state confidently that you have an inner feeling, an essence, an awareness, that is attached to this or that item. The essence itself can be reported. If it can be reported, then one must have decided that it is present.

Consider the process of decision-making. Philosophers and scientists of consciousness are used to asking, "What is consciousness?" A more precisely formulated question, relevant to the data that are given to us, would be, "How is it that, on introspection, we consistently decide that we are conscious?" All studies of consciousness, whether philosophical pondering, casual introspection, or formal experiment, depend on a signal-detection and decision-making paradigm. A person answers the question, "Is awareness of X present inside me?" Note that the question is not, "Is X present," or, "Is information about X present," but rather, "Is awareness present?"

Much has been learned recently about the neuronal basis of decision-making, especially in the case of perceptual decisions about visual motion (for review, see Gold & Shadlen, 2007; Sugrue, Corrado, & Newsome, 2005). The decision about visual motion depends on two processes. First, perceptual machinery in the visual system constructs signals that represent motion in particular directions. Second, those signals are received elsewhere in the brain by decision integrators that determine which motion signal is consistent or strong enough to cross a threshold such that a response can be triggered.

Translating the same framework to a decision about awareness, one is led to the following hypothesis. Answering whether you contain awareness depends on at least two processes. First, neuronal machinery (in the present proposal, social perceptual machinery with an emphasis in the right STS and TPJ) generates explicit neuronal signals that represent awareness. Second, a decision process (perhaps in the MPFC) receives and integrates the signals to assess whether the property of awareness is present at a strength above the decision threshold.

The realization that conscious report depends on decision-making provides enormous leverage in understanding what, exactly, awareness is. Everything that is known about decision-making in the brain points to awareness as explicitly generated neuronal signals that are received by a decision integrator. A similar point has been made with respect to reportable sensory experience by Romo and colleagues (de Lafuente & Romo, 2005; Hernandez et al., 2010).

A crucial property of decision-making is that, not only is the decision itself a manipulation of data but also the decision machine depends on data as input. It does not take any other input. Feeding in some *res cogitans* does not work on this machine. Neither will

chi. You cannot feed it ectoplasm. You cannot feed it an intangible, ineffable, indescribable thing. You cannot feed it an emergent property that transcends information. You can only feed it information. In introspecting, in asking yourself whether you have conscious experience, and in making the decision that, yes, you have it, what you are deciding on, what you are assessing, the actual stuff your decision engine is collecting and weighing, is information. The experienceness itself, the consciousness, the subjective feeling, the *essence* of awareness that you decide you have, must be information that is generated somewhere in the brain, and that is transmitted to a decision integrator. By considering Arrow B and working backward, we arrive at a conclusion: Consciousness must be information, because only information can be grist for a decision and we can decide that we have consciousness.

One might pose a counter-argument. Suppose consciousness is an emergent property of the brain that is not itself information; but it can affect the processing in the brain, and therefore information about it can be neuronally encoded. We have cognitive access to that information, and therefore ultimately the ability to report on the presence of consciousness. What is wrong with this account? In the theory presented here, nothing is wrong with this account. Indeed, it exactly matches the present proposal except in its labeling. In the present theory, the brain does contain an emergent property that is not itself an informational representation. That property is attention. The brain also does contain an informational representation of that property. We can cognitively access that informational representation, thereby allowing us to decide that we have it and report that we have it. However, the item that we introspectively decide that we have and report that we have is not attention; strictly speaking, it is the informational representation of attention. In circumstances when the informational representation differs from the thing it represents, we necessarily report the properties of the representation, because it is that to which we have access. The mysterious and semi-magical properties that we report, that we ascribe to consciousness, must be attributed to the informational representation.

As an analogy, consider an ordinary book that contains, printed in it, the words, "This is not a book. It is really a fire-breathing dragon." Some readers might announce the presence of a metaphysical mystery. "It seems to be made of simple material in its outer manifestation," they might say, "but in its inner manifestation, it is something else. How can it be a book and a dragon at the same time? How can we resolve the dualistic mystery between *res libris* and *res dragonis*?" The resolution is simple. It *is* a book. The book contains information on itself. That information is inaccurate.

In the same way, humans have strong intuitions about consciousness, awareness, soul, mind. We look introspectively; that is to say, we consult the information present in the brain that describes its own nature, and we arrive at an intuitive understanding in which awareness is a nonphysical thing, something like energy or plasma, that has a location vaguely inside the body but no clear physical substance, that feels like something, that is vivid in some cases and less vivid in other cases, that is a private essence, that is experience, that seizes on information. What we are doing, when introspecting in this way, is reading the information in the book. That information is not literally, physically accurate. In the present hypothesis, it is a useful informational representation, a depiction of the real physical process of attention. Attention is something the brain does, not something it knows. It is procedural and not declarative. Awareness, in the present hypothesis, is declarative. It is an informational representation that depicts, usefully if not entirely accurately, the process of attention.

We end with one final analogy. Recall Magritte's famous painting of a pipe with the words scrawled beneath it, "Ceci n'est pas une pipe." This is not a pipe. It is a representation of a pipe, an existentially deep realization. A distinction exists between a representation and the thing being represented. In the present proposal, consciousness is a representation of attention. But the representation has taken on a life of its own. As we examine it cognitively, abstract properties from it, and verbalize those properties, we find that the representation depicts magic, spirit, soul, inner essence. Since we cannot imagine how those things can be produced physically, we are left with a conundrum. How does the brain produce consciousness? Yet the brain can easily construct information, and information can depict anything, even things that are physically impossible.

Commentaries

Mirror-touch synaesthesia: A case of faulty self-modelling and insula abnormality

Michael J. Banissy[1], Vincent Z. Walsh[1], and Neil G. Muggleton[1,2]

[1]UCL Institute of Cognitive Neuroscience, London, UK

[2]Institute of Cognitive Neuroscience, National Central University, Jhongli, Taiwan

E-mail: m.banissy@ucl.ac.uk

Abstract: This commentary describes an additional line of evidence related to the suggestion that mechanisms of social perception facilitate the development of a perceptual model of one's own mind. Drawing on findings from mirror-touch synaesthesia, we highlight the importance of the insula in distinguishing between the perceptions of one's own experience versus someone else's.

In their discussion of the relationship between human consciousness and mechanisms of social perception, Graziano and Kastner provide a number of interesting proposals about how mechanisms of social perception may act to facilitate an understanding of one's own awareness. Their main assertion that, through the allocation of attention we perceive someone else's awareness in the same way that we perceive our own, is thoughtfully considered and a number of intriguing lines of evidence are provided to support this argument. One aspect of their proposal that caught our interest was their prediction that the neural mechanisms for social perception should construct a self-model (Prediction 3). Here, we consider additional relevant work related to this proposition by considering evidence from mirror-touch synaesthesia and discuss the potential role of the insula in distinguishing between one's own and another's awareness.

Mirror-touch synaesthetes experience a conscious tactile sensation on their own body when simply observing touch to another persons' body (Blakemore et al., 2005). These experiences are thought to occur because of heightened neural activity in the same network of brain regions that are active when we observe touch to another person. This network includes the primary and secondary somatosensory cortices (SI; SII), premotor cortex, and the temporo-parietal junction. This mirror-touch system therefore shows an overlap with the same regions of the brain as when we are physically touched (i.e. SI; SII) and has been linked more widely to our social perception abilities (Banissy & Ward, 2007; Banissy et al., 2011; Keysers, Kaas, & Gazzola, 2010).

In the context of mechanisms of social perception constructing a self-model, mirror-touch synaesthesia provides a unique experimental preparation to consider how our perception of others is intrinsically linked to the self because these synaesthetes activate the same affective neural system that we all do when perceiving another's tactile experience, yet have an additional tactile experience on their own body. The condition may therefore be useful in providing insights into how neural mechanisms of social perception can be modulated to construct a self-model.

We have previously suggested that mirror-touch synaesthesia is related to a misrepresentation of another's body onto the synaesthete's own body schema (Banissy et al., 2009). Extending this to Graziano and Kastner's argument, one may suggest that the condition reflects an error in constructing a perceptual model of the source of another's tactile awareness. Moreover, consistent with the authors' account of out-of-body-experiences reflecting a spatial mislocalization of the source one's awareness, it is possible that in mirror-touch synaesthesia an error in the neural systems distinguishing between self and other leads to the source of another persons' tactile awareness being mislocated onto the synaesthetes' own body.

In this context, it is interesting to note that the only brain region shown to differ between mirror-touch synaesthesia and non-synaesthetes when observing touch to others is the anterior insula (Blakemore et al., 2005). This region has been related to self-awareness (Critichley et al., 2004) and processing one's awareness of others (Craig, 2004; Lamm & Singer, 2010). The amount of neural activity within the insula has also been shown to increase by directing attention to one's own emotions when viewing affective stimuli (Straube & Miltner, 2011). The role of the insular cortex is overlooked in the current article, but it is likely to be a particularly important component for constructing a self-model from mechanisms of social perception and at least a potential candidate region

involved in modulating distinctions between the sources of one's own or another's tactile awareness within the mirror-touch system.

* * *

Functions of consciousness

Bruce Bridgeman

Dept. of Psychology, University of California, Santa Cruz, CA, USA

E-mail: bruceb@ucsc.edu

Abstract: Evolutionary theory indicates that consciousness has a function, if it is complex enough to be supported by genetically guided brain structures. Otherwise there would be no selective pressure against degrading it. Hints about its function come from word priming studies, where conscious awareness of a prime allows it to be avoided according to instructions. Consciousness, then, allows behavior to be driven by internal plans rather than primed contingencies from the momentary environment. Since consciousness is normally tested by memory of events, its mechanism may be a form of working memory.

Graziano and Kastner begin their brilliant analysis of consciousness by asking whether it performs some function that is beneficial to survival, or is just an inevitable by-product of a large, complex brain. The by-product argument doesn't get far – it would predict that any large complex information processing device should be conscious, such as the internet or the network of telephone landlines. Clearly, size alone doesn't support consciousness.

Evolutionary theory indicates that consciousness does have a function. The argument requires only the assumption that consciousness be something complex, instantiated in a brain organization guided by many genes. It isn't something developmentally incidental, such as the navel.

