



# The interactive brain hypothesis

Ezequiel Di Paolo<sup>1,2,3\*</sup> and Hanne De Jaegher<sup>2,3</sup>

<sup>1</sup> Department of Logic and Philosophy of Science, Ikerbasque - Basque Science Foundation, San Sebastian, Bizkaia, Spain

<sup>2</sup> Department of Logic and Philosophy of Science, University of the Basque Country (UPV/EHU), San Sebastian, Bizkaia, Spain

<sup>3</sup> Centre for Computational Neuroscience and Robotics, University of Sussex, Brighton, UK

## Edited by:

Ulrich Pfeiffer, University Hospital  
Cologne, Germany

## Reviewed by:

Scott Kelso, Florida Atlantic  
University, USA

Evan Thompson, University  
of Toronto, Canada

## \*Correspondence:

Ezequiel Di Paolo, Department of  
Logic and Philosophy of Science,  
Ikerbasque - Basque Science  
Foundation, Av De Tolosa 70,  
20018, San Sebastian,  
Bizkaia, Spain.  
e-mail: ezequiel@sussex.ac.uk

Enactive approaches foreground the role of interpersonal interaction in explanations of social understanding. This motivates, in combination with a recent interest in neuroscientific studies involving actual interactions, the question of how interactive processes relate to neural mechanisms involved in social understanding. We introduce the Interactive Brain Hypothesis (IBH) in order to help map the spectrum of possible relations between social interaction and neural processes. The hypothesis states that interactive experience and skills play enabling roles in both the development and current function of social brain mechanisms, even in cases where social understanding happens in the absence of immediate interaction. We examine the plausibility of this hypothesis against developmental and neurobiological evidence and contrast it with the widespread assumption that mindreading is crucial to all social cognition. We describe the elements of social interaction that bear most directly on this hypothesis and discuss the empirical possibilities open to social neuroscience. We propose that the link between coordination dynamics and social understanding can be best grasped by studying transitions between states of coordination. These transitions form part of the self-organization of interaction processes that characterize the dynamics of social engagement. The patterns and synergies of this self-organization help explain how individuals understand each other. Various possibilities for role-taking emerge during interaction, determining a spectrum of participation. This view contrasts sharply with the observational stance that has guided research in social neuroscience until recently. We also introduce the concept of readiness to interact to describe the practices and dispositions that are summoned in situations of social significance (even if not interactive). This latter idea links interactive factors to more classical observational scenarios.

**Keywords:** social interaction, enaction, participatory sense-making, autonomy, transitions in coordination, readiness to interact, interactive brain hypothesis

## INTRODUCTION: CHANGING SOME BACKGROUND ASSUMPTIONS IN SOCIAL NEUROSCIENCE

Research in social neuroscience has started to move away from its almost exclusive focus on the individual brain as a detached interpreter of social stimuli (see e.g., Van Overwalle, 2009) and to pay attention to neural mechanisms involved in embodied social interaction (Schilbach et al., 2006; Tognoli et al., 2007; Lindenberger et al., 2009; Dumas et al., 2010; Redcay et al., 2010; Pfeiffer et al., 2011). As a result, a series of fundamental questions concerning the function of the brain in social understanding become apparent. One option that is opened by investigating complex processes of social interaction is that brains might bear less of a cognitive load than assumed in modular and individualistic explanations of social cognition based on mindreading (the individual attribution and meta-cognitive processing of the “mental states” of others). This is due to the well-documented fact that processes of social interaction are complex, multi-layered, self-organizing, and can shape individual intentions, orient individual perception and guide the performance of individual action (Marsh et al., 2009; De Jaegher et al., 2010; Riley et al., 2011). A consequence of this is that the brain is potentially less involved in

reconstructing or computing the “mental state” of others based on social stimuli and more involved in participating in a dynamical process outside its full control, thus inviting explanatory strategies in terms of dynamical concepts such as synergies, coordination, phase attraction, (meta)stability, structural stability, transients, and stationarity, etc.

We articulate the tension between this possibility and conservative mindreading accounts by introducing the Interactive Brain Hypothesis (IBH), which is aimed at broadening the spectrum of possible explanations in social cognition research. As we shall see, the IBH is an overarching assumption from which different specific hypotheses may be derived. Its main contrasting perspective is the currently dominant assumption that gives priority to processes of mindreading. While the focus on mindreading has been criticized on various fronts, the alternatives have yet to coalesce into a well-defined research program. We believe that the IBH contributes to this end by developing an alternative set of starting assumptions for social cognition research.

Our proposal is framed within the enactive approach to life and mind. With roots in work by Francisco Varela and colleagues (Varela et al., 1991), this approach has seen a major theoretical

development since the turn of the century (e.g., Di Paolo, 2005; Thompson, 2007). The main focus in this approach is the living body, its autonomy as a self-organizing system, its precarious identity and its sense-making relation to the world (Di Paolo et al., 2010). As such the approach is nourished by dynamical systems concepts and by phenomenology, as well as ecologically plausible experiments and agent-based modeling work. For social cognition research, the central implications of this approach have been developed in the concept of *participatory sense-making* (De Jaegher and Di Paolo, 2007; De Jaegher, 2009), which breaks with several assumptions about social cognition, such as the spectatorial, individualistic view of the social cogniser or the hidden nature of intentions. In this perspective, interpersonal interaction dynamics play a central explanatory role in social understanding and this is what will be emphasized in this paper. The claims that we make here about the roles of interaction processes and individual mechanisms in social understanding are part of the larger theory of participatory sense-making (which includes several key elements apart from interaction, such as emergence, autonomy, agency, sense-making, and subjective experience) and should be considered in that context.

A premise of the enactive approach is that cognition is not exhaustively determined by neural processes (De Jaegher et al., 2010) but implies the embeddedness of such processes in a living body and the embeddedness of this body in a world. Having said that, we hasten to emphasize that the enactive approach is *not* an externalist perspective on cognition. Externalism, often contrasted with neuro-centric internalism, proposes that cognitive processes are to be found outside the brain and even the organism, and that intentions acquire their full meaning only when such external factors are taken into account. The enactive approach emphasizes the inherent *relational* nature of cognition and while it rejects neuro-centrism, it also sees the externalist position as wanting because merely pointing to external dependencies fails to articulate what makes a relation between agent and world meaningful or a process cognitive. Instead, enactivism conceives of cognitive agents as participants who enact a world, not as passive data collectors who model or represent the world. The key difference is in how the agent/world relation is explicitly or implicitly conceived. Because the enactive approach sees cognitive processes as inherently relational, and agents primarily as participants, it considers it crucial to elucidate different aspects of this relation, including what goes on in individual brains as a result of it. Our aim in this paper is to project this overarching framework onto the plane of neuroscience and explore implications for the study of brain processes in the context of intersubjectivity. If interaction processes are central to explaining social cognition, then how do we understand the neural mechanisms active during social engagements or social tasks? Our goal is to help promote and develop research on this question. We propose the IBH in the spirit of raising a series of questions and indicating research paths, which if taken will lead to specific interactive hypotheses about neural processes in concrete instances of social cognition. Our aim, therefore, is to describe the conceptual and empirical justification for the IBH, link it to debates in psychology and neuroscience, and explore its implications.

We have previously argued that social interaction can play roles in social cognition that are *more than contextual*. By this we mean that interaction dynamics are not data to be decoded and stored by information-processing mechanisms. Rather, the dynamical processes of interaction are complex and can themselves enable socio-cognitive performance or even be a constitutive part of it (De Jaegher et al., 2010—see Herschbach, in press, for a critical discussion of these claims). Such cases of social understanding enabled or constituted by interactive processes can be used to question the widespread assumption that subpersonal “mindreading” mechanisms fulfill a predominant role in *all* of social cognition (what we describe below as the priority of mindreading stance).

The IBH goes further than this questioning, but as an open overarching hypothesis, not a claim. It proposes that social interaction processes play enabling and constitutive roles in the development and in the ongoing operation of brain mechanisms involved in social cognition, *whether the person is engaged in an interactive situation or not*. Accordingly, when an individual interacts with others, the interaction processes would not function merely as perceptual input to ready-made mechanisms but they would also play a role in shaping those mechanisms. The IBH proposes that the neural mechanisms involved in social understanding acquire and sustain their current functionality thanks to past and present engagements in social interaction. In other words, the IBH states that the function of the neural mechanisms involved in social understanding is derivative of the functions of neural mechanisms used in skillful social interaction. It is derivative in the sense that the practice of interaction has forged social understanding mechanisms during development, allowing them to acquire functions that they would otherwise not have, and also in the sense that those mechanisms are in fact a specialization of brain mechanisms used during skillful interaction. This general hypothesis can be translated into specific forms when we consider particular mechanisms, performances, and contexts. It is conceivable even for different competing specific interactive hypotheses to fall within the broader assumptions of the IBH.

The proposal should not be interpreted as negating the existence of a kind of mindreading as a cognitive performance. We acknowledge that interaction is not always present and that people sometimes need to reflect on the behavior of others. Our position is that such reflective performances are not at play always or in general, not that they do not exist or are unimportant. We propose a hypothesis about the origins and function of the mechanisms involved in these and other forms of social understanding. We believe that reflective stances are likely to involve higher level mechanisms and are built upon a variety of embodied skills, including interactive ones, as the IBH proposes.

As hypotheses go, we acknowledge this is a bold one. However, ideas that point in this direction have been suggested before. For instance, Schilbach and colleagues hypothesise that the development of “mentalising” (reasoning about the attributed “mental states” of others) is a function of the “dynamic interplay of social interactions in which the contents of mental states (of oneself or an other) are experienced via *quasi* automatic attunement to others.” This attunement “may then constitute a basic and primary

form of intentionality which predisposes the dyadic nature also inherent in more detached mental representations” (Schilbach et al., 2006, pp. 727–728). The presence of a “dyadic nature” in the activity of detached mindreading, i.e., even when the other person is not in direct engagement with us, comes as no surprise to proponents of constructivist and/or (neo-)Vygotskian or Meadian approaches to the development of theory of mind (ToM) in children (Garfield et al., 2001; Carpendale and Lewis, 2004; Symons, 2004; Fernyhough, 2008; Stone et al., 2012). For these approaches it is the degree of socialization during development that best predicts the child’s capacity for social understanding. Garfield et al. (2001, p. 496) put the claim in no uncertain terms: “the development of language and the development of a set of social skills are prior to, jointly causally sufficient, and individually causally necessary for the acquisition of ToM, in contradistinction both to strongly modular theories of the genesis of ToM and ‘theory theory’ accounts.” Indeed, a large amount of evidence points to the importance of the quantity and quality of linguistic engagements at home and with peers for ToM capacities (Milligan et al., 2007). The IBH attempts to express what these perspectives imply for neural processes. One prediction is that the influence of linguistic social engagements could be manifested differentially in neural activity during social understanding depending on language and cultural background, which is what the evidence suggests (Kobayashi et al., 2007). Similar shaping roles for proto- or non-linguistic social engagements have been proposed to drive the development of shared attention (Racine and Carpendale, 2007), understanding the attention of others, and self-conscious emotions (Reddy, 2003, 2008). More generally, even very early forms of pro-social behavior, such as the maturing of babies’ vocalizations in terms of syllabic structure and faster consonant-vowel transitions, are shaped by social interaction involving contingent emotional feedback by their carers (Goldstein et al., 2003).

