

CHAPTER 3

The neuroscience of social understanding

John Barresi and Chris Moore

How do we understand and engage with the purposeful, emotional and mental activities of other people and how does this knowledge develop? What can recent work on mirror neurons in monkeys and human beings teach us about how the brain supports social understanding? According to Intentional Relations Theory (Barresi and Moore 1996), the understanding of the self-other equivalence requires concurrent knowledge of mind from both a first- and a third-person point of view and that any mental concept must directly match and link these two ways of knowing it. In this chapter we will argue that Intentional Relations Theory is consistent with and can help interpret recent neurophysiological findings on “mirror neurons” that fire equivalently for intentional relations (i.e., object-directed actions, emotions, and mental activities) of self and other.

1. Introduction

Human beings, like many other social animals, spend an enormous amount of time engaged in activities that require quick adjustments to socially transmitted information. By observing others we learn to adapt effectively to changes in the environment as well as to the actions and reactions of our social peers. How do we do it? To what extent do we need to *understand* the mental processes governing our own and others' actions or can we function socially based on simple mechanisms by which we come to *share* psychological states with others, without understanding them? In other words, to what extent does a skill at *mind sharing* function as a form of social understanding well before we come to a level of *mind understanding*? Furthermore, how do these two capacities – mind-sharing and mind-understanding – relate to each other?

In the Theory of Mind (ToM) approach to social understanding emphasis is placed on sophisticated abilities to understand *mental* states – in particular the ability to attribute representational mental states such as beliefs to self and other. It is the ability to attribute false beliefs that is taken as a hallmark of the specifically human form of mentalistic social understanding that characterizes a “theory of

mind". However, social understanding is a more general phenomenon that occurs both in many social species that seem to have no ToM of this kind, and in children well before the late preschool period when the understanding of false beliefs develops. Indeed, an early form of social understanding is evidenced essentially from birth as neonates show a particular sensitivity to human social stimuli. We suggest that the kinds of social sensitivity observed in infants as well as in many social animals should be seen as forms of *non-reflective* social understanding, dependent on an array of mechanisms that yield an ability to share mental states with others without necessarily recognizing that those shared mental states are in fact attributable to individual agents. A satisfactory account of the development of social understanding will require an explanation of how these original mechanisms that enable early social responsiveness combine with later developing skills to yield more sophisticated forms of intersubjectivity. In parallel, such an account must specify how engaging in shared understanding or shared mental activities with others facilitates the later more individualistic understanding of mind.

In the present chapter we will approach these problems with a focus on recent findings in the neuroscience of social understanding. With the discovery in monkeys of pre-motor "mirror neurons" that respond to the actions of others as well as to their own motor plans, there is reason to believe that even monkeys somehow understand actions of both self and others in a similar object-directed way. But should such a "common code" between perception and action be treated merely as an instance supporting the common coding hypothesis (Prinz 1997; Knoblich and Jordan 2002) or a more elaborate understanding for what we have called action intentional relations (Barresi and Moore 1996)? Even if it seems unlikely that monkeys represent these actions as full-blown mental events involving conscious intentions of the other, distinct from their own, it is still a question of how simple is their understanding here and how it connects to more elaborate forms of social understanding. Perhaps their understanding occurs more simply as sharing in the goal-directed nature of the activity of the other by entering into a comparable goal-directed pre-motor state, while not themselves engaging in the activity. Such a sub-personal level of understanding of the action of another would in effect convert it into a first-person representation of one's own actions, but it would not yet represent that action as what we call an intentional relation, involving a representation of an agent as well as the object-directed action. Nevertheless, such sub-personal matching between goal-directed actions of self and other provides a basis for eventual understanding of full blown intentional relations that can be applied to agent-oriented actions directed at objects at a personal level, whether of self or other. We believe that this is the way that these phenomena should be understood and that this matching between aspects of the observing monkey's intentional relations (IRs) and the IRs of others provides evidence for the *matching*

hypothesis that we have previously proposed as the basis of social understanding. We believe that our general account, which we have named Intentional Relations Theory (IRT) is superior to alternative accounts of the origins and development of intersubjectivity, and in the present chapter will bring neuroscientific evidence involving humans as well as monkeys to support our position.

2. The matching problem of social understanding and three approaches to intersubjectivity

A fundamental aspect of human social understanding is what we have previously referred to as ‘self-other equivalence’. Human beings understand self and other to be essentially the same kind of thing – namely a human agent or person that can engage in a variety of intentional relations with objects or states of affairs. This aspect of human social understanding is quite obvious and passes unnoticed in commonsense psychology and yet it hides a significant epistemic problem. How can we attribute the *same meaning* to actions of other individuals that we attribute to our own actions when the *third-person information* that we have of the actions of others is radically different from the *first-person information* that we have of our own actions? The information we get about others’ actions is apparently information about the overt aspects of behavior, while the object towards which the action is directed is often not obvious (or even opaque in the case of mental states such as beliefs). In contrast the information we get about our own actions is apparently information about our orientations towards the objects and events we witness or imagine but does not typically include information about ourselves as the actor or agent being so oriented. So how are these qualitatively different forms of information recognized to be tokens of the same type – expressions of intentional relations between an agent (self or other) and some object or state of affairs? In the recent history of research on social understanding, there are three fundamentally different answers to this question.

According to the ‘theory theory’ (TT) approach humans have innately, or acquire early in development, a ToM mechanism that can be applied uniformly to self and other based purely on inference from *behavior* (e.g., Gopnik 1993; Leslie 1987). Self-other equivalence in this account is based on the fact that one can interpret one’s own behavior in the same way that one can interpret the behavior of others. For instance, consider an example of what we have called an emotional intentional relation – the case of *love*. Since *love* is a public concept, whose main criterion of application is supposedly based on behavior, a person can know when she or another person is in love by noticing the same kinds of behavior of self and other directed toward the object of love.