Since consciousness is culturally universal, we can infer that it is supported by genetically influenced brain mechanisms. The corresponding genes must have been selected for during human evolution and are presently maintained by evolutionary selection. Most of the influence of natural selection is not to foster change, but to enforce retention of advantageous adaptations. Because brain tissue is energetically expensive the machinery of consciousness would be selected against

if it bestowed no benefit. An analogous process is more visible in organisms such as fish that become trapped in caves, living generation after generation in darkness. Released from the natural selection that normally maintains the structure of their eyes, these fish gradually evolve to become blind. The fact that we retain the capability for consciousness indicates that natural selection is maintaining it in us; it has a function. Evolutionary theory, though, does not tell us what that function is.

The idea that self-consciousness arose from mechanisms that evolved to assess the intentions of others is compelling and convincing. Once it appeared, there is evidence that the conscious mode of brain function facilitated other functions as well, functions that enhance the ability to plan sequences of actions and to execute those plans driven by an internal structure rather than by the vicissitudes of the sensory environment of the moment.

An empirical example reveals the advantages of such a mechanism. Priming can be used as a probe by giving a person a word stem, such as fi___, and asking the observer to generate a 5-letter word with that stem, such as 'final' or 'fists'. Now the observer is briefly exposed to target words, with each target word followed immediately by its stem (Merikle, Joordens & Stolz, 1995).

Instructed to complete the stem with a word other than the target word, observers sometimes correctly generated an alternative stem completion and at other times incorrectly blurted out the target word. If they were shown the target word for a relatively long time, such as a quarter of a second, they usually saw it and were able to complete the task successfully. After a shorter exposure, however, such as 50 msec, they often were unable to report the word, and were more likely to produce the target word than to choose another one. In fact, they picked the target word more often than would be expected at random. Picking that word indicates that information about it had entered the nervous system. Without a consciousness of that event, however, the observers were unable to prevent it from affecting their behavior as instructed.

One function of consciousness, then, is to prevent behavior from bending to the momentary availability of information in the environment. It allows behavior to be driven by one's own plans, escaping the tyranny of the environment. This and similar experiments lead to the conclusion that consciousness enables internal plans to overcome environmental nudges (Bridgeman, 2003, pp. 276–280). This line of evidence points to consciousness as important in cognitive processes. Though there is a convincing case for origins of

consciousness in the ability to infer the mental states of others, the cognitive advantages that consciousness affords extend beyond social- and self-perception.

A corollary of the social origin of consciousness is that other animals should also have it in some form. If an animal can predict what a conspecific will do based on inferring its moods and motivations, all of the cognitive machinery that Graziano and Kastner specify will function. A tragic counterexample for the relationship of social relationships to consciousnessis is severe autism, in which people with autism cannot use a theory of mind to infer the states of others; they seem conscious of events, but not of the self. In short, there is a mind but no theory of mind.

Consciousness as arising from perception and interpretation of social interactions can also be applied to speech, a uniquely conscious and uniqely human behavior. Hearing one's own speech, originally intended to communicate with others, allows a recursion where the speech stream re-enters the brain through auditory channels originally evolved to interpret the speech of others. The perceived self-speech can then make another pass through the brain, and thinking results. The process can be speeded with internal speech, using much of the same neurological machinery.

All of these processes also depend on memory, which is the normal test for conscious awareness. A teacher asking "Johnny, what did I just say?" is using a memory test for consciousness. A drunk might remember nothing of the night before. Was he conscious of those events? This leads to the idea that the mechanism of consciousness is a particular form of working memory, no more and no less.

* * *

Should damage to the machinery of social perception damage perception?

Peter Carruthers and Vincent Picciuto
University of Maryland, College Park, MD, USA
E-mail: pcarruth@umd.edu

Abstract: We argue that Graziano and Kastner are mistaken to claim that neglect favors their self-directed social perception account of consciousness. For the latter should not predict that neglect would result from damage to mechanisms of social perception. Neglect is

better explained in terms of damage to attentional mechanisms.

We are sympathetic to the main thesis defended by Graziano and Kastner (hereafter "G&K"). This is that phenomenal consciousness results when self-directed social perception binds a metarepresentational concept of awareness into the contents of globally broadcast, widely integrated, perceptual states. This view is quite close to that developed previously by some philosophers (Van Gulick, 2004; Kriegel, 2009; Picciuto, 2011). These are described as "self-representational" theories of consciousness because the resulting perceptual states present *themselves* to us (via the embedded meta-representational concepts) while also representing some aspect of the world. In addition to developing such an account within a neuroscientific framework, what G&K add is that the self-presenting perceptual state must also include a representation of the subject's own body and its location and orientation in space. This strikes us as a useful addition. However, they also provide neuroscientific arguments in support of their thesis. Some of these we find much less convincing. We will focus on neglect.

G&K point out that some patients with hemispatial neglect have suffered damage to right TJP and/or right STS, which are regions crucially implicated in the social cognition network. These patients seem unaware of stimuli in their left visual fields, as measured by failure to cross out lines in a line-crossing visual task. This finding is said to support the theory of conscious awareness under consideration. We disagree. The findings are more plausibly explained by damage to aspects of first-order (but top–down) attentional systems. These happen to be located close to brain areas crucially involved in social cognition (as G&K themselves acknowledge).

What should one predict would happen if the metarepresentational awareness-concept that is normally embedded into globally broadcast perceptual states was lost through damage to the social cognition system? One should *not* predict that the subject would become blind. For if attentional systems remain undamaged, the more-basic process of global broadcasting should continue as before, making the contents in question accessible to decision-making systems among others. So the subject should still (non-consciously) perceive stimuli and be capable of acting on them appropriately (such as crossing out all the lines on a page), at least when acting spontaneously and unreflectively. Nor should the result be similar to blindsight, with subjects declaring that nothing is present in the relevant portion of the visual field while nevertheless being capable of simple discriminations when

asked to guess. On the contrary, subjects should be capable of making first-order reports about the presence and nature of stimuli in the affected field (at least, provided the reports are generated swiftly and spontaneously—perhaps in a task requiring them just to answer "Yes" as fast as possible if a horizontal line appears somewhere in their visual field). For if first-order attentional systems are working normally, the relevant percepts will be globally broadcast, and should thus be accessible to inform a verbal report.

What G&K should predict is just that damage to the social cognition system would result in loss of the *subjective*, *phenomenal*, aspects of the relevant perceptual states. This ought to issue in a number of cognitive or behavioral deficits. In particular, subjects should be incapable of drawing the is–seems distinction with respect to stimuli in the affected portion of the visual field. While subjects should perceive the world, they should be unable to contrast what *is* present with what *seems* to them to be present, or with what they *experience as* being present. For the relevant self-representing *awareness*-representation will no longer be bound into the content of the perceptual states in question, and hence will not be accessible for report. Moreover, there should be a dissociation between swift spontaneous reporting of stimuli and what subjects are inclined to say on reflection, after noticing that they lack any appropriate conscious experience.

The only way in which blindness or neglect could result from damage to self-directed social cognition would be if the presence of a metarepresentational awareness-concept were the main determinant of the global broadcasting—or widespread integration—of the relevant perceptual representations. But this is quite implausible. First-order mechanisms of top–down attention are presumably much more basic than social cognition, and were determinants of global broadcasting long before the evolution of mechanisms for social cognition. They ought therefore to continue unabated when the latter is damaged.

In conclusion, while we applaud the theoretical framework that G&K provide, we think that they over-sell some of the evidence that is said to favor it.

* * *

Consciousness is for sharing

Chris D. Frith

Wellcome Trust Centre for Neuroimaging at University College London, UK, & Center of Functionally Integrative Neuroscience at Aarhus University, Denmark
E-mail: c.frith@ucl.ac.uk

Abstract: I welcome the idea that consciousness, rather than being private, has a crucial role in understanding other minds. However, I believe that the critical mechanism underlying consciousness is meta-cognition; the ability to reflect upon and report mental states, whether our own or those of others. This allows us to share experiences and create a more accurate picture of the world.

Graziano & Kastner put forward an ingenious proposal about the origin of consciousness. This will be counter-intuitive for some, since, rather than being a private place for the self, consciousness emerges from the neural mechanism that permits tracking the attentional focus of others. The same mechanism is also used to represent our own attention. G & K suggest that consciousness emerges from this ability to represent our own attention. In this scheme consciousness is the meta-cognitive representation of our own awareness.

I must immediately declare a strong bias in favor of these ideas. Like G & K I believe that, if we want to understand consciousness we need to ask what it is for and, like them, my answer is that its fundamental function is social (Frith, 2008). I also like very much their novel and innovative idea that the machinery for representing of the focus of attention is applied to others as well as to ourselves. A major clue to where others are focussing their attention comes from eye gaze direction and humans are very accurate at judging where people are looking. The human iris, uniquely among primates (Kobayashi & Kohshima, 1997), is surrounded by white sclera making it easier to discern eye gaze direction. Thus, the evolution of this aspect of the human eye was not directly for the benefit of the individual, but for the benefit of others. This is confirmatory evidence for G & K's suggestions about the social function of representations of the focus of attention.

However, I am unconvinced about the link with consciousness. People compute the direction of attention of others, where they are looking and what they see, automatically (Bayliss & Tipper, 2006; Qureshi, Apperly, & Samson, 2010). I conclude that we can take account of the attention of others without awareness. Also I don't see why G & K's account necessitates a subjective component. What does the feeling of being aware add to our ability to compute what people are attending to? The problem is to explain the function of this subjective component.

I believe we have all been misled by the long-standing idea that consciousness is private. On the contrary, conscious experience is the one outcome of our brain's information processing that can be shared with others. It is this sharing that provides the value of consciousness. By sharing their experiences, two people can gain a more accurate perception than the best person working on her own (Bahrami et al., 2010).

On this account the key component of consciousness is meta-cognition. I agree with G & K that a representation of our awareness, creating a feeling of being aware, is fundamentally important. But this meta-cognitive process can be applied to many things in addition to attention, including, for example, the feeling of agency. What we need to uncover is the neural mechanisms underlying meta-cognition. There is currently increasing interest in meta-cognition with developments in its empirical study (Lau, 2008) and quantification (Galvin, Podd, Drga, & Whitmore, 2003). Although, as yet, we know little of relevant neural mechanisms, the prefrontal cortex is strongly implicated (Del Cul, Dehaene, Reyes, Bravo, & Slachevsky, 2009; Fleming, Weil, Nagy, Dolan, & Rees, 2010).