The IBH thus fits within a view of social brain function that is neither pre-given nor fully unshaped, and a view of the social world that is not merely a data content for individual cognition. In line with a recent proposal, the IBH sees the brain primarily as an organ of relation (Fuchs, 2011) and less as an organ of detached cognizing. Humans are pre-disposed to engagement in interactions that include the material and social world (e.g., Trevarthen, 1979; Trevarthen and Aitken, 2001; Tucker and Ellis, 2001), which plastically (re-)shape the functionality and structure of the brain. From the IBH perspective we could make better sense of the evidence of the plasticity of the adult human mirror system (Catmur et al., 2007; Heyes, 2010), which is not buffered against radical reconfiguration even after relatively small amounts of training. Or similarly of the evidence indicating that perspective-taking in a visual ToM task can be improved after sensorimotor training in which participants are asked *not* to imitate a finger movement stimulus, thus suggesting that inhibiting imitation responses in one task can transfer to better capabilities to take the point of view of the other in a different task (Santesteban et al., 2012). Adult socio-cognitive mechanisms’ susceptibility to improvement or reconfiguration depending on experiences that are readily available in everyday interactions gives support to the idea that these mechanisms not only *develop* but also *sustain* their functionality in part through participation in social interaction.

Sections “Why Should Neuroscience Take Social Interaction (More) Seriously?,” “The Interactive Brain Hypothesis,” and “Examining the Evidence” of this paper are dedicated to examining the background of the IBH, its formulation and plausibility. Section “Towards a Neuroscience of Social Interaction” addresses different challenges that must be faced for investigating the IBH. In particular, we focus on the challenge of studying social interaction as a dynamical process. We break down the complexity of social interaction into relevant components that may be investigated empirically as independent or dependent variables. These aspects include dynamical transitions in coordination patterns, synergistic effects of interactional autonomy, the emergence of roles and dispositions to interact.

### WHY SHOULD NEUROSCIENCE TAKE SOCIAL INTERACTION (MORE) SERIOUSLY?

Social cognition has traditionally been defined as “information processing in a social setting” (Frith, 2008, p. 2033) and considered the result of a linear process starting from social stimuli, turning them into perceptions of the social world, leading to decisions, and followed by actions (Frith, 2008, p. 2033). Until recently, little attention has been paid to more realistic and widespread scenarios where this linear picture breaks down, i.e., where persons are involved in ongoing, multi-modal sensorimotor loops at various timescales and in addition these loops are modulated by coupling with the sensorimotor loops of other persons. Not only do actions, perceptions and decisions mutually depend on one another, and often happen concurrently within a single individual, but they also interconnect with the actions, perceptions and decisions of others. Similarly, it is also often implied that social information processing is at the basis of all aspects of social interaction from the basic (see e.g., Blakemore and Frith, 2004; Frith and Frith, 2007) to the most sophisticated (Forbes and Grafman, 2010). In this view, social information processing is what allows us to share a social world. Again, the link seems to be questionably unidirectional.

In order to examine the relation between social understanding and social interaction, it is useful first to make a small detour to introduce the distinction between a cognitive performance and the mechanisms that support it. Among other things, cognition involves an engagement of the full agent with the world involving intentions, actions, perceptions, affect, and meaning. This engagement at the personal level should be conceptually distinguished from the subpersonal mechanisms involved, much in the same sense that winning a car race is not the same as having a powerful engine (see Dennett, 1969; Bennett and Hacker, 2003). In this paper we intend this distinction in a pragmatic sense that will help us formulate our hypothesis.

Mechanisms may bear a variety of relations to cognitive performance. Some processes in the agent and in the environment may bear no relation to performance at all, others may be merely contextual, i.e., they introduce variations in the outcome without determining the result (e.g., changes in lighting conditions can affect how efficiently we solve a jigsaw puzzle). Others may enable performance, i.e., without them it would not be possible. These may be historical (e.g., having learned to manipulate the pieces of the puzzle, to group them together according to some

strategy) or contemporaneous (e.g., being in possession of an adequate perceptual system capable of distinguishing shapes and colors). Among the enabling mechanisms, we tend to isolate those that seem to have a higher relevance for the task (e.g., matching complementary patterns in shape and color seems more relevant to solving jigsaw puzzles than the capability to move the pieces around). But the question of which mechanism is more crucial for cognitive performance is often answered intuitively and the distinction is not always well grounded (grouping pieces together according to color or shape can in fact be as important as matching patterns). Answering this question well requires always a careful characterization of the object of investigation and the structure of its context (e.g., what alternative explanations we care about or how we can intervene in practice to alter the results, Garfinkel, 1981). An explanandum that is not well described leads to confusing explanations.

Having established the distinction between performance and mechanisms, we can now examine the current situation in social cognition research. Social cognition has been almost exclusively limited to some version of mindreading, i.e., to the interpretation of the intentions of another person. Mindreading as a performance has in turn been almost exclusively conflated with mechanisms in the individual brain in the form of computational modules postulated to be in charge of interpreting social stimuli and inferring the intentions of the other, assumed as not directly perceivable (Baron-Cohen, 1995). This general view is what we describe as the *priority of mindreading stance*. We may question first whether social cognition, in the larger sense of effective performance involving not only social understanding but also action and affective experience in social situations, must be based exclusively on mindreading. We may, separately, also question whether mindreading itself, as a performance, must be based exclusively on individual brain mechanisms that implement some form of subpersonal “interpreting,” “simulating” or “inferring” (see Gallagher, 2008a). In other words, we question both (1) the centrality of mindreading and (2) the mapping of the structure of performance onto the function of neural mechanisms.

To make sense of these two questions and explore how things could be otherwise, we must first do something that is not often done: say what we mean by *social interaction*. We define social interaction as the autonomous engagement that can emerge between two or more autonomous agents who are mutually regulating their dynamical coupling. We here mean coupling in a dynamical systems sense: i.e., two systems are said to be coupled when parametrical and other structural descriptions of the laws of transformation of states in one of them have a functional dependence on the state variables of the other, which may be non-linear, piece-wise state-dependent, and time-varying (in which case we call the coupling “dynamical”). Coupling can be unidirectional or mutual. When we speak about cognitive agents in interaction, the basis for such a coupling can take various shapes and involve various perceptual systems, sensorimotor flows, neural, and physiological processes, external objects, and technological mediation. Notice that we use the word “autonomous” to describe both the agents and their engagement. Autonomy here is meant in the operational sense used in the enactive literature, involving a self-maintaining organization (for technical definitions see De Jaegher

and Di Paolo, 2007). As such, social interaction goes beyond the mere co-presence, or even the mere dynamical coupling between agents. It requires that the processes of co-regulation of such a coupling become self-sustaining. This definition allows us to make sense of everyday situations where interaction seems to “take a life of its own” in spite of the individual intentions of the participants, and sometimes to their own frustration. This happens through synergistic effects (see, e.g., Kelso, 2009a,b) at the level of individual actions and intentions involving relational bodily variables, such as relative positions and timing between movements, coordination between perceptual systems, and neuro-physiological variables. Such effects can be unintentional, for instance, in the narrow corridor situation when people walking in opposite directions become stuck trying to get past each other, arguments that cannot seem to be avoided, telephone conversations that linger on after having already said goodbye, escalations in intensity of utterances or antagonistic actions, and so on.

This definition impacts on how we conceive social understanding. In developing the enactive approach to intersubjectivity, we have argued that the processes that make up interactive dynamics can be described as processes of *coordination*, *breakdown*, and *recovery of coordination* between the participants at various levels: physiological, bodily, affective, cognitive, etc. (De Jaegher and Di Paolo, 2007). We have proposed that such processes have complex relations to social understanding. In particular, individual “mental states” (those that “do” the understanding and those to be understood) are not fully independent or fully established prior to the interaction, but are instead affected, negotiated, and even created as a result of interaction dynamics. We describe this set of possibilities, much broader than social cognition based on mindreading, as *participatory sense-making*. Everyday social interactions do not bear out that social situations consist in figuring out the “mental states” of others, where these states are hidden, pre-existing, unaffected by the interaction, owned by each individual participant and opaque to the other.

Thus, by defining social interaction in operational terms and tapping into the wider notion of social understanding that is given in everyday experience, we propose a negative answer to the question of whether social cognition as a performance *must* be based on mindreading (question 1). It must not and alternatives exist. Considering social cognition in the light of participatory sense-making allows us to relax the assumptions of linear processing, individual cognitive load, and pre-given, hidden “mental states” that might make mindreading the main factor of social cognition if they were true. Understanding in a social situation can happen through a variety of possibilities, mindreading being only one of them (and according to phenomenological critiques, not even the main one, see Gallagher, 2008a,b, in press).

In addition, we also respond in the negative to the question of whether social cognition as a performance, even mindreading, must be based on subpersonal computational versions of “mindreading” instantiated in brain processes (question 2). The mechanisms of social understanding, enactively construed, involve being engaged in the dynamics of interaction. Collective, relational, and normative social processes instantiated during interactions can play enabling roles in socio-cognitive performance, and

even be a constitutive part of such performance, as recent experiments in perceptual crossing have demonstrated (Auvray et al., 2009; De Jaegher et al., 2010, see also Auvray and Rohde, in press; Lenay and Stewart, 2012). As interactions are studied empirically, such possibilities can be put to the test in order to measure to what extent engagement with others can more parsimoniously be an *explanans* and not always the *explanandum*.

This does not mean that *all* social understanding must be done in actual social interaction—this would also contradict our everyday experience. Neither does it mean that, in some circumstances, a reflective capacity of some kind, *conceived of at the level of performance*, is not a valid hypothesis. What remains open is how interaction and these *mechanisms* relate in such cases. The neural mechanisms presumed to be involved in mindreading scenarios might work differently in interactive situations—most likely not by performing tiny subpersonal inferences or simulations in homuncular fashion. More even, interaction may affect such mechanisms in a more than contextual manner. And even more strongly, we may raise the question of whether social neural mechanisms might not be *primarily interactive* in origin and function, and it is only as a special case that they are put to the task of mindreading. These possibilities will be the basis of the IBH.

## THE INTERACTIVE BRAIN HYPOTHESIS

The IBH is concerned mainly with the brain mechanisms of social understanding. It is formulated with the aim of contributing to experimental social neuroscience but its more general implications also help us re-think social understanding as an act that depends on interactive elements.

The IBH helps landmark an outpost in the logical space of possible explanations in social neuroscience. It describes an extreme possibility, namely that all social brain mechanisms depend on interactive elements either developmentally or in the present, even in situations where there is no interaction. This may turn out to be true only in some cases, or even in none.

The IBH comes in two versions:

**Developmental (DIBH):** The functions of individual brain mechanisms involved in social understanding have been shaped during development by skillful engagements in social interactions where interactive processes have been involved in social performance in a more than contextual way.

**Contemporaneous (CIBH):** Even in the absence of immediate interaction, the functions of brain mechanisms enabling social understanding are derived *contemporaneously* from functions used primarily in skillful social interactions where interactive processes are involved in social performance in a more than contextual way.