In contrast, Simulation theorists (ST) would take a different view from TT on how a person knows about her own love versus another person's love (e.g., Goldman 1992; Gordon 1986; Humphrey 1984; Harris 1989). On their view, *love* may have some behavioral consequences that can be used to identify it in another person, but it is fundamentally a subjective mental state, and without a personal appreciation of the "feeling state" that usually goes with the overt behavior, we cannot truly understand *love* as a psychological state. We understand *love* "directly" in our own case, but only indirectly and by simulation in the case of another person. We must *imagine* what someone else feels when we observe their behavior in context (e.g., around the object of love), in order to understand the psychological, intentional, and subjective meaning of their behavior. In our own case, our behavior is a consequence of this subjective state, so no inference is necessary from our own behavior to the mental state that we are in. Although we need to reflect on these states to categorize them, we do not need knowledge of comparable states in other people to form these categories and concepts.

A third kind of theory invokes the notion of matching or sharing attitudes or psychological states between self and other and is represented in a range of different accounts (e.g., Gallagher and Hutto this volume; Gallese, Keysers and Rizzolatti 2004; Hobson 1991, 1998; Hobson and Hobson this volume; Wilson and Knoblich 2005; Zlatev this volume). Although the various theories in this third group can all be considered to invoke some form of intersubjectivity – understood widely as involving matched or shared mental states between or among individuals – they vary on the extent to which they provide an account of the foundations or the origins of intersubjectivity and on the processes by which infants are hypothesized to move from forms of intersubjective sharing of mental states to understanding that self and other are persons or selves that might have distinct mental states. Several of these theories (e.g., Gallagher and Hutto this volume; Hobson and Hobson this volume) invoke Trevarthen's (e.g., Trevarthen and Hubley 1978) concepts of "primary intersubjectivity" and "secondary intersubjectivity" to describe early phases of development. However, while the capacity for mind-sharing is evident in these forms of intersubjectivity, what isn't clear is how the infant moves from sharing mental states with others to understanding mental phenomena as distinct and possibly different in self and other. In the case of Gallagher and Hutto, this latter form of understanding is thought to rely on the acquisition of language and of the differentiating roles of self and other in situated narratives, some of which involve folk psychological terms.

Our own Intentional Relations Theory (Barresi 2001, 2004; Barresi and Moore 1996; Moore 1999, 2006, 2007) does not differ substantially from these accounts in its interpretation of the early phases of development of social understanding that involve mind-sharing through processes that produce interpersonal matching

of self and other. However, it differs from these other accounts, as well as from ToM accounts, in explicitly addressing the genesis of the recognition of self-other equivalence and difference, as involving a developmental shift from mind-sharing to mind-understanding. The key notion in IRT is that the first-person information that we have about our own IRs (e.g., the “feeling” of love for someone) is distinctly different from the third-person information that we have about the IRs of others (e.g., another’s “behavior” toward the object of love), and that in order to develop uniform concepts or representations of IRs that can be applied equally, but distinctly, to self and other, we need to match these two types of information in a single concept or form of knowledge that contains both types of information. In Barresi and Moore (1996) we posited an “intentional schema” to integrate this multimodal combination of first- and third-person information initially derived from self and other. On this view, *being in love* should not be defined primarily as a private, subjective experience, as in the ST view, nor as a mental intentional state that can be inferred from behavior, as in the TT view, but as an embodied IR between the agent and object, that, in the case of *love*, involves both feelings and concomitant behavioral expressions. Moreover, in learning the concept of *love* or any other IR, it is supposed that we must learn both the first-person, “inner” aspect, of the IR, as well as the third-person, “outer” aspect; otherwise, we fail to have the concept. For instance, one can be in love, say for the first time, without knowing it, because all one knows about *love* is the outer aspect, and one does not recognize this outer aspect in one’s feelings for another until one’s concomitant behavior is pointed out to one. Of course, *love* in our culture is primarily a social concept and learned to a large extent through language. But other more basic IRs, like fearing, seeing, or picking up are more fundamental, and may be understood to some extent by an organism without the mediation of language.

In the rest of the chapter, we consider in more detail Intentional Relations Theory and specifically the issue of how 1st and 3rd person information about intentional relations are integrated. We go on to review the neuroscientific findings that support this approach to social understanding. We then consider autism as a case of failure to integrate 1st and 3rd person information in the understanding of self and other.

3. Matching of 1st and 3rd person information and their integration

In Barresi and Moore (1996) we developed a model of social understanding that focused on the origins of understanding of IRs. We distinguished 4 levels of understanding IRs and used these levels to interpret both developmental and phylogenetic differences in social understanding (cf. Zlatev this volume, for a

similar multilevel model). At level 1, the organism represents the activities of self and other in distinctly different ways and neither in terms of IRs. We suggested that most animals typically operate at this level and it may also characterize social understanding in certain forms of psychopathology such as autism. We will return to consider this level and the case of autism in Section 5. In the rest of this section we review 3 levels of social understanding in which first and third person information about IRs are integrated. We devote most attention to how such integration is possible in the first place.

3.1 Interactive routes to matching

In order to understand IRs at all, the organism must be able to combine first person information about IRs with third person information about IRs into integrated representations involving an agent, an intentional relation and an object that can be equally applied both to the IRs of self and the IRs of others or to the joint activity of self and other. This combination occurs at level 2 of our model when there is matched first- and third-person information about intentional action available to the organism. There are various ways in which such matching can come about. Our suggestion is that matching occurs normally in human development when infant and mother engage in interactions, initially dyadic and later triadic. These interactions are typically patterned in such a way that the infant and mother both express and experience similar psychological activity. For example, in dyadic interactions, infant and mother may smile and vocalize in close synchrony. Whether the synchrony between an infant and adult in interactions of this sort is based on innate contagious mechanisms, or occurs through a form of mimicry initiated at first by the adult, it seems clear that there is a matching in such cases, where first-person information about self can be experienced concurrently with matched third-person information about the other. We believe that in such early dyadic communicative interactions the infant acquires integrated knowledge of first- and third-person aspects of emotional expressions, though not yet of intentional relations involving those expressions directed at objects.

