



Illuminating the dark matter of social neuroscience: Considering the problem of social interaction from philosophical, psychological, and neuroscientific perspectives

Marisa Przyrembel^{1,2}, Jonathan Smallwood³, Michael Pauen¹ and Tania Singer^{3*}

¹ Berlin School of Mind and Brain, Institute for Philosophy, Humboldt-Universität zu Berlin, Berlin, Germany

² Center for Subjectivity Research, University of Copenhagen, Copenhagen, Denmark

³ Department of Social Neuroscience, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Edited by:

Leonhard Schilbach,
Max-Planck-Institute for Neurological
Research, Germany

Reviewed by:

Cristina Becchio, Università degli
Studi di Torino, Italy
Riitta Hari, Helsinki University of
Technology, Finland
Shaun Gallagher, University of
Memphis, USA

*Correspondence:

Tania Singer, Department of Social
Neuroscience, Max Planck Institute
for Human and Cognitive Brain
Sciences, Stephanstraße 1A, Leipzig
04103, Germany.
e-mail: singer@cbs.mpg.de

Successful human social interaction depends on our capacity to understand other people's mental states and to anticipate how they will react to our actions. Despite its importance to the human condition, the exact mechanisms underlying our ability to understand another's actions, feelings, and thoughts are still a matter of conjecture. Here, we consider this problem from philosophical, psychological, and neuroscientific perspectives. In a critical review, we demonstrate that attempts to draw parallels across these complementary disciplines is premature: The second-person perspective does not map directly to Interaction or Simulation theories, online social cognition, or shared neural network accounts underlying action observation or empathy. Nor does the third-person perspective map onto Theory-Theory (TT), offline social cognition, or the neural networks that support Theory of Mind (ToM). Moreover, we argue that important qualities of social interaction emerge through the reciprocal interplay of two independent agents whose unpredictable behavior requires that models of their partner's internal state be continually updated. This analysis draws attention to the need for paradigms in social neuroscience that allow two individuals to interact in a spontaneous and natural manner and to adapt their behavior and cognitions in a response contingent fashion due to the inherent unpredictability in another person's behavior. Even if such paradigms were implemented, it is possible that the specific neural correlates supporting such reciprocal interaction would not reflect computation unique to social interaction but rather the use of basic cognitive and emotional processes combined in a unique manner. Finally, we argue that given the crucial role of social interaction in human evolution, ontogeny, and every-day social life, a more theoretically and methodologically nuanced approach to the study of real social interaction will nevertheless help the field of social cognition to evolve.

Keywords: mentalizing, online/offline social cognition, second-person perspective, simulation, social interaction, social neuroscience, stimulus independent thoughts, theory-theory

INTRODUCTION

Whether searching for a cure for life-threatening disease, developing a hybrid engine that reduces carbon emission, or simply enjoying a barbecue in the park in the company of good friends, understanding the desires, beliefs, and intentions of other people is essential for almost every human endeavor. Despite the relative ease with which we interact with others, philosophers, psychologists, and, most recently, neuroscientists have puzzled over how exactly we gain sufficient access to the content of another's mind. Unlike other forms of mental content, such as the perception of objects, we cannot directly experience what is on the minds of others; most likely a process of social evolution or social learning is responsible for our species expertise at simulating or predicting how others will act, feel, or think. What remains less clear is

precisely how humans perform the feat of mental dexterity known as mental state attribution.

One possibility is that social interactions involve qualitatively different processes than those perceptual, cognitive and motor computations that subservise the processing of information about the objective physical reality (Adolphs, 2010). Such a modular view of the mind (Fodor, 1983) would postulate that in the case of mental state attribution in social contexts, specific forms of knowledge and particular brain modules exist that serve these explicit social information processing functions. In cognitive neuroscience, it has been debated whether certain brain regions respond selectively to social stimuli, as in the case of faces, the fusiform-face area (FFA) (Kanwisher et al., 1997; Kanwisher, 2000) or in the case of body parts, the extrastriate body area

(EBA; Downing et al., 2001). Similarly, it has been argued that the right temporal parietal junction (rTPJ) performs computations specifically related to the mental states of other individuals (Saxe and Kanwisher, 2003). Such domain-specific views of the mind and brain has been criticised by advocates of a more distributed theory of mental processes in which information coding and processing emerges from the brain's dynamic and distributed organization in space and time (e.g., Haxby et al., 2000; Mitchell, 2008 in the field of social neuroscience).

In this paper, we evaluate the evidence for exclusive social information processing by critically reviewing the philosophical, psychological, and neuroscientific evidence regarding how we understand other people's thoughts, feelings, and actions. More specifically, we first demonstrate that the philosophical concept of second-person perspective taking can be used to point out specific features of social cognition, as this notion of second-person perspective taking is distinct from both the acquisition of objective knowledge about the world and subjective knowledge about the individual. We will then consider simulation, interactive, narrative, and theory-theory (TT) accounts of mind reading and summarize neuroscientific findings suggesting different neural networks underlying the ability to mentalize, empathize, and understand the actions of others. Although superficial similarities exist between the different modes of analysis, significant problems emerge when the philosophical, psychological, and neuroscientific findings are simply mapped onto to each; we suggest that attempts to unify these different levels of social cognition may be premature. The last section (1) outlines different types of models of social interaction that would be necessary to shed light on the mechanisms of social interaction, (2) summarizes the neuroscientific studies to date which have focused on social interaction, and (3) discusses whether these findings can really shed light on the "dark matter" of social neuroscience or whether new paradigms are necessary to fill this gap. Here, we borrow from the field of physical astronomy where "dark matter" is a term for matter that cannot be directly detected via the existing scientific instruments. Astrophysicists assume that this intangible matter constitutes a stupefying 73% minimum of the total matter in the universe (Lahanas and Nanopoulos, 2003)¹. The physical rules and elements of our universe are thus largely unexplored—might this also hold for the neuroscientific investigation of social interaction?

FIRST, SECOND, AND THIRD PERSON PERSPECTIVES

Psychologists and philosophers differ in their approaches, even when they deal with the same phenomena. Broadly speaking, psychologists usually focus on behavioral differences and the underlying cognitive and emotional mechanisms, while philosophers tend to concentrate on conceptual and normative issues. When it comes to social cognition, the picture is sometimes a bit more complicated, partly due to an intensive cooperation

between philosophers and psychologists. Still, psychologists typically investigate the relevant mechanisms with experimental methodologies, while philosophers try to find out, for example, whether social cognition creates a specific sort of knowledge—a second-person perspective that is systematically different from the third-person perspective of the world and our own first-person knowledge about ourselves.

The idea that an intersubjective epistemic perspective has to be added to the subjective first-person and the objective third-person perspective is by no means new. The basic idea can already be traced back to the beginning of the last century in the work of Heidegger (1927/1975, 1927/1986) and Mead (1925, 1926, 1934). More recent versions that use the metaphor of a "second-person perspective" can be found in Varela and Shear (1999); Bohman (2000); Davidson (2004); Habermas (2004); and Reddy (2008), for an overview see Lindemann (2006).

The idea has gained momentum in recent years with the advance of social neuroscience. Research in this area has resulted in an increasing demand for a clarification as to what kind of knowledge understanding other minds is, how this knowledge is acquired, and whether or not it can be separated from other kinds of knowledge acquisition.

This requires a closer investigation of how the notion of the "second-person" is currently utilized in the literature. Varela and Shear (1999) as well as Petitmengin (2006) introduced the term for an interview method by which subjective first-person experiences are gathered as "data" with the help of another "second" person. This approach, especially its underlying idea of the physical presence of a second-person, may have influenced other fields like developmental psychology (Reddy, 2008) and neuroscience (Schilbach, 2010; Schilbach et al., 2010a; Wilms et al., 2010). Still, there is unclear usage of a number of similar terms. For example, the confusing use of similar terms such as "second-person account," "second-person engagements," "second-person experiences" (Schilbach et al., in press), and "second-person perspective" demonstrate a lack of a common language. Here we will focus on the concept of second-person perspective, as it seems to be essential in social cognition.

One crucial question in defining second-person perspective is whether social interaction with a verbatim second-person plays a decisive role. This seems, for example, to be Schilbach's (2010) view. He postulates, "*social cognition is fundamentally different when an individual is actively and directly interacting with others. In such cases, an individual adopts a "second-person perspective" in which interaction with the other can be thought of as essential or even constitutive for social cognition, rather than merely observing others and relying on a "first- (or third-) person grasp" of their mental states.*" (p. 1).

Wilms et al. (2010) use the term "online interaction" instead of "online social cognition." Assumingly, what they wish to express is social cognition which is in place during real-life interaction. Moreover, they use this as a synonym for second-person perspective taking: "*'Online' interaction crucially involves [...] establishing reciprocal relations where actions feed directly into the communication loop [...]. This has been referred to as adopting a "second-person-perspective" [...] which can be taken to suggest that awareness of mental states results from being psychologically engaged with someone and being an active participant of reciprocal*

¹We thank one of the reviewers for pointing toward a paper with a similar title (Zhang and Raichle, 2010). The authors refer to the same equation (75% of the matter in the universe counts as dark energy, and we quote Lahanas and Nanopoulos, 2003, saying that 73% does). However, Zhang and Raichle (2010) use the term "dark energy" in the context of brain metabolism and not like us to refer to social interaction.

interaction thereby establishing a subject-subject (“Me-You”) rather than a subject-object (“Me-She/He”) relationship.” (p. 1).

Others (De Jaegher et al., 2010) define social cognition without any reference to the second-person perspective. Although stressing the relevance of a comparatively strong version of interaction for the development of social cognition, De Jaegher et al. (2010) concede that social cognition may occur in the absence of interaction, e.g., in remote observation of social scenes.

