Reconceptualizing second-person interaction

ASSIGN BUSTER

Introduction

One of the main challenges for contemporary neuroscience has been to uncover the neural correlates of social cognition. Research in this area has been dominated by two main theories: the Theory Theory and the Simulation Theory. According to the Theory Theory, social cognition depends on a " Theory of Mind"—a psychological theory about how beliefs, desires, and intentions are interrelated and inform actions (Fodor, 1992 ; Gopnik and Meltzoff, 1997 ; Carruthers, 2009). Simulation Theory claims that social cognition involves " putting ourselves in the shoes of others" by simulating the mental states we would have in their situation (Goldman, 2006 ; Hurley, 2008 ; Gallese and Sinigaglia, 2011).

Despite the fact that they are often portrayed as rivals, most versions of the Theory Theory and the Simulation Theory share an important assumption. They take it for granted that social understanding (usually) involves " mindreading," i. e., the capacity to attribute mental states such as beliefs, desires, and intentions to others in order to predict or explain their behavior (Nichols and Stich, 2003 ; Apperly, 2011). Mindreading does not require us to interact with other people: we may simply speculate about their mental states while standing at the margins of the situation. As a result, proponents of the Theory Theory and the Simulation Theory have primarily investigated the neural correlates of social cognition by means of a " third-person" (3P) approach in which participating subjects are not actively engaging with other agents but merely observe them. Most studies on the neural correlates of Theory of Mind, for example, require subjects to make inferences about how the protagonist of a story would behave or feel (for review, see: Mar, 2011).

These experiments are usually devoid of any interaction between the subjects and the protagonist whose mental states they are supposed to read. This lack of interaction is also characteristic of neuroimaging research conducted in the Simulation Theory framework. Studies of the mirror neuron system (MNS), for instance, typically involve a condition in which subjects observe another agent who performs an action, and a condition in which they perform the same action themselves. However, there is no interaction between the subjects and the agent in either condition.

Recently this 3P paradigm has been challenged by researchers who call for a " second-person" (2P) approach to social cognition. These " interaction theorists," as we will label them, argue that the Theory Theory and the Simulation Theory are fundamentally flawed because they fail to recognize the importance of our dynamic interactions with others 1 . What is needed, according to them, is an " interactive turn" in social cognition research (de Jaegher et al., 2010). Some interaction theorists suggest that a 2P approach will shed new light on the neural mechanisms that underlie social cognition (Schilbach et al., forthcoming). Others claim that a 2P approach does justice to the phenomenology of our everyday encounters with others (Ratcliffe, 2007 ; Gallagher and Zahavi, 2008). Then there are those who think that a 2P approach to social cognition will allow us to solve the problem of other minds, i. e., the problem of how we can access the mind of others (Gallagher, 2004 ; Reddy, 2008). Besides these different motivations for advocating a 2P approach to social cognition, interaction theorists also have different conceptions of what such an approach precisely entails, and how it

should be spelled out in contrast to the 3P approach endorsed by the Theory Theory and the Simulation Theory.

The first aim of the present article is to critically analyze the 2P challenge to social neuroscience, and assess the various ways in which interaction theorists have articulated the distinction between 2P versus 3P modes of social cognition. In Section " Against the Idea of an ' Observational Stance'", we argue that interaction theorists are right to oppose the idea of an " observational stance." Drawing a parallel with recent criticism on the two-systems model of visual perception, we will show that there actually is no such thing as passive observation—every perceiver, no matter how detached, is actively involved in what she perceives. Although we take this to be a strong argument for interaction theory, it also shows that we cannot use the difference between active engagement and passive observation to ground a strong distinction between 2P and 3P modes of social cognition. In Section " Social Interaction Versus Social Cognition," we discuss another way of explicating this distinction. Some interaction theorists not only make a case for the intertwinement of perception and action, but also claim that this may be *constitutive* of social cognition. According to them, there are situations in which social cognition is nothing over and above social interaction. However, we argue not only that social cognition often does entail more than social interaction, but also that the proposed contrast between social interaction and social cognition does not provide a good basis for the distinction between 2P versus 3P modes of social cognition.

The second aim of the article is to put forward an alternative conceptualization of the distinction between 2P versus 3P modes of social

cognition—one that gives pride of place to the notion of *reciprocity* . In Section " Reconceptualizing 2P Interaction," we argue that what distinguishes 2P from 3P modes of social cognition is not their interactive or non-cognitive nature, but rather the fact that they involve reciprocal interaction. On our view, 2P modes of social cognition may and often do recruit capacities that interaction theorists take to be characteristic of 3P modes of social cognition—as long as the demand for reciprocity is met. Finally, in Section " The Real Challenge to Social Neuroscience," we briefly compare our proposal to [Frith and Frith's (2011)](#) " signaling model" of social cognition, and discuss its implications for neuroscientific experiments on social cognition.

Against the Idea of an " Observational Stance"

Interaction theorists often criticize the 3P stance toward others that is presupposed by the Theory Theory and the Simulation Theory for being a specialized and relatively rare mode of social cognition—one that is characterized by a lack of actual interaction and a reliance on passive observation. They argue that, in everyday life, we find ourselves constantly engaged in dynamic interactions with others: we buy a ticket at the counter of the station, travel by train with our fellow passengers, have a coffee with our colleagues and discuss new plans in a meeting. These 2P modes of social cognition do not require us to adopt an observational stance. In what follows, we will provide a further argument against the idea of a " pure" observational stance by drawing a parallel with recent criticism on the two-systems model of visual perception.

According to the two-systems model, visual perception depends on two different streams that are both functionally and neurally segregated. Dating back to the early work of Leslie Ungerleider and popularized by Milner and Goodale, this influential model distinguishes a ventral processing stream dedicated to " vision-for-perception" from a dorsal stream that is involved in " vision-for-action" (Milner and Goodale, 2008). The ventral processing pathway projects from early visual areas to the inferior temporal lobe, while the dorsal processing pathway projects to the parietal lobe. Neuropsychological support for this distinction is provided by patients with visual form agnosia, such as patient DF, who was unable to report the orientation of a bar that he was able to grasp in a correct way (Goodale et al., 1991). By contrast, patients with optic ataxia show preserved object recognition abilities, while having difficulties with directing actions toward these same objects. Other evidence for a dissociation between the dorsal and ventral visual stream has been obtained by studies on visual illusions, showing for instance that grasping kinematics in the Ebbinghaus illusion are insensitive to the illusory percept accompanying the mere perception of the stimulus (Smeets and Brenner, 2006). The two-systems model has been further corroborated by neuroimaging studies showing that visual information is processed differently depending on whether the information is used for subsequent action or perception (e. g., Valyear et al., 2006).