Dyadic interactions do not revolve around objects so the intentionality of the shared psychological activity is at best implicit. However, in the triadic interactions that develop at about 9 months of age, the patterned interaction is now object-focused so that both infant and mother may share psychological activity to a particular object – they may look at the same object or produce similar object-directed actions through imitation. We have argued that such interactive experiences are crucial for the development of understanding IRs because it is

in these interactions that the infant's first-person experience of her own object-directed psychological activity is coordinated reliably with their corresponding third-person experience of the mother's object-directed psychological activity. Reliable coordination of the available first- and third-person information allows the construction of representations of intentional activity that integrate both forms of information and are thereby applicable to the joint activity of self and other, and subsequently with further development to individual activities of either self or other.

3.2 Noninteractive routes to matching

Although dyadic and triadic interactions provide the normal context for the sharing of psychological activity in human development, it is probably not necessary for there to be joint engagement of either dyadic or triadic kinds for a degree of matching of intentional relations to occur. For instance, as indicated earlier, research on monkeys seems to show that they can represent the goal-directed actions of another organism in the same manner as they represent their own actions (Gallese, Fadiga, Fogassi and Rizzolatti 1996; Rizzolatti, Fadiga, Gallese and Fogassi 1996). The pre-motor 'mirror' neurons mediating these representations fire in the planning and execution of the monkey's own actions, but also in perceiving comparable goal-directed actions in another animate being. While we do not wish to exclude the possibility of innate forms of matching between self and other, for instance in emotional expressive domains where unlearned forms of mimicry may be the basic mechanism for matching, in the case of action understanding a learning mechanism needs to be involved. Matching between perception and action may come about because for certain forms of psychological activity such as object-directed reaching, the organism gains information about its own action via more than one perceptual modality (Keysers and Perrett 2004). When a monkey reaches for objects, it is reliably provided with both visual and proprioceptive information about its own reaching, and an integrated multimodal representation of the action will result. Then vision may mediate the connection to the action of others. The same multimodal representation will later be activated by only the relevant visual information and thereby can be applied to the experience of seeing another organism perform the action. Vision here serves as a third-person 'bridging' modality that can be applied to both self and other, thus linking the strictly first-person information of proprioception to the available third-person information about goal-directedness. In this way a representation of action that is similarly applicable to the actions of both self and other may be achieved. However,

it should be noted that all that is involved here is the understanding of the action, *per se*, not of an agent performing the action. Thus the representation is at a sub-personal rather than at a personal, or agent, level of representation. Hence, an organism does not here understand intentional relations involving agents, but only sub-personal actions directed at objects. There is evidence that such a process may also operate in early human development. Woodward (1998) has shown that infants are able to recognize the goal-directed reaches of others at about the same time as they themselves engage in visually guided reaching. Importantly, teaching infants to make object-directed reaches at an early age is correlated with their representation of similar reaching actions of another person (Sommerville, Woodward, and Needham 2005). Thus, at least for simple actions, it seems that learning to succeed at an action, which involves coordination of first-person (e.g., proprioceptive) and typically third-person (e.g., visual) information of one's own action, may be correlated with representing the similar actions of others.

3.3 Sub-personal and interpersonal forms of understanding IRs

It will be recognized that the latter route to representations of actions that are equally applicable to self and other will only serve for those actions, such as manual reaching, for which the same perceptual information is available for both self and other. It is in such circumstances that a common code for the perception and production of action can bear fruit both in monkeys and humans, with a sub-personal level of understanding of goal-directed actions. However, in the understanding of intentional relations more is required. The difference between the human case and the cases of monkeys is that the dyadic and triadic interactive contexts of early human development provide multiple instances in which there are richly elaborated structures of shared intentional relations. For example, in a typical episode of a joint attentional (triadic) interaction, there may be shared emotional experience (e.g., smiling), shared object-directed action (e.g., object exchange) and shared epistemic activity (e.g., gaze following). These interactive structures therefore provide not just experiences in which a particular, simply observable, action intentional relation is shared but experiences in which a variety of different yet complementary intentional relations of various types are shared. As a result, there is the opportunity for infants to acquire complex representations of intentional activity that combine and integrate the first-person information pertaining to their own activity and the third-person information pertaining to the activity of others across a range of intentional relations. This difference between the human and animal cases, such as monkeys, is important because it may explain why humans step onto the path of development that leads ultimately to an agent level

form of social understanding, whereas monkeys appear not to.¹ To see why, it is important to examine whether the earliest forms of integrated representations of intentional relations are recognized to be at a 'personal' or at a 'subpersonal' level. Some authors (e.g., Tomasello 1999) have argued that the phenomena of triadic interactions arising at about 9 months signal the development of a concept of an 'intentional agent' that can be applied equally to self and other agents. However, a plausible alternative is that concepts of intentionality are initially acquired in a more piecemeal way. For example, Woodward and her colleagues' research (for a review see Woodward 2005) has shown that infants represent the object-directedness of different actions at different points in development. Whereas reaching is represented as object-directed before 6 months, gaze is not represented as object-directed until the end of the first year. Furthermore, when such intentional relations are first being acquired, the acquisition does not appear to be correlated so that infants who represent gaze as object-directed may not represent pointing as object-directed and vice versa. To explain this pattern of results, Moore (2006) proposed the notion of 'intentional islands' (cf. Tomasello 1992, on 'verb islands' in language acquisition), whereby intentional representations start out as separate sub-personal 'islands' relevant to particular object-directed actions and are only gradually integrated into more complex concepts at a personal level relevant to goal-directed agents. We suggest that it is the richly structured patterns of intentional relations that occur in triadic interactions, which allows the generation of the more complex representations of goal-directed agents. In contrast, while other animals such as monkeys may acquire sub-personal integrated representations of object-directed actions, such as reaching, without experience of rich combinations of shared intentional relations, they do not proceed to construct representations of goal-directed agents.