But this critique of G & K concerns just minor details. I am very pleased to join them in promoting the idea that consciousness is not a private place. It has a major role in enabling the richness of human of social interactions.

* * *

Hybrid social cognitive models, meta-consciousness, and representations

Marco Iacoboni

Ahmanson-Lovelace Brain Mapping Center, Dept. of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, Brain Research Institute, David Geffen School of Medicine at UCLA, CA, USA
E-mail: iacoboni@ucla.edu

Abstract: Hybrid social cognitive models, based not only on simulation, are desirable, especially in complex scenarios. It is unclear, however, which neural systems support non simulative processes. Furthermore, the proposal that consciousness is a form of representation of attention, while novel, relies on a traditional concept of representation. Recent findings

suggest an ontological priority of representations, where the representation and the thing being represented are not distinct at all.

Graziano and Kastner propose a novel model of consciousness, according to which consciousness (or at least one prominent aspect of it, awareness) is a declarative process about something else. Indeed, according to their model, consciousness is a representation of an eminently procedural phenomenon, the deployment of attention.

This is a rather interesting hypothesis, and definitely a novel one. I can't recall of any existing model of consciousness in the literature that resembles this one. This is already a merit, in a quite crowded literature.

My comments will focus on three aspects of the paper. First, the idea that social cognitive models can be hybrid or mixed, based on both highly specialized systems for social perception and simulative processes supported by forms of neural mirroring. Second, the idea that our reports, verbal or otherwise, are faithful measures of what is being reported. Third, the very notion of representation, the way it has been classically construed and still is prevalently conceived, and possible alternative ways of thinking about it based on recent neurophysiological observations.

It makes sense that flexible social cognition is based on more than one functional process, and that these processes interact in some way, as Graziano and Kastner propose. An exclusively simulative mechanism seems unlikely, especially when complex scenarios require more than simulation, including cognitive control, inferential reasoning, etcetera. What is unclear is the role of some of the structures used by Graziano and Kastner in Figure 3. While existing data on STS make it clear that this functional region is anatomically well defined and functionally linked to perception of biological motion, both TPJ and MPFC have been associated with a wide variety of processes and tasks.

Furthermore, they are anatomically less well defined. With regard to MPFC, many imaging studies have also shown a peculiar pattern of activity with less de-activation for the 'experimental' task and more de-activation for the control task. This pattern boils down to less deviation from baseline for the experimental task. In physiology, what generally is considered the most meaningful response is the response that deviates from baseline MORE, not LESS. This interpretive anomaly for the MPFC pattern has been justified by the high level activity of this region at rest. In physiology, however, baseline activity does not typically affect the interpretation of the data in this fashion. A neuron with high baseline firing rate is still considered to respond more to, say, stimulus A rather than B, when its firing rate is suppressed MORE for stimulus A than for B. It is unclear

why the imaging literature is reversing a time honored tradition of interpreting physiological responses in cases such as the MPFC. At some point the neuroscience community will have to face this puzzling anomaly.

The underlying, although left mostly unsaid, assumption of the manuscript is that subjective reports are fundamentally correct. Plenty of evidence suggests otherwise. Humans are not too good at monitoring their own cognition (Schooler, 2002). While I don't see this as a major problem for the model proposed by Graziano and Kastner, it is something to be considered when specific predictions of the model will be tested.

Finally, a comment on the issue of representations. This is obviously a big issue in cognitive science, and dates back to Plato. Indeed, since Plato, the dominant view of representation is the same advocated by Graziano and Kastner. "A distinction exists betweena representation and the thing being represented." The word itself, re-presentation, embodies this whole concept of representations being both ontologically and temporally secondary to the things they represent. In the Graziano and Kastner model, consciousness is a representation of attention. The very discovery of mirror neurons, however, suggests that perhaps representations themselves may not work the way we think they work. Mirror neurons suggest an almost unsettling ontological priority of representations, where the representation and the thing being represented are not distinct at all (Iacoboni, in press). This is a concept that is difficult to face, because our idea of representation is so entrenched in our minds, and is such a foundational aspect of how we think, that I bet very few are willing to even consider the possibility of rethinking it. And yet, rethinking it may make us better equipped at dealing with the theoretical implications of recent scientific advances (Thiagarajan, Lebedev, Nicolelis, & Plenz, 2010).

* * *

Autism and perception of awareness in self and others: Two sides of the same coin or dissociated abilities?

Rogier A. Kievit and Hilde M. Geurts
Faculty of Psychology, University of Amsterdam, The Netherlands
E-mail: r.a.kievit@uva.nl

Abstract: Graziano and Kastner propose a theoretical framework suggesting that the same cognitive machinery underlies computation and inferences concerning (the content of) awareness in others as underlies the perception of the contents of our own awareness. We draw from this hypothesis a strong prediction: Individuals who have deficiencies in one of these abilities must also be impaired in the other. We discuss evidence supporting this prediction from the literature on autism spectrum disorder, but also discuss tentative evidence for a possible dissociation between these two abilities. We conclude that these lines of evidence form crucial empirical tests of the theory.

In their target article, Graziano and Kastner suggest that the same machinery underlies computation and inferences concerning (the content of) awareness in others as underlies the perception of the contents of our own awareness. We take their hypothesis and deduce two empirical predictions: 1) individuals with deficiencies in inferring the content of their own mental states should have deficiencies in perceiving awareness of others and 2) individuals who have known deficiencies in inferring the content of awareness in others are probably also impaired in inferences concerning their own mental states. It therefore makes sense to examine cases where this ability goes awry: People with autism spectrum disorder. Difficulties in predicting the behavior and (emotional) content of others are thought to be a central aspect of the social difficulties seen in people with ASD. If consciousness is indeed a construct of social perceptual machinery, such cases offer a critical test of the theory.

Research on ASD has shown well-documented impairments in inferring mental states of others, in adults and children (Frith, 2004), known in the literature as the Theory-of-Mind hypothesis. If the hypothesis in the target article holds, people with ASD should also have difficulties with self-awareness. Several lines of evidence suggest that this is indeed the case.

First, severe degrees of alexithymia (difficulties in experiencing one's own emotions) have been found to affect 50% of populations of individuals with ASD, compared to 10% in the general population (Hill, Berthoz, & Frith, 2004). Second, Ben Shalom et al. (2006) showed that children with ASD displayed normal physiological responses when experiencing emotions, but *differed* from controls in the conscious report of these emotions, suggesting a deficit in the conscious awareness of their emotions. Finally, Silani et al. (2008) used fMRI to show that self-reported poor awareness of own and others' feelings was associated with decreased activity in the insula. They concluded: "We hypothesize that the same neural architecture underlies the conscious representation of emotion in

the self and in others. We suggest that the insula subserves this function" (p. 106). These findings are in line with the hypothesis that deficiencies in inferring the mental state of others co-occurs with difficulties in self-related perception.

The implications of these findings go beyond support for the hypothesis of the target article: If accuracy of self-perception is impaired in individuals with, for instance, ASD, this means clinicians should be careful in basing clinical decisions on the content of patients' self-report, a conclusion supported by recent findings (Mazefsky, Kao, & Oswald, 2011).

However, that these two abilities are correlated does not, in itself, prove that they are subserved by the same machinery. Evidence for a dissociation would go directly against the claims in the target article, and require either an alternate explanation or a revision of the core hypothesis.

For instance, Silani et al. (2008, p. 37) caution against overgeneralizing their findings: "particular difficulties in emotional awareness in individuals with HFA/AS are not related to their impairments in self-reflection/mentalizing". Furthermore, Magnée, de Gelder, Van Engeland, and Kemner (2007, p. 1122) showed that people with ASD had 'enhanced sensitivity to facial cues at the level of reflex-like emotional responses'. Finally, Bird et al. (2008, p. 1515) show that the correlation between ASD and deficiencies in affect sharing is modulated by alexithymia: "Importantly, there was no difference in the degree of empathy between autistic and control groups after accounting for alexithymia."

To summarize, the hypothesis in the target article predicts co-occurring deficiencies in perception of self and others. Literature on ASD tentatively supports this prediction. However, evidence is equivocal, and the precise nature of deficiencies should be further studied, as it provides a critical test of the theory of social perceptual machinery.

* * *

Not all information is created equal

Christof Koch

Division of Biology, California Institute of Technology, Pasadena, CA, USA
E-mail: koch.christof@gmail.com

Abstract: Granziano & Kastner's account of consciousness as a process based on knowledge of what others are attending to does explain conscious access to spatial location. However, it fails to account for consciousness of the many object attributes that an external observer cannot infer from the subject's gaze.

I completely agree with the viewpoint that consciousness is information. However, it just can not be any type of information in the Shannon sense (Shannon & Weaver, 1949). For example, my liver enzymes represent a vast amount of information about metabolites and their end products circulating in the blood stream; yet I do not have conscious access to the state of my liver. Nor am I conscious of the neuronal representations in my retina; and a plethora of experiments have demonstrated that subjects can be unaware of millions of neurons in primary visual cortex firing vigorously to some perceptually suppressed stimulus. Crick and Koch (2001) termed the many rapid sensory-motor behaviors that function in the absence of consciousness (e.g. typing on a keyboard; Logan & Crump, 2009) zombie systems. Where is the informational difference between zombie and conscious systems?

Rather, one has to distinguish between two types of information. This is why the *Integrated Information* account of Tononi is such an attractive candidate for a theory of consciousness (Tononi, 2008; Balduzzi & Tononi, 2008). The same criticism applies when the authors frame the problem of consciousness as computing the answer to the decision of whether or not I have an inner experience. Why can I answer in the affirmative when it comes to the question "what is the color of that surface patch I'm looking at?" but am not able to answer the question "what is the pH of the content of my stomach?" even though the enteric nervous system, with about 100 million neurons, monitors the acidity of my digestive tract?

I do agree with Graziano and Kastner's viewpoint that attention and consciousness/awareness are distinct processes. Indeed, I have vehemently argued for such a two-way distinction (Koch & Tsuchiya, 2007). It would be interesting to see whether their theory can explain the remarkable experimental finding that attention and consciousness can sometimes have opposing effects (van Boxtel, Tsuchiya, & Koch, 2010).