However, the two-systems model has not gone unchallenged. First, within the neuroscience community an ongoing debate concerns the interpretation of the evidence in favor of the two streams hypothesis (for recent discussion, see: Schenk and McIntosh, 2010). For instance, several studies have shown

that grasping and pointing movements are affected by visual illusions as well (see for instance: Skewes et al., 2011). In addition, in a recent paper it has been shown that patient DF's differential performance on the action and perception task can largely be accounted for by the effects of haptic feedback (i. e., only after grasping she gets feedback about the correctness of the movement; Schenk, 2012). At a neural level there is strong evidence for reciprocal interactions between dorsal and ventral stream areas at several levels in the processing hierarchy (Himmelbach and Karnath, 2005 ; Pisella et al., 2006). For example, it has been shown that the ability to consciously see an object and identify its " Gestalt" depends on both ventral and dorsal processing streams (Huberle and Karnath, 2011). Finally, the errors displayed by patients with optic ataxia or visual form agnosia cannot always easily be interpreted as evidence in favor of damage to one specific visual stream. For instance, patient DF, with supposed damage to ventral stream areas showed action planning deficits as well, such as a failure to anticipate the fingertip forces required for object grasping or displaying action semantic errors such as grasping objects in a functionally incorrect way (Carey et al., 1996). These considerations have led to a revision of the original two-systems model, such that the distinction between dorsal and ventral processing streams should be considered as reflecting a relative rather than an absolute functional specialization (Schenk and McIntosh, 2010).

In recent philosophical debates, the basic assumptions underlying the two-systems model have also been contested. Proponents of the sensorimotor approach to visual cognition, for example, have argued that the strict

distinction between " vision-for-perception" and " vision-for-action" is misguided, because there is no such thing as pure " vision-for-perception" (O'Regan and Noe, 2001 ; Noë, 2004). They argue that the problems with the two-systems model described above testify to the fact that perception involves the employment of sensorimotor skills, and cannot be fully separated from action 2 . Whenever we see a tomato, for example, our eyes only take into the fovea the plane orthogonal to the vector of the eyes' focus. However, our sensorimotor capacities let us perceive the tomato as a three-dimensional solid object—one that can be grasped, and which appearance changes as we move around it. On the view advocated by the sensorimotor approach, the visual system has evolved in order to enable us to act in the surrounding world (Wheeler, 2006). As a result, the way in which we perceive the world depends on our bodily capabilities.

The idea that perception and action are intimately linked is not new and dates back to the ideomotor principle put forward by William James, who noted that " every representation of a movement awakens in some degree the actual movement which is its object" (James, 1890/1981). More recently this principle has seen renewed interest in the so-called " theory of event coding," according to which perception and action share a common representational format (Hommel et al., 2001). Support for this idea is found in behavioral experiments for instance, in which it is shown that the presentation of an action effect (e. g., a sound) results in the reactivation of the motor program associated with achieving the action effect (e. g., making a button press; e. g., Hommel, 1996). These findings reflect that based on training we have acquired strong associations between specific actions and

their resultant effects. The ideomotor principle accounts for a wide range of behaviors in which perception and action are tightly linked, such as imitation, observational learning and joint action. In the case of imitation, for instance, observing a specific movement, such as lifting a finger, activates in the observer the corresponding motor program required for achieving the effect and thereby facilitates imitative behavior (e. g., Brass et al., 2000). At a neural level, this perception-action coupling is likely mediated by visuomotor neurons in premotor and parietal areas (Koski et al., 2002 ; Kilner et al., 2004 ; Newman-Norlund et al., 2007). A complementary line of evidence for the idea that perception is directly coupled to action can be found in the " selection-for-action" principle, according to which the sole purpose of the perceptual system is to gather information for interaction with the environment (Allport, 1987). It has been found for instance, that one's action intention determines the way in which sensory information is processed already at an early stage in the visual system, as reflected in a modulation of early visual evoked potentials when one intends to grasp compared to point toward a target (van Elk et al., 2010). Together these studies highlight the close link of perception and action, and suggest that any attempt to demarcate perception- from action-related processes in a principled way is arbitrary.

The criticism of the two-systems model of visual perception can be extended to the debate on social cognition to illustrate that there is no such thing as a pure observational stance toward others. For example, Schilbach et al. (2008) have shown that when we see a smiling face we automatically tend to mimic this smile, at least in terms of specific muscle activation. Therefore,

the authors conclude that " the process of perceiving faces always includes an ' enactive' element through which we engage with and respond to stimuli instead of a mere ' passive' perception of face-based cues." Another illustration is provided by the MNS studies: if one takes a closer look at the neural processes involved in cases where subjects " passively" observe another agent's action from a 3P point of view, one notices that there is only a short amount of time (30–100 ms) between the activation of the visual cortex and the activation of the pre-motor cortex (Gallagher, 2007).

Although this casts doubt on the possibility to draw a strict demarcation line between action and perception, it does not imply that we cannot differentiate between observation and action conditions. The MNS studies, for example, show that during the observation of another agent's action, our motor system becomes active " as if" we were executing the action ourselves (Gallese, 2001). Some argue that in the case of action observation the actual execution of the action is inhibited (Schutz-Bosbach et al., 2009). Others claim that the absence of an efference copy of the motor command signals that the event is externally generated (Wolpert et al., 1995). However, what is agreed upon is that we can sensibly distinguish between observation and action conditions.

Social Interaction Versus Social Cognition

Interactivists often claim that 2P interactions rather than 3P observations are the backbone of social cognition. More in particular, they argue that 2P modes of social cognition are *primary* to 3P modes of social cognition, not only in the sense that (1) they involve capabilities that come earlier in development and are likely to be partially innate, but also in the sense that

(2) they remain the default way how we understand others (Gallagher, 2001 , 2011).

The first claim about the developmental primacy of 2P modes of social cognition might look problematic in the light of recent studies on " implicit" false belief understanding in early infancy. Several " spontaneous-response" false belief tests, in which infants' understanding of false belief is inferred from the behavior they spontaneously produce (e. g., anticipatory looking, longer looking times), seem to indicate that infants at a very young age are already able to adopt a 3P observational stance toward other agents in order to anticipate their behavior (see Baillargeon et al., 2010 for an overview).

However, even without taking into account these findings, proponents of the Theory Theory and the Simulation Theory could maintain that the claim about developmental primacy is compatible with the idea that social development basically comes down to a *transition* from 2P to 3P modes of social cognition. As Currie (2008 , p. 212) sees it, for instance, the abilities for 2P modes of social cognition " underpin early intersubjective understanding, and make way for the development of later theorizing or simulation [i. e., 3P modes of social cognition]" (see Spaulding, 2010 for a discussion). However, this is certainly not what most interaction theorists have in mind. They argue that 2P interaction does not " make way" for purportedly more sophisticated mindreading processes, but instead *continues* to characterize our everyday encounters even as adults. This is where the second claim about the dominance of 2P interaction comes in. If we look at the " phenomenological evidence" and pay attention to our " everyday experience," so the argument goes, we will find that 2P

interactions rather than 3P observations are pervasive in our social life (see, e. g., Ratcliffe, 2007 ; Gallagher and Zahavi, 2008).