1. Great apes provide evidence that they stepped onto a new path similar to, but not the same as, our own. Chimpanzees, and probably other apes, engage in intense social interactions that promote an understanding of other's actions on an individual level, through what Zlatev (this volume), ascribes to dyadic mimesis and which, we (Barresi and Moore 1996) originally hypothesized was associated with their general imitative ability. Recent research suggests that the evolutionary path taken here may be different from our own in that while learning in dyadic interactions between infant and mother chimpanzees involves an apprenticeship relationship (Matsuzawa 2007) human dyadic and triadic relationships between human infant and adults is much more intensely communicative and collaborative (Tomasello et al. 2005). A consequence of this latter form of interaction results in what Zlatev calls triadic mimesis, which is roughly similar to level 2 interactions transforming to level 3 interactions in our own model.

3.4 Individualistic understanding of IRs

So far we have advanced from a sub-personal understanding of the simple actions of self and other that do not explicitly code for agent to the capacity for understanding *shared* IRs evident at level 2 of our model. This sharing entails the existence of representations of IRs that are interpersonal, though probably not explicitly represented as interpersonal. Rather the interrelated and similar IRs of self and other are understood using a uniform representational form that codes for the concurrent identity between first-person information of self and third-person information of the other. But it is not yet the case that agents are recognized to be individual centres of intentional activity. The next level of understanding IRs (Barresi and Moore 1996) requires the ability to reflect on, or imagine IRs as properties of individual agents. According to IRT this requires the use of imagination to fill in the third-person information for IRs of self and first-person information for IRs of others. Without this ability it would not be possible to represent diversity of intentional relations across self and other when the same object is involved.

In the developmental account given in Barresi and Moore (1996), children attain level 3 of understanding IRs during the second year of life. A variety of phenomena evidence this change (see Moore 2007). On the one hand the child becomes capable of recognizing the self as an individual agent as seen by phenomena such as mirror self-recognition. On the other hand, children become able to appreciate that others may have a different intentional orientation to an object from the self. For example, 18-month-olds understand that someone else may like something that they do not and vice versa (Repacholi and Gopnik 1997) and they understand that they may see something that someone else does not and vice versa (Moll and Tomasello 2005). At this point in development, therefore, children are able to attribute some forms of mental states, those exhibited in present activities, to individual agents, both self and other.

This level of understanding goes beyond mind-sharing toward a conceptual understanding of individuals as embodied agents with points of views that may differ from each other. In some respects our account here is similar to the simulation account. However, whereas ST proposes that we simulate the mental state of the other through imaginative substitution of our own mental states, we here suggest that only the first-person aspect of the intentional relation of the other requires imaginative construction, as the third-person aspect is pragmatically available in the situation. Moreover, we suggest that at this same time the infant acquires the skill to understand its own intentional relations by imagining the third-person aspect that goes with the currently first-person experience of the intentional relation, something the ST does not even attempt to explain. Our account also differs from Gallagher and Hutto, since we do not think that language

alone mediates the conceptual development that occurs at this time, which allows one to distinguish one's own from the other's embodied mental states. Indeed, their narrative interpretation of how children distinguish mental states of self and other, seems to focus on only representational mental states such as false beliefs, a capacity for which we provide a separate account in the next section.

3.5 Representation of mental agents

In the fourth year, pre-school children achieve yet another level of social understanding, when they can imagine both first- and third-person properties of a mental state. This results in children developing knowledge of mental representation as such, which allows them to show evidence of the conceptual understanding of mind seen in traditional ToM tasks. However, according to IRT, the levels of intentional understanding at which there is an *understanding of individual minds* derive from previous shared intentional activities where first- and third-person information originally became associated. It is the derivation from shared psychological activity that enables the concepts of mind that humans have, yielding notions like *love* having both internal bases involving feelings and external bases involving behavior. All levels of social understanding which depend originally on the integration of first- and third-person information are held to be different from Level 1 forms of understanding of self and other, which rely separately on first-person information alone to understand self and third-person information alone to understand others. Consideration of level 1 will become important later in the chapter when we discuss autism (see Section 5). We turn now to research on the neuroscience of social understanding to see to what extent there is support for the model of social understanding we have outlined here. We should note, however, that whereas the evidence from neuroscience indicates a particular pattern of brain organization underlying social understanding in adult human beings as well as nonhuman primates, there is of course no guarantee that the same organization exists at all earlier stages of human development.

4. Neuroscience and social understanding

In reviewing research on the neuroscience of social understanding, we will organize the initial review into sections dealing with action IRs, emotion IRs, and epistemic IRs, respectively. In these sections our concern will be to identify brain regions and processes that deal primarily with first- and third-person information separately, from areas where first- and third-person information meet and

where their integration makes possible relatively uniform application of these representations to both self and others. Where first- and third-person information is separated we would expect them to apply differentially to self and other, with first-person information tending to apply mostly to self and third-person information mostly to other. Where they are integrated, the question becomes how we use this integrated information to distinguish between self and other. We will also identify regions in which lower level perceptual processing can be distinguished from higher level metacognitive processing. Finally, we identify research indicating that first- and third-person information is sometimes represented independently, in particular in the case of autistic individuals. Figure 1 depicts essential components of IRT along with possible anatomical correlates that will be described in subsequent sections of this chapter.

4.1 Action intentional relations

Since the discovery of mirror neurons in the premotor cortex in monkeys that respond to the goal-directed actions of others (Rizzolatti et al. 1996), studies have investigated whether evidence can be found for similar neural structures in humans. A standard paradigm used in a number of these studies is to compare an *observation condition*, where participants watch the activity of another person, an *execution condition* where participants perform the action on cue, and an *imitation condition*, where participants perform the action that they observe another person perform. Transitory Magnetic Stimulation (TMS) studies affecting processing in the relevant neural systems have attempted either to facilitate/produce actions in observation conditions, or to interfere with actions in action or imitation conditions (see Iacoboni 2005, for a review). Taken together these studies affirm that premotor and parietal cortices in humans show mirror properties similar to those in individual neurons of monkeys. Both of these areas are active when performing the actions or observing the actions of others, and more active than in either of these conditions when these actions are both observed and imitated. In contrast to the additional activation found in these two regions (premotor and parietal) when imitating compared to mere observing, a third region, the Superior Temporal Sulcus (STS), tends to show the same level of activation in both observation and imitation conditions but is inactive in the action-only condition.