With these observations in mind, we can start to characterize social cognition as the acquisition of knowledge² about other persons’ mental states, i.e., their beliefs, desires, and intentions and also insight about the meaning of their utterings. It would follow that social cognition includes at least two essential features that should be accounted for with any definition. First, social cognition is a means of *knowledge acquisition*. We suggest that this aspect can be specified by referring to the distinction between the second-person perspective, on the one hand, and the first- and third-person perspective on the other. Second, social cognition occurs in *social contexts*. One way to specify this aspect is to ask whether or not the subjects involved interact, i.e., whether they are engaged in online or offline social cognition. Taking interaction as a reciprocal pattern of action and reaction³ between at least two agents affecting each other, we assume that knowledge about other persons can be acquired without interacting with them, for example, when one reads a letter or watches a movie describing another person’s mental states. Consequently, we argue that second-person perspective taking can happen without direct interaction and that this perspective is, therefore, not synonymous with being engaged in interaction or online social cognition. Rather, treating interaction and perspective taking as two different aspects of social cognition results in a much more differentiated and suitable view.

Perspectives, in a nutshell, are ways of acquiring knowledge (for more details see Pauen, 2012). Perspectival distinctions answer questions like: (1) “What is this knowledge about?” and (2) “How do we acquire this knowledge?” First-person perspective taking provides self-knowledge. So, reflecting on the questions posed above, first-person knowledge (1) is about the subject’s own mental states and (2) is acquired directly via those very mental states that are directly accessible only for the subject him or herself. It can, thus be characterized as subjective because it is acquired by and is about the subject’s own mental states (Pauen, 2010). Third-person knowledge, by contrast, (1) is about all kinds of objective (and mostly external) facts, both scientific and non-scientific and (2) it is acquired by all kinds of objective evidence that is accessible to everyone, among them external observation and scientific methods. As a consequence, the third-person perspective can be characterized as “objective.”

²We do not only refer to “justified true beliefs,” but to a broader definition of knowledge.

³Normally, this includes several cycles of action and reaction, but it seems unreasonable to talk about interaction if there is not at least one such cycle by every agent involved. This is in line with Wilms et al. (2010) who state that “online” interaction [rather: online social cognition; authors’ note] crucially involves [...] establishing reciprocal relations where actions feed directly into the communication loop.”

But why is it necessary to add a second-person perspective to the first- and third-person perspective to begin with? In order to see this, imagine that you are locating a restaurant in an unfamiliar city. In this endeavor the third-person perspective is helpful because the eatery has a definite location in space that can be assessed by consulting a map. By contrast, you apply the first-person perspective when you wonder whether it is worth stopping for a pretzel to slake your hunger before completing the journey: Are you really *that* hungry? But when you reach your destination a quarter of an hour late because of this detour, and being full yourself, you need to find out how your companion feels. Is she still hungry? Is she angry because you are late? Or would she like to go to another place? In assessing our companion’s state the first-person perspective provides no information because, unless they also stopped for dinner, their mental state is different to our own. Likewise, the third-person perspective cannot be of assistance because there are no objective facts upon which to assess the person’s thoughts and feelings.

Thus, our capacity to infer our companions feelings is a paradigmatic case of social cognition which is set apart both from third- and first-person perspective taking by at least two distinctive features. First, unlike first-person perspective taking, it is not about one’s own mental states. Second, unlike the third-person perspective, it is not just about facts. Rather, social cognition is a question regarding another person’s *mental states*; i.e., it is about what our companion thinks, what she feels, and what her intentions are.

But how does social cognition relate to our capacity to acquire knowledge? Social cognition is neither about pure objective data as in third-person perspective taking, nor is it the application of our subjective mental states, as in first-person perspective taking. Instead, social cognition is a means of knowledge acquisition that involves a combination of both. Just as in first-person perspective taking, we draw on our own feelings and experiences during social cognition in order to access the *other* person’s feelings and experiences. Likewise, social cognition is like third-person perspective taking when we draw on our general background knowledge as well as on the person’s behavior, gestures, and facial play to understand why they are acting as they are. It is clear that knowledge that we gain by taking the second-person perspective is neither purely objective nor subjective; it is *inter-subjective* because it requires that we understand the other *as* a person with their own thoughts, feelings, and experiences⁴. In other words, the second-person perspective is set apart from the first- and the third-person perspective both in terms of its relation to (1) knowledge content and (2) knowledge acquisition (Pauen, 2012).

Note, first, that only first- and second-person perspective taking are restricted regarding their objects; third-person perspective taking is not. As a consequence, you can take the third-person perspective regarding your own or another person’s pain experience, for example by drawing on objective fMRI data or skin-conductance measures. Second, as already indicated above,

⁴So even if we draw on our own feelings and thoughts in order to understand another person’s feelings and thoughts, we have to understand that we are referring to the other person. Thus, empathizing would, but emotional contagion would not count as full-fledged second-person perspective taking.

the present notion of second-person perspective taking does not require interaction—even though interaction certainly plays an important role in the ontogenetic and phylogenetic development of social cognition in general and second-person perspective taking more specifically. Still, interaction is not an epistemic feature itself. That is why epistemic access might be completely identical, regardless of whether or not there is interaction. In order to see this, think about someone who tries to figure out whether another person is angry and does so by taking the other person's perspective. This can happen if one is interacting with someone who is (1) physically present, (2) the person can be seen in a movie, or (3) is a character in a novel. Epistemic access to the other person's thoughts and feelings might be identical in all three of these cases. What differs here are non-epistemic features; for example that the other person reacts in the first case but does not in the second and the third cases. Conversely, interactions can take place without second-person perspective taking, for example if the epistemic subject interacts with another person for whom only objective information is available. Given that second-person perspective taking (like first- and third-person perspective taking) is an *epistemic* feature, these differences do not matter for an assessment of the perspective that an individual adopts during the interaction, even if such issues are of great importance in other respects. For this reason, we suggest that differences regarding interaction should be denoted by the distinction between offline and online social cognition and not by perspectival distinctions. This is in line with similar considerations by Mead, Habermas, and Bohman who understand second-person perspective taking—either implicitly or explicitly—as a way of interpreting others, regardless of whether or not they are present.

SOCIAL COGNITION: THEORIES FROM PHILOSOPHY AND PSYCHOLOGY

Other theories discussed in philosophy, psychology, and recently also in neuroscience have rather focused on explaining the mechanisms underlying our ability to understand other minds, feelings, and actions. These theories try to find answers to the common questions: How can we tell what another person's mental states are? How can we predict and explain the behavior of others, i.e., what are the psychological processes that allow for mindreading? First, we discuss the debate of Simulation-Theory (ST) and Theory-Theory (TT) vs. interactive and narrative accounts and then we turn to clarifying different accounts of on- vs. offline social cognition.

SIMULATION-THEORY vs. THEORY-THEORY

Two prominent approaches to mindreading commonly described in the literature are TT and ST. TT and ST are not simply psychological theories, but are similarly rooted and largely debated in philosophy, as well as in neuroscience (e.g., see Gallese and Goldman, 1998; Keysers and Gazzola, 2007). According to TT (e.g., Sellars, 1956; Gopnik, 1993), the psychological process that enables us to understand others' minds consists of theorizing, as there is no direct access to the mental states of others. Instead, mental states of others are concealed entities, which, while unobservable, can be calculated implicitly or explicitly. If

shortly after meeting our friend in the restaurant we saw signs of uneasiness, we would be in a position to infer that our companion was still hungry. To do so, we draw on common sense knowledge about the signs of hungriness but, more importantly, on our knowledge about social norms, i.e., good manners. Here we might rely on a general rule or societal norm. In this case, it is that not being on time is impolite and thus causes disapproval.

According to ST (e.g., Goldman, 2005), mental state attribution is a process-driven rather than a theory-driven mechanism that allows us to understand other minds. We are able to understand others as we generate (or embody) states in ourselves that are similar to the other's mental states. We simulate what we would experience if we were the other person. Unlike the TT viewpoint, this process relies far less on explicit knowledge, and instead depends upon the capacity of the individual to put oneself in the other's mental shoes. In the example of the dinner date, ST would argue that we might find out about our friend's state directly based on imagining ourselves in our friend's situation. So, if our companion ate rapidly as soon as the waiter brought her food to the table, we could translate this non-verbal enthusiasm into a state of hunger in our companion.

INTERACTION THEORY

Phenomenologists recently introduced the Interaction-Theory (IT) as an alternative to ST and TT (see Gallagher, 2008). Following Husserl's and Scheler's tradition, IT postulates that most of the mental states of others are incorporated and visible in the "Leib," the "lived body." According to Gangopadhyay and Schilbach (2012), there is plenty of empirical evidence that experiencing others' mental states, i.e., having an immediate perceptual access to the perception of their embodied intentionality is possible due to the tight coupling of action and perception. Hence, the problem of understanding others minds depends neither on explicit theorizing nor simulation, but on direct interaction embedded in a concrete interpersonal realm. The mental states of others are not "hidden" per se and do not always have to be consciously inferred. The question is, however, how exactly do we perceive other minds in direct interaction? What does "direct" mean in the first place (see Zahavi, 2011, for more details)? Furthermore, understanding others feels qualitatively different than having an experience from the first-person perspective. But what the IT yields is a more contextual and embodied look at the problem of other minds.

THE NARRATIVE PRACTICE HYPOTHESIS

The Narrative Practice Hypothesis (NPH; Hutto, 2007) is yet another approach to social cognition. The NPH postulates that being told stories about others' mental states from an early age allows children to understand other persons' inner lives in particular contexts. There is a lot of empirical evidence for the linguistic and narrative competence in the development of a theory of mind (ToM) (Woolfe et al., 2002). However, because a basic understanding of implicit rules and theories is necessary for narrative comprehension and ToM, the NPH appears to be a legitimate refinement of TT rather than a novel approach to the understanding of other minds (Przyrembel, thesis in preparation). Therefore, being told stories certainly broadens the ability to understand

others, but it is not a completely theory-independent explanation for understanding other minds.