As we have argued elsewhere (de Bruin et al., 2011 ; de Bruin and Kästner, 2012), the claim that 2P interactions remain the default way how we understand others is problematic insofar it depends on an appeal to phenomenology. The question which mode of social cognition is characteristic of our everyday encounters with others is an empirical one, and cannot be decided on the basis of a " simple phenomenological argument" (Gallagher, 2004). Overgaard and Michael (under review) rightly criticize the idea of having a single " everyday stance" toward other people: in the course of any one day, we not only interact with others in various ways, but we also, and not infrequently, simply observe people. Ultimately, the question about the dominance of 2P versus 3P modes of social cognition might simply boil down to a question about the commonality of a certain type of personality, for instance, extrovert (as in " interacting") versus introvert (as in " observing") (McCrae and Costa, 1987).

Claims about the developmental primacy and phenomenological pervasiveness of 2P versus 3P modes of social cognition also face a more general worry. If interaction theorists spell out the difference between 2P and 3P modes of social cognition in terms of active engagement versus passive observation, then it becomes unclear how to draw a line between 2P and 3P modes of social cognition. For, as we have argued in the previous section, the distinction between active engagement and passive observation appears to be gradual rather than absolute. And this, in turn, undermines the

claims about the developmental primacy and phenomenological pervasiveness of 2P interactions.

Some interaction theorists, however, spell out the difference between 2P and 3P modes of social cognition in a different way. They claim that 2P modes of social cognition are " direct" in the sense that they do not require *cognitive processes* to mediate between our perception of others and our actions toward them. Gallagher (2008 , p. 540), for instance, maintains that " what we call social cognition is often nothing more than social interaction. What I perceive in these cases does not constitute something short of understanding. Rather my understanding of the other person is constituted within the perception–action loops that define the various things that I am doing with or in response to others." Gallagher proposes a rich notion of enactive perception, which is meant to obviate the kind of cognitive processes postulated by the Theory Theory and the Simulation Theory. He argues that " in seeing the actions and expressive movements of the other person in the context of the surrounding world, one already sees their meaning; no inference to a hidden set of mental states (beliefs, desires, etc.) is necessary" (ibid., p. 542).

In a recent article, de Jaegher et al. (2010) explain in more detail how social cognition can be equivalent to social interaction. The authors distinguish between *constitutive* and *enabling* conditions for social cognition. In contrast to an enabling condition, according to which the ability must have been acquired at some point in development, a constitutive condition requires that the ability is exercised at the very moment we are trying to make sense of others. de Jaegher et al. (2010) argue that, in some cases, 2P interactions

can be a constitutive and not merely an enabling condition for social cognition.

It is not our aim here to argue against this modest claim. Rather, we would like to point out that interaction theorists still have to account for those cases in which social cognition clearly is something over and above social interaction. Take interaction theory's criticism of the 3P paradigm employed by the Theory of Mind approach and the Simulation Theory, for example. As Overgaard and Michael (under review) argue, if interaction theorists agree that this paradigm puts subjects in the role of detached spectators rather than interacting agents—and their complaint shows that they do agree with this—then the results of these experiments clearly show that social cognition is possible without social interaction. Or consider empirical studies of cases in which social interaction is completely lacking but a capacity for social cognition remains. Patients suffering from a total " locked-in-syndrome" (Bauer et al., 1979), for example, are no longer able to engage in real-time interaction with others, but they are still able to understand them to some degree (Laureys et al., 2005) 3 .

We can find similar dissociations between enabling and constitutive conditions in other domains as well. For example, the development of a body image, i. e., a (cognitive) system of perceptions, attitudes, and beliefs pertaining to oneself (Cash and Brown, 1987 ; Powers et al., 1987 ; Gardner and Moncrieff, 1988), depends on a *body schema* —a system of sensorimotor capacities that functions without reflective or perceptual monitoring in an immediate and close to automatic fashion (Gallagher, 2005). Although a body schema is an enabling condition for a body image, it is not

constitutive condition. Patients with deafferentation, such as Ian Waterman (Cole, 1995 ; Gallagher and Cole, 1995), suffer from certain impairments in their body schema (loss of tactile and proprioceptive input), but their body image remains intact and even allows them to compensate their disabilities to some extent. Another interesting dissociation between enabling and constitutive conditions has been found in relation to the use of linguistic concepts. Whereas there is a clear correlation between action verbs like " kick," " pick" and " lick" and pre-motor cortex activation (Pulvermüller and Fadiga, 2010), this is not the case for abstract verbs such as " think" (Rueschemeyer et al., 2007). In other words, although understanding action verbs may be a necessary step for understanding more abstract psychological verbs, it is certainly not a constitutive condition.

What these examples show is that it is not hard to come up with cases in which social cognition is something over and above social interaction. The question is to what extent interaction theorists are able to account for these often more advanced forms of social cognition. According to de Jaegher and Froese (2009 , p. 439), the biggest challenge for interaction theorists is " to show how an explanatory framework that accounts for basic biological processes [i. e., enactivism] can be systematically extended to incorporate the highest reaches of human cognition." This is what they call " the cognitive gap" 4 .

A more important question for our purpose here, however, is whether the proposed contrast between social interaction and social cognition provides us with a good basis for the distinction between 2P versus 3P modes of social cognition. For most interaction theorists, the main target in the debate on

social cognition has been the so-called " sandwich model" of the mind, which regards " perception as input from the world to the mind, action as output from the mind to the world, and cognition as sandwiched in between" ([Hurley, 2008](), p. 2). According to the sandwich model, cognition is required in order to " translate" visual input into motor output, since there is no direct interaction between perception and action. Because of their commitment to this model, many proponents of the Theory Theory and the Simulation Theory have simply assumed that our social engagements require us to engage in a cognitive process of mental state attribution (by means of either theory or simulation or both).

On the one hand, we agree with interaction theorists that the sandwich model should not be presupposed as a general model underlying all forms of social cognition (as mindreaders tend to do). At the same time, however, from this it does not automatically follow that one has to reject the cognitive capacities that are thought to be representative of the sandwich model. Some of these capacities might actually play an important role in 2P modes of social cognition as well. In the next section, we will substantiate this idea by proposing an alternative conceptualization of the distinction between 2P and 3P modes of social cognition.

Reconceptualizing 2P Interaction

We propose that what distinguishes 2P from 3P modes of social cognition is their *reciprocal* nature. That is, 2P modes of social cognition feature agents who coordinate their actions with one another—what is sometimes called " attunement" ([Fuchs and de Jaegher, 2009]() ; [de Jaegher et al., 2010)]() . Importantly, we take the capacity for reciprocal interaction to be an

ontogenetic achievement and not something that human beings are simply born with. Following Sebanz et al. (2006) , we can identify several important developmental stepping stones.