Iacoboni, Kaplan and Wilson (in press) have proposed a model incorporating IRT in accounting for these findings. They propose that the STS provides third-person visual information of the action that is being performed. This information is transferred to the Posterior Parietal, where it is matched with first-person information on the kinesthetic, kinematic and somatosensory properties that might go

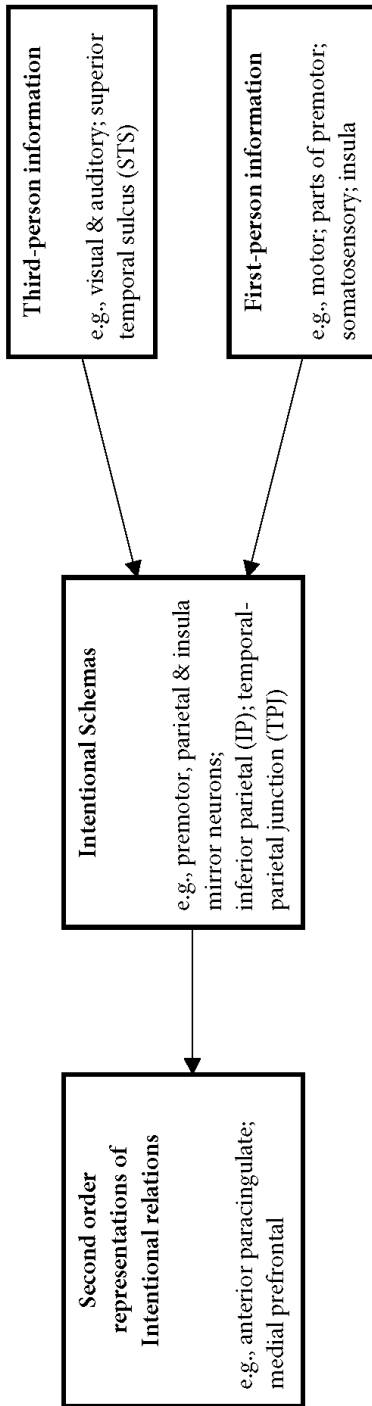


Figure 1. Essential components of Intentional Relations Theory and possible anatomical correlates. Third-person information involves extero-senses, which tend to apply more to others than to self; first-person information involves action intentions and interior senses, which tend to apply more to self than to others. Intentional schemas are posited to involve multimodal association areas where first- and third-person information get integrated. Although our main focus is on object-directed intentional relations, non-object directed integration are expected to occur at body-schema levels as well. The central site for the integration of first- and third-person information involving agents in intentional relations is hypothesized to be the temporal-parietal junction (TPJ) and/or inferior parietal (IP), which is hypothesized to involve an egocentric or first-person representation of the agent in space in the right hemisphere and a connected allocentric or third-person representation of the agent in the left. Second order, or reflective representations of intentional relations are hypothesized to occur in the prefrontal cortex. The directions of arrows represent the dominant direction of information processing, though feedback and other connections also occur between anatomical regions both within and between boxes of the model.

with the action—information provided by internal first-person sources of information integrated in the inferior parietal (see Figure 1). This matched representation of embodied action is then forwarded to the pre-motor area where alternative action plans can be compared to this input. This feed forward mechanism, is then matched to information being fed back from alternative pre-motor plans, and an interpretation is made, in the inferior parietal, between alternative interpretations. In their model, both the pre-motor area and the inferior parietal areas involve matching between first- and third-person properties and so are attributed to involve integration of first- and third-person information by intentional schemas. One way to conceive of the relationship between these two areas is that the inferior parietal (and/or nearby Temporal/Parietal Junction – TPJ) provides an egocentric, body-centered representation of the source of action of an agent-in-world, while the pre-motor area represents the goal or object of the action. Both require matching of first- and third-person information and together provide a full representation of the action intentional relation.

From the point of view of IRT, the more important area of integration of first- and third-person information is the inferior parietal or TPJ, rather than the pre-motor area, particularly as this area seems to reappear on complex ToM tasks, and may be crucial for distinguishing self and other as intentional agents. Whereas mirror neurons in the pre-motor area may be insensitive to the difference between self and other and focus mainly on the goals of actions, something that monkeys and young infants can represent, we would hypothesize that left and right parietal regions represent agents in intentional relations, and might be used to distinguish self from other as intentional agents. Studies by Decety and his colleagues (see Decety and Grezes 2006 for a review) provide support for the idea that the TPJ is the locus of a body-centered integration of first- and third-person information that applies both to self and to other but that may also be used to distinguish self from other. In these studies, imitations of other-by-self or self-by-other are compared. The general finding is that TPJ (they include studies citing inferior parietal as well as posterior STS) is more active on the right side when other imitates self, but more active on the left side, when self imitates other. One way to interpret this difference is that *left* TPJ is more active, when a third-person representation of a human body in space is more dominant than a first-person representation, and that the reverse is true for *right* TPJ. In other words, when the participant is the original source of the action, the right hemisphere is dominant and when the participant is imitating the other, the left hemisphere is dominant. More typically, we would suggest that when left TPJ is dominant, another person is being represented, where third-person information is perceived but first-person information is imagined (what might be called an allocentric representation of a person in space). However, when right TPJ is dominant, it is the self that

is typically being represented, where first-person information is perceived and third-person information imagined (what might be called an egocentric representation). Independent support for this idea comes from studies of brain damage on these two sides. As we shall see, damage at the left TPJ is found to be associated with failure at false belief tasks involving representations of others, whereas other studies have demonstrated that damage at the right TPJ is associated with spatial neglect, a distortion of egocentric or first-person perspective of space (see Halligan, Fink, Marshall and Vallar 2003). Furthermore, damage at the TPJ (or IP) has recently been shown to produce autoscopic hallucinations – seeing oneself – with right-sided damage associated with a non-egocentric out-of-body experience of self, and left-sided damage associated with an egocentric seeing of one's double (Blanke and Mohr 2005).