The authors rightly stress the importance of Abel knowing where Bill attends to. However, this demand could be satisfied by a sort of saliency map (Itti & Koch, 2001) that indicates that - at the moment - Bill looks over there. However, what is missing from this account is that Abel is not just aware where Bill is attending to (i.e. spatial location) but that Abel is aware of many of the attributes of objects at that location. Consciousness

is replete with detailed content, most of which has nothing to do with spatial location. Abel can be aware of the surface texture, orientation and color of the cup, not just of the fact that Bill is attending to the cup of *there*. Interestingly, it is this ability to be conscious of the content of one's focus that remains in severe forms of bilateral neglect. Patients suffering from Balint's syndrome (Rafal, 1997) or simultanagnosia (Jackson et al., 2009) have lost all ability to locate items within the visual scene but they are still conscious of whatever is at the center of their gaze.

For some conscious modalities, the need to pinpoint the spatial focus of attention of an outside observer does not arise. A strong odor has no particular localization (if one prevents the subject from moving about), yet one can be highly conscious of the smell. Diffuse visceral pain can often only be spatially pinpointed following a detailed medical examination. Conversely, a tooth ache can be quite precisely located to within a single tooth, yet without an external observer having access to information that would allow her to make this sort of accurate spatial inference. Yet all these states can be consciously experienced, sometimes painfully so.

So while the social aspects of eye movements may be able to explain the spatial location of attention and awareness, such an account fails for a vast swath of conscious content.

* * *

What is it like to be a human?

David A. Leopold

Unit on Cognitive Neurophysiology and Imaging, National Institute Mental Health, National Institutes of Health, Bethesda, MD, USA

E-mail: leopoldd@mail.nih.gov

Abstract: The explicit link between awareness and sociality put forward in the accompanying article opens new doors to thinking about the evolutionary origins of consciousness. Human subjective experience undoubtedly has some features that are species-specific and others that are shared over a broad phylogenetic base. The authors' proposal that consciousness depends on high-level neural circuits evolved for social perception begs the question whether animals lacking such circuitry experience a fundamentally different form of consciousness from humans. It also highlights the need for comparative

work elucidating neural mechanisms by which animals other than primates perceive and respond to their conspecifics.

The neural basis of consciousness is arguably the most mysterious scientific puzzle facing us. Human subjective experience is a product of the brain, yet cannot be neatly pinned down to a particular neural system or circuit. As far as we can tell, other animals partake consciously in the world as we do, albeit sometimes drawing upon very different perceptual adaptations (Nagel, 1974). This raises the question of whether consciousness, and the neural systems that promote it, can be studied in a comparative and evolutionary context.

The accompanying article focuses almost exclusively on humans, drawing upon data from nonhuman primates when necessary. Yet by linking consciousness to social perception, the proposed framework raises the possibility that the study of consciousness via comparative social neuroscience can provide valuable new insights. The article argues that high-level social perception, turned inwardly, gives rise to conscious awareness. In other words, human self-awareness draws upon the same neural computations used to assess the mental states of others. While this explanation, like most explanations of consciousness, retains a mystical tinge, it does unambiguously link awareness to neural mechanisms of social perception. What is meant by awareness? The authors are clear on this point: it pertains not only to metacognition or access consciousness, but also to phenomenal consciousness – as they put it, "awareness of anything, awareness of a color, or a sound, or a smell."

The notion that social neuroscience can advance our understanding of consciousness is an optimistic and welcome view. Let us briefly consider what is known about social perception in animals. First, major taxa such as mammals and birds are composed of species that range from being solitary to highly gregarious. Second, social complexity is a critical factor in shaping species intelligence and encephalization (Humphrey, 1976; Shultz & Dunbar, 2010). Third, behavioral and ethological testing has shown that socially oriented species across a range of taxa are capable of monitoring and interpreting the behaviors of their conspecifics (Byrne & Bates, 2010; Leopold & Rhodes, 2010). Thus there is good reason to believe that high-level social perception is common in the animal world, that its level of expression varies widely among species, and that it has co-evolved with measurable changes in the brain.

At the physiological level, we know almost nothing about neural structures in other animals that may be homologous to those implicated in human social cognition such as the superior temporal sulcus (STS) and right

temporoparietal junction (TPJ). This is in some ways surprising, since comparative research has gleaned considerable information about other, related aspects of brain function. For example, electrophysiological studies have revealed detailed, multimodal sensory maps of the cerebral cortex in diverse mammals (Kaas, 2002). Comparative work has also identified molecules and primitive neural circuits that contribute to affiliative behaviors across a range of vertebrates (Insel, 2010). Nonetheless, the question of whether high-level social circuits evolved recently to fit the specific needs of primate, or possibly human, society, or whether they shared features of the mammalian brain, is presently a matter of speculation.

It is worth pointing out that there is one other mammalian species whose cortical mechanisms of social perception have been studied: sheep. Sheep are social animals that readily recognize faces and interpret facial expressions of their conspecifics, even when the faces presented on a computer monitor. Single-unit electrophysiological recordings have revealed that neurons in the sheep temporal cortex exhibit selectivity for faces that bears a striking similarity to that observed in monkeys and humans (Tate, Fischer, Leigh, & Kendrick, 2006). While this observation does not imply that sheep possess a human-like potential for mind reading, it does at least raise the question whether the sheep homologs of the human STS and TPJ might be involved in some form of social monitoring.

Could advances in social neuroscience ultimately bear on our understanding of the evolution of consciousness? Possibly so – if consciousness and social perception are indeed fundamentally interlinked. The fact that we can presently only speculate on the cortical systems supporting high-level social perception in animals calls attention to the fact that more work needs to be directed toward comparative social neuroscience. As we learn more about the types of neural specializations in different animals, we may slowly gain insight into the evolution of our own social brain, and, if the hypothesis in the accompanying article is correct, into the neural circuits that lie at the heart of our conscious experience.

* * *

Stimulus independence, social cognition and consciousness

Jonathan Smallwood

Department of Social Neuroscience, The Max Planck Institute for Human and Cognitive Brain Sciences, Leipzig, Germany
E-mail: jonsmallwood2004@yahoo.com

Abstract: A consensus emerging from neuroscience is that certain brain regions show activity correlated with stimulus independent (e.g. private) conscious thought and yet are also implicated in public social processes. The fact that systems supporting social processing also exhibit activity with no obvious perceptual referent, can be seen as support for the framework suggested by Graziano and Kerber (this volume) once it is recognized that the property of stimulus independence is also an important feature of consciousness. Understanding the social basis behind private stimulus independent thought, therefore, may provide an important assessment of the validity of the Graziano and Kerber hypothesis.

The proposal of Graziano and Kastner (this issue) argues that our phenomenal awareness is a product of our capacity to understand and represent the mental states of another individual; a process known as social cognition (SC). As consciousness pervades our private waking life, a role of SC in consciousness would imply that in a fundamental way: I can't get you out of my head. While the social basis of consciousness is not new (e.g. Dennet, 1992), the current hypothesis goes substantially beyond previous suggestions by implicating SC in phenomenal awareness (a critical element of what is known as the hard problem of conscious experience, Chalmers, 1996). In particular, because social processes can drive natural selection (e.g. Tomasello, 1999), Graziano and Kastner provide a hypothesis for how phenomenal awareness could arise through an evolutionary process.

On the face of it the role of SC in conscious thought seems to differ from accounts of consciousness that emphasize processes that are chiefly internal in nature. For example, in 1995 Francis Crick suggested an experimental definition of consciousness that emphasized that the neural process associated with conscious experience could be differentiated from those that represent code the raw features of sensory events. This astonishing hypothesis made the problem of consciousness experimentally tractable because it made a testable prediction – those experiences associated consciousness would be neural processes that are relatively stimulus independent (SI) in nature.

Soon after Crick and colleagues published their thesis on consciousness, interest in SI neural activity increased due to a seemingly unrelated observation that a coordinated system of brain regions continue to exhibit activity when participants are not constrained by an external task.

Whether this activity constituted a coherent system was once a matter of debate, however, now the so-called default mode network (DMN, Raichle, 2001) and the more general 'resting state' is a widely accepted element of the neuro-science paradigm. Because activity in the DMN persists in situations when external task processes are negligible, this system was initially defined by its capacity to exhibit *SI neural activity*. In the last few years studies have gone on to implicate the DMN directly in the production of the stream of internally generated thoughts we experience, for example, when we daydream (e.g. Mason et al., 2007).

One surprising feature of the DMN is that many of the structures involved in SI thought are also related to SC (e.g. mPFC, Mitchell, 2009). Elements of the DMN system are, for example, active in explicit tasks requiring predicting the decisions of another individual (such as the prisoners dilemma, e.g. Rilling et al., 2008), implicated in empathic processes (e.g. Lamm, Decety & Singer, 2011) and plays a role in estimating another individuals' mind (e.g. Gallagher & Frith, 2003; Spreng & Grady, 2010).

That the tell-tale signs of the DMN are private SI mentation and public SC seems, on the face of it, a contradiction. Such an association between the internal world of the mind and the external world of social interaction is, however, precisely what would be expected if consciousness was a social phenomenon (Frith, 2008). As such the SI activity performed by the DMN provides an important conceptual link between the proposal of Graziano and Kastner and the arguments made Crick; however, as the precise processes that DMN plays remain mysterious, alternative explanations are a definite possibility. First, theories of conscious experience known as global workspace models (e.g. Dehaene, et al., 2006) argue that information only becomes conscious when it gains access to a general purpose system that allows information to become globally available to the system. Based on this logic, it is important to identify whether global availability is necessary for the DMN network to lead to conscious SI experience (Smallwood, 2010).

Second, although Graziano and Kestner argue that consciousness is produced when I "put yourself in my place", other theorists argue that the directionality is reversed (e.g. "knowing me, knowing you", Mitchell, 2009). As the Graziano and Kebsters proposal is specifically uni-directional, experimental paradigms that explore whether representations of self or others are the prime movers in the process of simulation are needed to assess this specific element of the hypothesis in question.

Third, while the DMN is an important contribution to the neural process in operation during the resting state, recent studies have documented many other brain networks show SI activity when participants are given no external task to perform (e.g. Smith et al., 2010). It is unlikely, that all of these systems are directly concerned with conscious thought (e.g. Raichle, 2010) and so understanding the computations that these other networks perform at rest is, therefore, necessary before the links between SI, SC and consciousness are completely understood.