ONLINE vs. OFFLINE SOCIAL COGNITION

Most social neuroscience studies to date have focused on understanding the effects of socially relevant stimuli on the mind of an individual, i.e., an isolated understanding of our own thoughts and feelings. In contrast, the study of social interaction involves a bidirectional relation between two or more agents as well as the impact of the social context in which they emerge. It is concerned with understanding how two minds mutually shape each other through reciprocal interactions (see Frith, 2003; Singer et al., 2004c). An investigation of social interaction also needs to understand how we communicate thoughts and feelings to another mind to enable this person to build an appropriate representation of our thoughts and feelings that will ultimately be fed back to ensure there has not been any misunderstanding. In a keynote lecture Frith (2003) referred to such a mechanism underlying this kind of real-life social interaction as “Neural Hermeneutics.” Based on this view, it has recently been suggested that social cognition involves two distinct modes, which are also known as the “offline” or “online mode”; whereas the former refers to agents passively viewing another agent during social interaction, the latter refers to an reciprocal interaction in which two or more agents are involved in real-life social engagement and in which the behavior of one leads to a change in another person’s behavior (Schilbach et al., 2006; Wilms et al., 2010).

Note however, that in current papers, we find a quite heterogeneous usage of the terms. Mojzisch et al. (2006, see p. 185) as well as Schilbach et al. (2006, see p. 718) speak of on- vs. offline Theory of Mind (ToM). Wilms et al. (2010) refer sometimes to on- vs. offline mentalizing (p. 1), and sometimes to on- vs. offline social cognition (p. 8). While social cognition is, however, generally used as an umbrella term for all socially relevant processes and thus includes also action intention understanding, affective resonance and empathy, face recognition, social memory, and many others, mentalizing is usually reserved to specifically denote cognitive perspective taking processes and the underlying ToM network. Therefore, we prefer to use the term of on- or offline social cognition in the context of the present paper.

It is furthermore worth noting that in these papers, the terms online/offline with respect to social cognition are used in a way that is contrary to the way that these processes are generally understood and discussed in cognitive neuroscience. According to this view, states of offline, or decoupled cognition, tend to emerge in situations in which the mind generates streams of thoughts that have minimal direct correlation to ongoing perceptual events and are often defined as stimulus independent thoughts (SIT). These SIT can also subserve either inferences about other people’s minds, or, alternatively, reasoning about the self and the world (Smallwood et al., 2008, 2011; Barron et al., 2011; Kam et al., 2011). The offline mode of social cognition proposed by Schilbach and colleagues, in contrast, does not refer to SIT, as the subjects in the scanner do receive social stimuli from direct online perception; these subjects are simply not addressed by these stimuli or engaged in the social encounter, and this is why this kind of social cognition is called offline social cognition.

FIRST, SECOND, AND THIRD PERSON PERSPECTIVES vs. OFFLINE AND ONLINE SOCIAL COGNITION

Now what about the relation between perspectival distinctions on the one hand and on- and offline social cognition on the other? To evaluate these definitions of online and offline social cognition, it is necessary to examine how they compare to the philosophical definitions of different perspectives of knowledge that are involved in any attempt to understand another mind. While interaction is essential for the difference between off- and online social cognition, it does not play an important role in second-person perspective taking. Given that perspectives are means of epistemic access, it should be epistemic features that are decisive for perspectival distinctions. Even if interaction plays an important role in many cases of second-person perspective taking, as well as in the ontogenetic and phylogenetic development of social skills, it is not an epistemic feature itself. That is why epistemic access or, more specifically, evidence (one’s own mental states, social norms) and type of knowledge (another person’s mental states) set the second-person perspective apart from first- and third-person perspectives. On the other hand, it is interaction rather than epistemic access that makes the difference between on- and offline social cognition. So, even though interactions certainly have important neurobiological effects, this does not constitute evidence of a unique epistemic perspective. One of the reasons why this distinction is important is that epistemic access to the mental states of the other person might be completely identical, regardless of whether or not there is interaction.

NEUROSCIENTIFIC EVIDENCE FOR MULTIPLE SYSTEMS UNDERLYING SOCIAL COGNITION

In the last decades, social neuroscience has made progress in refining models of social cognition. These studies have revealed that there are several neural routes to the understanding of another person’s actions, feelings, and thoughts. Three major routes have reliably been identified as being crucial for our ability to understand others, namely (1) motor actions and motor intentions—the so-called mirror neuron system (MNS), (2) beliefs, desires, and thoughts—the so-called ToM or mentalizing system, and (3) emotional and bodily states—relating to our ability to empathize with others.

Each of these abilities is associated with different brain circuits. Early research on the discovery of the mirror neurons in macaque monkeys (di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 1996; for a review see Rizzolatti and Fabbri-Destro, 2010) suggested that the same cells, which are activated when a monkey is performing a particular grasp, also fire if the same monkey merely observes another during the same action. Later research in humans, mostly using fMRI, demonstrated “shared networks” between self-performed and vicariously perceived actions activate similar regions in the human brain. The identified neural network comprised the inferior parietal lobule (IPL), the ventral premotor cortex, and the caudal part of the inferior frontal gyrus (IFG) (Dinstein et al., 2007; Gazzola et al., 2007; Etzel et al., 2008, for reviews see Blakemore and Decety, 2001; Grèzes and Decety, 2001).

This research was subsequently expanded to the domain of emotions and empathy (Gallese and Goldman, 1998; Gallese,

2001; Preston and deWaal, 2002) culminating in the emergence of empathy research in the field of social neuroscience (for reviews see Decety and Jackson, 2004; de Vignemont and Singer, 2006; Keysers and Gazzola, 2006, 2009; Singer and Lamm, 2009). A multitude of imaging experiments in humans in the domain of empathy for pain, disgust, taste, and touch revealed that, in contrast to mirror neuron networks, in the domain of motor actions, sharing sensations and feelings with others engages somatosensory cortices, as well as anterior parts of the insula and anterior cingulate cortex (ACC; Keysers et al., 2010; Fan et al., 2011; Lamm et al., 2011). In addition to this affective route, researchers have distinguished a cognitive route which is helpful for understanding the beliefs, thoughts, and desires of other people. This “mentalizing,” “ToM,” or “cognitive perspective taking” (Premack and Woodruff, 1978; Wimmer and Perner, 1983; Frith and Frith, 1999, 2003; Baron-Cohen et al., 2000) network typically comprises areas in the medial prefrontal cortex (mPFC), precuneus, superior temporal sulcus (STS), and rTPJ (for reviews see Frith and Frith, 1999, 2003; Gallagher and Frith, 2003; Saxe et al., 2004; Amodio and Frith, 2006; Saxe, 2006; Mitchell, 2009).

Many of these neural systems are also recruited when individuals are not interacting in a social context. For example, the mPFC is activated by tasks involving the evaluation of the personality of the self and others (Kelley et al., 2002; Macrae et al., 2004; Mitchell et al., 2006), as well as by the task of assessing the likelihood of enjoyment of activities that will occur in the future (Tamir and Mitchell, 2011). Likewise, regions including the TPJ, the mPFC, and the PCC are recruited when thinking becomes decoupled from the events in the here and now and stimulus independent mental contents form the cornerstone for consciousness (Mason et al., 2007; Christoff et al., 2009; Smallwood et al., 2011; Stawarzyck et al., 2011). Taken together, the fact that similar neural processes are engaged during self-referential processes and social interactions, as well as internally generated thoughts with no explicit external referent, suggests that many forms of social cognition are likely to be involved in a more general set of processes that allow the mind to devote processing resources to make predictions necessary for navigation through social life (Frith, 2007).

Finally, an enormous amount of work has been performed in the domain of the neural networks underlying the recognition of facial emotional and non-emotional expressions. For this social cognitive ability, brain regions such as the amygdala, secondary somatosensory cortices, and FFA seem to be particularly relevant (Adolphs, 2002; Kanwisher and Yovel, 2006).

INTEGRATION OF THESE DIVERSE APPROACHES

As is clear from the review of philosophical, psychological, and neuroscientific approaches, the problem of understanding others minds has been addressed by different disciplines using several different methods. There seems to be an implicit assumption among many scholars studying social cognition that despite differences in these approaches, there is a close link between the third-person perspective, TT, and offline social cognition on the one hand and the second-person perspective, IT, ST, and the online mode, including empathy on the other. Moreover, it is often argued that these different modes depend on different

neural networks. The next section of this paper critically reviews the extent to which such a dichotomous view of social interaction across the different domains of philosophy, psychology, and neuroscience is realistic.

ST AND TT vs. FIRST, SECOND, AND THIRD PERSON PERSPECTIVE

As we have already argued, a theory of second-person perspective taking should describe a specific sort of knowledge. More precisely, it must specify the object of this knowledge (another person's mental states) and the relevant evidence (one's own experiences, feelings, social norms, etc.), and finally it tells us something about the relational status of this knowledge (inter-subjective). Again, these specifications do not encompass the underlying psychological or neurobiological mechanisms. They only specify the criteria any mechanism needs to meet in order to realize second-person perspective taking.

For the same reason, second-person perspective taking is not just another word for simulation: It describes an epistemic position rather than a psychological process. Simulation is *one* implementation of second-person perspective taking; however, second-person perspective taking may include automatic and subpersonal replications of another person's mental state or explicit logical theorizing regarding what they must be thinking or feeling. The latter might be involved when we try to account for perspectival differences between our own point of view and the perspective of the person that we are trying to understand, particular if we are not familiar with these differences. Openness with respect to the psychological and neurobiological mechanisms is of special importance. As we will argue below, second-person perspective taking might be realized by a multitude of psychological processes beyond simulation and theorizing, as presented in the NPH (Hutto, 2007) or the IT (Gallagher, 2008). There is no reason to accept the dichotomous view outlined above. Instead, we endorse Hybrid-Theories that incorporate elements of ST and TT, as well as IT and NPH. Still, this openness has its limits. Mere theorizing, combined with external observation of a person's behavior, does not constitute second-person perspective taking; it is third-person perspective taking, because one's own mental states are not accessed in order to understand someone else's beliefs, desires, or feelings.