Taken together, these findings support Iacoboni et al.'s application of the IRT to their imitation studies, and their attribution of our notion of intentional schema to the inferior parietal, as they provide independent evidence that the inferior parietal or TPJ is the main center for an integrated representation of a person in space, whether it is self or other. But these findings also highlight how we can distinguish self from other through the source of information that drives the representation, third-person if it is other and first-person if it is self. These findings also provide a basis for connecting the more complex human activities involved in traditional false belief tasks, which have been shown also to require representations involving the TPJ and more mundane actions that are investigated in imitation tasks.

However, in considering imitative tasks, it should be noted that imitation of novel actions requires skills that do not appear in monkeys, and only appear in humans in a full blown state during the second year of life, when the infant is forming its concept of an intentional agent. Indeed, two-year-olds find it particularly fascinating to engage in mutual imitation, where they take turns leading and following each other in novel intentional actions, in a manner analogous to the Decety studies. This play behavior can be interpreted as working out possibilities made available at this time by developments in the use of the intentional schema, both to understand self and other individually and to discriminate self from other even in contexts, where both actors are performing similar actions.

4.2 Affective and motivational intentional relations

Typically, when dealing with action IRs, first-person information directly involves motor plans, proprioception, and kinesthetic feedback, while third-person information directly involves visual and auditory information. The integration of these

sources of information yields representations of a body acting in space with these first- and third-person resources integrated into a representation that can be applied to self or other, possibly through the use of vision and audition as bridging modalities that provide information about actions of self as well as other. Even so, there is a residual motor component, including a readiness to act (see, e.g., Ramnani and Miall 2004), as well as the sense of agency previously discussed that tends to distinguish self from other. When it comes to affective and motivational IRs, the focus is more on sensation than on action. So, the distinction between first-person information and third-person information and their integration, will tend to focus more on internal states within the body rather than on external appearances and expressions. Research involving such affective intentional relations has been consistent in showing the importance of integrated somatic representations of internal feeling states of a person whether such representations are applied to self or other. Generalizing such representations of internal states to another person occurs even when there is no social judgment involved in the task and where the participant merely observes the other. Recent research on pain has been particularly revealing. With respect to pain in self and other, single cells in the Cingulate Cortex (CC) have been found to respond not only to own pain, but also to the appearance of pain in another (Hutchinson et al. 1999). This response occurred even though no instructions to empathize were involved. In an fMRI study of empathy for another's pain, where again no instructions to empathize were involved, Singer, et al. (2004) had female participants and their partners receive mild shocks following a signal which indicated who was to receive the shock. The participants could see the hands of self and other as well as the signals while they were in the magnetic resonance chamber. It was found that certain primary somatosensory areas responded only to pain in self, but that the Anterior Insula, and the CC responded to the shock signal and anticipated pain both in self and in other. It has been hypothesized by Craig (2003) and Damasio (1999) that the anterior portion of the Insula, particularly on the right side, is a recently evolved region of the brain that represents a "feeling self". This region and the CC may both be involved in conscious representation of pain, in contrast to the primary sensory cortex, which may measure the intensity and sensory quality of the pain stimulus, but which may not always contribute to consciousness of pain. Part of the evidence for the distinction is that placebo effects, where perception of pain is induced, produce activations in Anterior Insula but not in the primary sensory areas (Wager et al. 2004). It seems then that, like mirror regions in the pre-motor and parietal areas, this 'feeling self' level of representation of pain is responsive, not only to one's own feeling of pain, but to the expressed, or merely inferred, pain of another person.

What the Singer et al. (2004) study seems to show is that the areas involved in conscious perception or feeling of one's own pain, are also active for the anticipated pain of another. Without instructions to do so, the participants seem to participate empathically in the anticipated pain of the other, thus sharing in it, and presumably being aware of their pain by sharing in it. Further support for this interpretation comes from the fact that dispositional measures of empathic ability were obtained in this study and a correlation between degree of dispositional empathy and degree of activation in the Insula and CC for the observer-other condition was found. Therefore, not only does the third-person perception of the other's behavioral situation apparently result in a conceptual understanding of the feeling state of the other, but it actually induces a comparable feeling state in the observer, which may be the ground upon which conceptual understanding is based. The degree to which this internal feeling state is induced seems to depend on the capacity for empathy, or sympathetic imagination, of the observer. However, as a subsequent study shows (Singer et al. 2006), it also depends on how one feels about the other person. If one has reason to like the other, then there is a stronger tendency to show an empathic response to the other's pain, than if one has reason to dislike the other person. In the latter case, men, but not women, were shown not to have this empathic response to the other's pain, but instead showed evidence of personal pleasure at seeing the other in pain. So the story here is fairly complex. Unlike the action mirror system in the pre-motor area, which seems to depend only on attention to the activity of the other, the degree of identification with or caring for the other may matter in representing the feeling states of the other in the same mode as one's own feeling states.

In the original Singer, et al. (2004) study, as well as in similar studies on observing touch (Keysers et al. 2004), and disgust (Wicker et al. 2003) in others, primary sensory areas could be used to provide first-person information that distinguished between self and other. However, subsequent research on observation of localized pain inducing stimuli on another person raises the issue of whether primary sensory areas are immune to empathically induced responses. For instance, Avenanti et al. (2005) had participants observe needles being pierced into the hand of another person and found TMS motor cortex induced inhibitory responses of hand muscles in observers that matched those that occurred in their own case. Based on this and other findings, Singer and Frith (2005) have suggested that whether one is attending to – or imagining – the emotional response of the other person or the sensory quality of the pain may be what distinguishes these two kinds of results. The implication of this is that to the extent that one can project oneself into the particular situation and experiential state of the other to that extent will one tend to display a matching embodied state. According to IRT, it is the fact that one has at one's disposal this personal shared experiential base upon

which to understand the state of the other person that one succeeds in accurately imagining that state. But to elicit such an internal state that typically applies to self when observing another, a matching must occur between the expressed state of the other and one's own associated experience of being in a comparable state, or be elicited by attending to the situation that the other is in as if it were shared. In the case of an expressed state this requires matching of first-person information about the appropriate internal state to third-person information about expressed state. So motor aspects of the behavior of others may be a mediating factor in situations where we have no direct personal experience of emotional responses in those situations, or where we would respond differently from the other person.