First of all, reciprocal interaction depends on the ability to share representations of objects and events with others. Visual habituation studies indicate that 5-month-old infants already respond selectively to the goals of another agent rather than the physical details of their actions (Woodward, 1998 , 2005). However, it is not until 9–12 months of age that they begin to engage in shared attention, and their interactions with others begin to have a reference to the things that surround them (Hobson, 2002 ; Tomasello et al., 2005). Shared attention creates a " perceptual common ground" insofar it requires that the attending of infant and agent has a common focus. This allows infants to direct another agent's attention to outside objects in which they are interested in themselves. The pointing gesture, for example, enables them to declare their interest in specific objects in their surroundings (Phillips et al., 2002 ; Woodward and Guajardo, 2002 ; Sodian and Thoermer, 2004). More importantly, however, shared attention also allows infants to coordinate their actions with those of another agent. Meltzoff (1995) showed that 18-month-olds are capable of completing an unfinished action of another agent, such as pulling apart miniature dumbbells.

Although shared attention provides interacting agents with a focal point of interest, it is grounded in a more basic system for sharing representations: the MNS. The MNS matches action observation and action production (Rizzolati and Craighero, 2004 ; Rizzolatti and Sinigaglia, 2010), and

facilitates a " common coding" of perception and action (see Section " Against the Idea of an ' Observational Stance'") 5 . MNS activation has been investigated in early infancy as well (Kanakogi and Itakura, 2010), and research on infant imitation has been cited as evidence for the fact that the MNS is an innate mechanism (e. g., Iacoboni et al., 1999 ; Decety et al., 2002 ; Grezes et al., 2003 ; Iacoboni, 2005 ; Iacoboni and Dapretto, 2006) 6 .

What is important is that the MNS facilitates *action anticipation* , which is considered a second prerequisite for coordinating one's actions with those of another agent according to Sebanz et al. (2006) . Knowing what the other will do next is crucial for coordinating one's actions with those of another agent. Falck-Ytter et al. (2006) , for example, showed that 12-month-old infants are capable of anticipating an agent's action toward an object (picking up and placing it in a container) by making eye movements ahead of the moving hand. The experimenters argued that these findings provide direct support for the idea that action anticipation depends on a MNS which is triggered by the infant's perception of another agent's goal-directed behavior. More direct support for the involvement of the MNS in action prediction was obtained in a study by Meyer et al. (2011) , which showed a stronger anticipatory motor-related brain response when 3-year old children observed the action of a partner they were actively interacting with compared to the action of an outsider.

We can elucidate the role of the MNS in action anticipation by mapping the neural circuit of the MNS onto an inverse-forward model (Iacoboni, 2003 , 2005). The superior temporal sulcus (STS) is responsible for the visual representation of an observed action. An inverse model then feeds this visual

representation into the fronto-parietal MNS and converts it into a motor plan. In a next step, this motor plan is sent back from the fronto-parietal mirror neuron to the STS and converted into a predicted visual representation (a sensory outcome of action) by means of a forward model. This two-step process explains how infants (and adults, see Flanagan and Johansson, 2003 ; Ambrosini et al., 2011) are able to track another agent's goal-directed behavior toward objects with predictive eye-movements.

The MNS might also play a role in the initiation and execution of complementary actions. Newman-Norlund et al. (2007) found that mirror neuron areas (right inferior frontal gyrus and bilateral inferior parietal lobes) are more active when observers are simultaneously preparing a complementary action than when they are preparing an imitative action. However, as Sebanz et al. (2006) point out, the ability to prepare complementary actions cannot be fully explained in terms of shared representations. Motor resonance might enable action anticipation, but this (1) crucially depends on action perception and (2) does not explain how we become capable of choosing an appropriate complementary action at an appropriate time. In order to address the first point, Sebanz et al. (2006) appeal to studies on *shared task representations* , in which two agents have to covertly represent each other's task requirements without observing each other's action. For instance, in a study by Ramnani and Miall (2004) , participants acquired stimulus–response mappings, and were then presented with stimuli indicating whether they should respond, a co-actor in another room should respond, or a computer should respond. Although the other's actions could not be observed, participants anticipated the co-actor's

actions. This was associated with activity in motor areas, including ventral premotor cortex, as well as areas typically involved in mindreading. According to Sebanz et al. (2006) , these results suggest that the mechanisms underlying mental state attribution might be triggered by shared task representations (cf. Sebanz and Frith, 2004). In order to deal with the second point, Sebanz et al. (2006) postulate a third prerequisite for action coordination: the ability to integrate the predicted effects of own and others' actions. They discuss this ability in relation to a number of studies that show how individuals incorporate others' action capabilities into their own action planning (Richardson et al., 2007), and how temporal feedback about another agent's action is used in anticipatory action control (Knoblich and Jordan, 2003 ; Jordan and Knoblich, 2004).

Sebanz et al. (2006) pay relatively little attention to what we take to be another crucial prerequisite for reciprocal interaction: *perspective taking* . In order to engage in reciprocal interaction, agents have to be able to account for differences in perspective. Elsewhere, we have proposed a developmental model in which we distinguish three modes of perspective taking (de Bruin and Newen, 2012):

1. Motor perspective taking, which allows infants to understand another agent on the basis of her movements (e. g., Woodward, 1998 , 2003 , 2005).
2. Visual perspective taking, which allows infants to understand another agent on the basis of what she (visually) perceives (e. g., Onishi and Baillargeon, 2005 ; Southgate et al., 2007).

3. Cognitive perspective taking, which allows children to understand another agent on the basis of propositional attitudes such as beliefs and desires (e. g., Wimmer and Perner, 1983 ; Baron-Cohen et al., 1985 ; Rakoczy et al., 2007).

The development of perspective taking is important insofar as reciprocal interaction requires that agents are on " the same level." For example, classic versions of the false belief test show that children under 4 years of age fail to verbally predict the behavior of another agent on the basis of her false belief (cognitive perspective taking). Of course this does not mean that they are unable to engage in reciprocal interaction. As Gallagher (2005) has pointed out, for example, although these children fail to predict the behavior of the agent they observe, they have no difficulty understanding the experimenter. But it does show that they are not yet able to reciprocally interact with other agents in terms of their (false) beliefs—at least not on a verbal level 7 . More advanced modes of perspective taking allow children to engage in more advanced modes of social interaction.