If the capacity for SI is linked to the processes of SC, then understanding DMN activity could illuminate the process of consciousness. Although understanding consciousness simply in terms of spontaneous brain activity is a dangerous game, fMRI and other techniques allow the assessment of otherwise unobservable SI mental activity. If Graziano and Kaster are correct, rather than being light years away, an understanding of the enigma of consciousness could well be closer than either you or I think.

ACKNOWLEDGMENTS

* * *

Bodily self-consciousness, and the primacy of self related signals such as the 1st person perspective and self-location

Michiel van Elk and Olaf Blanke
Laboratory of Cognitive Neuroscience, Ecole Polytechnique Fédérale de Lausanne,, Lausanne, Switzerland
E-mail: Olaf.blanke@epfl.ch

Abstract: G&K provide a new and interesting perspective on consciousness, by arguing that consciousness is a product of social perception. Based on the overlap between the neural mechanisms underlying spatial awareness of other people and oneself, out-of-body experiences (OBEs), and social perception the authors argue that consciousness is based on the brain signals that represent other people and their spatial awareness. Although we generally

welcome the authors efforts, we (1) would like to emphasize that consciousness has two distinct spatial aspects, namely self-location and first-person perspective, (2) cite evidence about distinct developmental and brain mechanisms concerning first- versus third-person perceptions and cogitations and (3) argue for a primacy of multisensory own body perception over social perception and awareness as a neurobiological foundation of consciousness.

Citing experimental data in patients (Blanke, Ortigue, Landis, & Seeck, 2002) and healthy subjects (Lenggenhager, Tadi, Metzinger, & Blanke, 2007) G&K highlight the spatial character of consciousness and link this to OBEs. Consciousness, however, has two distinct spatial aspects: self-location (SL) *and* first-person perspective (1PP; Blanke & Metzinger, 2009). SL refers to the location where subjects experience themselves to be ("Where am I in space?") and 1PP refers to the perspective and directedness of consciousness ("From where do I perceive the world?"). SL and 1PP are open to experimentation and can be dissociated in neurological disease, yet G&K only discuss SL. Thus, whereas most OBEs are characterized by a spatially congruent SL and 1PP, this is not always the case and in some cases SL and 1PP can be experienced at separate spatial positions (De Ridder, Van Laere, Dupont, Menovsky, & Van de Heyning 2007; see also Blanke et al., *Science E-letter*, 2008). Moreover, it was recently shown that the experienced direction of 1PP and SL can be manipulated independently (Ionta et al., 2011). G&K's proposal does discuss the importance of the spatial character of consciousness, but we argue that their consciousness proposal also needs to integrate distinct mechanisms for 1PP and SL.

We further note that several lines of evidence have shown that during development the ability to take a 1PP precedes the ability to take a 3PP - defined by some as having a theory of mind (Wimmer & Perner, 1983). Does a similar lPP-over-3PP primacy also exist for perception and consciousness? For example, awareness of the bodily self and associated aspects of 1PP develops early within the first few months of life (Rochat, 1998), whereas theory of mind abilities only gradually emerge from the 4th year of life onwards. In addition, neuroimaging studies in both children and adults suggest that taking a 1PP versus a 3PP recruits different neural networks (Dosch, Loenneker, Bucher, Martin, & Klaver, 2010; Kockler et al., 2010). Together, these findings argue against the view that self-awareness and consciousness build upon awareness of other people, that the neural mechanisms supporting self- and other-awareness can be equated, and against a primacy of social perception.

A number of studies revealed that consciousness is based on the integration of visual, tactile, and vestibular mechanisms. We argue that this multi-sensory integration of bodily signals is more fundamental and develops earlier in the human brain than brain representations of consciousness of others (Rochat, 1998). Independent of attention or awareness we have proposed that bodily information accounts for what G&K refer to as "the inner essence, the feeling of consciousness, that seems to be attached to the information" and that we have called minimal phenomenal selfhood (Blanke & Metzinger, 2009). When applied to one of the examples given by G&K – the feeling of being stared at (or the feeling of a presence) – we predict that such experiences are not "built on lower level sensory cues such as subtle shadows or sounds" and thus social cues, but rather on multisensory signals originating from the subject's own body and processed in the TPJ/STS region. The controlled induction of the latter illusion in neurology did indeed reveal that that different positions and states of the subject's own body lead to differently experienced positions and states of the staring other's body (Arzy, Seeck, Ortigue, Spinelli, & Blanke, 2006). These considerations and empirical human data argue for a primacy of self-related compared over other-related signals with respect to awareness and consciousness.

* * *

Reply to Commentaries

Awareness as a perceptual model of attention

Michael S. A. Graziano and Sabine Kastner

Department of Psychology, Princeton University, Princeton, NJ, USA

We proposed a theory of consciousness in which the machinery for social perception constructs awareness, and awareness is a perceptual model of the process of attention. One can attribute awareness to others or to oneself. Awareness of X is the brain's perceptual metaphor for the deep attentive processing of X. A set of ten comments on our hypothesis are included in this issue. Each comment raises specific points some of which directly challenge the hypothesis. Here we respond to these specific points and challenges.

We thank all the authors who contributed commentaries. The many thoughts and criticisms are well informed, insightful, useful to us, and undoubtedly interesting to readers. Here we address each of the commentaries in brief.

Leopold asks an age-old question: what species are conscious? Few species have a full-blown, human-like ability for theory of mind. Yet in our hypothesis, awareness does not depend on social cognition in general, but instead on one specific aspect of it, the perceptual reconstruction of attention. In our hypothesis, any animal that can construct a rich model of another's attentional state knows what awareness is; and any animal that maintains a perceptual model of its own attentional state is aware. Whether a particular species has these abilities is an empirical question, but the bar is lower than for advanced social cognition such as solving the false belief task. Leopold is correct: our hypothesis suggests that consciousness is widespread in the animal kingdom.

Koch is a proponent of the view that consciousness is related to complex, bound information. He agrees with some of the points raised in our article. The commentary, however, mainly challenges our theory. We suggest that Koch criticizes the theory for failings that it does not have.

First, at the start of the commentary, Koch states, "I completely agree with the viewpoint that consciousness is information. However, it just can't be any information in the Shannon sense." He notes that the liver contains information but isn't conscious. Our article does not suggest that consciousness is any

information such as that found in the liver. Rather we suggest that it is a specific type of information computed in a specific brain system. The first half of the commentary therefore elaborates on an irrelevant criticism.

Second, Koch interprets our theory in an overly specific manner. He suggests that the theory deals only in spatial location — in perceiving the spatial source and spatial target of attention. He states, "what is completely missing from this account is that Abel is not just aware where Bill is attending to (i.e. spatial location) but he is aware of many of the attributes of objects at that location." Yet this criticism does not pertain to our article or our theory. As we point out, "In the present hypothesis, awareness is the perceptual reconstruction of attention, and therefore anything that can be the subject of attention can also be the subject of awareness." One can attend to color, to shape, to smell, to a thought, to a great diversity of items. Spatial location is a relevant part of the issue; it is one of the many aspects of attention that must be captured in a perceptual model of the process of attending to something; but it is not the entire story.

Carruthers and Picciuto are a proponent of the link between consciousness and social cognition. He takes a positive view of our approach. He does raise a specific point about hemispatial neglect with which we disagree.

Carruthers and Picciuto (C&P) argue that neglect is not a true deficit of awareness. In neglect, the patient fails to notice, orient to, or verbally report items in the affected region of space. Some unconscious processing such as priming survives. According to C&P, however,

if a person has a pure deficit in awareness, that person should be able to react to a stimulus, point to a stimulus, and talk about a stimulus, and should lack only the inner awareness of the stimulus.

To counter this argument, we begin with a hypothetical case that we do not support. Suppose that awareness is an epiphenomenon. It is an inner feeling that serves no function and has no outward impact on behavior. In that case, C&P would be right. A loss of awareness should produce a zombie that acts normally but has no inner experience. C&P's view of neglect, we suggest, comes from inadvertently thinking of awareness as mainly an epiphenomenon.

The difficulty with the epiphenomenon view is that it is logically impossible. We can report that we are conscious. Indeed everyone participating in this print discussion has acknowledged the presence of consciousness. Therefore whatever it is, it has an effect on human behavior.

In our theory, consciousness is a perceptual model. Like all perceptual models, it serves the purpose of guiding behavior. Awareness may help to guide attention, cognitive analysis, and behavioral choice. It certainly has an impact on verbal report. Yet even beyond the normal influence on behavior that any perceptual model may have, awareness has a property that makes it unique. In our theory, awareness is a representation of the enhancer of representations. A natural resonance, a positive feedback loop, must exist between awareness (a representation of attention) and attention (the enhancer of representations). Awareness and attention are situated something like two mirrors facing each other. Because of this resonance, awareness must profoundly influence attention and therefore signal processing and behavioral response.

We argue that without awareness, behavior would be crippled, and that at least some forms of neglect do indeed match the deficit one expects from damaging the mechanism of awareness. Hemispatial neglect is, classically, a loss of awareness of anything in the affected half of space, as well as a loss of awareness that there is such a thing as that half of space.

Frith was one of the first to propose a link between consciousness and social intelligence. We are therefore delighted that he takes a positive view of our approach. He also raises several specific concerns about our theory, and here we address a particularly important one.

Frith notes that when a person processes someone else's attentional state, that processing can sometimes occur automatically, without awareness. How can an unconscious process be the source of awareness?

We believe this criticism stems from lumping together two items that in our theory are dissociable.

There is a distinction between a perceptual model of someone else's attentional state (assigning awareness to someone else) and a perceptual model of one's own attentional state (assigning awareness to oneself). The two may depend on similar neuronal machinery, but they are two different perceptual models. The fact that we can process someone else's attentional state without ourselves being aware of it does not strain the theory. In our hypothesis, to process someone else's attentional state, and therefore to assign awareness to someone else, and at the same time to be aware that we are doing it, requires an extra layer. It requires a perceptual model of how one's own attentional state is focused on someone else's attentional state. The theory predicts that much of the perceptual processing of other people's attentional state, and thus much of the attribution of awareness to other people, occurs outside of one's own awareness. The underlying logic is straightforward, even if its application to specific circumstances sometimes requires recursive complexities.