Several other studies conducted by Iacoboni and his colleagues indicate that mirroring of expressed affective states may be an important basis for understanding emotions in others. In these studies fMRI brain imaging of participants occurred either while they were engaged in observing or imitating a variety of emotional expressions depicted in photos (Carr et al. 2003; Dapretto et al. 2006). In the study reported by Carr et al. (2003), observing and imitating emotional expressions in others activated regions involved in those emotional expressions for self, in particular the amygdala and insula were involved, but also the pre-motor area and STS. Again these results can be interpreted as eliciting from third-person information (STS) the matching first-person action information necessary to understand the internal state of the other individual. Because the observation and imitation condition had similar pre-motor findings to action studies, this suggests that implicit if not explicit matching of emotional expression is involved in emotional empathy, which may feed into the representation of the feeling self in the insula.

So far we have seen that matching between first- and third-person information seems to occur when observing another person's affective state, and it may not require active use of imagination to feel and understand another's affective IRs in that a form of affective sharing may occur directly in response to the situation or the other's expression. Indeed, from a phylogenetic as well as developmental perspective contagion of emotional states from one organism to another is the original basis of emotional sharing (cf. Zlatev this volume). However, as we have argued, sharing a psychological state is not the same as understanding that state. Other evidence suggests that understanding affective states in the sense of attributing emotions and other affective states to individuals as well as discriminating one's own from another's emotional state, likely requires frontal activity, and occurs later in human development. It appears necessary to have the involvement of frontal areas, in particular, the Medial Prefrontal Cortex (MPFC), in order to reflect on and understand the mental state as either one's own, or another's.

The role of the MPFC in understanding at a reflective level pain states in self and other is highlighted in another recent study directly comparing imagination of self and other in pain as compared to damage to a manikin figure (Jackson, Brunet, Meltzoff and Decety 2006). While in a magnet, participants viewed images of arms and legs apparently from a first-person perspective in situations likely to be painful or neutral. They were told to imagine the body part as their own, or another person's, or that of a manikin. In line with the notion that the MPFC is involved in representing second order IRs, there was a strong response in this region only for the humans, but not for the manikin. In addition, there was differential activation in the posterior cingulate, which responded to pain in self and other, and to the inferior parietal. As in previous studies the insula and ACC were responsive to both self and other in a comparison between pain and non-pain conditions. But differences between self and other also occurred. The comparison between self and other found several regions of difference, indicating different routes to representing the same pain state in self and other, and the ability to distinguish between our own and another's pain.

Taken together the results on emotional processing show that matching can occur not only in the motor system where actions or expressions of others are mimicked, perhaps subpersonally, but that feeling states that are connected to those expressions in ourselves are often also active when observing others or in inferring their emotional states in conditions where sympathetic contagion or empathy might be elicited. These internal feeling states are then processed further in frontal areas when we are attempting to understand the emotional state of the other as distinguished from our own emotional response. Both the matching in the premotor area and in the feeling self can be viewed as first-person aspects of emotional IRs, while the visual expressions can be viewed as third-person aspects. However, for second-order representations of these IRs, frontal activity is necessary.

4.3 False belief and complex social inference tasks

A considerable amount of research has been devoted to establishing the brain basis of the understanding of the more complex intentional relations characteristic of "theory of mind". The focus of studies using ToM tasks is on determining brain regions functionally involved in the interpretation of complex stories of social interaction that are visually or verbally presented and in attributing mental states to individuals in these stories. Two brain regions have been shown to be most active in brain imaging studies using various techniques, when compared to control conditions involving comparable processing of non-ToM stimuli: (1) The Temporal/Parietal Junction (TPJ); including neighboring Superior Temporal regions

incorporating STS as well as Inferior Parietal regions, cf. Decety and Grezes 2006); (2) The Medial Prefrontal Cortex (MPFC).

The TPJ is believed to be an area in which complex visual stimuli, often involving biological motion and social interaction, are analyzed or represented perceptually and semantically. In the section on action IRs our discussion of the TPJ focused only on intentional actions of a single agent, but the TPJ is also crucial for social interactions and for interpreting more complex mental states than actions. Hence, in terms of IRT the TPJ, at least on the left side, can be understood as representing the third-person information about IRs of one or more organisms, involved in simple or complex object and interpersonal interactions. For instance, even in monkeys this area has individual neurons that are sensitive to eye direction of a person being observed by the monkey and the congruence with the person's behavior involving another object, with their direction of gaze. Comparable findings with humans, involving more complex IRs, for instance, involving intentions, have been made using fMRI (see, e.g., Pelphery et al. 2004). So this region can pick up epistemic as well as action IRs and is also involved in emotion IRs, involving multiple agents. The second region of importance for the ToM tasks is the MPFC. This region appears to be important for "decoupling" (Leslie 1994), or creating second order representations of IRs that can be attributed to individuals. Reflective or conceptual understanding of the intentionality of the behavior seems an important activity for this region. Indeed, merely noting a stimulus as an act of an intentional agent rather than a machine seems sufficient to involve this region (Ramani and Miall 2004). But this region has a number of other functions of a metacognitive, or executive, sort, and there appear to be subregions with specialized functions, some of which we will consider shortly.