In both versions, the IBH concerns the functionality of individual neural mechanisms and their role in social understanding. The IBH is not a hypothesis about performances.

We use the term *contextual* in the sense indicated in Section “Why Should neuroscience Take Social Interaction (More) Seriously?” For interactive processes to be involved in cognitive performance in at most a contextual way means that their role in the outcome is one of “data”: i.e., the most we can establish

is that variations in interactive factors introduce variations in *outcome* of the cognitive performance without affecting its *functionality*. To play more than a contextual role means that the interactive elements are more than data and become *enabling* (necessary) factors in determining not only the outcome but also the functionality of the processes underlying performance. The hypothesized involvement of interactive factors, both developmentally and contemporaneously, in shaping or sustaining neural mechanisms involved in social understanding is thus not trivial (as it would have been if we were concerned with merely contextual influences on the performance outcome introduced by aspects of interaction).

To avoid misinterpretations, it is important to emphasize what the IBH *does not* say. It does not claim that individual brain mechanisms play no important role in social understanding, even during engaged social interactions. On the contrary, it is concerned with the interactive origins and aspects of the functioning of neural mechanisms, because it recognizes them as essential. The involvement of these mechanisms in the explanation of various socio-cognitive phenomena is always something to be determined case by case rather than assumed *a priori* as is usually done. Neither does the IBH suggest that among the developmental precursors or among the current functional components of a given neural mechanism involved in social understanding we will find *only* interactive elements as enabling or constitutive factors. Several non-interactive factors and functional elements will undoubtedly also be required for the explanation of the functionality of brain mechanisms in social cognition. The IBH simply makes the non-trivial proposal that among the necessary factors, we will *always* find some enabling or constitutive interactive elements.

Let's consider the context for the IBH. As we mentioned, we denote the assumption that some form of mindreading has priority over all forms of social cognition as the priority of mindreading stance. In particular, this position holds “that mindreading facilitates interaction rather than the other way around. On this view, mindreading has priority in the logic of how we interact with others; we first observe, then infer the other's beliefs, and only then, on this basis, engage in interaction” (Gallagher, in press). This follows the staged view that we have mentioned at the beginning of Section “Why Should Neuroscience Take Social Interaction (More) Seriously?”: “Signals arising from the environment impinge upon us. Sensations [...] are turned into perceptions [...]. Then, decisions are made about what should best be done in response to these perceptions [...]. Actions are planned and finally output is initiated in the form of motor movements [...]. Within this general framework of stimulus and response, we can have a subset of processes concerned with social stimuli (e.g., reading facial expressions), social decisions (Should I trust this person?) and social responses (making facial expressions)” (Frith, 2008, p. 2033). This position is seldom argued for explicitly and yet it is uncritically adopted very often. It can be found in the opening lines of many studies adopting the passive observer's view: “To successfully navigate the social world, we need to decode a dynamic stream of complex information [to] infer other humans' mental states, such as desires, intentions, emotions, and thoughts” (Wolf et al., 2010,

p. 894). “Understanding and predicting other people’s mental states and behavior are important prerequisites for social interactions” (David et al., 2008, p. 279). In spite of the observational stance adopted by the experimental evidence supporting the mindreading perspective (e.g., false-belief tests), “theory theorists have always emphasized that the primary use of mindreading is in interaction with others” (Carruthers, 2009, p. 167).

The position relies on an argument by default: “there is simply no other way (than using theory-driven computations of underlying mental states) of explaining our competence in this domain” (Carruthers, 2009, p. 47). The position also implies a representationalist view according to which mutual understanding involves the sharing of hidden “mental representations” (without attempting to specify what they are): “For successful interactions [we] need to share representations of the world” (Frith and Frith, 2007, p. R727). “Human social interactions crucially depend on the ability to represent other agents’ beliefs even when these contradict our own beliefs, leading to the potentially complex problem of simultaneously holding two conflicting representations in mind” (Kovács et al., 2010, p. 1830). The priority of mindreading stance also sees interaction as a discrete chain of informational exchanges subserving the goal of passing “representations” “from transmitter to receiver” (Singer et al., 2003, p. xxii).

Given that, in sharp contrast to this picture, dynamic, concurrent, multi-timescale, and non-staged social interaction processes can sometimes be shown to be the main explanatory factor for social cognitive performance with no mindreading involved (see Auvray et al., 2009; De Jaegher et al., 2010) and that not all of social understanding involves sharing “representations” of hidden “mental states” (Gallagher, 2008a, in press), the priority of mindreading claim is demonstrably false. However, the stance may still survive as a general heuristic for research.

The purpose of the IBH is to investigate the alternative possibility that neural mechanisms subserving and shaped by interaction have (developmental and/or functional) priority over those subserving mindreading. The DIBH states that having the capacity to interact with others skilfully and having (perhaps specific) experiences of interaction is necessary for the development of all kinds of social cognition, without putting any conditions on the kind of individual mechanisms involved during acts of social cognition in the present. The CIBH states that, whatever the developmental path taken, neural mechanisms supporting social cognition in the present functionally depend on mechanisms that are used during interactive engagement also in the present; in other words, social cognition, including non-interactive mindreading, makes crucial use of interactive mechanisms. Both versions question the priority of mindreading.

It would be possible for the DIBH to be true without the CIBH being true. The developmental paths that lead to brain mechanisms involved in social understanding could depend crucially on interactive experiences and yet the mechanisms themselves could function in the present without involving interactive elements. And the opposite is also the case, at least logically: the CIBH could be true without the DIBH being true. Social cognition mechanisms in general could always involve mechanisms that are used in interactive situations, only that their development

has not depended crucially on having undergone interactive experiences where interaction plays more than a contextual role. In other words, this would describe a situation in which the function of socio-cognitive mechanisms, all of them involving interactive elements, has not been shaped by past interactions (i.e., interactive experiences have played at most a contextual role as data). Admittedly, this possibility seems less plausible and would involve a strictly nativist perspective. It is not so controversial to claim that at least some aspects of how we engage with others are developed precisely as we perfect the skill of interacting in actual encounters, and that the corresponding individual mechanisms are shaped accordingly. If this is accepted, then holding the CIBH true also implies accepting the DIBH. In practice, the CIBH is the stronger of the two versions.

The IBH in both its forms is open to empirical scrutiny and falsifiable. Given that social understanding can have multiple components, it is also possible to distinguish general and particular versions of the IBH. The general version holds the IBH true of all forms of social understanding. The particular version concerns the role of interactive mechanisms in specific capabilities (e.g., understanding at different levels the actions of others, their beliefs, their expressions, their goals, their relations to others, their personality traits, taking the other’s perspective, etc.).

## EXAMINING THE EVIDENCE

Is there any evidence supporting the plausibility of the IBH? Let’s consider the extreme case where the hypothesis would seem least applicable: non-interactive mindreading. The question of the developmental and operational relation between mechanisms that support mindreading and those that support skillful interaction is very much under-studied. The majority of research is concerned with mindreading capabilities outside interactive contexts and in so far as mechanisms are discussed, research in cognitive psychology or social neuroscience is mostly focused on postulating functionality based on computational requirements (e.g., social contingency detection modules, Gergely and Watson, 1999), or based on localization of neural activation in the absence of interaction.

Certain neural and developmental evidence hints at the plausibility of the IBH, without being conclusive. A few studies have indicated that different neural circuits seem to be activated depending on the presence or absence of specific interactive elements, such as situations of conflict, being addressed, or elicitation to interact. Such elements seem to modulate “interpretational” mechanisms like those involved in mindreading. In a study of monkeys with different levels of dominance in the social hierarchy sharing a social space, Fujii et al. (2007) have found differential activation in the parietal cortex, in circuits supposedly involved in understanding the actions of others, according to whether the configuration presents a conflict of interests or not. During imaging studies in humans using virtual characters, Schilbach et al. (2006) have found that the interpretation of social content in the stimuli relates specifically to the activation of the ventral medial prefrontal cortex, whereas the experience of self-involvement (present in interactive situations but not generally in passive interpretation) recruits in addition neural activity in a more dorsal part of the medial prefrontal cortex, which has

been suggested to be involved in more general tasks requiring self-reference. And different imaging studies looking at the response to stimuli that do or do not display an elicitation to interact, in gestures (Lotze et al., 2006), vocalizations (Dietrich et al., 2007), facial and bodily expressions (Lawrence et al., 2006), and visual scenes (Tylén et al., 2009) also reported differential activation in brain regions normally associated with verbal language when communicative intent was apparent.

These studies suggest that the function of mechanisms involved in observational social understanding is modulated by the presence of interactive elements such as conflict in a shared space, self-involvement or communicative intent. If this evidence had been otherwise, then at first sight it might be compatible with some version of the priority of mindreading stance, given that the mechanisms used to interpret social stimuli would have been unaffected by interactive elements. Indeed, the unspoken reliance on this priority is apparent in interpreting cases that show a *coincidence* of activation in brain regions due to interactive stimuli (calls to attention involving the self in observed gaze direction in static images or upon hearing one's own name) and mindreading (Kampe et al., 2003). The fact that a coincidence has been found in this case between interactive and mindreading neural circuits is readily seen as evidence that we need to mentalise in order to understand social stimuli like a gaze directed toward the self. The authors suggest that: "mentalizing is involved in understanding the signals that a sender emits to initiate communication with someone. It is likely that we attribute mental states such as beliefs, desires, and intentions to the sender while guessing the meaning of these signals" (Kampe et al., 2003, p. 5262). However, the conclusion is unwarranted since it is clear that the same evidence can equally be interpreted in IBH terms, perhaps even more parsimoniously as there is no experiential evidence that stimuli directed to the self are accompanied by guesses about their meaning and "mental state" attributions (Gallagher, 2008b, in press). Confronted with this evidence, the IBH perspective would suggest that, on the contrary, it is mindreading that recruits functionality which is otherwise used in interactive contexts, such as, for instance, understanding person-oriented attitudes in others based on our own experience of having been the object of their attention.

It is clear, then, that activation of the same brain areas cannot differentiate between the two interpretations (IBH and priority of mindreading). Evidence of differential activation between interpretational and interactional situations is at least suggestive of mindreading mechanisms not being necessarily involved or not playing the same roles during interaction as they do during detached interpretation. This evidence should moderate the readiness to interpret data and design experiments putting mindreading first. The priority of mindreading, uncritically assumed in the last example, is what the IBH questions and at least this questioning is supported by evidence of differential activation. However, such evidence by itself cannot count in favor of the positive part of the IBH which establishes a priority for mechanisms that support interactive capabilities. For this, it would have to be established that interaction is involved in shaping mindreading mechanisms or that such mechanisms are a specialized case of interactive mechanisms *also* in situations where

interaction is absent. To the best of our knowledge such evidence has not been produced so far and will have to wait for more systematic investigation involving actual interactions as well as a clear proposal in terms of the functional interrelation between the mechanisms involved going beyond mere correlational activation. Nevertheless, some existing support pointing in this direction may be found by considering the developmental aspects of both mindreading and interactive capabilities.