ST AND TT vs. NEUROSCIENCE

As mentioned above, scholars from different fields have recently argued for a direct mapping between certain models, terms, and theories in philosophy, developmental psychology, and social neuroscience. Thus, it has been suggested that mirror neuron networks and “shared networks” underlie empathic understanding and can be taken as evidence for ST or IT. TT approaches, by contrast, are mapped to ToM processes and their underlying neural networks.

There are several problems to such an approach. First, it is questionable how ST, IT, NPH, and TT accounts could ever be translated into the language of neural processes and the brain. What in terms of neuronal computations would simulation or using a theory about the world actually mean? The difficulty of mapping high-level constructs like these on brain organization and functions is quite evident here. Many cognitive neuroscientists prefer to take a more cautious approach and refer to “mirror

neuron cells”, which have the property of processing one’s own and others observed movements if there is access to single cell recording. Alternatively, when referring to fMRI studies, so-called “shared brain networks” are assumed to underlie the representation of emotions or actions in first-person and second-person experience only when similar brain regions respond under both conditions. Although the notion of mirror neurons and shared network share many things in common at a gross theoretical level, neuroscientists maintain the awareness that functional imaging techniques reflect the activation of large assemblies of cells, while single cell recording reveal the computations performed by a single neuron. That is why these activations neither allow any inference about the properties of the single cells nor about the real computations subserved by these networks (e.g., Singer and Lamm, 2009). In line with substantial evidence for predictive coding in the human brain (e.g., Schultz et al., 1997; Seymour et al., 2004; O’Doherty et al., 2006; Frith, 2007), it is more likely that, for example, activation in anterior insula when empathizing with the pain of others rather reflects predictive models about the potential negative effects of pain (Singer et al., 2009; Lamm and Singer, 2010), which are also activated when we anticipate the effects of impending pain in ourselves (Ploghaus et al., 1999).

Accordingly, rather than using the term simulation, it would probably be closer to the biological reality of the brain to use terms such as vicarious prediction or the activation of cortical representations that have been generated through the performance and experience of similar movement or affective experience in the self.

Second, even if we set these problems aside, recent findings suggest that mapping TT to mentalizing of cognitive processes and ST or IT to empathy or action understanding is incorrect. Jason Mitchell, for example, has presented neuroscientific evidence that we “simulate” others, even when we are in the domain of mentalizing about cognitive states or abstract knowledge, like political attitudes. This follows from a series of ToM or mentalizing studies, suggesting an important role of mPFC when reflecting on one’s own as well as other peoples’ mental states (Mitchell et al., 2005). These studies also demonstrated functional differences between judging the mental states of similar and dissimilar others, with the former activating parts of the ventral mPFC and the latter dorsal parts of the prefrontal cortex (Mitchell et al., 2006). Furthermore, Waytz and Mitchell (2011) stated that simulation consists not only of mirroring (“a vicarious response in which a perceiver experiences the same current mental state as that of another person,” p. 197), but also of self-projection (i.e., “imagining oneself in the same situation as another person, predicting one’s thoughts and feelings in that hypothetical scenario and assuming that the other would think and feel the same way,” p. 197). The latter again involves the mPFC. This suggest that even in the domain of tasks that may seem as if these require an outright rule-governed, intellectual stance, we apparently use cortical representations underlying the inference of such attributes for ourselves to derive knowledge about the other—a process which would map to ST rather than to TT.

On the other hand, empathy research has clearly shown that when we empathize we only activate parts of the entire neural networks elicited when experiencing a certain emotion in

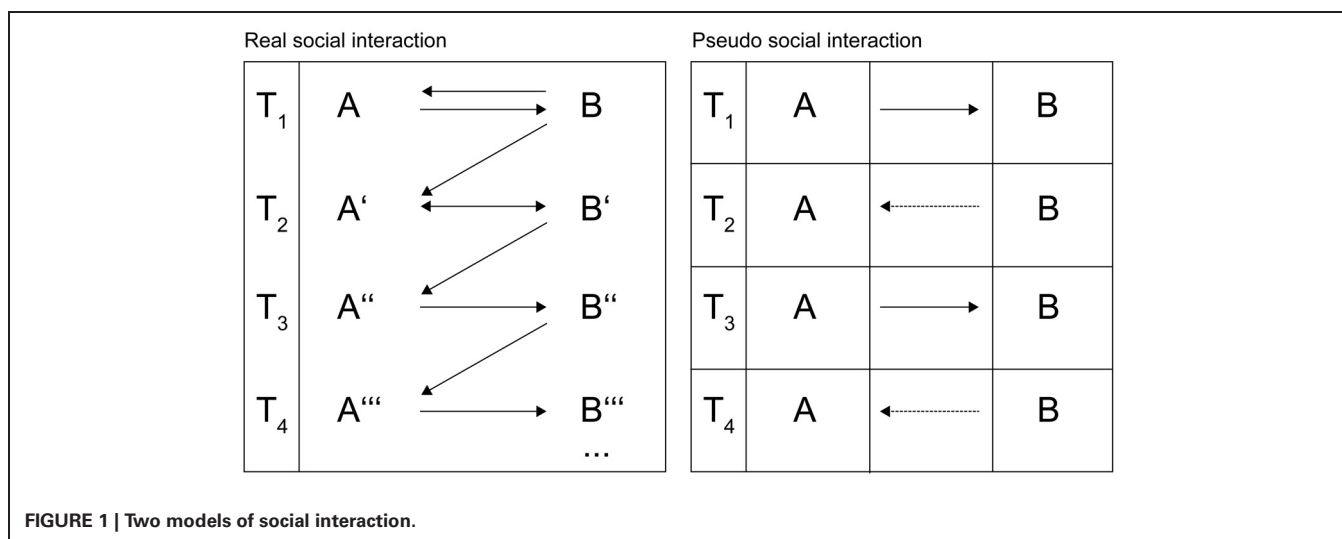
ourselves. As these representations in the anterior insula are also observed in empathy for other unpleasant experiences, such as disgust (Wicker et al., 2003) or obnoxious tastes (Jabbi et al., 2007), and are modulated by contextual factors as well as person-specific factors (for an overview see e.g., Hein and Singer, 2008), it has been suggested that these activities stand for higher-level representations of subjective feelings that have already integrated both contextual information and information from the body into global feeling states (Craig, 2009; Singer et al., 2009; Lamm and Singer, 2010). This higher-level coding of information would probably better map to information processing of abstract content than to simulation based on an automatic activation of primary sensory networks.

Together, these results suggest that a direct match between ST and the MNSs or empathy-networks vs. ToM-networks to TT is problematic. Consequently, an unproductive “either/or logic” concerning simulation and theorizing should be avoided, as Mitchell (2005, p. 363) has suggested (see also Keysers and Gazzola, 2007).

MODELS OF SOCIAL INTERACTION IN PSYCHOLOGY AND SOCIAL NEUROSCIENCE

In the previous section, we demonstrated that equating epistemic perspectives, cognitive processes, and neural mechanisms underlying social cognition is problematic. Based on this analysis we will now consider whether online social cognition is necessarily the dark matter of social neuroscience. To address this issue we must answer two further questions. First, what do we actually mean when we talk about social interaction? And second, do we really need to assume that there are neural networks specifically dedicated to only processing the social world or social interactions?

What does it take for an action to be real social interaction? The experimental aim in social neuroscience, following Wilms et al. (2010) who, in turn, follow Frith, (2007) consists in “closing the loop between interaction partners” (p. 1). This means that the action of one subject (henceforth A) should trigger a response of her partner, a sentient being (henceforth B), which in turn influences A and A’s reaction. This particular feedback has a specific effect on B resulting in a reaction in B that subsequently changes A’s mental state and so on (see **Figure 1**, left panel). In every iteration, each partner’s mental state is changed by his/her partner and these new states form the basis of the next iteration of social interaction. One basic feature of such a reciprocal interaction is the occurrence of emergent qualities, i.e., the largely unpredictable rearrangement of the already existing entities, namely A and B’s possible reactions. Such emergence is only possible if none of the involved subjects responses are controlled. Without the essential unpredictability that occurs in natural social interaction reciprocal changes in behavior would not occur. Along these lines Schippers et al. (2010) stated that it is, therefore, difficult to assess when one action ends and another starts (p. 9388; please note that for this reason, the time of measurement in the left panel of **Figure 1** should be seen as no more than a formal orientation). It is only when the design of the experiment allows for an action possessing four specific criteria (dynamic interplay, a virtually unlimited range of responses, living and uncontrolled



partners, and emergent qualities) that we can speak of real social interaction.

Just consider an example of a fundamentally social form of behavior such as tickling. Blakemore et al. (1998) have argued that tickling, certainly a low-level, bottom-up, mostly pre-reflective phenomenon is quintessentially social because, at least under normal conditions, the sensation can only arise when another individual delivers the touch. It has been proposed that the reason we cannot tickle ourselves is that during self-generated movements the brain produces a forward model that allows us to predict the effects of tickling and thus cancel these effects out in advance. When the touch is delivered by another, this model is absent, making the touch unpredictable that, hence, leads to the sensation of being tickled (Blakemore et al., 1998, 2000).

In contrast, the right panel in **Figure 1** illustrates other ways of modeling social interaction, which are relevant in the context of present social neuroscience research. In these types of social interaction, A's behavior is studied under the full control of B's response (in experimental research, the presence of the subject's interaction partner B is often feigned and therefore entirely controlled by the experimenter). Strictly speaking, because A's behavior does not cause a novel and unpredictable response in B, it is similar to a blind alley. B, being just an algorithm, would always react independent of A's action. This is indicated by the dotted lines in the right panel of **Figure 1**.

Although A repeatedly reacts to fixed "actions" made by B, the model that guides B's behavior does not change over time and so there is "no closing of the loop." Accordingly, the emergent qualities of such interactions are limited; hence, no reciprocal transformation can happen from T1 until T4. While temporally dynamic, such controlled interchanges are no more than "pseudo social interaction."