Importantly, the various capacities described above can be recruited in 2P as well as 3P modes of social cognition. They are not to be classified as 2P or 3P because of their interactive or perceptual nature, or because they do or do not involve cognitive processing. What counts instead is whether they are recruited for reciprocal (2P) or non-reciprocal (3P) interaction. On our view, therefore, 2P modes of social cognition may involve a lot of observation and only a minimal amount of action (see, for example, Schilbach et al. (2010) on interactive gaze following). Furthermore, 2P modes of social cognition may involve cognitive processes such as mental state attribution. Imagine that I

am playing an online chess-game with a friend who lives in the US. I'm staring at my computer screen and from time to time I click on my left mouse button. There is a lot of mindreading going on: I am trying to find out what my friend's next move will be, and whether I can capture his queen in the next turn. This scenario qualifies as a 2P mode of social cognition—even though it involves a lot of mindreading and only a minimal amount of bodily movement—because there is reciprocal interaction between us. Now imagine that I am helping someone who is drunk walk home 8 . I am practically dragging him forward, but he is too drunk to realize this. I am not thinking about whether he believes he is drunk, or whether he still desires beer; all my attention is focused on preventing him from stumbling. On our view, this scenario should not be classified as a 2P mode of social cognition. Despite the fact that it features a very active agent who is not engaged in mindreading, there is no reciprocity between the agents and hence no 2P interaction.

These examples show that capacities that are usually associated with (non-reciprocal) 3P modes of social cognition, such as perspective taking, actually play a crucial role in (reciprocal) 2P modes of social cognition as well. Developmental studies show that this is not only true for adult human beings, but also for infants. Buttelmann et al. (2009) , for example, provides an excellent illustration of how infants manage to engage in reciprocal interaction with an experimenter by taking into account his visual perspective. In the experiment, infants watched as a toy was transferred from box A to box B while an experimenter either witnessed the transfer of the toy (true belief condition) or not (false belief condition). Then the

experimenter attempted unsuccessfully to open box A—the empty box. In the true belief condition, infants could follow their natural tendency to help the experimenter by opening box A for him. In the false belief condition, if infants understood the experimenter's false belief, they had to understand that he wanted the toy he thought was in there. In this case they should not simply help him to open box A, but rather go to box B and retrieve the toy for him. The results indicated that, by 18 months of age, infants were able to actively assist the experimenter in his search for the toy. What this shows is that perspective taking is not limited to non-reciprocal 3P modes of social cognition, but instead plays a constitutive role in 2P modes of social cognition as well.

According to our reconceptualization, 2P modes of social cognition *can* but do not necessarily *have* to be cooperative in nature. Competitive interactions can still be reciprocal. Think, for example, of a tennis game or a soccer match. Furthermore, 2P modes of social cognition are not only about understanding other agents but also about misunderstanding them. As de Jaegher (2009) suggests " misunderstandings are the pivots around which the really interesting stuff of social understanding revolves. In these instances where coordination is lost, we have the potential to gain a lot of understanding" (p. 540).

The Real Challenge to Social Neuroscience

Let us briefly summarize our line of argument. So far we argued against two ways in which the distinction between 2P versus 3P modes of social cognition can be articulated: as active engagement versus passive observation, and as social interaction versus social cognition. Instead, we have proposed an

alternative conceptualization of this distinction—one that gives pride of place to the notion of *reciprocity* . Accordingly, capacities that interaction theorists take to be characteristic of 3P modes of social cognition play an important role in 2P modes of social cognition as well.

Thus, on our view, 2P modes of social cognition may involve mindreading. However, this does not mean that we take mindreading to be a necessary ingredient of 2P modes of social cognition. Consider the " signaling" model of social cognition recently put forward by Frith and Frith (2011) . This model distinguishes between involuntary signaling and ostensive signaling. Involuntary signaling is automatically triggered by bodily movement. Frith and Frith point out that the perception of biological movements elicits activity in the STS, especially the posterior part (Allison et al., 2000), and suggest that this is likely to be a very basic and universal brain mechanism. Ostensive signaling, by contrast, is done deliberately (e. g., by making eye contact or calling someone by name). This type of signaling is needed for " closing the loop" in 2P modes of social cognition, where both sender and receiver need mutual knowledge that signals are being exchanged deliberately. Furthermore, Frith and Frith propose that a critical role in establishing mutual knowledge between sender and receiver is played by anterior rostral medial prefrontal cortex (MPF) or arMPFC (see also Amodio and Frith, 2006). And because activity in the arMPFC is elicited by mentalizing tasks, they argue that mindreading is very important for closing the loop between minds.

We would like to propose that what is required for closing the loop is *reciprocal interaction* rather than mutual knowledge. This proposal is less

problematic as well as less demanding. It is less problematic than the requirement of mutual knowledge because, in order for knowledge between agents to be mutual, each agent has to know what the other agent knows and also know that the other agent knows that the first agent knows etc. This leads to an infinite regress (Lewis, 1969 ; Clark and Marshall, 1981 ; Sperber and Wilson, 1995). It is less demanding because it does not necessarily involve mindreading (since mindreading is only necessary as long as we assume that mutual knowledge is required to close the gap). Our discussion of the various forms of perspective taking (see " Reconceptualizing 2P Interaction" section) showed that there is more than one way to close the loop between minds. For example, visual perspective taking closes the loop insofar it enables agents to represent whether a given object is seen by another agent—without requiring them to attribute mental states to others (Hutto, 2011). Cognitive perspective taking, by contrast, enables agents to represent another agent's belief about a given state of affairs. This way of closing the gap *does* involve mental state attribution.

What are the implications of our view for neuroscientific research on social cognition? First, our reconceptualization of 2P interaction is meant to encourage researchers to take into account *both* observational and enactive conditions when studying the neural correlates of reciprocal interaction. For example, it would be interesting to contrast observational 2P conditions in which subjects are following the gaze of a virtual avatar (Schilbach et al., 2010) with more enactive 2P conditions in which subjects are throwing a ball with a virtual avatar (David et al., 2006). This would make clear to what extent these conditions recruit common resources or are neurally

differentiated. Second, our proposal invites a closer look at the role of cognitive processing in reciprocal interaction. So far, a lot of research in social neuroscience has focused on non-reciprocal modes of social cognition, in which subjects have to attribute mental states to another agent. We know that mental state attribution in such conditions is associated with a Theory of Mind network, consisting of the MPF, the temporoparietal junction (TPJ), the STS and the temporal poles (Frith and Frith, 2003 ; Amodio and Frith, 2006). However, what we also want to know is to what extent this network is recruited during reciprocal interactions, in which subjects have to attribute mental states *to each other* . The new field of neuro-economics, for example, uses paradigms from game theory and behavioral economics to study the neural correlates of social interactions and preferences, e. g., for fairness, cooperation and trust (e. g., Singer, 2012). Most studies in this field involve reciprocal interactions in which subjects attribute mental states to each other, for instance when playing some version of the prisoner's dilemma game. It would be interesting to see if these reciprocal interactions share common (neural) resources with the non-reciprocal modes of social cognition mentioned above. Similar questions can be raised about the role of the MNS in reciprocal interactions. Most MNS studies still employ non-reciprocal paradigms, in which subjects either observe another agent's action or perform the same action themselves. The real challenge to social neuroscience would be to transform both Theory of Mind and MNS studies into full-blown dynamical studies involving reciprocal 2P interactions. This might not be as hard as it looks. For example, one could take a classic version of the false belief test, in which infants have to attribute false belief to another agent, as a starting point, and add reciprocal elements like gaze

interaction between the infant and the agent in a stepwise manner. Such an experiment might also put the findings on false belief understanding in a new perspective.