Evidence of mindreading in infants under 3 or 4 years of age is still controversial in spite of recent studies indicating behavioral and attentional sensitivities in false-belief situations indicating violation of expectation or anticipated looking in infants as young as 7–17 months of age (Onishi and Baillargeon, 2005; Surian et al., 2007; Song et al., 2008; Kovács et al., 2010; Southgate et al., 2010—see Gallagher, in press for parsimonious behavioral and enactive interpretations according to which this evidence does not require "mental state" attribution). In contrast, the capability of infants to skillfully engage in affective, richly rhythmic and intentional interaction from birth or very early in life has been undisputed for some time (Bullowa, 1979; Trevarthen, 1979; Murray and Trevarthen, 1985). This difference notwithstanding, support for the IBH must be sought in the potential interactive roots or nature of the postulated precursors for mindreading, as we shall attempt next, and not simply in the fact that interactive abilities appear earlier developmentally.

In what follows it is important not to understand the idea of a precursor as a necessary stage in a developmental progression that is later overcome and does not continue to play important roles once more sophisticated skills have been established. Instead a "precursor" may involve mechanisms that remain active throughout the lifespan (Gallagher, in press).

One proposed precursor of mindreading, the capability to understand the attention of others (Baron-Cohen, 1995), is believed to be the result of the development of joint attention toward the end of the first year. However, it has been argued to have much earlier roots in the infant's understanding of the other's attention directed toward the self during interaction (Reddy, 2003). Infants of about 2 months of age are able to respond with smiles or coyness and become more expressive when adults make eye contact with them, and show the opposite emotional responses when adults stop attending. By 4 months, they attempt to engage the adult's attention with vocalizations and by initiating "games." After 6 months, infants are able to specifically regulate their responses with respect to the attention others give to their actions, engaging in exaggerated performances to attract attention, eliciting praise, laughter and challenging the expectations of others by teasing them. As Reddy argues, the infant's grasp of the relation between the other's gaze and the object of visual attention is enabled by their intimate experience of having been themselves the object of the other's attention in the past. This experience happens in interaction (see also Reddy and Morris, 2004; Reddy, 2008). Similarly, in an extensive review of the literature on the development of shared attention, Racine and Carpendale (2007) conclude that capacities such as pointing and social referencing are evidenced in interactive shared practices earlier than the infant's purported understanding of others' "mental states." The developmental evidence identifies the

practices that are shaped through the infant's interactions with others, in particular in their affective engagements, as prior to the development of shared attention. However, as Racine and Carpendale recognize, this identification does not explain *how* interactions shape these capacities, which is what research into the IBH could help uncover.

Another proposed precursor for mindreading is the capability for imitation (Meltzoff and Moore, 1998). Although imitation in infancy remains the topic of ongoing debate (see Hurley and Chater, 2005a,b for representative positions; also Ray and Heyes, 2011), in the context of the IBH it is interesting to note that imitation has primarily been discussed in terms of motor-based accounts of how we understand the goal-directed actions of others to produce our own. The IBH would predict that such accounts are likely to involve mechanisms whose functionality has been shaped during interaction (Froese et al., *in press*). Recent evidence indeed suggests that mirror system functionality in humans is forged by experience. Catmur et al. (2007) have shown that the mirror system in adults is easily re-adjusted plastically to produce “counter-mirror” responses after training with incompatible sensorimotor stimuli. This demonstrates that the human mirror system is highly plastic (effects were measured after three training sessions lasting 45 min each)—(see also Catmur et al., 2008, 2009, 2011).

The plasticity of the mirror system responses is in itself not unexpected. They depend on such factors as the level of performance skill in the actions being observed, as shown in the case of ballet dancers (Calvo-Merino et al., 2006), the use of chopsticks (Järveläinen et al., 2004) and other tools (Ferrari et al., 2005). However, it could be claimed that the effect of experience on mirror neurons is merely *contextual* according to our definition (the intensity of neural responses co-varies with the intensity of experience, e.g., the amount of training). But the data of Catmur and colleagues indicates a stronger role for experience. One would expect for incompatible experience to at most diminish the strength of mirroring responses, not to reverse their “meaning” after a short training (i.e., not, as in their case, for the observation of a foot-lift to elicit motor responses associated with lifting the hand). This indicates that sensorimotor experience functionally re-shapes the mirror system, thus playing an *enabling* role in determining its involvement in social cognition.

These results give support to claims that the mirror system is the outcome of associative learning involving correlated observation and execution of actions either through spatio-temporal contiguity (Keysers and Perrett, 2004; Del Giudice et al., 2009) or through sensorimotor contingency (Heyes, 2001, 2010). The proposal would seem to remain plausible even in the light of counter-arguments involving evidence of newborn imitation (Ray and Heyes, 2011). If fulfilling a functional role, the plasticity of mirror responses is suggestive of a system able to adapt to social engagements that are potentially changing rapidly. This would imply that, to function effectively, such mechanisms are constantly being adjusted by interactive experience.

Does this evidence support the IBH? The experience necessary for enabling the development of “mirror” mechanisms is clearly available in the social world of the infant. What is not immediately clear is whether this experience is primarily interactive or

merely observational. Evidence points in the interactive direction. The behavioral effects of counter-mirror training are stronger in the presence of contingency between stimulus and action than in cases that also involve a neutral, non-contingent stimulus (Cook et al., 2010). This indicates that the most likely and reliable source of experience shaping mirror function are situations of social contingency involving close links between observations and actions, i.e., situations of social interaction. Not only this, but interactive situations in combination with the associative learning hypothesis can also explain the development of “mirror” responses demanding complementary actions (Newman-Norlund et al., 2007). These are likely to be more present than imitative action matching in general, and increasingly so as the infant develops what we call a readiness to interact (Section “Towards a Neuroscience of Social Interaction”) and interactions acquire more complexity.

As a point of clarification, we repeat that the IBH does not claim that non-interactive factors play no role in the development of socio-cognitive neural function. As such, while it gains support from the associationist perspective on “mirror” responses, the IBH is not immediately contradicted by counter-hypotheses based on evidence of mirror activity very early in life (Lepage and Théoret, 2007; Gallese et al., 2009). The IBH can accommodate these alternatives, as long as they indicate not a nativist position, but the otherwise undisputed presence of pro-social pre-dispositions *in utero*. And even then interactive factors are not easily discarded. For instance, in a study of twin pregnancies, Castiello et al. (2010) have found that the kinematic profiles of limb movements already at the 14th week of gestation is different depending on whether they are aimed at the wall of the uterus or at the other twin and that the proportion of movements directed at the sibling increases in the following weeks. While the authors suggest that this is evidence of a “pre-wiring” for social interaction, the evidence is inconclusive. The role of interactive experience cannot be easily discounted even in this case; after all there is another twin also moving and touching the self. Moreover, the *increase* of movements directed toward the sibling may indicate the presence of interactive learning (a pre-wired mechanism would predict an equal intensity from the start).

As regards the developmental evidence, what matters is whether interactive experience plays a forging, enabling role—but not necessarily a fully determining one—in shaping the functionality of socio-cognitive neural mechanisms. It is not necessary, then, to interpret the IBH in an externalist way, but rather as describing a dialectical scenario involving social dynamics in the cognitive-emotional development and sustaining of social understanding. In this scenario, interactive experience and the mechanisms involved in interaction co-develop with non-interactive mechanisms. They mutually shape each other's development and efficacy, resulting in an integrated set of social skills that could not have existed without interaction.

## TOWARD A NEUROSCIENCE OF SOCIAL INTERACTION

The complexity of social interaction makes its study potentially very rich, but also methodologically challenging. The patterns and structures of social interaction have long been the focus of numerous studies in social psychology (Kendon, 1990), sociology (Goffman, 1963, 1967; Garfinkel, 1967), conversation



analysis (Sacks et al., 1974; Goodwin, 1981), psychiatry and psychotherapy (Bateson et al., 1956; Watzlawick et al., 1967). Though perhaps not immediately applicable, a lot of the accumulated expertise in these fields will still be very relevant for studies of interaction in social neuroscience. Particularly relevant are studies of dynamical patterns in interpersonal coordination (Richardson et al., 2007; Marsh et al., 2009; Shockley et al., 2009; Riley et al., 2011), joint action (Sebanz et al., 2006), mother infant interaction (Bullowa, 1979; Trevarthen, 1979; Jaffe et al., 2001) and agent-based computational modeling (Di Paolo, 2000; Quinn et al., 2003; Di Paolo et al., 2008; Froese and Di Paolo, 2010).

Social interaction is “a co-regulated coupling between at least two autonomous agents, where: (1) the co-regulation and the coupling mutually affect each other, constituting an autonomous self-sustaining organization in the domain of relational dynamics and (2) the autonomy of the agents involved is not destroyed (although its scope can be augmented or reduced)” (De Jaegher et al., 2010, p. 442–443). Social interaction is a complex dynamical pattern of different forms of coordination between the activity of two or more agents mutually affecting each other. Accordingly, the most challenging aspect of studying interaction in controlled experiments is its unpredictability, rendering it seemingly more amenable to naturalistic observation and analysis. Nevertheless, it is possible to identify aspects of interaction that can be empirically manipulated in systematic ways. Here we review some of them.

## COORDINATION TRANSITIONS

Interpersonal coordination can happen at the level of bodily movement (Richardson et al., 2007; Marsh et al., 2009; Shockley et al., 2009, etc.), posture (Varlet et al., 2011), physiological variables, such as heart rates and breathing patterns (McFarland, 2001; Müller and Lindenberger, 2011), autonomic responses (Ebisch et al., 2012), and EEG patterns (Tognoli et al., 2007; Lindenberger et al., 2009; Dumas et al., 2010; Naeem et al., 2012). It happens spontaneously and sometimes, as expected from the autonomy of social interaction, even against the individual intention not to coordinate (Schmidt and O’Brien, 1997; Issartel et al., 2007). Coordination may involve the performance of similar movements (rocking chairs, tapping) or the timing of more complex actions, not necessarily similar to each other (Newman-Norlund et al., 2007; Riley et al., 2011). Interestingly, it may also be aperiodic, as in the case of two people reading from a text together (Cummins, 2011). It may be absolute (perfect entrainment) or relative (more inconstant and fluid distribution of variables over time exhibiting coherence and phase attraction). A case of relative coordination would be an adult and a child walking together at the same speed despite their natural differences in stride length (Kelso, 1995).

From the enactive perspective, in which we create and transform meaning together (we participate in each other’s sense-making), what is particularly interesting are *transitions* in coordination. It is at the transitions, like coordination breakdowns and recoveries, that our interactions take a different direction, and we with it. This goes together with changes in individual involvement, and in making sense of the situation. The different ways in which these transitions can take place describe a spectrum of possibilities for participating in sense-making. They

range from orientation of individual attention and affect to joint sense-making where actions and intentions are co-constructed in the interaction (rendering the interactor’s intentions anything but opaque to each other).