It has repeatedly been criticized that classic ToM tests allow for a bystander or spectator stance toward others rather than involving a stance of participation and involvement (e.g., Schilbach et al., 2006; Reddy, 2008). This critique raises the question whether current neuroscientific paradigms succeed in investigating the neural basis of minds that truly interact, as illustrated in

the left panel of **Figure 1**. Despite the astute variety of creative and visionary approaches to operationalize social cognition and interaction, it goes without saying that the degrees of freedom for real-life social responses and interactions in the scanner environment are limited. In the next paragraph, we will briefly review the social neuroscience literature, discuss some exemplary types of social interaction, and present some of the results. Please note, however, that we cannot provide a full review of social neuroscience literature, as this would go far beyond the scope of this article.

PARADIGMS IN SOCIAL NEUROSCIENCE TO STUDY SOCIAL COGNITION

One major challenge in social neuroscience is creating real social cognition and social interaction within the non-social environments associated with neuroscientific methods such as fMRI, EEG, MEG, or TMS. To confront these challenges, neuroscientists have used a rather large diversity of methods and paradigms. One specifically promising category of such paradigms is based on the use of game-theory derived from economic research and now widely used in neuroeconomics. In such economic game paradigms, one subject, typically lying in the scanner, engages in monetary exchange with real or pretended playing partners situated either outside of the scanner in another room (e.g., McCabe et al., 2001; Rilling et al., 2002, 2004, 2008; Gallagher and Frith, 2003; Sanfey et al., 2003; Singer et al., 2004a; Fehr and Camerer, 2007; Baumgartner et al., 2008, 2011; for an overview see Glimcher et al., 2009) or in another scanner in the context of hyper-scanning experiments (Montague et al., 2002).

Among other research questions, one focus of these studies was to examine whether brain responses of subjects differ as a function of whether they believe that they are interacting with a real human partner or simply with a computer or a non-intentional playing partner (e.g., McCabe et al., 2001; Gallagher et al., 2002; Singer et al., 2004b). McCabe et al. (2001) were among the first to conduct experiments on playing a two-person game in the scanner. The subjects played either with another alleged person or a computer and were asked to make choices in the game tree.

They all cooperated less in the condition under which they thought that they were playing with an algorithm. Moreover, the mere belief that they were playing with a human being resulted in the activation of specific regions of the prefrontal cortex.

Gallagher et al. (2002) introduced the “stone-paper-scissors” game in a PET paradigm. Their goal was to “investigate the neural substrates of ‘on-line’ mentalizing” (p. 814). Again, the subjects believed that they were either playing with another person (whom they met shortly before) or a computer, while in fact all responses were constantly generated by a computer program. As the respective neural substrate of the alleged social encounter, Gallagher et al. (2002) recorded activity in the anterior portion of paracingulate cortex bilaterally.

Similarly, Sanfey et al. (2003) scanned subjects playing the Ultimatum Game (UG) who had to respond to fair as well as unfair offers. Unfair offers elicited anger and rejection in the participants when another person to whom they had been introduced beforehand made them but not when these unfair offers were made by a computer. In the case of the illusion of interacting with a conspecific, the anterior insulae, the dorsolateral prefrontal cortex, and the ACC showed higher activation.

Rilling et al. (2004) tried to create a paradigm that allowed for the “immersion of participants in real social interaction that have personally meaningful consequences” (p. 1694) by scanning subjects while playing the UG and the Prisoner’s Dilemma Game (PDG) using both, assumed human and computer partners, outside the scanner. These studies demonstrated that these tasks only activated the ToM network (including the anterior paracingulate cortex, the posterior, and mid-STS, as well as the hippocampus and regions of the hypothalamus) when the subjects believed that they were playing with real human beings. Just like in the studies conducted by Gallagher et al. (2002) and Sanfey et al. (2003), the illusion that a real partner B was present elicited different brain patterns than if the partner was assumed to be artificial.

Finally, Singer et al. (2004b) involved participants in sequential trust games with intentional or non-intentional playing partners and revealed, in line with the findings above, that only when subjects believed to play with intentional agents, emotion-related brain activation (e.g., in the left amygdala, the insulae, and reward-related areas) were induced when perceiving intentional co-operator or defector faces as compared to neutral players.

Another set of paradigms focused on measuring the effects of directed or averted gaze on neural processes. For example, in an early PET study, Wicker and colleagues (1998) investigated the neural activation of mutual gaze (“a psychological process during which two persons have the feeling of a brief link between their two minds”; p. 221). Subjects were shown videos of persons looking toward them (in a mutual gaze condition) or away (in an averted gaze condition). This study revealed that eye contact activates the occipital part of the fusiform gyrus, the right parietal lobule, the right inferior temporal gyrus, and the middle temporal gyrus in both hemispheres. Further effects of eye contact were presented by Kampe et al. (2001) who demonstrated that the perceived attractiveness of an unfamiliar face depicted on still photographs augmented activity in the ventral striatum when the viewer met the person’s gaze, whereas it decreased in the absence of direct

eye contact. Central reward-related brain areas seem, thus, to be engaged during direct but not averted gaze when presented with still pictures of human faces.

In more recent eye-gaze paradigms, Schilbach et al. (2006, 2010a,b) sought to characterize neural correlates of being personally involved in social interaction by introducing more dynamic virtual-reality technologies in the scanner environment. Virtual characters were created that gazed at and greeted others—either the subjects (who were lying passively in the scanner) or a bystander. One of their main neuroscientific findings was that the vmPFC underlies the perception of social communication and feeling of personal involvement (Schilbach et al., 2006). Moreover, when the sharing of attention with the avatar was self-initiated by the participants, this led to an increase of neural activity in the ventral striatum (Schilbach et al., 2010a). Similarly, Wilms et al. (2010) tried to study social encounters in a truly interactive manner. They asked subjects in the scanner to respond or to probe the gaze of another person, depicted as an anthropomorphic avatar (who was actually computer-operated). The subject’s goal was to establish eye contact with the avatar and to jointly attend to one of three objects on a screen as a function of the subject’s eye-gaze. This method of interactive eye-tracking reflects the attempt to close the interaction loop between A and B. More precisely, Wilms et al. (2010) were interested in the neural differences of successfully initiating joint attention compared to mere gaze following. In this regard, they reported a main effect of joint attention that resulted in the activation of the mPFC, PCC, and the anterior temporal poles.

Another type of social neuroscience paradigms involves the presence of real people present in the scanner environment. For example in empathy for pain research, subjects in the scanner were coupled with either their loved ones (Singer et al., 2004a) or unfamiliar persons who differed in important aspects such as perceived fairness or group membership (Singer et al., 2006; Hein et al., 2010) who sat outside of the scanner room but visible to the subjects. In these paradigms, brain responses are elicited by creating the experience of painful shocks in the subjects. This first-personal pain response is compared with those brain responses elicited when watching the real person present in the same room suffering from pain.

Finally, some recent innovative paradigms have started to use cross-correlational statistics to compare brain activity of two individuals involved in a task together (Schippers et al., 2010; Anders et al., 2011). For example, Schippers et al. (2010) asked couples of participants to play charades in the scanner in order to examine brain activity during longer streams of social communication. Both of their brains were measured at separate times in the same scanner (thus, the authors did not draw on hyper-scanning technology). In one session, brain activity was recorded during the gesturing of a word for the partner. This gesturing was videotaped so that their partners could guess it in the subsequent scanning session. Results show that during guessing, the subject’s brain activity in the putative MNS and the vmPFC is caused by fluctuations in activity in the pMNS of the gesturing partner.

All the above-mentioned studies have been conducted with imaging methods. What about EEG-studies investigating social processes? Lindenberger et al. (2009) investigated interbrain

synchronization when eight pairs of guitar players performed a short piece together. They found significant between-brain oscillatory couplings during the preparatory period of metronome tempo setting in especially the fronto-central connections in the frequency range between 2 and 10 Hz, as well as after the play onset in the frequency range between 0.5 and 7.5 Hz. According to the authors, this coupling can be interpreted as a sign for social attunement.

In another EEG-study, Kourtis et al. (2010) registered the brain couplings of two persons who “interacted,” i.e., passed each other an object and then put it back on its original place. The authors then compared these EEG-measures with the brain activity of a third-person who only watched this interaction. The two interacting persons showed more motor activation during action anticipation than the “loner” who was not involved in the action. This motor activation was measured via the amplitude of the contingent negative variation (CNV), known to reflect motor preparation and activity in both the supplementary motor area (SMA) and the primary motor cortex (MI). Kourtis et al. (2010) suggest that this data indicates that social interaction modulates action simulation.

HAS SOCIAL NEUROSCIENCE REALLY YET SUCCEED IN STUDYING SOCIAL INTERACTION?

The previous section reviewed social neuroscience paradigms that have tried to investigate the neural basis of social cognition and social interaction using real-life experimental set-ups or cross-brain correlational methods rather than the presentation of static pictures such as faces or stories. However, when these are considered in light of our definition of real social interaction, a closer analysis reveals that none of these paradigms demonstrate a pattern of actions and reactions in which living and uncontrolled partners engage in behavior that leads to reciprocal impact on each other’s behavior. For example, experiments relying on game theory paradigms have had complete control over playing partner B’s responses (see right panel of **Figure 1**). McCabe et al. (2001) used the game tree, which offers limited response options: participants were presented to a dichotomous choice (they had to select either the left or the right branch of the tree). Moreover, player A was made to play with a partner who was fully controlled. Sanfey et al. (2003) and Rilling et al. (2004) were also unable to establish real social interaction. Even though the subjects met their partner (B) before scanning started and saw B’s photo in the scanner, there was no opportunity for real interaction. These “interactional” degrees of freedom offered just two options—either acceptance or rejection of B’s (again fully pre-fabricated) offer.

Similarly, the paradigms created by Schilbach and colleagues do not constitute real social interaction because even though B reacts to A’s behavior, B’s response is programmed by the experimenter and so lacks the unpredictability of a real person. Despite its novelty relative to the previous social neuroscience paradigms, the use of still pictures of faces as outlined above offers no emergent qualities of real interaction. Schippers et al. (2010) presented a different and ambitious approach to the “*information flow across brains during social interactions*” (p. 9391). Nevertheless, each trial of guessing the action ended without the other’s spontaneous feedback, and the subjects gestured the word-to-be-guessed

in a video camera instead of directly toward the partner’s face, therefore we can conclude that, although this experiment captures naturalistic complex symbolic and non-verbal behavior, the “loop” was never fully closed within a single interaction. The sought-after information flow from one brain to another was not really flowing.