In this article we have argued for an understanding of 2P modes of social cognition in terms of reciprocity. What distinguishes 2P from 3P modes of social cognition is not the amount of action involved or the absence of cognitive processing, but rather the fact that they involve reciprocal interaction. In the end, this is what the interactive turn in social cognition research should be about.

Conflict of Interest Statement

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

Footnotes

1. ^ This narrow approach to social cognition is probably partly the result of certain methodological problems that enter the picture when one tries to investigate dynamic second-person interactions (see Schilbach et al., forthcoming). We thank one of the reviewers for bringing this to our attention.

2. ^ We believe this claim is sound, even though we acknowledge that there are serious problems with Noë's theory of object perception (Schlicht and Pompe, 2007).

3. ^ Total locked-in syndrome is a version of locked-in syndrome where the eyes are paralyzed as well.

4. ^ See de Bruin and de Haan (forthcoming) for a more detailed discussion of this problem and a thorough evaluation of recent proposals that try to bridge this cognitive gap.

5. ^ Overlapping MNS activation has also been found when subjects listen to action-related sounds (Aglioti and Pazzaglia, 2010), observe another person being touched (Keysers et al., 2010) or observe emotional expressions (Wicker et al., 2003).

6. ^ However, we would like to point out that there are still many open questions about the role of the MNS in infant development (Gerson and Woodward, 2010 ; Meltzoff, 2006). It is also not clear whether the MNS should indeed be seen as an inherited adaptation for action understanding (an evolved system), or rather as a byproduct of associative learning that is shaped through interaction with others and which is basically the result of social experience (Heyes, 2010).

7. ^ Our notion of cognitive perspective-taking is rather demanding, in the sense that it requires children to be sensitive to beliefs and desires as propositional attitudes with propositional content. Elsewhere we have argued that studies on " implicit" false belief understanding in early infancy do not meet this constraint (e. g., de Bruin et al., 2011 ; Section 4; Strijbos and de Bruin, forthcoming, Section 6). Although we realize that this is a controversial issue, we do not have enough space to discuss it in more detail.

8. ^ We thank one of the reviewers for bringing this interesting example to our attention.

References

Aglioti, S. M., and Pazzaglia, M. (2010). Representing actions through their sound. *Exp. Brain Res* . 206, 141–151.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

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

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Allport, A. (1987). " Selection for action: some behaviorial and neurophysiological considerations of attention and action," in *Perspectives on Perception and Action* , eds H. Heuer and A. F. Sanders (Hillsdale, NJ: Lawrence Erlbaum Associates), 395–419.

Ambrosini, E., Constantini, M., and Sinigaglia, C. (2011). Grasping with the eyes. *J. Neurophysiol* . 106, 1437–1442.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

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

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Apperly, I. (2011). *Mindreaders: The Cognitive Basis of ' Theory of Mind'* . New York, NY: Psychology Press.

Baillargeon, R., Scott, R. M., and Zijing, H. (2010). False-belief understanding in infants. *Trends Cogn. Sci* . 14, 110–118.

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

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

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Bauer, G., Gerstenbrand, F., and Rumpl, E. (1979). Varieties of the locked-in syndrome. *J. Neurol* . 221, 77–91.

[Pubmed Abstract](#) | [Pubmed Full Text](#)

Brass, M., Bekkering, H., Wohlschläger, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain Cogn* . 44, 124–143.

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Buttelmann, D., Carpenter, M., and Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112, 337–342.

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Carey, D. P., Harvey, M., and Milner, A. D. (1996). Visuomotor sensitivity for shape and orientation in a patient with visual form agnosia. *Neuropsychologia* 34, 329–337.

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Cash, T. F., and Brown, T. A. (1987). Body image in anorexia nervosa and bulimia nervosa: a review of the literature. *Behav. Modif* . 11, 487–521.

Pubmed Abstract | Pubmed Full Text

Carruthers, P. (2009). How we know our own minds: the relationship between mindreading and metacognition. *Behav. Brain Sci* . 32, 121–138.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Clark, H. H., and Marshall, C. R. (1981). " Definite reference and mutual knowledge," in *Elements of Discourse Understanding* , eds A. K. Joshe, B. L. Webber, and I. A. Sag (Cambridge: Cambridge University Press), 10–63.

Cole, J. D. (1995). *Pride and a Daily Marathon* . Cambridge, MA: MIT Press; originally London: Duckworth, 1991.

Currie, G. (2008). Some ways to understand people. *Philos. Explor* . 11, 211–218.

David, N., Bewernick, B. H., Cohen, M. X., Newen, A., Lux, S., Fink, G. R., Shah, N. J., and Vogeley, K. (2006). Neural representations of self versus other: visual-spatial perspective taking and agency in a virtual ball-tossing game. *J. Cogn. Neurosci* . 18, 898–910.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

de Bruin, L. C., and Kästner, L. (2012). Dynamic embodied cognition. *Phenomenol. Cogn. Sci* . doi: 10. 1007/s11097-011-9223-1

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

de Bruin, L. C., and Newen, A. (2012). An association-based account of false belief understanding. *Cognition* 123, 240–259.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

de Bruin, L. C., Strijbos, D. W., and Slors, M. (2011). Early social cognition: alternatives to implicit mindreading. *Rev. Philos. Psychol* . 2, 499–517.

Decety, J., Chaminade, T., Grezes, J., and Meltzoff, A. (2002). A PET exploration of the neural mechanisms involved in reciprocal imitation. *Neuroimage* 15, 265–272.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

de Jaegher, H. (2009). Social understanding through direct perception? Yes, by interacting. *Conscious. Cogn* . 18, 535–542.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

de Jaegher, H., Di Paolo, E., and Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends Cogn. Sci* . 14, 441–447.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

de Jaegher, H., and Froese, T. (2009). On the role of social interaction in individual agency. *Adapt. Behav* . 17, 444–460.

Falck-Ytter, T., Gredebäck, G., and von Hofsten, C. (2006). Infants predict other people's action goals. *Nat. Neurosci* . 9, 878–879.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Flanagan, J. R., and Johansson, R. S. (2003). Action plans used in action observation. *Nature* 424, 769–771.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Fodor, J. (1992). A theory of the child's theory of mind. *Cognition* 44, 283–296.