In view of this, simply measuring coordination should not be the end goal of interactive neuroscience. We propose that a fruitful approach will be to study the transition patterns in their relation to meaning, affect and intention, either as they occur spontaneously or by experimental manipulation. In particular, not much data exist about the neural involvement and effects that precede, co-occur and follow these transitions (as opposed to periods of established coordination). The stability of coordination patterns and their transitions can be studied experimentally by introducing perturbations to the coupling between the interaction participants and analysing the effects on social understanding. Such perturbations may range from the manipulation of noise and time delays along the channels of interaction, to more sophisticated methods along the lines of Virtual Partner Interactions (Kelso et al., 2009) or animated virtual characters (Pfeiffer et al., 2011) where various parameters influencing coupling strength may be varied. These kinds of investigation will bear a direct relevance to the IBH because they will help identify the relation between aspects of the interaction and social understanding and the corresponding role for brain mechanisms.

The empirical investigation of transitions in coordination can be done via two kinds of approaches. Dynamically, transitions in coordination can occur between coordination patterns and the absence of coordination, or they may involve changes between absolute and relative coordination. These qualitative differences can be measured with traditional dynamical systems techniques (Kelso, 1995, 2009a,b; Riley et al., 2011) as well as with measures of long-term correlations able to reveal different qualities of interaction couplings (van Orden et al., 2003, 2005; Kello et al., 2010). It is important to bear in mind the need to identify the relevant collective variables or order parameters at a given level of description. In parallel, in terms of the significance in the interactive context, transitions can involve changes in relation between the interactors, for instance, changes from imitative to complementary action, or between symmetric and asymmetric roles. Dual EEG (e.g., Dumas et al., 2010; Naeem et al., 2012) could be used to explore these questions by measuring the fine temporal structure of neural events prior to and just after behavioral breakdowns and the re-establishment of coordination.

Even in complex unconstrained interactions, coordination-breakdown-recovery patterns and their relation to sense-making can be studied, for instance, by measuring transitions in non-verbal synchrony. Using motion energy analysis to study inter-bodily synchrony in psychotherapy in over 100 recorded interactions, Ramseyer and Tschascher (2011) have found that high levels of non-verbal coordination correlate with patients reporting good relationship quality and experiencing high self-efficacy as well as higher symptom reduction as evaluated by the therapist. Thermal IR imaging could track autonomic responses in similarly unconstrained situations and our proposal can be easily adapted to verify whether *trends* in these therapeutic variables are predicted by the amount of transitions in and out of bodily and physiological synchrony.

Coordination can be modulated through different dimensions of an interaction. We find an example of this in the narrow corridor situation, where two persons walking in opposite directions get stuck in a symmetrical situation in their attempts to get past each other. The unintended coordination is sustained at the level of body movements (moving left or right at the same time). In order to get out of this unintended coordination they can wait until it spontaneously breaks down, or they can try to unlock it from another level, so to speak, by intervening with something like saying “after you” or a gesture to that effect. Here, the spoken or gestural coordination intervenes and breaks down the movement coordination. Indeed, it has been found that non-social multimodal situations (involving proprioception, touch and sound) induce a higher likelihood of transitions in coordination (for instance, in the case of rhythmic tasks, transitions between phase locking and phase drift) (Lagarde and Kelso, 2006). Experimental designs that study multimodal coordination can be adapted to interpersonal situations in social tasks that involve different channels of sensorimotor coupling.

### AUTONOMY OF INTERACTION

During interaction, periods of coordination can orient individual attitudes, actions and intentions with a trend to sustaining a relational configuration. In turn the configuration can facilitate certain forms of coordination but not others. As a result, periods of engagement can have a distinct dynamical signature and it is at this point that interaction patterns can play important roles in social understanding and coordinate and shape individual mechanisms. Not recognizing the point at which an interaction “gets going” can lead to methodological problems. For instance, evolutionary robotics models suggest that a plausible explanation of the infant’s lack of interest in the delayed video image of her mother in Murray and Trevarthen’s (1985) double TV monitor experiment is the dynamical stability of the engaged interaction pattern during the live condition (Di Paolo et al., 2008; De Jaegher et al., 2010). Rochat et al. (1998) reported not being able to replicate the original results, a failure that was likely due to the fact they recorded the first minute of the live condition for use during the replay whether or not engagement had been established. Allowing engagement to develop in the live condition, however, leads to a confirmation of the original findings (Nadel et al., 1999) as the interactive explanation predicts. Simply putting participants in an interactive configuration is no guarantee that a social interaction will take place.

How to recognize and understand the effects of engagement? One aspect of the self-organization of social interaction is the presence of synergistic effects. These effects result from the relational configuration of attitudes, intentions, and actions of the participants and may be promoted by the situation and past history of interactions. Their dynamical signature is often a reduction of dimensionality in the system (Kelso, 2009a; Riley et al., 2011) and increased mutual predictability between inter-personal variables. Such effects may or may not be in line with individual intentions. It is often the case that participants are not aware of such synergies and may misattribute the origin of these effects to the other participants.

Examples of synergistic effects involve situations of escalation (often found in arguments that recur to everyone’s frustration). Force escalation is demonstrated in a simple interactive experiment (Shergill et al., 2003). Participants are asked to activate in turn a lever that produces pressure on the other person’s hand with the same amount of force as perceived in the previous round. Due to a systematic bias in underestimating one’s own force, participants perceive the force exerted by the other as stronger and respond by increasing their pressure in the next round, resulting in an unwanted escalation.

Synergistic effects thus rely on individual mechanisms that find themselves paired into self-sustaining, sometimes paradoxical relations. We already mentioned the narrow corridor situation where the very act of trying to get past the person coming from the opposite direction provokes simultaneous symmetrical moves to the side resulting in the perpetuation of the configuration. Or the rules of politeness that sometimes overcome the interactors’ individual intentions to end the interaction. A striking example is the perceptual crossing experiment by Auvray et al. (2009, see also Auvray and Rohde, *in press*). In this case, while the strategies used by the participants in trying to discriminate virtual objects controlled by the other are individually insufficient to solve the task, their collective pairing achieves the intended result.

### THE SPECTRUM OF PARTICIPATION

The self-organization of interaction has two sides depending on whether we focus on the collective pattern or on an individual participant. From the latter’s point of view, synergistic effects are often experienced as demands for specific forms of participation and the (not always intended) taking-up of specific roles. A key aspect for neuroscientific research is that a participant is different from an observer. A participant cannot fully control her own flow and timing of perceptions and actions and has to respond to demands made by the actions of other participants. Otherwise the interaction breaks down. Of course, a participant also places constraints and demands on others, resulting in a situation of mutual influence and co-adaptation. This is a very unusual situation in terms of what has been investigated under the observational paradigm (Van Overwalle, 2009).

Participation is rarely strictly symmetric and depends on social context, task, norms, and history. Interactive roles (e.g., leader and follower) may be pre-established, but it is often the case that they emerge during interaction and vary at different points during the engagement (together with transitions in modes of coordination). The emergence of roles does not require explicit channels of meta-communication. In a study of haptically coupled cooperating dyads moving a heavy crank toward a target, Reed et al. (2006) found that “dyads specialized such that one member contributed more to acceleration and the other to deceleration” of the movement (p. 366) without any interaction channel other than the movement itself.

Even well defined roles (like a pre-established agreement on who is going to lead) require subtle and ongoing mutual confirmation in the form of a sustained engagement. If the follower cannot or will not follow, the leader’s role immediately vanishes.

The different possibilities ranging from pre-established fixed roles, to emerging temporary or durable roles all the way to

symmetrical situations mark what we call a *spectrum of participation*. Interactive situations can occupy different positions along this spectrum and the quality of the interactive patterns will depend on this (Konvalinka et al., 2010; Noy et al., 2011). Effects on social performance and on the function of individual neural mechanisms are also likely to depend on this factor. In the context of the IBH this suggests that studying interaction as an element that could shape or involve individual neural mechanisms for social understanding cannot be done without awareness of the position within this spectrum of participation, which may be manipulated or at least measured using statistical tools to determine the influence between interpersonal variables.

### READINESS TO INTERACT

Once we start taking social interaction seriously, it is possible to interpret some evidence in terms of what we call a *readiness to interact*. We characterize this as a disposition to engage or participate in socially meaningful situations, which range from perceiving a stimulus that presents another person (e.g., a portrait, a film, a voice on the radio), to full-blown interactions. Dispositions to interact can result from previous interactions. This is shown in finger tapping studies, where natural individual frequencies for tapping in the absence of interaction are moved closer together after a period of visual interaction involving synchronized tapping (see Oullier et al., 2008—similar effects have been found in evolutionary robotics models of social coordination, Di Paolo, 2000). Manifestations of readiness to interact include expectancies of social contingencies and anticipatory dispositions during communication (Jordan, 2009), the “eye-contact effect” by which perceived direct gaze from others modulates socio-cognitive performance (Senju and Johnson, 2009), understanding the possibilities that fit the collective affect of a social situation (like the “mood” of a meeting), or the embodied and affective “pull” to respond that we experience when a direct demand is made on us.

As a disposition elicited by a socially meaningful situation, readiness to interact can play roles in social understanding analogous to the role played by the mastery of the law of sensorimotor contingencies in the sensorimotor theory of (non-social) perceptual experience (O'Regan and Noë, 2001). Accordingly, the possibilities for action afforded by an object and our bodies is sedimented in dispositions that depend on how bodily movement and sensation co-vary when the object is skilfully engaged with. We directly perceive, for instance, the “meaning” of a cup's handle because it affords holding for raising the cup for drinking and this perception results from past experience, our bodily structure, and (often ignored in the sensorimotor approach), social and cultural normativity. According to O'Regan and Noë, the elicitation of these predisposed potentialities co-constitutes the perceptual act together with the actual actions taken by the perceiver. In a similar (though not identical) fashion, socially meaningful situations can be understood in terms of the potential for interactive involvement they elicit, even in situations where such involvement is not actualized.

Social understanding thus becomes intertwined with social interactive skills (McGann and De Jaegher, 2009). Their mastery, however, is subject to different laws than those of sensorimotor

engagement with an object. The co-variation of perception and action in the social realm is regulated by at least two autonomous agents, and so mastering the sensorimotor laws involved is likely to require more flexibility than in the case of inanimate objects. The difference lies in that objects can generally be perceptually determined by sensorimotor engagement, whereas other persons always remain to some extent indeterminable. It is precisely this indeterminacy that help us recognize them as autonomous persons. The mastery of the regularities in skillful interaction with others comes from patterns of transitions in coordination that leave a mark throughout our interaction history. We want to highlight that, since even the manipulation of objects is influenced by socio-cultural normativity, it is likely that there is a continuum of flexibility in “law-like” couplings, from how we engage with objects to how we interact with other subjects.