Bridgeman emphasizes the behavioral impact of awareness. He also emphasizes that, because it has behavioral consequences, it is subject to evolutionary pressure. This emphasis is, in our view, exactly right. By searching for a neuronal basis for consciousness one is necessarily accepting a scientific, biological framework. In that framework, traits evolve and are retained because they have some specific impact on an organism's survival. Our hypothesis provides a possible behavioral utility to consciousness, thereby providing at least a theoretical account of its evolutionary path.

Van Elk and Blanke discuss the issue of first person perspective and the out of body experience. We agree with most of their points and find their ideas on different types of first person perspective to be of great interest. However, we do not agree that a body schema, by itself, is a form of consciousness. The brain clearly computes a body schema, which contains information on the spatial instantiation of the self. The body schema is complex and probably multifaceted. But an informational model of the body, constructed in the brain, cannot by itself account for awareness. That is where our theory becomes useful. In our theory, awareness is information. It is an informational model, a schema computed by the brain. But it is specifically a schema that describes what it means to attentively process information. It is the brain's metaphor for the deep processing of information. Without that metaphor added to the mix, the body schema by itself would be merely a simulation of the body without awareness.

Kievit and Geurts offer an insightful comment on autism. To put this comment in a broader perspective,

here we point out a challenge sometimes directed at social theories of consciousness. Autistic people have impaired social perception. If consciousness is linked to social perception, then shouldn't autistic people show some deficit in consciousness? Kievit and Geurts answer the question: they do. Autism may indeed be associated with some loss of awareness of at least some mental processes. Consciousness normally encompasses only a small fraction of the processes in the brain; continuing to study how autistic and non-autistic people differ in this respect may be of great interest.

Banissy, Walsh, and Muggleton discuss the phenomenon of mirror-touch synesthesia. A mirror-touch synesthete will actually feel a touch on his or her own body when merely watching someone else being touched. This phenomenon suggests that the ability to build a perceptual model of the experiences of others is somehow related to one's own perceptual awareness. The phenomenon of mirror-touch synesthesia does not necessarily support the specific theory of awareness that we propose in our target article, but it does add support to the general approach. The authors also correctly point out that social perception is likely to involve many brain structures in addition to the few emphasized in our article.

Smallwood discusses an important area of the literature left out of our target article: the default mode network and its possible role in stimulus independent thought. The commentary is informative and compelling. Smallwood suggests that research on the default mode network supports our proposal of an overlap between regions of the brain involved in modeling others, in modeling the self, and in awareness.

Iacoboni comments mainly on the issue of mirror neurons in social perception and their relation to the brain regions discussed in our article: STS, TPJ, and DMFC. Iacoboni is quite correct that STS is the only one of these regions to be specifically experimentally linked to the mirror neuron system. The other areas thus far are less well characterized and have not been associated with the mirror neuron system. We would point out, however, a potentially limiting aspect of the work on mirror neurons. The experiments almost always involve the perception of motor acts performed by the hand, whereas social perception obviously encompasses a much larger domain. Perhaps one reason for the apparent anatomical focus of the mirror neuron system – in STS, posterior parietal lobe, and premotor cortex – is that the experiments are limited in scope. Whether TPJ, DMFC, and other areas implicated in social perception really have a specific role in informing simulation mechanisms such as mirror neurons is of course speculation on our part, and only further experiments will clarify the issue.

We end by considering a question pertinent to all of the commentaries and to our target article. What experimental predictions does our theory make? The most direct predictions involve compromising awareness by damaging specific brain regions. Certainly a large lesion to the social machinery should, by hypothesis, also affect one's own awareness. We argue that this relationship explains why lesions to the right TPJ and STS can result in hemispatial neglect.

But more specific hypotheses can be formulated. If the task of social attention – of building a perceptual model of someone else's attentional state – is emphasized in specific neuronal structures, then by hypothesis damage to those specific structures, or targeted reversible disruption of them, should lead to a deficit in one's own awareness.

Likewise in neglect in which the patient's awareness is impaired in one spatial hemifield, the patient should also be impaired in perceiving when someone else is directing attention to that same hemifield.

The specificity of these predictions is important. The theory does not predict that neglect patients should be generally lacking in social cognitive abilities. Likewise, the theory does not predict that awareness should fade given any neural damage or autism-like disability related to social perception. The proposed relationship between awareness and social perception is more specific and therefore more experimentally approachable.

In summary, we are encouraged by the many useful comments from colleagues, and we see potential for future experiments. The proposed theory of consciousness is specific, conceptually simple, and testable.

References from the Discussion Paper, the Commentaries, and the Reply

Aflalo, T. N., & Graziano, M. S. A. (2011). The organization of the macaque extrastriate visual cortex re-examined using the principle of spatial continuity of function. *Journal of Neurophysiology*, *105*, 305–320.

Alvarado, C. S. (2009). Late 19th- and early 20th-century discussions of animal magnetism. *International Journal of Clinical and Experimental Hypnosis*, *57*, 366–381.

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: Neuropsychological evidence from a false-belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience*, *16*, 1773–1784.

Arzy, S., Seeck, M., Ortigue, S., Spinelli, L., & Blanke, O. (2006). Induction of an illusory shadow person. *Nature*, *443*(7109), 287.

Astafiev, S. V., Shulman, G. L., & Corbetta, M. (2006). Visuospatial reorienting signals in the human temporo-parietal junction are independent of response selection. *European Journal of Neuroscience*, *23*, 591–596.

Baars, B. J. (1983). Conscious contents provide the nervous system with coherent, global information. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (pp. 41–79). New York, NY: Plenum Press.

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science*, *329*(5995), 1081–1085.

Balduzzi, D., & Tononi, G. (2008). Integrated information in discrete dynamical systems: Motivation and theoretical framework. *PLoS Comp. Biol.*, *4*(4), 1–18.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1–26.

Banissy, M. J., Cohen Kadosh, R., Maus, G. W., Walsh, V., & Ward, J. (2009). Prevalence, characteristics and a neurocognitive model of mirror-touch synaesthesia. *Exp Brain Res.*, *198*, 261–272.

Banissy, M. J., Garrido, L., Kusnir, F., Duchaine, B., Walsh, V., & Ward, J. (2011). Superior facial expression, but not identity recognition, in mirror-touch synesthesia. *J Neurosci*, *31*, 1820–1824.

Banissy, M. J., & Ward, J. (2007). Mirror-touch synesthesia is linked with empathy. *Nat Neurosci*, *10*, 815–816.

Barraclough, N. E., Xiao, D., Oram, M. W., & Perrett, D. I. (2006). The sensitivity of primate STS neurons to walking sequences and to the degree of articulation in static images. *Progress in Brain Research*, *154*, 135–148.

Baumeister, R. F., & Masicampo, E. J. (2010). Conscious thought is for facilitating social and cultural interactions: How mental simulations serve the animal–culture interface. *Psychological Review*, *117*, 945–971.

Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments: Should eye trust you? *Psychol Sci*, *17*(6), 514–520.

Ben Shalom, D., Mostofsky, S. H., Hazlett, R. L., Goldberg, M. C., Landa, R. J., Faran, Y., et al. (2006). Normal physiological emotions but differences in expression of conscious feelings in children with high-functioning autism. *JADD*, *36*, 395–400.

Beauregard, M. (2007). Mind does really matter: Evidence from neuroimaging studies of emotional self-regulation, psychotherapy, and placebo effect. *Progress in Neurobiology*, *81*, 218–236.

Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research*, *49*, 1154–1165.

Bird, C. M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on "theory of mind" and cognition. *Brain*, *127*, 914–928.

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain*, *133*, 1515–1525.

Birmingham, E., & Kingstone, A. (2009). Human social attention: A new look at past, present, and future investigations. *Annals of the New York Academy of Sciences*, *1156*, 118–140.

Blakemore, S. J., Boyer, P., Pachot-Clouard, M., Meltzoff, A., Segebarth, C., & Decety, J. (2003). The detection of contingency and animacy from simple animations in the human brain. *Cerebral Cortex*, *13*, 837–844.

Blakemore, S. J., Bristow, D., Bird, G., Frith, C., & Ward, J. (2005). Somatosensory activations during the observation of touch and a case of vision-touch synaesthesia. *Brain*, *128*, 1571–1583.

Blanke, O., & Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends Cogn Sci*, *13*(1), 7–13.

Blanke, O., Mohr, C., Michel, C. M., Pascual-Leone, A., Brugger, P., Seeck, M., et al. (2005). Linking out-of-body experience and self processing to mental own-body imagery at the temporoparietal junction. *Journal of Neuroscience*, *25*, 550–557.

Blanke, O., Ortigue, S., Landis, T., & Seeck, M. (2002). Stimulating illusory own-body perceptions. *Nature*, *419*, 269–270.

Block, N. (1996). How can we find the neural correlates of consciousness? *Trends in Neurosciences*, *19*, 456–459.

Brain, W. R. (1941). A form of visual disorientation resulting from lesions of the right cerebral hemisphere. *Proceedings of the Royal Society of Medicine*, *34*, 771–776.

Bridgeman, B. (2003). *Psychology and Evolution: The Origins of Mind*. Thousand Oaks, CA: Sage Publications.

Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, *46*, 369–384.

Brunet, E., Sarfati, Y., Hardy-Baylé, M. C., & Decety, J. (2000). A PET investigation of the attribution of intentions with a nonverbal task. *NeuroImage*, *11*, 157–166.

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

Byrne, R. W., & Bates, L. A. (2010). Primate social cognition: Uniquely primate, uniquely social, or just unique? *Neuron*, *65*(6), 815–830.

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

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, *32*, 121–182.

Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*, 200–219.

Chalmers, D. (1996). *The Conscious Mind*: *In Search of a Fundamental Theory*. Oxford University Press.

Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B. G., et al. (2007). The intentional network: How the brain reads varieties of intentions. *Neuropsychologia*, *45*, 3105–3113.

Coover, J. E. (1913). The feeling of being stared at. *American Journal of Psychology*, *24*, 570–575.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, *3*, 292–297.

Cottrell, J. E., & Winer, G. A. (1994). Development in the understanding of perception: The decline of extramission perception beliefs. *Developmental Psychology*, *30*, 218–228.

Craig, B. (2004). Human feelings: Why are some more aware than others? *Trends Cogn Sci.*, *8*, 239–241.

Creutzfeldt, O. (1987). Inevitable deadlocks of the brain-mind discussion. In B. Gulayas (Ed.), *The brain–mind problem: Philosophical and neurophysiological approaches* (pp. 3–27). Leuven: Leuven University Press.