Pubmed Abstract | Pubmed Full Text

Frith, U., and Frith, C. D. (2003). " Development and neurophysiology of mentalising," in *Philosophical Transactions, Series B* , 58, eds C. D. Frith and D. Wolpert (Special issue on Mechanisms of social interaction), 459–473.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Frith, U., and Frith, C. (2011). The social brain: allowing humans to boldly go where no other species has been. *Philos. Trans. R. Soc. Lond. B Biol. Sci* . 365, 165–176.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Fuchs, T., and de Jaegher, H. (2009). Enactive intersubjectivity: participatory sense-making and mutual incorporation. *Phenomenol. Cogn. Sci* . 8, 465–486.

Gallagher, S. (2001). The practice of mind: theory, simulation or primary interaction? *J. Conscious. Stud* . 8, 83–108.

Gallagher, S. (2004). Understanding interpersonal problems in autism: interaction theory as an alternative to theory of mind. *Philos. Psychiatry Psychol* . 11, 199–217.

Gallagher, S. (2005). *How the Body Shapes the Mind* . New York, NY: Oxford University Press.

Gallagher, S. (2007). " Logical and phenomenological arguments against simulation theory," in *Folk Psychology Re-assessed* , eds D. Hutto and M. Ratcliffe (Dordrecht: Springer Publishers), 63–78.

Gallagher, S. (2008). Direct perception in the intersubjective context. *Conscious. Cogn* . 17, 535–543.

[Pubmed Abstract](#) | [Pubmed Full Text](#) | [CrossRef Full Text](#)

Gallagher, S. (2011). Strong interaction and self-agency. *Hum. Mente* 15, 55–76.

Gallagher, S., and Cole, J. (1995). Body image and body schema in a deafferented subject. *J. Mind Behav* . 16, 369–390.

Gallagher, S., and Zahavi, D. (2008). *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Sciences* . London/New York: Routledge.

Gallese, V. (2001). The ' shared manifold' hypothesis: from mirror neurons to empathy. *J. Conscious. Stud* . 8, 33–50.

Gallese, V., and Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends Cogn. Sci* . 15, 512–519.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Gardner, R. M., and Moncrieff, C. (1988). Body image distortion in anorexics as a non-sensory phenomenon: a signal detection approach. *J. Clin. Psychol* . 44, 101–107.

Pubmed Abstract | Pubmed Full Text

Gerson, S., and Woodward, A. L. (2010). " Building intentional action knowledge with one's hands," in *Neo-constructivism* , ed S. P. Johnson (New York, NY: Oxford University Press), 295–313.

Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading* . New York, NY: Oxford University Press.

Goldman, A. I. (2009). Mirroring, simulating and mindreading. *Mind Lang* . 24, 235–252.

Gopnik, A., and Meltzoff, A. N. (1997). *Words, Thoughts, and Theories* . Cambridge, MA: MIT Press.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Goodale, M. A., Milner, A. D., Jakobson, L. S., and Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature* 349, 154–156.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Grezes, J., Armony, J. L., Rowe, J., and Passingham, R. E. (2003). Activations related to ' mirror' and ' canonical' neurones in the human brain: an fMRI study. *Neuroimage* 18, 928–937.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Heyes, C. (2010). Mesmerising mirror neurons. *Neuroimage* 51, 789–791.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Himmelbach, M., and Karnath, H. O. (2005). Dorsal and ventral stream interaction: contributions from optic ataxia. *J. Cogn. Neurosci* . 17, 632–640.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Hobson, R. P. (2002). *The Cradle of Thought* . London: Macmillan.

Hommel, B. (1996). The cognitive representation of action: automatic integration of perceived action effects. *Psychol. Res* . 59, 176–186.

Pubmed Abstract | Pubmed Full Text

Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci* . 24, 849–937.

Pubmed Abstract | Pubmed Full Text

Huberle, E., and Karnath, H. O. (2011). The role of temporo-parietal junction (TPJ) in global Gestalt perception. *Brain Struct. Funct* . 214, 411–417.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Hurley, S. L. (2008). The shared circuits model (SCM): how control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behav. Brain Sci* . 31, 1–58.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Hutto, D. (2011). " Elementary mind minding, enactivist-style," in *Joint Attention: New Developments in Philosophy, Psychology, and Neuroscience* , ed A. Seemann (Cambridge, MA: MIT Press), 307–341.

Iacoboni, M. (2003). " Understanding intentions through imitation," in *Taking Action: Cognitive Neuroscience Perspective on Intentional Acts* , ed S. H. Johnson (Cambridge, MA: MIT Press), 107–138.

Iacoboni, M. (2005). Neural mechanisms of imitation. *Curr. Opin. Neurobiol* . 15, 632–637.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Iacoboni, M., and Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci* . 7, 942–951.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science* 286, 2526–2528.

Pubmed Abstract | Pubmed Full Text

James, W. (1890/1981). *The Principles of Psychology* . Cambridge, MA: Macmillan/Harvard University Press. (Original work published 1890).

Jordan, J. S., and Knoblich, G. (2004). Spatial perception and control. *Psychon. Bull. Rev* . 11, 54–59.

Pubmed Abstract | Pubmed Full Text

Kanakogi, and Itakura. (2010). The link between perception and action in early infancy: from the viewpoint of the direct-matching hypothesis. *Jpn. Psychol. Res* . 52, 121–131.

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

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Kilner, J. M., Vargas, C., Duval, S., Blakemore, S.-J., and Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nat. Neurosci* . 7, 1299–1301.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Knoblich, G., and Jordan, J. S. (2003). Action coordination in groups and individuals: learning anticipatory control. *J. Exp. Psychol. Learn. Mem. Cogn* . 29, 1006–1016.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Koski, L., Wohlschlager, A., Bekkering, H., Woods, R. P., and Dubeau, M. C. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cereb. Cortex* 12, 847–855.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Laureys, S., Perrin, F., Schnakers, C., Boly, M., and Majerus, S. (2005). Residual cognitive function in comatose, vegetative and minimally conscious states. *Curr. Opin. Neurol* . 18, 726–733.

Pubmed Abstract | Pubmed Full Text

Lewis, D. (1969). *Convention: A Philosophical Study* . Cambridge, MA: Harvard University Press.

Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annu. Rev. Psychol* . 62, 103–134.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

McCrae, R. R., and Costa, P. C. Jr. (1987). Validation of the five-factor model across instruments and observers. *J. Pers. Soc. Psychol* . 52, 81–90.

Pubmed Abstract | Pubmed Full Text

Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Dev. Psychol* . 31, 838–850.