Readiness to interact can contribute to explaining the development of social understanding. In a study of false belief understanding in 2-year-olds, an experimenter put an interesting toy in one of two containers on a high shelf (out of reach of the child), either in the presence of his parent or in their absence (O'Neill, 2005). It was found that when the parent was absent while the toy was put away, upon her return the children named the toy and its location significantly more often and gestured in its direction more intensely when asking the parent to retrieve the toy than when the parent had been present. Results like this are often interpreted in terms of mindreading: the child infers a knowledge state on the basis of a perceptual state, and this informs his subsequent actions. However, these findings can also be interpreted in terms of readiness to interact, which predicts that the disposition to interact differs in each situation. In the case where the parent has been absent, the infant has individually attended to the toy, and when the parent arrives, he can now interact with the parent. What has to happen is for the child to orient the parent's sense-making toward the toy. This involves vocalizations and gestures that extend his interactional possibilities so that the parent finds the toy and gives it to him. In the case where the parent is present during the putting away of the toy, infant and parent have both and together attended to what happened. They, therefore, experienced a disposition to interact that already involved the object (the infant could potentially, as in similar past situations, convey a desire to play with the toy and the parent could bring it closer). The infant does not have to orient the parent to the toy after it is hidden because both, in interacting together, experience the change in the interaction potentialities as the experimenter hides the toy. For this reason, the gestures and vocalizations are less intense. This explanation does not involve postulating hidden internal “mental states,” but is based on what is available to each participant, namely their mutual attention, what happens to the toy, and the interactive potentialities and actualities that change in the situation. Similar enactive explanations of very young infants' response to false-belief situations are given by Gallagher (in press). They involve, in our view, the concepts of readiness to interact and of changing possibilities for interaction.

An empirically relevant possibility that is raised by taking readiness to interact into account is that dispositions vary not only with the *actual* stimulus but with the *potential* interactive

possibilities that the situation affords. Thus, observing a picture of a person gazing at us is not the same as observing a real person gazing at us. There is a difference in general in the intensity of the readiness to interact in each case (the image is understood as an image and so elicits a readiness to interact that is generally less intense—though not fully absent—than observing a real person who could actually, and not just potentially, be looking at us). This could explain the differences in neural response to direct or averted gaze when presented with a picture or with a live person stimulus (Pönkänen et al., 2011) or the difference in discriminating between human and object movements when presented live or via a TV screen (Shimada and Hiraki, 2006). The use of virtual characters that may or may not be controlled by a real person (Pfeiffer et al., 2011) could prove a useful technique for investigating variations in readiness to interact brought about by the presence of interactive contingency.

Having emphasized that readiness to interact can be modulated by the actual presence of an interactor, and in general by the ecological and social significance of the experimental situation, it is worth noting that of the different interactive factors mentioned so far, it is more amenable than others to being investigated using the passive observer paradigm. Dispositions to interact may be evidenced in the excitability of motor neurons. In a recent study that we argue shows support for the IBH, Sartori et al. (2011), have investigated how social context modulates corticospinal excitability indicating a covert disposition to respond to a social gesture. Using transcranial magnetic stimulation (TMS) and recording motor-evoked potentials (MEPs) in the hand muscles of participants during the passive observation of a video sequence, this study demonstrates two important aspects of readiness to interact: (1) its contextually varying time-course and (2) the significance of a pre-attuned social context. The video sequence shows a person extending her arm to grasp either an apple with the whole hand or an almond with a precision grip, then moving the object to a plate and following this action by extending the arm with an open hand toward the viewer. In some cases, another object is present in the direction of the extended hand but out of reach and the gesture can readily be interpreted as asking the viewer to hand over the object (either an almond or an apple). The amplitude of the MEPs in different hand muscle groups indicates a disposition to imitate the observed grip during the first part of the video (the grasping). This is followed, when an out-of-reach object is present, by a disposition toward executing the grip corresponding to the out-of-reach object in the second part, in preparation to potentially handing the object over. If instead of an outstretched arm, the viewer is shown an arrow indicating the out-of-reach object (and no person in view), the MEPs are significantly lower in amplitude than if shown the social gesture, and not distinguishable from the case when the objects are shown by themselves with no arrow. This suggests that the change from imitative to complementary action dispositions is contingent on perceiving the social gesture toward an object and not likely to be mediated by inferential mechanisms. The latter would presumably act in a similar fashion in the presence of the arrow, a well-known indexical pointer (the action admittedly would be understandable as “grasping” only given the context of the experiment since the same participants were also

exposed to the outstretched arm request condition). Thus, the result implicitly supports the IBH over a mindreading alternative.

The disposition to interact tracks the time-varying social and ecological context and is strongly modulated by the pre-existing social significance and sensitive to the potential for interactive engagement. Readiness to interact, the result of previous interactive practice, parsimoniously explains the data in a situation where inferential mechanisms, because they amount to cognitive overkill, are unable to generate the differential effect. Based on what we have suggested above, we can predict that the strength of the disposition to interact, which increases when the stimulus is changed from an arrow indicating an object to the image of a request gesture for the same object, will further increase if the gesture is made by a contingently animated virtual character or indeed by a real person.

Readiness to interact can also be measured indirectly by looking at interference effects when participants are instructed either to perform imitative or complementary actions to those passively observed (Newman-Norlund et al., 2007; van Schie et al., 2008). When the context of imitation or complementarity is disrupted by a different cue (e.g., a number or a color indicating the performance of one specific action, whatever the context), then difficulty in responding to out-of-context commands can indicate the strength of the disposition to perform the contextually suggested action. However, unlike the study by Sartori et al. (2011), the tasks performed in many of these studies rely less on socially significant gestures and are likely to uncover dispositions that are formed during the training and execution of the experimental task itself.

Our readiness to interact with others connects interactive experiences and skills on the one hand, and situations with a social meaning on the other. Therefore, of the different interactive factors we have mentioned in this section, it might bear most directly on the dependence of social understanding mechanisms on interactive elements postulated by the IBH. As we interpret the behavior of someone we are observing but not interacting with (e.g., a character in a film), we could, given the appropriate circumstances, be in a situation where such an interaction was possible. Readiness to interact marks our sensitivity to this potentiality. Our dispositions in such cases are embodied, they are sometimes even bodily felt, and can modulate our social understanding. We conjecture that an enactive explanation of reflective social understanding as a performance is likely to draw significantly on the concept of readiness to interact.

Ideally, none of the aspects of interaction mentioned in this section and their roles in social understanding should be studied exclusively in a one-person paradigm. Interactions should be studied live using methodologies like hyperscanning or thermal cameras that allow the simultaneous recording of neurophysiological activity during relatively unconstrained engagements. However, manipulating interaction dynamics can still be methodologically challenging. For this reason, we would like to emphasize that, at least in connection with the DIBH, the effects of interactive experience on individual neural mechanisms can also be investigated “after the fact” by more traditional comparative methods (for instance, applying the methods used by Cook et al., 2010, but using interactive situations as training)

and that several of the above aspects of interactions may be approached in this manner.

## CONCLUSION

Some recent discussions on embodied approaches to social cognition have reduced the role of the body to that of formatting so-called neural “representations” although no effort is made to clarify what this term could mean (Goldman and de Vignemont, 2009; Gallese and Sinigaglia, 2011). Such approaches remain neuro-centric and individualistic. In contrast, the enactive approach foregrounds a different notion of the living body (of which the brain is a part) in its ongoing sense-making relation to the world. According to this approach, the brain is understood as embedded, not in a protective and nourishing casing, but in ongoing circular processes of sense-making that pass through it, the body and the world; in other words it is understood as a mediating organ (Fuchs, 2011) with all the implications that this view has for the study of brain function.

In the case of social understanding the human body lives in the social world and among embodied others. The multiple phenomena in this social world serve not just as the objects, but also as the sustainers of different forms of sense-making and modes of participation. The enactive approach does not neglect the brain; it emphasizes the living body and the world of significance that enable, shape and constrain brain function (Cosmelli and Thompson, 2010). We have proposed the IBH as an attempt to articulate these relations between world, body and brain for the case of social understanding. Our focus has been on a concept that mediates between all these elements: social interaction.

We have shown that until recently social interaction has been neglected in mainstream social neuroscience in everything but in name and that the majority of research in this field has adopted the priority of mindreading stance. Consequently, to propose that social interaction can play shaping or constitutive roles in social understanding and, more strongly, to hypothesize that interactive elements shape and may even constitute socio-cognitive neural mechanisms is, in our opinion, far from trivial. The traditional picture is turned on its head and the reflective performances that were thought to be the fertile basis for all of social understanding are now recast as dependent on interactive skills and mechanisms. We have indicated some of the empirical directions that derive from taking the IBH seriously. They include investigating transitions in coordination, the autonomy and synergies of interaction patterns, the emergence of and transitions between different modes of participation and the role of social dispositions, skills and readiness to interact.

## REFERENCES

- Auvray, M., Lenay, C., and Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New Ideas Psychol.* 27, 32–47.
- Auvray, M., and Rohde, M. (in press). Perceptual crossing: the simplest online paradigm. *Front. Hum. Neurosci.*
- Baron-Cohen, S. (1995). *Mind-blindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.
- Bateson, G. (1972). *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. Chicago: University Of Chicago Press.
- Bateson, G., Jackson, D. D., Haley, J., and Weakland, J. (1956). Towards a theory of schizophrenia. *Behav. Sci.* 1, 251–264.
- Benjamin, J. (1988). *The Bonds of Love: Psychoanalysis, Feminism and the Problem of Domination*. New York, NY: Pantheon.
- Bennett, M. R., and Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. Oxford: Wiley-Blackwell.
- Blakemore, S. J., and Frith, U. (2004). How does the brain deal with the social world? *Neuroreport* 15, 119–128.
- Brüne, M. (2003). “Social cognition and behaviour in schizophrenia,” in *The Social Brain: Evolution and Pathology*, eds M. Brüne, H. Ribbert,
- For reasons of space, in making the case for the IBH we have focused on certain socio-cognitive phenomena (understanding the actions of others) to the neglect of other important aspects. Among these we can briefly mention social affect. Self-conscious emotions, such as shame and guilt, make little sense in the absence of the experience of the other as someone capable of recognizing us as autonomous agents (Reddy, 2008)—in analogy to the act of giving, which cannot be completed by a single person alone. Recognition is manifested in interactions, as are neglect, admiration, desire, pity, love, and hatred. These affective phenomena are not “carried” over the interaction channels, but are themselves modes of the interactive experience of connectedness, as well as ways in which interaction dynamics vary. They are also a consequence of the dialectics of recognition and domination that emerge from the potential conflict between individual autonomies at the heart of our definition of social interaction. To the extent that this is the case, engaging with others is key to the development and sustaining of our emotional lives (Benjamin, 1988).
- Other areas that may benefit from investigating the IBH include research into the social etiological aspects of psychopathologies like schizophrenia (see Bateson et al., 1956; Bateson, 1972; Brüne, 2003; Burns, 2006) and autism (Hobson, 2002), and the role of language and dialogical processes (including implications for cognitive functions, such as planning and formal reasoning) (Garfield et al., 2001; Carpendale and Lewis, 2004; Symons, 2004; Fernyhough, 2008).
- Our focus on social interaction does not mean, to say it once more, that we suggest that non-interactive factors play no important role in explanations of social understanding. Our enactive proposal is participatory and dialectical: there cannot be interaction without individual participants whose roles, skills and higher forms of autonomy and cognition could not exist without and are shaped by social interaction.

## ACKNOWLEDGMENTS

This work is supported the Marie-Curie Initial Training Network, “TESIS: toward an Embodied Science of InterSubjectivity” (FP7-PEOPLE-2010-ITN, 264828). Hanne De Jaegher is supported by the Marie-Curie Fellowship “INDYNAUTS, Interaction dynamics and autonomy in social cognition,” (FP7-PEOPLE-2009-IEF, 253883) and by the Spanish Government project: *El concepto de autonomía en bioética e investigación biomédica*, MICINN (FFI2008-06348-C02-02/FISO). Thanks to the reviewers and to Marieke Rohde, Marek McGann, and Jacqueline Nadel for useful comments that helped improve the manuscript.