Finally, the empathy paradigms (Singer et al., 2004a, 2006; Hein et al., 2010), even though integrating authentic uncontrolled (and uncontrolled) partners in the same scanner environment, can again not be considered as real social interaction paradigms, as the response of B is not important for the analyses or claims of this investigation. The only thing that matters is the brain response of A.

Do the EEG-studies capture real social interaction? Although Lindenberger et al. (2009) studied coordinated action, this again do not fulfill our criteria of studying real social interaction. The guitar players follow a common goal (performing a specific piece together over 60 trials), preventing opportunities for creative actions and responses. For real musical interaction to be studied, it would be necessary to record the brain activities while two individuals, for example, improvise. Imagine for example jazz improvisation, where the tunes emerges through the reciprocal impact that one players melody has on the other player’s tune and by doing so capture real elements of social interaction. During improvisation, repetition, and predictable responses are nearly impossible because each individual’s contribution emerges from the process of listening and replying to the other (Seddon, 2005). An additional worry concerning these studies is that observed between-brain oscillation can also be explained by the fact that the musicians are merely following the same synchronizing stimulus (i.e., the metronom or the melody they play). This would then again not be an example of real social interaction but similar to two persons watching the same movie, and whose brain activities consequently are synchronized by this same visual material.

Also, in Kourtis et al.’s (2010) EEG-study two persons “interact,” but only in the sense that they pass each other an object. Although this counts as interaction because the subjects are real living persons who can act jointly in a face-to-face setting, the paradigm restricted the type of interaction that was possible and so the two minds do not mutually shape each other by the unpredictability of their responses. Due to the limited range of behavior and the restricted possibility for emergent qualities, we do not see the closing of the loop here, neither.

In sum, this short review has shown that even though some of these studies provided innovative ways to study social cognition, they still fail to capture *real* social interaction and so fall short of revealing the neural processes that occur during online social cognition. However, what *could* be investigated were indeed several forms of second-person perspective taking.

IS THERE EVIDENCE FOR PURELY SOCIAL PROCESSES AND COMPUTATIONS IN THE HUMAN BRAIN?

Our short review about some social neuroscience studies with associated neuronal findings point to another important question raised in this paper. Can we assume that the brain contains specific modules only devoted to the processing of social stimuli or,

even more radically, specific brain regions or specific cells tuned to online social cognition. More precisely, we are distinguishing between two separate questions: (a) To what extent is the so-called social brain specifically social? and (b) Are there neuronal networks, computations or single cells specifically tuned to process only online social interactions?

A closer look into the brain areas involved in the social neuroscience experiments reviewed suggest that the answer to both questions is no. To answer question (a), the social brain is not exclusively social; rather, each of these brain regions has been shown to also be involved in other non-social tasks. For example, even though “*the meeting of minds*” (Amodio and Frith, 2006) certainly has hedonic qualities and may “feel good” (see Schilbach et al., 2010a), the ventral striatum or other reward-related brain areas cannot be seen as specific neural correlates of the second-person perspective or of mutual social interaction because these brain areas are known to be sensitive to all kinds of rewards, be it social or non-social (e.g., Schultz, 2002; O’Doherty et al., 2006). Hence, these regions will also be equally strongly involved in pleasant *non-social* activities such as indulging in a glass of high-class Bordeaux. Similarly, the consistent involvement of anterior insula and mACC in empathy for pain paradigms (see Lamm et al., 2011) does not make these brain regions specific “empathy regions.” On the contrary, these “shared activation” studies reveal that these regions are also involved in processing negative affective experiences in the self or in other non-social context (see also Singer, 2012).

The phenomenon of ticklishness when being tickled by others but not by oneself, as mentioned earlier, can be taken as a good example to discuss the question to whether a specific neural system is responsible for social interaction. This phenomenon does not occur because of unique neural computations specialized for being tickled but rather emerges from more general predictive properties of our sensory and motor system and the fact that we can predict the effects of our own actions but not those of others⁵. Accordingly, online social cognition could be explained by understanding how different basic cognitive, emotional, and motor processes and their underlying brain mechanisms cooperate to produce representations of the actions, sensations, or mental state of another individual. Hence, a review of the present social neuroscience literature does not support the claim that neural processes or computations that specifically subserve the processing of social stimuli exist. In sum, the answer to question (a) would be that to the best of our knowledge so far, no neural computations or neuronal networks specifically dedicated to social stimuli or social cognition alone have been identified.

Would this conclusion have to change if eventually neuroscientists succeed in bringing real social interactions into neuroscience paradigms? The hope of identifying neuronal networks, neural computations, or even single cells, that would selectively only react when we are involved in online social cognition, seems—given the hitherto existing neuroscientific evidence—simply highly unlikely. For this to be the case, our brains would

need to contain cells or perform computations that are only sensitive to the interactive nature of two intentional living agents but are silent when merely one agent is concerned, even if this involves the processing of social stimuli. Following this line of reasoning, the answer to question (b) would again have to be no.

Having said this, we would, however, expect that social interaction might activate neuronal *patterns* (but not hitherto unknown areas) different from situations where subjects are not personally being involved. Still, this effect would come as no surprise, as it may just stem from mere attention and saliency effects known to modulate activation patterns in the brain. The brains of subjects being half asleep will obviously show a different activation pattern than a highly engaged subject irrespective of whether this subject is engaged in social or non-social information processing.

Note also, that even though we have argued that social neuroscience has not yet succeeded in implementing real social interaction paradigms, the so-called social brain circuitries have nevertheless been discovered and repeatedly described on the basis of subjects merely believing in the presence of another interaction partner.

Even though the implementation of real social interaction paradigms may not reveal novel brain mechanisms exclusively devoted to social cognition or social interaction, it is important to stress at this point that the implementation of real-life social interactive paradigms can, nonetheless, inform our understanding of social dynamics and the psychological phenomena that emerge in these conditions. Social interaction is a central and enormously important factor for human evolution, ontogeny, and daily life, for example, in the development of individual personhood or self-consciousness. In evolutionary terms, for example, it has been argued that the demands of interacting with group members have been a vital and a powerful influence on the size of the neo-cortex of the brain (as measured by the ratio between the volume of medulla oblongata and neocortex Dunbar, 1998). According to this *social brain hypothesis*, the need to understand other minds, as well as the related processes of communication and self-control, drove an increase in neo-cortical volumes in mammals, particularly in primates. Social environments are also important at an ontogenetic level. Anecdotal evidence from the eighteenth century indicates that isolation from the social group, such as that experienced by feral children like the Wildboy of Aveyron, leads to problems with developing more than rudimentary skills essential for social interaction (Zingg, 1940). More recent studies showed that human psychosocial development is largely influenced by the quality of parent-child interactions (Beebe et al., 2008). Furthermore, developmental psychologists have stated that it is through interactive sharing that children ontogenetically acquire the capacity of taking and confronting intersubjective perspectives, i.e., understanding other minds (Moll and Meltzoff, 2011). We argue that face-to-face encounters are a necessary condition for social cognitive abilities to evolve, but that, once in place, other minds can also be understood when the persons are not engaging in real social interaction. Therefore, we have challenged the currently widespread narrow definition of second-perspective taking. In sum, at both the level of evolution and of ontogeny, immersion in a complex social environment seems necessary for the human mind to develop normally.

⁵Note that this does of course not imply that unpredictability is a sufficient condition for being a person or for an action to be social interaction.

Future investigations of online social cognition could be inspired by paradigms developed in developmental and attachment psychology (as conducted by Tronick et al., 1978; Tronick, 2003, and, more recently, Beebe et al., 2008). The subjects in these studies, mainly mothers and their babies, are given time and space to interact naturally and face-to-face with one another. They play with each other, mirror, and validate the other's expression, misunderstand one another, undergo communicative disruptions, and by return engage in so-called repair processes of these mismatches—or they do not. A detailed micro-analytic decomposition of the interplay may reveal features of individual, natural interaction that are important for further investigations in developmental, social, or clinical psychology, but will perhaps not be that useful for the identification of specifically social neuronal computations in particular. This is, of course, not to say that social interaction should not be investigated in the context of social neuroscience. Thus, even though its study may not be about understanding new brain processes, it is about understanding how existing brain processes are deployed and influenced in a particular dynamic during social interaction. Furthermore, investigating real interaction might enable social neuroscience to shed more light on some of the classical problems of social psychology like conformity (Asch, 1951; Milgram, 1963) or decision-making in groups (Surowiecki, 2004).

CONCLUSION

Moving toward an understanding of the neural basis of social interaction is one of the main goals in social neuroscience. In this paper, we asked why social interaction is accredited with such a central significance. We showed that its definition, as well as the definition of the terms on- vs. offline social cognition and second-person perspective taking is imprecise leading to confusion. Furthermore, significant problems emerge when the philosophical, psychological, and neuroscientific investigations concerning the understanding of other minds are simply mapped to each

other. We have also reviewed relevant neuroscientific studies that focus on social interaction and have demonstrated that none to date have investigated real social interaction, understood as the emergent qualities of an encounter that occur through the reciprocal interaction of two real individuals. In this sense, true social interaction remains the “dark matter” of social neuroscience. However, this is not as daunting as it may seem at a first glance. First, understanding other minds is not bound to interaction because it is an epistemic perspective rather than a process tied to online social cognition. Second, the specific neural correlates of reciprocal interaction are unlikely to differ from those that have already been identified by prior work. Third, even though our short review has shown that no studies have captured online social cognition during real social engagement, they have captured important features of second-person perspective taking and so do reveal important information on how we make sense of other minds. We suggest that this should be the focus of future work investigating mental states attribution. Rather than seeking neural substrates for computations that can only be performed during social interaction, research on how we understand other minds would be more likely to be informative if it examined how basic cognitive and affective processes are deployed to cope with the demands placed on the mind by the complex interactions that make the social lives of our species so remarkable. In this way neuroscience can help us understand the way that social interaction continues to shape our evolution, ontogeny, and every-day lives.