Crick, F. (1995). The Astonishing Hypothesis: The Scientific Search For The Soul (Scribner reprint edition, 1995).

Crick, F., & Koch, C. (1990). Toward a neurobiological theory of consciousness. *Seminars in the Neurosciences*, *2*, 263–275.

Critchley, M. (1953). *The parietal lobes*. London, UK: Hafner Press.

Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A., & Dolan, R. J. (2004). Neural systems supporting introceptive awareness. *Nat Neurosci.*, *7*, 189–195.

Damasio, A. R. (1990). Synchronous activation in multiple cortical regions: A mechanism for recall. *Seminars in Neuroscience*, *2*, 287–296.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, *10*, 204–211.

de Lafuente, V., & Romo, R. (2005). Neuronal correlates of subjective sensory experience. *Nature Neuroscience*, *8*, 1698–1703.

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., & Slachevsky, A. (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain*, *132*(Pt 9), 2531–2540.

Descartes, R. (1984 [1641]). Meditations on first philosophy. In *The philosophical writings of René Descartes* (Vol. 2, pp. 1–62, Trans. J. Cottingham, R. Stoothoff, & D. Murdoch). Cambridge, UK: Cambridge University Press.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *4*, 2051–2062.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

Dennet, D. C. (1992). *Consciousness Explained*. Back Bay Books.

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

Dosch, M., Loenneker, T., Bucher, K., Martin, E., & Klaver, P. (2010). Learning to appreciate others: Neural development of cognitive perspective taking. *Neuroimage*, *50*(2), 837–846.

Ehrsson, H. H. (2007). The experimental induction of out-of-body experiences. *Science*, *317*, 1048.

Engel, A. K., König, P., Gray, C. M., & Singer, W. (1990). Stimulus-dependent neuronal oscillations in cat visual cortex: Inter-columnar interaction as determined by cross-correlation analysis. *European Journal of Neuroscience*, *2*, 588–606.

Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences*, *5*, 16–25.

Ferber, S., & Karnath, H. O. (2001). How to assess spatial neglect-line bisection or cancellation tasks. *Journal of Clinical and Experimental Neuropsycholology*, *23*, 599–607.

Filimon, F., Nelson, J. D., Hagler, D. J., & Sereno, M. I. (2007). Human cortical representations for reaching: Mirror neurons for execution, observation, and imagery. *NeuroImage*, *37*, 1315–1328.

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science*, *329*(5998), 1541–1543.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., et al. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*, 109–128.

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

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*, 694–724.

Frith, C. (1995). Consciousness is for other people. *Behavioral and Brain Sciences*, *18*, 682–683.

Frith, C. (2002). Attention to action and awareness of other minds. *Consciousness and Cognition*, *11*, 481–487.

Frith, C. D. (2008). The social functions of consciousness. In L. Weiskrantz, & M. Davies (Eds.), *Frontiers of Consciousness* (pp. 225–244). Oxford: Oxford University Press.

Frith, U. (2004). Confusions and controversies about Asperger syndrome. *JCCP*, *45*, 672–686.

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

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of Theory of Mind. *Trends in Cognitive Sciences*, *7*, 51–96.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia*, *38*, 11–21.

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

Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychon Bull Rev*, *10*(4), 843–876.

Gazzaniga, M. S. (1970). *The bisected brain*. New York, USA: Appleton-Century-Crofts.

Goel, V., Grafman, J., Sadato, N., & Hallett, M. (1995). Modeling other minds. *Neuroreport*, *6*, 1741–1746.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535–574.

Graziano, M. S. A. (2010). *God, soul, mind, brain: A neuroscientist's reflections on the spirit world*. Teaticket, MA: Leapfrog Press.

Graziano, M. S. A., & Aflalo, T. N. (2007). Rethinking cortical organization: Moving away from discrete areas arranged in hierarchies. *Neuroscientist*, *13*, 138–147.

Graziano, M. S. A., & Botvinick, M. M. (2002). How the brain represents the body: Insights from neurophysiology and psychology. In W. Prinz & B. Hommel (Eds.), *Attention and performance*: *Vol. XIX. Common mechanisms in perception and action* (pp. 136–157). Oxford University Press, Oxford, UK.

Graziano, M. S. A., & Sabine, K. (2011). Human consciousness and its relationship to social neuroscience: A novel hypothesis. *Cognitive Neuroscience*.

Gross, C. G. (1999). The fire that comes from the eye. *Neuroscientist*, *5*, 58–64.

Gross C. G., Bender, D. B., & Rocha-Miranda, C. E. (1969). Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science*, *166*, 1303–1306.

Gross, C. G., & Graziano, M. S. A. (1995). Multiple representations of space in the brain. *Neuroscientist*, *1*, 43–50.

Grossberg, S. (1999). The link between brain learning, attention, and consciousness. *Consciousness and Cognition*, *8*, 1–44.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., et al. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, *12*, 711–720.

Halligan, P. W., Fink, G. R., Marshall, J. C., & Vallar, G. (2003). Spatial cognition: Evidence from visual neglect. *Trends in Cognitive Sciences*, *7*, 125–133.

Halligan, P. W., & Marshall, J. C. (1992). Left visuo-spatial neglect: A meaningless entity? *Cortex*, *28*, 525–535.

Heilman, K. M., & Valenstein, E. (1972a). Frontal lobe neglect in man. *Neurology*, *22*, 660–664.

Heilman, K. M., & Valenstein, E. (1972b). Mechanism underlying hemispatial neglect. *Annals of Neurology*, *5*, 166–170.

Hernandez, A., Nacher, V., Luna, R., Zainos, A., Lemus, L., Alvarez, M., et al. (2010). Decoding a perceptual decision process across cortex. *Neuron*, *66*, 300–314.

Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience & Biobehavioral Reviews*, *34*, 575–583.

Hill, E., Berthoz, S., & Frith, U. (2004). Cognitive processing of own emotions in individuals with autistic spectrum disorders and in their relatives. *JADD*, *34*, 229–235.

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

Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson, & R. A. Hinde (Eds.), *Growing Points in Ethology* (pp. 303–317). London: Cambridge University Press.

Humphrey, N. (1983). *Consciousness regained: Chapters in the development of mind*. Oxford, UK: Oxford University Press.

Iacoboni, M. (in press). The ontological priority of representations: The case of mirror neurons and language. *Language and Dialogue*.

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

Insel, T. R. (2010). The challenge of translation in social neuroscience: A review of oxytocin, vasopressin, and affiliative behavior. *Neuron*, *65*(6), 768–779.

Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., et al. (2011). Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. *Neuron*, *70*(2), 363–374.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Neuroscience Reviews*, *2*, 194–204.

Jackson, G. M, Swainson, R., Mort, D., Husain, M., & Jackson, S. R. (2009). Attention, competition, and the parietal lobes: Insights from Balint's syndrome. *Psychol. Research*, *73*(2), 263–270.

Jellema, T., & Perrett, D. I. (2003). Cells in monkey STS responsive to articulated body motions and consequent static posture: A case of implied motion? *Neuropsychologia*, *41*, 1728–1737.

Jellema, T., & Perrett, D. I. (2006). Neural representations of perceived bodily actions using a categorical frame of reference. *Neuropsychologia*, *44*, 1535–1546.

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 17048–17052.

Kaas, J. (2002). Convergences in the modular and areal organization of the forebrain of mammals: Implications for the reconstruction of forebrain evolution. *Brain Behav Evol*, *59*(5–6), 262–272.

Kant, I. (1966 [1781]). *Critik der reinen Vernunft*. Trans. F. M. Muller. New York, NY: Anchor Books.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.

Karnath, H. O., Ferber, S., & Himmelbach, M. (2001). Spatial awareness is a function of the temporal not the posterior parietal lobe. *Nature*, *411*, 950–953.

Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (2004). Spatial attention speeds discrimination without awareness in blindsight. *Neuropsychologia*, *42*, 831–835.

Keysers, C., Kaas, J. H., & Gazzola, V. (2010). Somatosensation in social perception. *Nat Rev Neurosci.*, *11*, 417–428.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processes*, *8*, 159–566.

Kinsbourne, M. (1970). A model for the mechanism of unilateral neglect of space. *Transactions of the American Neurological Association*, *95*, 143–146.

Kinsbourne, M. (1977). Hemi-neglect and hemisphere rivalry. *Advances in Neurology*, *18*, 41–49.

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature*, *387*(6635), 767–768.

Koch, C., & Crick, F. C. (2001). On the zombie within. *Nature*, *411*, 893.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences*, *11*, 16–22.

Kockler, H., Scheef, L., Tepest, R., David, N., Bewernick, B. H., Newen, A., et al. (2010). Visuospatial perspective taking in a dynamic environment: perceiving moving objects from a first-person-perspective induces a disposition to act. *Conscious Cogn*, *19*(3), 690–701.

Kriegel, U. (2009). *Subjective Consciousness: A Self-Representational Theory*. Oxford University Press.

Lamm, C, Decety, J., & Singer, T. (2010). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy, Neuroimage, *54*(3), 2492–2502.

Lamm, C., & Singer, T. (2010). The role of the anterior insula cortex in social emotions. *Brain Struct Funct.*, *214*, 579–591.

Lamme, V. A. (2004). Separate neural definitions of visual consciousness and visual attention: A case for phenomenal awareness. *Neural Networks*, *17*, 861–872.

Lamme, V. A. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, *10*, 494–501.

Lau, H. C. (2008). Are we studying consciousness yet? In L. Weiskrantz & M. Davies (Eds.), *Frontiers of Consciousness* (pp. 245–258). Oxford: Oxford University Press.

Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video ergo sum: Manipulating bodily self-consciousness. *Science, 317*(5841), 1096–1099.

Leopold, D. A., & Rhodes, G. (2010). A comparative view of face perception. *J Comp Psychol*, *124*(3), 233–251.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431–461.

Llinas, R., & Ribary, U. (1994). Perception as an oneiric-like state modulated by the senses. In C. Koch & J. Davis (Eds.), *Large-scale neural theories of the brain* (pp. 111–124). Cambridge, MA: MIT Press.

Logan, G. D., & Crump, M. J. C. (2009). The left hand doesn't know what the right hand is doing: The disruptive effects of attention to the hands in skilled typewriting. *Psychol. Sci.*, *20*, 1296–1300.