- and W. Schiefenhövel (Chichester: Wiley), 277–314.
- Bullowa, M. (ed.). (1979). *Before Speech. The Beginning of Interpersonal Communication*. Cambridge: Cambridge University Press.
- Burns, J. (2006). The social brain hypothesis of schizophrenia. *World Psychiatry* 5, 77–81.
- Castiello, U., Becchio, C., Zoia, S., Nelini, C., Sartori, L., Blason, L., D'Ottavio, G., Bulgheroni, M., and Gallese V. (2010). Wired to be social: the ontogeny of human interaction. *PLoS ONE* 5:e13199. doi: 10.1371/journal.pone.0013199
- Calvo-Merino, B., Grezes, J., Glaser, D. E., Passingham, R. E., and Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Curr. Biol.* 16, 1905–1910.
- Carpendale, J. I. M., and Lewis, C. (2004). Constructing an understanding of mind: the development of children's social understanding within social interaction. *Behav. Brain Sci.* 27, 79–151.
- Carruthers, P. (2009). Mindreading underlies metacognition. *Behav. Brain Sci.* 32, 164–182.
- Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M., and Heyes, C. (2008). Through the looking glass: counter-mirror activation following incompatible sensorimotor learning. *Eur. J. Neurosci.* 28, 1208–1215.
- Catmur, C., Mars, R., Rushworth, M., and Heyes, C. (2011). Making mirrors: premotor cortex stimulation enhances mirror and counter-mirror motor facilitation. *J. Cogn. Neurosci.* 23, 2352–2362.
- Catmur, C., Walsh, V., and Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Curr. Biol.* 17, 1527–1531.
- Catmur, C., Walsh, V., and Heyes, C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 2369–2380.
- Cook, R., Press, C., Dickinson, A., and Heyes, C. (2010). Acquisition of automatic imitation is sensitive to sensorimotor contingency. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 840–852.
- Cosmelli, D., and Thompson, D. (2010). "Embodiment or envatment? Reflections on the bodily basis of consciousness," in *Enaction: Towards a New Paradigm for Cognitive Science*, eds J. Stewart, O. Gapenne, and E. Di Paolo (Cambridge, MA: MIT Press), 361–385.
- Cummins, F. (2011). Periodic and aperiodic synchronization in skilled action. *Front. Hum. Neurosci.* 5:170. doi: 10.3389/fnhum.2011.00170
- David, N., Aumann, C., Santos, N. S., Bewernick, B. H., Eickhoff, S. B., Albert Newen, A., Shah, N. J., Fink, G. R., and Vogeley, K. (2008). Differential involvement of the posterior temporal cortex in mentalizing but not perspective taking. *Soc. Cogn. Affect. Neurosci.* 3, 279–289.
- De Jaegher, H. (2009). Social understanding through direct perception? Yes, by interacting. *Conscious. Cogn.* 18, 535–542.
- De Jaegher, H., and Di Paolo, E. (2007). Participatory sense-making: an enactive approach to social cognition. *Phenom. Cogn. Sci.* 6, 485–507.
- De Jaegher, H., Di Paolo, E., and Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends Cogn. Sci.* 14, 441–447.
- Dennett, D. (1969). *Content and Consciousness*. London: Routledge.
- Del Giudice, M., Manera, V., and Keysers, C. (2009). Programmed to learn? The ontogeny of mirror neurons. *Dev. Sci.* 12, 350–363.
- Di Paolo, E. (2000). Behavioural coordination, structural congruence and entrainment in a simulation of acoustically coupled agents. *Adaptive Behav.* 8, 27–48.
- Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenol. Cogn. Sci.* 4, 97–125.
- Di Paolo, E. A., Rohde, M., and Izuka, H. (2008). Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing. *New Ideas Psychol.* 26, 278–294.
- Di Paolo, E. A., Rohde, M., and De Jaegher, H. (2010). "Horizons for the enactive mind: values, social interaction, and play," in *Enaction: Towards a New Paradigm for Cognitive Science*, eds J. Stewart, O. Gapenne, and E. A. Di Paolo (Cambridge, MA: MIT Press), 33–87.
- Dietrich, S., Hertrich, I., Alter, K., Ischebeck, A., and Ackermann, H. (2007). Semiotic aspects of non-verbal vocalizations: a functional imaging study. *Neuroreport* 18, 1891–1894.
- Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Inter-brain synchronization during social interaction. *PLoS ONE* 5, 278–288. doi: 10.1371/journal.pone.0012166
- Eibisch, S. J., Aureli, T., Bafunno, D., Cardone, D., Romani, G. L., and Merla, A. (2012). Mother and child in synchrony: thermal facial imprints of autonomic contagion. *Biol. Psychol.* 89, 123–129.
- Fernyhough, C. (2008). Getting Vygotskian about theory of mind: mediation, dialogue and the development of social understanding. *Dev. Rev.* 28, 225–262.
- Ferrari, P. F., Rozzi, S., and Fogassi, L. (2005). Mirror neuron responding to observation of actions made with tools in monkey ventral premotor cortex. *J. Cogn. Neurosci.* 17, 212–226.
- Forbes, C. E., and Grafman, J. (2010). The role of the human prefrontal cortex in social cognition and moral judgment. *Annu. Rev. Neurosci.* 33, 299–324.
- Frith, C. D. (2008). Social cognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 2033–2029.
- Frith, C. D., and Frith, U. (2007). Social cognition in humans. *Curr. Biol.* 17, R724–R732.
- Froese, T., and Di Paolo, E. (2010). Modeling social interaction as perceptual crossing: an investigation into the dynamics of the interaction process. *Connect. Sci.* 22, 43–68.
- Froese, T., Lenay, C., and Ikegami, T. (in press). Imitation by interaction: a new proposal based on the dynamics and phenomenology of social interaction. *Front. Hum. Neurosci.*
- Fuchs, T. (2011). The brain—a mediating organ. *J. Conscious. Stud.* 18, 196–221.
- Fujii, N., Hihara, S., and Iriki, A. (2007). Dynamic social adaptation of motion-related neurons in primate parietal cortex. *PLoS ONE* 2:e397. doi: 10.1371/journal.pone.0000397
- Gallagher, S. (2008a). Inference or interaction: social cognition without precursors. *Philos. Explor.* 11, 163–173.
- Gallagher, S. (2008b). Direct perception in the intersubjective context. *Conscious. Cogn.* 17, 535–543.
- Gallagher, S. (in press). In defense of phenomenological approaches to social cognition: interacting with the critics. *Rev. Philos. Psychol.*
- Gallese, V., Rochat, M., Cossu, G., and Sinigaglia, C. (2009). Motor cognition and its role in the phylogeny and ontogeny of intentional understanding. *Dev. Psychol.* 45, 103–113.
- Gallese, V., and Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends Cogn. Sci.* 15, 512–519.
- Garfield, J. L., Peterson, C. C., and Perry, T. (2001). Social cognition, language acquisition and the development of theory of mind. *Mind Lang.* 16, 494–541.
- Garfinkel, A. (1981). *Forms of Explanation: Rethinking the Questions in Social Theory*. New Haven: Yale University Press.
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
- Gergely, G., and Watson, J. S. (1999). "Early socio-emotional development: contingency perception and the social-biofeedback model," in *Early Social Cognition*, ed P. Rochat (Hillsdale NJ: Erlbaum), 101–137.
- Goffman, E. (1963). *Behavior in Public Places: Notes on the Social Organization of Gatherings*. New York, NY: The Free Press.
- Goffman, E. (1967). *Interaction Ritual: Studies in Face-to-Face Behavior*. Chicago: Aldine.
- Goldman, A., and de Vignemont, F. (2009). Is social cognition embodied? *Trends Cogn. Sci.* 13, 154–159.
- Goldstein, M. H., King, A. P., and West, M. J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8030–8035.
- Goodwin, C. (1981). *Conversational Organization: Interaction between Speakers and Hearers*. New York, NY: Academic Press.
- Heyes, C. (2001). Causes and consequences of imitation. *Trends Cogn. Sci.* 5, 253–261.
- Heyes, C. (2010). Where do mirror neurons come from? *Neurosci. Biobehav. Rev.* 34, 575–583.
- Hersbach, M. (in press). On the role of social interaction in social cognition: a mechanistic alternative to enactivism. *Phenom. Cogn. Sci.*
- Hobson, R. P. (2002). *The Cradle of Thought*. London, Macmillan.
- Hurley, S., and Chater, N. (2005a). *Perspectives on Imitation. Volume 1: Mechanisms of Imitation and Imitation in Animals*. Cambridge, MA: MIT Press.
- Hurley, S., and Chater, N. (2005b). *Perspectives on Imitation. Volume 2: Imitation, Human Development and Culture*. Cambridge, MA: MIT Press.
- Issartel, J., Marin, L., and Cadopi, M. (2007). Unintended interpersonal coordination: 'Can we march to the beat of our own drum?' *Neurosci. Lett.* 411, 174–179.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., and Jasnow, M. D. (2001). *Rhythms of Dialogue in Infancy: Coordinated Timing in Development*. Oxford: Blackwell.
- Järveläinen, J., Schürmann, M., and Hari, R. (2004). Activation of the human primary motor cortex