ACKNOWLEDGMENTS

This work was supported by the VolkswagenStiftung grant “Das Gehirn—Ein Beziehungsorgan” to Marisa Przyrembel and Michael Pauen. We also thank the participants of the philosophical colloquium at the Center for Subjectivity Research, University of Copenhagen, for valuable comments on an earlier draft of this paper.

REFERENCES

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr. Opin. Neurobiol.* 12, 169–177.
- Adolphs, R. (2010). Conceptual challenges and directions for social neuroscience. *Neuron* 65, 752–767.
- Amodio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.
- Anders, S., Heinzle, J., Weiskopf, N., Ethofer, T., and Haynes, J. D. (2011). Flow of affective information between communicating brains. *Neuroimage* 54, 439–446.
- Asch, S. E. (1951). “Effects of group pressure upon the modification and distortion of judgements,” in *Groups, Leadership, and Men*, ed H. Guetzkow (Pittsburgh, PA: Carnegie Press), 177–190.
- Baron-Cohen, S., Tager-Flusberg, H., and Cohen, D. (eds.). (2000). *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*. Oxford, UK: Oxford University Press.
- Barron, E., Riby, L. M., Greer, J., and Smallwood, J. (2011). Absorbed in thought: the effect of mind wandering on the processing of relevant and irrelevant events. *Psychol. Sci.* 22, 596–601.
- Baumgartner, T., Götze, L., Gügler, R., and Fehr, E. (2011). The mentalizing network orchestrates the impact of parochial altruism on social norm enforcement. *Hum. Brain Mapp.* 33, 1452–1469.
- Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., and Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron* 58, 639–650.
- Beebe, B., Jaffe, J., Buck, K., Chen, H., Cohen, P., Blatt, S., Kaminer, T., Feldstein, S., and Andrews, H. (2008). Six-week postpartum maternal depressive symptoms and 4-month mother-infant self- and interactive contingency. *Infant Ment. Health J.* 29, 442–471.
- Blakemore, S. J., and Decety, J. (2001). From the perception of action to the understanding of intention. *Nat. Rev. Neurosci.* 2, 561–567.
- Blakemore, S., Wolpert, D., and Frith, C. (1998). Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640.
- Blakemore, S., Wolpert, D., and Frith, C. (2000). Why can't you tickle yourself? *Neuroreport* 11, R11–R16.
- Bohman, J. (2000). “The importance of the second person: interpretation, practical knowledge, and normative attitudes,” in *Empathy and Agency. The Problem of Understanding in the Human Sciences*, eds H. H. Kögler and K. R. Stueber (Boulder, CO: Westview), 222–224.
- Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., and Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8719–8724.
- Craig, A. D. (2009). How do you feel – now? the anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70.
- Davidson, D. (2004). *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.
- Decety, J., and Jackson, P. L. (2004). The functional architecture of human empathy. *Behav. Cogn. Neurosci. Rev.* 3, 71–100.
- De Jaegher, H., Di Paolo, E., and Gallagher, S. (2010). Can social

- interaction constitute social cognition? *Trends Cogn. Sci.* 14, 441–447.
- de Vignemont, F., and Singer, T. (2006). The empathic brain: how, when and why? *Trends Cogn. Sci.* 10, 435–441.
- Dinstein, I., Hasson, U., Rubin, N., and Heeger, D. J. (2007). Brain areas selective for both observed and executed movements. *J. Neurophysiol.* 98, 1415–1427.
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91, 176–180.
- Downing, P. E., Jiang, Y., Shuman, M., and Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science* 293, 2470–2473.
- Dunbar, R. I. M. (1998). The social brain hypothesis. *Evol. Anthropol.* 6, 178–190.
- Etzel, J. A., Gazzola, V., and Keysers, C. (2008). Testing simulation theory with cross-modal multivariate classification of fMRI data. *PLoS ONE* 3:e3690. doi: 10.1371/journal.pone.0003690
- Fan, Y., Duncan, N. W., de Greck, M., and Northoff, G. (2011). Is there a core neural network in empathy? an fMRI based quantitative meta-analysis. *Neurosci. Biobehav. Rev.* 35, 903–911.
- Fehr, E., and Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* 11, 419–427.
- Fodor, J. A. (1983). *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.
- Frith, C. D. (2003). “Neural hermeneutics: how brains interpret minds,” in *Presented at the 9th Annual Meeting of the Organization of Human Brain Mapping*, (New York, NY).
- Frith, C. D. (2007). *Making Up the Mind. How the Brain Creates our Mental World*. Oxford, UK: Blackwell.
- Frith, C. D., and Frith, U. (1999). Interacting minds – a biological basis. *Science* 286, 1692–1695.
- Frith, U., and Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. B Biol. Sci.* 358, 459–473.
- Gallagher, H. L., and Frith, C. D. (2003). Functional imaging of ‘theory of mind’. *Trends Cogn. Sci.* 7, 77–83.
- Gallagher, H. L., Jack, A. I., Roepstorff, A., and Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage* 16, 814–821.
- Gallagher, S. (2008). Inference or interaction: social cognition without precursors. *Philos. Explor.* 11, 163–174.
- Gallese, V. (2001). The ‘shared manifold’ hypothesis. From mirror neurons to empathy. *J. Conscious. Stud.* 8, 33–50.
- Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609.
- Gallese, V., and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends Cogn. Sci.* 2, 493–501.
- Gangopadhyay, N., and Schilbach, L. (2012). Seeing minds: a neurophilosophical investigation of the role of perception-action coupling in social perception. *Soc. Neurosci.* 20, 410–423.
- Gazzola, V., Rizzolatti, G., Wicker, B., and Keysers, C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage* 35, 1674–1684.
- Glimcher, P. W., Camerer, C. F., Fehr, E., and Plodrack, R. A. (eds.). (2009). *Neuroeconomics. Decision Making and the Brain*. London: Academic Press.
- Goldman, A. (2005). “Imitation, mind reading, and simulation,” in *Perspectives on Imitation: From Neuroscience to Social Science, Vol. 2: Imitation, Human Development, and Culture*, eds S. Hurley and N. Chater (Cambridge, MA: MIT Press), 79–93.
- Gopnik, A. (1993). How we know our minds: the illusion of first-person knowledge of intentionality. *Behav. Brain Sci.* 16, 90–101.
- Grèzes, J., and Decety, J. (2001). Functional anatomy of execution, mental simulation, observation and verb generation of actions: a meta-analysis. *Hum. Brain Mapp.* 12, 1–19.
- Habermas, J. (2004). Freiheit und Determinismus. *Deutsche Zeitschrift für Philosophie* 52, 871–890.
- Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233.
- Heidegger, M. (1927/1975). *Die Grundprobleme der Phänomenologie. Gesamtausgabe Bd. 24*. Frankfurt: Klostermann.
- Heidegger, M. (1927/1986). *Sein und Zeit*. Tübingen: Mohr.
- Hein, G., Silani, G., Peuschoff, K., Batson, C. D., and Singer, T. (2010). Neural responses to the suffering of ingroup- and outgroup members predict individual differences in costly helping. *Neuron* 68, 149–160.
- Hein, G., and Singer, T. (2008). I feel how you feel but not always: the empathic brain and its modulation. *Curr. Opin. Neurobiol.* 18, 153–158.
- Hutto, D. (2007). “The narrative practice hypothesis: origins and applications of folk psychology,” in *Narrative and Understanding Persons*, ed D. Hutto (Cambridge, UK: Cambridge University Press), 43–68.
- Jabbi, M., Swart, M., and Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage* 34, 1744–1753.
- Kam, J. W. Y., Dao, E., Farley, J., Fitzpatrick, K., Smallwood, J., Schooler, J. W., and Handy, T. C. (2011). Slow fluctuations in attentional control of sensory cortex. *J. Cogn. Neurosci.* 23, 460–470.
- Kampe, K. K. W., Frith, C. D., Dolan, R. J., and Frith, U. (2001). Reward value of attractiveness and gaze. *Nature* 413, 589.
- Kanwisher, N. (2000). Domain specificity in face perception. *Nat. Neurosci.* 3, 759–763.
- Kanwisher, N., McDermott, J., and Chun, M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.
- Kanwisher, N., and Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Biol. Sci.* 361, 2109–2128.
- Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., and Heatherton, T. F. (2002). Finding the self? an event-related fMRI study. *J. Cogn. Neurosci.* 14, 785–794.
- Keysers, C., and Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Prog. Brain Res.* 156, 379–401.
- Keysers, C., and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends Cogn. Sci.* 11, 194–196.
- Keysers, C., and Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Curr. Opin. Neurobiol.* 19, 1–6.
- Keysers, C., Kaas, J. H., and Gazzola, V. (2010). Somatosensation in social perception. *Nat. Rev. Neurosci.* 11, 417–428.
- Kourtis, D., Sebanz, N., and Knoblich, G. (2010). Favouritism in the motor system: social interaction modulate action simulation. *Biol. Lett.* 6, 758–761.
- Lahanas, A. B., and Nanopoulos, D. (2003). WMAPing out super-symmetric dark matter and phenomenology. *Phys. Lett. B* 568, 55–62.
- Lamm, C., Decety, J., and Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage* 54, 2492–2502.
- Lamm, C., and Singer, T. (2010). The role of anterior insular cortex in social emotions. *Brain Struct. Funct.* 241, 579–951.
- Lindemann, G. (2006). “Die du-perspektive in der hirnforschung,” in *Akteur Gehirn—oder das vermeintliche Ende des handelnden Subjekts. Eine Kontroverse*, eds J. Reichertz and N. Zaboura (Wiesbaden: VS Verlag für Sozialwissenschaften), 263–285.
- Lindenberger, U., Li, S., Gruber, W., and Müller, V. (2009). Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neurosci.* 10, 1–12.
- Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F., and Kelley, W. M. (2004). Medial prefrontal activity predicts memory for self. *Cereb. Cortex* 14, 647–654.
- Mason, M., Norton, M., Van Horn, J. D., Wegner, D. W., Grafton, S. T., and Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science* 315, 393–395.
- McCabe, K., Houser, D., Ryan, L., Smith, V., and Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11832–11835.
- Mead, G. H. (1925). The genesis of the self and social control. *Int. J. Ethics* 35, 251–277.
- Mead, G. H. (1926). “The objective reality of perspectives,” in *Proceedings of the Sixth International Congress of Philosophy* (New York, NY), 75–85.
- Mead, G. H. (1934). *Mind Self and Society*. Chicago, IL: Chicago University Press.
- Milgram, S. (1963). Behavioral study of obedience. *J. Abnorm. Psychol.* 67, 371–378.
- Mitchell, J. P. (2005). The false dichotomy between simulation and theory-theory: the argument’s error. *Trends Cogn. Sci.* 9, 363–364.
- Mitchell, J. P. (2008). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cereb. Cortex* 18, 262–271.
- Mitchell, J. P. (2009). Inferences about mental states. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1309–1316.
- Mitchell, J. P., Banaji, M. R., and Macrae, C. N. (2005). The link between social cognition and self-referential thought in the medial