Magnée, M. J., de Gelder, B., van Engeland, H., & Kemner, C. F. (2007). Facial electromyographic responses to emotional information from faces and voices in individuals with pervasive developmental disorder. *JCCP*, *48*, 1122–1130.

Mason, M. F. et al. (2007). Wandering minds: The default network and stimulus-independent thought. *Science*, *315*, 393–395.

Mazefsky, C. A., Kao, J., & Oswald, D. P. (2011). Preliminary evidence suggesting caution in the use of psychiatric self-report measures with adolescents with high-functioning autism spectrum disorders. *RAD*, *5*, 164–174.

Meister, I. G., Wienemann, M., Buelte, D., Grünewald, C., Sparing, R., Dambeck, N., et al. (2006). Hemiextinction induced by transcranial magnetic stimulation over the right temporo-parietal junction. *Neuroscience*, *142*, 119–123.

Merikle, P. M., Joordens, S., & Stolz, J. A. (1995). Measuring the relative magnitude of unconscious influences. *Consciousness and Cognition*, *4*, 422.

Mesulam, M. M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of Neurology*, *10*, 309–325.

Mesulam, M. M. (1999). Spatial attention and neglect: Parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *354*, 1325–1346.

Mitchell, J. P. (2009). Social psychology as a natural kind. *Trends in Cognitive Sciences*, *13*(6), 246–251.

Mitchell, L. P. (2008). Activity in the right temporo-parietal junction is not selective for theory-of-mind. *Cerebral Cortex*, *18*, 262–271.

Mort, D. J., Malhotra, P., Mannan, S. K., Rorden, C., Pambakian, A., Kennard, C., et al. (2003). The anatomy of visual neglect. *Brain*, *126*, 1986–1997.

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science*, *13*, 416–424.

Nagel, T. (1974). What is it like to be a bat? *The philosophical review*, *83*(4), 435–450.

Newman, J., & Baars, B. J. (1993). A neural attentional model for access to consciousness: A global workspace perspective. *Concepts in Neuroscience*, *4*, 255–290.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259.

Nummenmaa, L., & Calder, A. J. (2008). Neural mechanisms of social attention. *Trends in Cognitive Sciences*, *13*, 135–143.

Ochsner, K. N., Knierim, K., Ludlow, D. H., Hanelin, J., Ramachandran, T., Glover, G., et al. (2004). Reflecting upon feelings: An fMRI study of neural systems supporting the attribution of emotion to self and other. *Journal of Cognitive Neuroscience*, *16*, 1746–1772.

Pasquali, A., Timmermans, B., & Cleeremans, A. (2010). Know thyself: Metacognitive networks and measures of consciousness. *Cognition*, *117*, 182–190.

Passingham, R. E., Bengtsson, S. L., & Lau, H. C. (2010). Medial frontal cortex: From self-generated action to reflection on one's own performance. *Trends in Cognitive Sciences*, *14*, 16–21.

Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience*, *16*, 1706–1716.

Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. *Cerebral Cortex*, *15*, 1866–1876.

Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., et al. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction.

Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, *223*, 293–317.

Picciuto, V. (2011). Addressing higher-order misrepresentation with quotational thought. *Journal of Consciousness Studies*, *18*(3–4).

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

Ptak, R., & Schnider, A. (2010). The dorsal attention network mediates orienting toward behaviorally relevant stimuli in spatial neglect. *Journal of Neuroscience*, *30*, 12557–12565.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, *18*, 2188–2199.

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition*, *117*(2), 230–236.

Rafal, R. D. (1997). "Balint syndrome." In Feinberg, T. E., & Farah, M. J. (Eds.), *Behavioral Neurology and Neuropsychology* (pp. 337–356). New York: McGraw-Hill.

Raichle, M. E. (2010). Two views of brain function. *Trends in Cognitive Sciences*, *14*(4), 180–190.

Raichle, M. E., Macleod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *PNAS 2001*, *98*(2), 676–682.

Rilling, J. K., Dagenais, J. E., Goldsmith, D. R., Glenn, A. L., & Pagnoni, G. (2008). Social cognitive neural networks during in group and out group interactions. *Neuroimage*, *41*(4), 1447–1461.

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

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*, 264–274.

Rochat, P. (1998). Self-perception and action in infancy. *Exp Brain Res*, *123*(1–2), 102–109.

Rorden, C., Fruhmann Berger, M., & Karnath, H. O. (2006). Disturbed line bisection is associated with posterior brain lesions. *Brain Research*, *1080*, 17–25.

Rosenthal, D. M. (2000). Consciousness, content, and meta-cognitive judgments. *Consciousness and Cognition*, *9*, 203–214.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology. Human Perception and Performance*, *36*, 1255–1266.

Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience*, *7*, 499–500.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: fMRI investigations of theory of mind. *NeuroImage*, *19*, 1835–1842.

Saxe, R., Moran, J. M., Scholz, J., & Gabrieli, J. (2006). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. *Social Cognitive and Affective Neuroscience*, *1*, 229–234.

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

Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E. N., & Saxe, R. (2009). Distinct regions of right temporo-parietal junction are selective for theory of mind and exogenous attention. *PLoS One*, *4*, e4869.

Schooler, J. W. (2002). Re-representing consciousness: Dissociations between experience and meta-consciousness. *Trends in Cognitive Sciences*, *6*(8), 339–344.

Searle, J. R. (2007). Dualism revisited. *Journal of Physiology, Paris*, *101*, 169–178.

Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

Shultz, S., & Dunbar, R. (2010). Encephalization is not a universal macroevolutionary phenomenon in mammals but is associated with sociality. *Proceedings of the National Academy of Sciences, USA*, *107*(50), 21920–21924.

Shulman, G. L., Pope, D. L., Astafiev, S. V., McAvoy, M. P., Snyder, A. Z., & Corbetta, M. (2010). Right hemisphere dominance during spatial selective attention and target detection occurs outside the dorsal frontoparietal network. *Journal of Neuroscience*, *30*, 3640–3651.

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., & Frith, U. (2008). Levels of emotional awareness and autism: An fMRI study. *Social Neuroscience*, *3*, 97–112.

Smallwood, J. (2010). Why the global availability of mind wandering necessitates resource competition: Reply to McVay and Kane. *Psychological Bulletin*, *36*(2), 202–207.

Smith, S. M., Fox, P. T., Glahn, D. C., Mackay, C. E., Fillipini, N., Watkins, K. E., Toro, R., Laird, A. R., & Bekman, C. F. (2010). Correspondence of the brain's functional architecture during activation and rest. *PNAS 2009*, *106*(31), 13040–13045.

Spreng, R. N., & Grady, C. L. (2010). Patterns of brain activity supporting autobiographical memory, prospection and theory of mind, and their relationship to the default mode network. *Journal of Cognitive Neuroscience*, *22*(6), 1112–1123.

Straube, T., & Miltner, W. (2011). Attention to aversive emotion and specific activation of the right insula and right somatosensory cortex. *NeuroImage*, *54*, 2534–2538.

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nature Reviews Neuroscience*, *6*, 363–375.

Szczepanski, S. M., Konen, C. S., & Kastner, S. (2010). Mechanisms of spatial attention control in frontal and parietal cortex. *Journal of Neuroscience*, *30*, 148–160.

Tate, A. J., Fischer, H., Leigh, A. E., & Kendrick, K. M. (2006). Behavioural and neurophysiological evidence for face identity and face emotion processing in animals. *Philos Trans R Soc Lond B Biol Sci*, *361*(1476), 2155–2172.

Thiagarajan, T. C., Lebedev, M. A., Nicolelis, M. A., & Plenz, D. (2010). Coherence potentials: Loss-less, all-or-none network events in the cortex. *Plos Biol*, *8*(1), e1000278. doi:10.1371/journal.pbio.1000278.

Thompson, J. C., Hardee, J. E., Panayiotou, A., Crewther, D., & Puce, A. (2007). Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *NeuroImage*, *37*, 966–973.

Titchner, E. B. (1898). The feeling of being stared at. *Science*, *8*, 895–897.

Tononi, G., & Edelman, G. M. (1998). Consciousness and complexity. *Science*, *282*, 1846–1851.

Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biol. Bull.*, *215*, 216–242.

Treisman, A. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *40A*, 201–237.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 11656–11661.

Vallar, G. (2001). Extrapersonal visual unilateral spatial neglect and its neuroanatomy. *NeuroImage*, *14*, S52–S58.

Vallar, G., & Perani, D. (1986). The anatomy of unilateral neglect after right-hemisphere stroke lesions. A clinical/CT-scan correlation study in man. *Neuropsychologia*, *24*, 609–622.

Van Boxtel, J. J. A., Tsuchiya, N., & Koch, C. (2010). Opposing effects of attention and consciousness on afterimages. *Proc. Natl. Acad. Sci. USA*, *107*(19), 8883–8888.

Van Gulick, R. (2004). Higher-order global states (HOGS): An alternative higher-order model of consciousness. In R. Gennaro (Ed.), *Higher-Order Theories of Consciousness*, Philadelphia: John Benjamins.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., et al. (2001). Mind reading: Neural mechanisms of theory of mind and self-perspective. *NeuroImage*, *14*, 170–181.

Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., & Fink, G. R. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *Journal of Cognitive Neuroscience*, *16*, 817–827.

Von der Malsburg, C. (1997). The coherence definition of consciousness. In M. Ito, Y. Miyashita, & E. Rolls (Eds.), *Cognition, computation, and consciousness* (pp. 193–204). Oxford, UK: Oxford University Press.

Weed, E., McGregor, W., Feldbaek Nielsen, J., Roepstorff, A., & Frith, U. (2010). Theory of mind in adults with right hemisphere damage: What's the story? *Brain and Language*, *113*, 65–72.

Wicker, B., Michel, F., Henaff, M. A., & Decety, J. (1988). Brain regions involved in the perception of gaze: A PET study. *NeuroImage*, *8*, 221–227.

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.

Wyk, B. C., Hudac, C. M., Carter, E. J., Sobel, D. M., & Pelphrey, K. A. (2009). Action understanding in the superior temporal sulcus region. *Psychological Science*, *20*, 771–777.

Zacks, J. M., Vettel, J. M., & Michelon, P. (2003). Imagined viewer and object rotations dissociated with event-related fMRI. *Journal of Cognitive Neuroscience*, *15*, 1002–1018.