- during observation of tool use. *Neuroimage* 23, 187–192.
- Jordan, J. S. (2009). Forward-looking aspects of perception–action coupling as a basis for embodied communication. *Discourse Process*. 46, 127–144.
- Kampe, K. K. W., Frith, C. D., and Frith, U. (2003). “Hey John”: signals conveying communicative intention toward the self activate brain regions associated with “mentalizing,” regardless of modality. *J. Neurosci.* 23, 5258–5263.
- Kello, C. T., Brown, G. D. A., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., and van Orden, G. C. (2010). Scaling laws in cognitive sciences. *Trends Cogn. Sci.* 14, 223–232.
- Kelso, J. A. S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge, MA: MIT Press.
- Kelso, J. A. S. (2009a). “Synergies: atoms of brain and behavior,” in *Progress in Motor Control: A Multidisciplinary Perspective*, ed D. Sternad (New York, NY: Springer), 83–91.
- Kelso, J. A. S. (2009b). “Coordination Dynamics,” in *Encyclopedia of Complexity and System Science*, ed R. A. Meyers (Springer: Heidelberg), 1537–1564.
- Kelso, J. A. S., de Guzman, G. C., Reveley, C., and Tognoli, E. (2009). Virtual partner interaction (VPI): exploring novel behaviors via coordination dynamics. *PLoS ONE* 4:e5749. doi: 10.1371/journal.pone.0005749
- Kendon, A. (1990). *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge: Cambridge University Press.
- Keyser, C., and Perrett, D. I. (2004). Demystifying social cognition: a hebbian perspective. *Trends Cogn. Sci.* 8, 501–507.
- Kobayashi, C., Glover, G. H., and Temple, E. (2007). Cultural and linguistic effects on neural bases of ‘theory of mind’ in american and japanese children. *Brain Res.* 1164, 95–107.
- Konvalinka, I., Vuust, P., Roepstorff, A., and Frith, C. (2010). Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Q. J. Exp. Psychol.* 63, 2220–2230.
- Kovács, Á. M., Téglás, E., and Endress, A. D. (2010). The social sense: susceptibility to others’ beliefs in human infants and adults. *Science* 330, 1830–1834.
- Lagarde, J., and Kelso, J. A. S. (2006). Binding of movement, sound and touch: multimodal coordination dynamics. *Exp. Brain Res.* 173, 673–688.
- Lawrence, E. J., Shaw, P., Giampietro, V. P., Surguladze, S., Brammer, M. J., and Davis, A. S. (2006). The role of “shared representations” in social perception and empathy: an fMRI study. *Neuroimage* 29, 1173–1184.
- Lenay, C., and Stewart, J. (2012). A minimalist approach to perceptual interactions. *Front. Hum. Neurosci.* 6:98. doi: 10.3389/fnhum.2012.00098
- Lepage, J. F., and Théoret, H. (2007). The mirror neuron system: grasping other’s actions from birth? *Dev. Sci.* 10, 513–529.
- Lindenberger, U., Li, S., Gruber, W., and Müller, V. (2009). Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neurosci.* 10, 22.
- Lotze, M., Heymans, U., Birbaumer, N., Veit, R., Erb, M., Flor, H., and Halsband, U. (2006). Differential cerebral activation during observation of expressive gestures and motor acts. *Neuropsychologia* 44, 1787–1795.
- McFarland, D. H. (2001). Respiratory markers of conversational interaction. *J. Speech Lang. Hear. Res.* 44, 128–143.
- McGann, M., and De Jaegher, H. (2009). Self-other contingencies: enacting social perception. *Phenomenol. Cogn. Sci.* 8, 417–437.
- Marsh, K. L., Richardson, M. J., and Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Top. Cogn. Sci.* 1, 320–339.
- Meltzoff, A. N., and Moore, M. K. (1998). “Infant intersubjectivity: broadening the dialogue to include imitation, identity and intention,” in *Intersubjective Communication and Emotion in Early Ontogeny*, ed S. Braten (Cambridge: Cambridge University Press), 47–62.
- Milligan, K., Astington, J. W., and Dack, L. A. (2007). Language and theory of mind: a meta-analysis of the relation between language ability and false-belief understanding. *Child Dev.* 78, 622–646.
- Müller, V., and Lindenberger, U. (2011). Cardiac and respiratory patterns synchronize between persons during choir singing. *PLoS ONE* 6:e24893. doi: 10.1371/journal.pone.0024893
- Murray, L., and Trevarthen, C. (1985). “Emotional regulations of interactions between two-month-olds and their mothers,” in *Social Perception in Infants*, eds T. M. Field and N. A. Fox (Norwood: Alex), 177–197.
- Nadel, J., Carchon, I., Kervella, C., Marcelli, D., and Réserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. *Dev. Sci.* 2, 164–173.
- Naeem, M., Prasad, G., Watson, D. R., and Kelso, J. A. S. (2012). Electrophysiological signatures of intentional social coordination in the 10–12 Hz range. *Neuroimage* 59, 1795–1803.
- Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M., and Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nat. Neurosci.* 10, 817–818.
- Noy, L., Dekel, E., and Alon, U. (2011). The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20947–20952.
- Onishi, K. H., and Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science* 308, 255–258.
- O’Neill, D. K. (1996). Two-year-old children’s sensitivity to a parent’s knowledge state when making requests. *Child Develop.* 67, 659–677.
- O’Regan, J., and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 939–1031.
- Oullier, O., de Guzman, G. C., Jantzen, K. J., Lagarde, J., and Kelso, J. A. S. (2008). Social coordination dynamics: measuring human bonding. *Soc. Neurosci.* 3, 178–192.
- Pfeiffer, U. J., Timmermans, B., Bente, G., Vogeley, K., and Schilbach, L. (2011). A non-verbal Turing test: differentiating mind from machine in gaze-based social interaction. *PLoS ONE* 6:e27591. doi: 10.1371/journal.pone.0027591
- Pönkänen, L. M., Alhoniemi, A., Leppänen, J. M., and Hietanen, J. K. (2011). Does it make a difference if I have an eye contact with you or with your picture? An ERP study. *Soc. Cogn. Affect. Neurosci.* 6, 486–494.
- Quinn, M., Smith, L., Mayley, G., and Husbands, P. (2003). Evolving controllers for a homogeneous system of physical robots: structured cooperation with minimal sensors. *Philos. Trans. R. Soc. Lond. A* 361, 2321–2343.
- Racine, T., and Carpendale, J. I. M. (2007). The role of shared practice in joint attention. *Br. J. Dev. Psychol.* 25, 3–25.
- Ramseyer, F., and Tschascher, W. (2011). Nonverbal synchrony in psychotherapy: coordinated body-movement reflects relationship quality and outcome. *J. Consult. Clin. Psychol.* 79, 284–295.
- Ray, E., and Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. *Dev. Sci.* 14, 92–105.
- Redcay, E., Dodell-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D., and Saxe, R. (2010). Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. *Neuroimage* 50, 1639–1647.
- Reddy, V. (2003). On being the object of attention: implications for self-other consciousness. *Trends Cogn. Sci.* 7, 397–402.
- Reddy, V. (2008). *How Infants Know Minds*. Cambridge, MA: Harvard University Press.
- Reddy, V., and Morris, P. (2004). Participants don’t need theories: knowing minds in engagement. *Theory Psychol.* 14, 647–665.
- Reed, K., Peshkin, M., Hartmann, M. J., Grabowczyk, M., Patton, J., and Vishton, P. M. (2006). Haptically linked dyads: are two motor-control systems better than one? *Psychol. Sci.* 17, 365–366.
- Richardson, M. J., Marsh, K. L., Isenhower, R., Goodman, J., and Schmidt, R. C. (2007). Rocking together: dynamics of intentional and unintentional interpersonal coordination. *Hum. Mov. Sci.* 26, 867–891.
- Riley, M. A., Richardson, M. J., Schokley, K., and Ramenzoni, V. C. (2011). Interpersonal synergies. *Front. Psychol.* 2:38. doi: 10.3389/fpsyg.2011.00038
- Rochat, P., Neisser, U., and Marian, V. (1998). Are young infants sensitive to interpersonal contingency? *Infant. Behav. Dev.* 21, 355–366.
- Sacks, H., Schegloff, E. A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language* 50, 696–735.
- Santisteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., and Bird, G. (2012). Training social cognition: from imitation to theory of mind. *Cognition* 122, 228–235.
- Sartori, L., Cavallo, A., Bucchioni, G., and Castiello, U. (2011). Corticospinal excitability is specifically modulated by the social dimension of observed actions. *Exp. Brain Res.* 211, 557–568.
- Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., 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.

- Schmidt, R. C., and O'Brien, B. (1997). Evaluating the dynamics of uninvited interpersonal coordination. *Ecol. Psychol.* 9, 189–206.
- Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends Cogn. Sci.* 10, 70–76.
- Senju, A., and Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends Cogn. Sci.* 13, 127–134.
- Shergill, S. S., Bays, P. M., Frith, C. D., and Wolpert, D. M. (2003). Two eyes for an eye: the neuroscience of force escalation. *Science* 301, 187.
- Shimada, S., and Hiraki, K. (2006). Infant's brain responses to live and televised action. *Neuroimage* 32, 930–939.
- Shockley, K., Richardson, D. C., and Dale, R. (2009). Conversation and coordinative structures. *Top. Cogn. Sci.* 1, 305–319.
- Singer, T., Wolpert, D. M., and Frith, C. D. (2003). "Introduction: the study of social interactions," in *The Neuroscience of Social Interaction: Decoding, Imitating, and Influencing the Actions of Others*, eds C. D. Frith and D. M. Wolpert (Oxford: Oxford University Press), xiii–xxvii.
- Song, H. J., Onishi, K. H., Baillargeon, R., and Fisher, C. (2008). Can an agent's false belief be corrected by an appropriate communication? Psychological reasoning in 18-month-old infants. *Cognition* 109, 295–315.
- Southgate, V., Chevallier, C., and Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Dev. Sci.* 13, 907–912.
- Stone, J. E., Carpendale, J. I. M., Sugarman, J., and Martin, J. (2012). A meadian account of social understanding: taking a non-mentalistic approach to infant and verbal false belief understanding. *New Ideas Psychol.* 30, 166–178.
- Surian, L., Caldi, S., and Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychol. Sci.* 18, 580–586.
- Symons, D. K. (2004). Mental state discourse, theory of mind, and the internalization of self-other understanding. *Dev. Rev.* 24, 159–188.
- Thompson, E. (2007). *Mind in Life Biology, Phenomenology and the Sciences of Mind*. Cambridge, MA: Harvard University Press.
- Tognoli, E., Lagarde, J., DeGuzman, G. C., and Kelso, J. A. (2007). The phi complex as a neuromarker of human social coordination. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8190–8195.
- Trevarthen, C. B. (1979). "Communication and cooperation in early infancy: a description of primary intersubjectivity," in *Before Speech*, ed M. Bullowa (Cambridge: Cambridge University Press), 321–348.
- Trevarthen, C. B. (2005). First things first: infants make good use of the sympathetic rhythm of imitation, without reason or language. *J. Child Psychother.* 31, 91–113.
- Trevarthen, C. B., and Aitken, K. J. (2001). Infant intersubjectivity: research, theory, and clinical applications. *J. Child Psychol. Psychiatry* 42, 3–48.
- Tucker, M., and Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Vis. Cogn.* 8, 769–800.
- Tylén, K., Wallentin, M., and Roepstorff, A. (2009). Say it with flowers! An fMRI study of object mediated communication. *Brain Lang.* 108, 159–166.
- van Orden, G. C., Holden, J. G., and Turvey, M. T. (2003). Self-organization of cognitive performance. *J. Exp. Psychol. Gen.* 132, 331–350.
- van Orden, G. C., Holden, J. G., and Turvey, M. T. (2005). Human cognition and 1/f scaling. *J. Exp. Psychol. Gen.* 134, 117–123.
- Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 30, 829–858.
- van Schie, H. T., van Waterschoot, B. M., and Bekkering, H. (2008). Understanding action beyond imitation: reversed compatibility effects of action observation in imitation and joint action. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1493–1500.
- Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Varlet, M., Marin, L., Lagarde, J., and Bardy, B. G. (2011). Social postural coordination. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 473–483.
- Watzlawick, P., Beavin, J., and Jackson, D. D. (1967). *Pragmatics of Human Communication. A Study of Interactional Patterns, Pathologies, and Paradoxes*. New York, NY: Norton.
- Wolf, I., Dziobek, I., and Heekeren, H. R. (2010). Neural correlates of social cognition in naturalistic settings: a model-free analysis approach. *Neuroimage* 49, 894–904.

**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: 27 February 2012; accepted: 22 May 2012; published online: 07 June 2012.

Citation: Di Paolo E and De Jaegher H (2012) The interactive brain hypothesis. *Front. Hum. Neurosci.* 6:163. doi: 10.3389/fnhum.2012.00163

Copyright © 2012 Di Paolo and De Jaegher. 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.