- prefrontal cortex. *J. Cogn. Neurosci.* 17, 1306–1315.
- Mitchell, J. P., Macrae, C. N., and Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron* 50, 655–663.
- Mojzisch, A., Schilbach, L., Helmert, J. R., Pannasch, S., Velichkovsky, B. M., and Vogeley, K., (2006). The effects of self-involvement on attention, arousal, and facial expression during social interaction with virtual others: a psychophysiological study. *Soc. Neurosci.* 3–4, 184–195.
- Moll, H., and Meltzoff, A. N. (2011). “Perspective-taking and its foundation in joint attention,” in *Perception, Causation, and Objectivity. Issues in Philosophy and Psychology* eds N. Eilan, H. Lernerman, and E. Roessler (Oxford, UK: Oxford University Press), 286–304.
- Montague, P. R., Berns, G. S., Cohen, J. D., McClure, S. M., Pagnoni, G., Dhamala, M., Wiest, M. C., Karpov, I., King, R. D., Apple, N., and Fisher, R. E. (2002). Hyperscanning: simultaneous fMRI during linked social interactions. *Neuroimage* 16, 1159–1164.
- O’Doherty, J. P., Buchanan, T. W., Seymour, B., and Dolan, R. J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron* 49, 157–166.
- Pauen, M. (2010). How privileged is first-person privileged access? *Am. Philos. Q.* 47, 1–15.
- Pauen, M., (2012). The second person perspective. *Inquiry* 55, 33–49.
- Petitengin, C. (2006). Describing one’s subjective experience in the second person, an interview method for the science of consciousness. *Phenomenol. Cogn. Sci.* 5, 229–269.
- Ploghaus, A., Tracey, I., Gati, J. S., Clare, S., Menon, R. S., Matthews, P. M., and Rawlins, J. N. (1999). Dissociating pain from its anticipation in the human brain. *Science* 284, 1979–1981.
- Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526.
- Preston, S. D., and deWaal, F. B. N. (2002). Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–20.
- Reddy, V. (2008). *How Infants Know Minds*. Cambridge, MA: Harvard University Press.
- Rilling, J. K., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., and Kilts, C. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.
- Rilling, J. K., Goldsmith, D. R., Glenn, A. L., Jairam, M. R., Elfenbein, H. A., Dagenais, J. E., Murdock, C. D., and Pagnoni, G. (2008). The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia* 46, 1256–1266.
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage* 22, 1694–1703.
- Rizzolatti, G., and Fabbri-Destro, M. (2010). Mirror neurons: from discovery to autism. *Exp. Brain Res.* 200, 223–237.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., and Fazio, F. (1996). Localization of grasp representations in human by PET: 1. Observation versus execution. *Exp. Brain Res.* 111, 246–252.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.
- Saxe, R. (2006). Uniquely human social cognition. *Curr. Opin. Neurobiol.* 16, 235–239.
- Saxe, R., Carey, S., and Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annu. Rev. Psychol.* 55, 87–124.
- Saxe, R., and Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in “theory of mind”. *Neuroimage* 19, 1835–1842.
- Schilbach, L. (2010). A second-person approach to other minds. *Nat. Rev. Neurosci.* 11, 449.
- Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G. B., Shah, N. J., Fink, G. R., and Vogeley, K. (2010a). Minds made for sharing. Initiating joint attention recruits reward-related neurocircuitry. *J. Cogn. Neurosci.* 22, 2702–2715.
- Schilbach, L., Eickhoff, S. B., Cieslik, E., Shah, N. J., Fink, G. R., and Vogeley, K. (2010b). Eyes on me: an fMRI study of the effects of social gaze on action control. *Soc. Cogn. Affect. Neurosci.* 6, 393–403.
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., and Vogeley, K. (in press). Toward a second-person neuroscience. *Behav. Brain Sci.*
- Schilbach, L., Wohlschlaeger, A. M., Kraemer, N., Newen, A., Shah, N. J., Fink, G. R., and Vogeley, K. (2006). Being with virtual others: neural correlates of social interaction. *Neuropsychologia* 44, 718–730.
- Schippers, M., Roebroek, A., Renken, R., Nanetti, L., and Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9388–9393.
- Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron* 36, 241–263.
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Seddon, F. A. (2005). Modes of communication during jazz improvisation. *Br. J. Music Educ.* 22, 47–61.
- Sellars, W. (1956). “Empiricism and the philosophy of mind,” in *Minnesota Studies in the Philosophy of Science, Vol. I*, eds H. Feigl and M. Scriven (Minneapolis, MN: University of Minnesota Press), 253–329.
- Seymour, B., O’Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K. J., and Frackowiak, R. S. (2004). Temporal difference models describe higher-order learning in humans. *Nature* 429, 664–667.
- Singer, T. (2012). The past, present and future of social neuroscience: a European perspective. *Neuroimage* 61, 437–449.
- Singer, T., Critchley, H. D., and Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* 13, 334–340.
- Singer, T., Seymour, B., O’Doherty, J., Kaube, H., Dolan, R. J., and Frith, C. D. (2004a). Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162.
- Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., and Frith, C. D. (2004b). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662.
- Singer, T., and Lamm, C. (2009). The social neuroscience of empathy. The year in cognitive neuroscience 2009. *Ann. N.Y. Acad. Sci.* 1156, 81–96.
- Singer, T., Seymour, B., O’Doherty, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469.
- Singer, T., Wolpert, D., and Frith, C. D. (2004c). “Introduction: the study of social interactions,” in *The Neuroscience of Social Interaction: Decoding Imitating and Influencing the Actions of Others*, eds C. Frith and D. Wolpert (Oxford: Oxford University Press), xiii–xxvii.
- Smallwood, J., Beech, E. M., Schooler, J. W., and Handy, T. C. (2008). Going AWOL in the brain – mind wandering reduces cortical analysis of the task environment. *J. Cogn. Neurosci.* 20, 458–469.
- Smallwood, J., Schooler, J. W., Turk, D. J., Cunningham, S. J., Burns, P., and Macrae, C. N. (2011). Self-reflection and the temporal focus of the wandering mind. *Conscious. Cogn.* 20, 1120–1126.
- Stawarczyk, D., Majerus, S., Maquet, P., and D’Argembeau, A. (2011). Neural correlates of ongoing conscious experience: both task-unrelatedness and stimulus-independence are related to default network Activity. *PLoS ONE* 6:e16997. doi: 10.1371/journal.pone.0016997
- Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economics, Societies, and Nations*, 1st Edn. New York, NY: Doubleday.
- Tamir, D. I., and Mitchell, J. P. (2011). The default network distinguishes construals of proximal versus distal events. *J. Cogn. Neurosci.* 23, 2945–2955.
- Tronick, E. Z. (2003). Things still to be done on the still-face effect. *Infancy* 4, 475–482.
- Tronick, E. Z., Als, H., Adamson, L., Wise, S., and Brazelton, T. B. (1978). The infant’s response to entrapment between contradictory messages in face-to-face interaction. *J. Am. Acad. Child Psychiatry* 17, 1–13.
- Varela, F., and Shear, J. (eds.). (1999). *The View from Within: First-Person Methodologies in the Study of Consciousness*. London: Imprint Academic.
- Waytz, A., and Mitchell, J. P. (2011). Two mechanisms for simulating other minds: dissociable neural bases for self-projection and mirroring. *Curr. Dir. Psychol. Sci.* 20, 197–200.
- 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.
- Wicker, B., Michel, F., Henaff, M. A., and Decety, J. (1998). Brain regions involved in the perception of gaze: a PET study. *Neuroimage* 8, 221–227.
- Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., and Vogeley, K. (2010). It's in your eyes – using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Soc. Cogn. Affect. Neurosci.* 5, 98–107.
- 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.
- Woolfe, T., Want, S. C., and Siegal, M. (2002). Signposts to development: theory of mind in deaf children. *Child Dev.* 73, 768–778.
- Zahavi, D. (2011). Empathy and direct social perception: a phenomenological proposal. *Rev. Philos. Psychol.* 2, 541–558.
- Zhang, D., and Raichle, M. E. (2010). Disease and the brain's dark energy. *Nat. Rev. Neurol.* 6, 15–28.
- Zingg, R. M. (1940). Feral man and extreme cases of isolation. *Am. J. Psychol.* 53, 487–517.
- 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.
- Received: 29 February 2012; accepted: 07 June 2012; published online: 21 June 2012.
- Citation:* Przyrembel M, Smallwood J, Pauen M and Singer T (2012) *Illuminating the dark matter of social neuroscience: Considering the problem of social interaction from philosophical, psychological, and neuroscientific perspectives.* *Front. Hum. Neurosci.* 6:190. doi: 10.3389/fnhum.2012.00190
- Copyright © 2012 Przyrembel, Smallwood, Pauen and Singer. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.