Meltzoff, A. N. (2006). The " like me" framework for recognizing and becoming an intentional agent. *Acta Psychol. (Amst.)* 124, 26–43.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Milner, A. D., and Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia* 46, 774–785.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Meyer, M., Hunnius, S., van Elk, M., van Eede, F., and Bekkering, H. (2011). Joint action modulates motor system involvement during action observation in 3-year-olds. *Exp. Brain Res* . 211, 581–591.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Newman-Norlund, R. D., Noordzij, M. L., Meulenbroek, R. G. J., and Bekkering, H. (2007). Exploring the brain basis of joint attention: co-ordination of actions, goals and intentions. *Soc. Neurosci* . 2, 48–65.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Nichols, S., and Stich, S. (2003). *Mindreading. An integrated Account of Pretence, Self-Awareness, and Understanding of Other Minds* . Oxford: Clarendon Press.

Noë, A. (2004). *Action in Perception* . Cambridge, MA: MIT Press.

Onishi, K. H., and Baillargeon, R. (2005). Do 15-month old infants understand false beliefs? *Science* 308, 255–258.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

O'Regan, J. K., and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci* . 24, 939–973; discussion 973–1031.

Pubmed Abstract | Pubmed Full Text

Phillips, A. T., Wellman, H. M., and Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition* 85, 53–78.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Pisella, L., Binkofski, F., Lasek, K., Toni, I., and Rossetti, Y. (2006). No doubledissociation between optic ataxia and visual agnosia: multiple sub-streams for multiple visuo-manual integrations. *Neuropsychologia* 44, 2734–2748.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Powers, P. S., Schulman, R. G., Gleghorn, A. A., and Prange, M. E. (1987). Perceptual and cognitive abnormalities in bulimia. *Am. J. Psychiatry* 144, 1456–1460.

Pubmed Abstract | Pubmed Full Text

Pulvermüller, F., and Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci* . 11, 351–360.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Rakoczy, H., Warneken, F., and Tomasello, M. (2007). " This way!", " No! That way!" – 3-year-olds know that two people can have mutually incompatible desires. *Cogn. Dev* . 22, 47–68.

Ramnani, N., and Miall, R. C. (2004). A system in the human brain for predicting the actions of others. *Nat. Neurosci* . 7, 85–90.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Ratcliffe, M. (2007). *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation* . Basingstoke: Palgrave Macmillan.

Reddy, V. (2008). *How Infants Know Minds* . Cambridge, MA: Harvard University Press.

Richardson, M. J., Marsh, K. L., and Baron, R. M. (2007). Judging and actualizing intrapersonal and interpersonal affordances. *J. Exp. Psychol. Hum. Percept. Perform* . 33, 845–859.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Rizzolati, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci* . 27, 169–192.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Rizzolatti, G., and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci* . 11, 264–274.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Rueschemeyer, S.-A., Brass, M., and Friederici, A. D. (2007). Comprehending prehending: neural correlates of processing verbs with motor stems. *J. Cogn. Neurosci* . 19, 855–865.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Schenk, T. (2012). No dissociation between perception and action in patient DF when haptic feedback is withdrawn. *J. Neurosci* . 32, 2013–2017.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Schenk, T., and McIntosh, R. (2010). Do we have independent visual streams for perception and action? *Cogn. Neurosci.* 1, 52–62.

Schilbach, L., Eickhoff, S. B., Cieslik, E., Shah, N. J., Fink, G. R., and Vogeley, K. (2010). Eyes on me: an fMRI study of the effects of social gaze on action control. *Soc. Cogn. Affect. Neurosci* . 6, 393–403.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Schilbach, L., Eickhoff, S. B., Mojzisch, A., and Vogeley, K. (2008). What's in a smile? Neural correlates of facial embodiment during social interaction. *Soc. Neurosci* . 3, 37–50.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Schlicht, T., and Pompe, U. (2007). Review of Alva Noë action in perception. *Z. Philos. Forsch* . 61, 250–254.

Schutz-Bosbach, S., Avenanti, A., Aglioti, S. M., and Haggard, P. (2009). Don't do it! Inhibition and self-attribution during action observation. *J. Cogn. Neurosci. 21* , 1215–1227.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint actions: bodies and minds moving together. *Trends Cogn. Sci* . 10, 70–76.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Sebanz, N., and Frith, C. (2004). Beyond simulation? Neural mechanisms for predicting the actions of others. *Nat. Neurosci* . 7, 5–6.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Singer, T. (2012). The past, present and future of social neuroscience: a European perspective. *Neuroimage* 61, 437–449.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Skewes, J., Roepstorff, A., and Frith, C. (2011). How do illusions constrain goal-directed movement: perceptual and visuomotor influences on speed/accuracy trade-off. *Exp. Brain Res* . 209, 247–255.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Smeets, J. B. J., and Brenner, E. (2006). 10 Years of Illusions. *J. Exp. Psychol. Hum. Percept. Perform* . 32, 1501–1504.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Sodian, B., and Thoermer, C. (2004). Infants' understanding of looking, pointing, and reaching as cues to intentional action. *J. Cogn. Dev* . 5, 289–316.

Southgate, V., Senju, A., and Csibra, G. (2007). Action anticipation through attribution of false belief by two-year-olds. *Psychol. Sci* . 18, 587–592.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Spaulding, S. (2010). Embodied cognition and mindreading. *Mind Lang* . 25, 119–140.

Spelke, E. S., Phillips, A. T., and Woodward, A. L. (1995). " Infants' knowledge of object motion and human action," in *Causal Cognition: A Multidisciplinary Debate* , eds A. J. Premack, D. Premack, and D. Sperber (Oxford: Clarendon Press), 44–77.

Sperber, D., and Wilson, D. (1995). *Relevance: Communication and Cognition* . Oxford: Blackwell.

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

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Valyear, K. F., Culham, J. C., Sharif, N., Westwood, D., and Goodale, M. A. (2006). A double dissociation between sensitivity to changes in object identity and object orientation in the ventral and dorsal visual streams: a human fMRI study. *Neuropsychologia* 44, 218–228.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

van Elk, M., van Schie, H. T., Neggers, S. F., and Bekkering, H. (2010). Neural and temporal dynamics underlying visual selection for action. *J. Neurophysiol* . 104, 972–983.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

van Schie, H. T., Mars, R. B., Coles, M. G., and Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nat. Neurosci* . 7, 549–554.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron* 40, 655–664.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

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

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Wheeler, M. (2006). *Reconstructing the Cognitive World: The Next Step* . Cambridge, MA: The MIT Press.

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1–34.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Dev. Sci* . 6, 297–311.

Woodward, A. L. (2005). " The infant origins of intentional understanding," in *Advances in Child Development and Behavior* , ed R. V. Kail (Oxford: Elsevier), 229–262.

Pubmed Abstract | Pubmed Full Text

Woodward, A. L., and Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cogn. Dev* . 17, 1061–1084.