Some recent elegant research using simple false belief tasks presented in stories and in videos, along with a number of important controls, to brain damaged patients with frontal and/or temporal-parietal lesions (Apperly, Samson, Chiavarino and Humphreys 2004; Samson, Apperly, Chiavarino and Humphreys 2004; Samson, Apperly, Kathirgamanathan, and Humphreys 2005) has provided evidence in partial congruence with these imaging studies. They found that damage to the left TPJ produces a fairly specific deficit in false belief reasoning about others, but that damage in the frontal regions does not. So it appears that a functional TPJ at least on the left (no tested patients had right TPJ damage) is necessary for false belief reasoning. By contrast, it appears that the impact of brain damage in frontal regions is less specific and more diverse, including effects on performance on tasks involving executive function but not on ToM tasks. Indeed, in one of their patients with frontal damage, there was evidence that problems occurred only on false belief tasks that required the inhibition of first-person knowledge of the real location but not on false belief tasks for which the participant

did not have knowledge of the real location (Samson et al. 2005). This result is congruent with other findings which suggest that executive function associated with frontal activity may be necessary to differentiate between mental states of self and other, and thus for attributing distinct mental states to individuals. In these circumstances a single mental state that is shared between self and other that might be used in cases of passive observation or empathic responding, will not be sufficient for mental state attribution.

If we consider just the two main regions involved in research with complex ToM tasks, these results fit well with what we would expect based on the “theory theory” (TT) approach to social understanding. The TPJ provides third-person behavioral analysis of animate activity or apparently animate activity, while the MPFC decouples or represents abstractly IRs, presumably in a theoretical or conceptual format. That the same behavioral analysis and conceptual representation could be applied to self and other is suggested by the fact that the MPFC shows overlap in activity for a variety of tasks involving self and other (e.g., see Decety and Sommerville 2003, for a summary of this research). It is possible that, in line with TT, TPJ analyzes and represents animate activity and IRs based *mostly* on visual or third-person information. As such the matching problem may not arise if the IRs of self and other are both analyzed in a behavioristic (or third-person) mode. MPFC could then provide “decoupled” (second order) representations of intentional relations of agents, whether they are of self or other (or jointly self and other).

However, the fact that the MPFC (and perhaps the TPC, particularly on the right side) is activated in cases of self-representation that *seem* not to be based entirely on third-person information about the self suggests that integrated representations involving both first- and third-person information of the kind postulated by IRT are involved. Furthermore, the frontal region and other regions along the midline have been postulated to be part of a system for representation and regulation of self (Northoff and Bermpohl 2004). So, perhaps, the MPFC generates a second order representation of another’s mental states, through prior association between a third-person behavioral analysis mainly from the left TPJ that applies more often to another person than to self and a simulation of first-person components of mental states found in the rest of the typically right-sided self-system. This latter interpretation is consistent with studies showing differential responses for self and other in high level processing of social stimuli (e.g., Lou et al. 2004).

The main conclusion to be drawn from these studies is that complex ToM tasks involve two main regions of the brain, a posterior one associated with perceptual representation of IRs and an anterior one associated with metarepresentation of these perceptual representations. Furthermore, there is a good deal of

overlap between the regions involved in representing self and other. Nevertheless, differences that occur suggest a mapping of third-person information typical of what we have from others to first-person information more typically associated with self. While the need for perceptual and metarepresentational processes for understanding individual IRs in complex ToM tasks is congruent with TT, the overlap between self and other, and use of first-person information as well as third-person information in these tasks fits better with the IRT approach to social understanding.

5. Level 1 understanding of intentional relations – the case of autism

Finally, it is worth mentioning some imaging research that supports the notion that representations of intentional relations can occur in distinct forms. Dapretto et al. (2006) studied high functioning autistic children and matched controls using the same imitation task as in the Carr et al. (2003) study mentioned earlier. However, in addition to observing and imitating emotional expressions, the autistic participants were measured on severity of autism, using several standardized scales. The behavioral findings were that the autistic participants were as able to imitate emotional expressions as other children, but the imaging findings suggested that the means that they used were different. The typically developing children replicated the results of the adult study, where mirror neuron pre-motor and insula areas were involved in observation and imitation of emotions, along with other areas. But in autistic children these mirror neuron areas were not as involved, and degree of involvement of these areas during imitation was inversely related to severity of autism in the social domain. Furthermore, other areas, the left anterior parietal and the right visual association areas, were more involved for autistic than for typical children. It was suggested that these latter areas served as an alternative route to imitation in this group instead of the usual one involving the mirror neuron system.

These results, combined with other findings, support the notion put forward by Barresi and Moore (1996) that the main reason why autistic people have difficulty in ToM tasks as well as emotion understanding and imitation is that they do not match and integrate first- and third-person information through an intermodal intentional schema, hence that they acquire and deploy independent first-person (or egocentric) and third-person (or allocentric) theories of mind. At the time that we wrote our article we had no idea how the notion of intentional schema might relate to brain activity. However, with the discovery of mirror neurons at about the same time, we, as well as others (e.g., Iacoboni et al. in press) have

been able to make the connection. The Dapretto et al. study probably provides the best confirmation for the view that it is the lack of matching of these two types of information through an intentional schema that is at the heart of problems in social understanding of autistic individuals. The inability to readily transform third-person perceptions into first-person matching experiences, as well as to make the reverse mapping, and thus to engage in mind sharing, makes it difficult for autistic individuals to make sense of mind, because of the absence of a direct connection between the two necessary, inseparably tied aspects of all mental phenomena, an externally available bodily expressive component, and an internally available feeling component. As a result of this deficiency in ability to share mind with others, they lose interest in other people, and have difficulty learning from them. Eventually, if they do attempt to reflect on and understand mind in self and others, they form two radically different accounts: on the one hand they develop rather complex TT-like accounts of mind from a third-person view of their own and other people's behavior; and on the other hand they over generalize in apparent simulation their own egocentric first-person perspective to others (cf. Frith and de Vignemont 2005). Because of lack of mind sharing during infancy and beyond, they are faced with intractable problems in understanding mind beyond those that appear as purely third-person TT types, or purely first-person ST types, instead of integrated theories where matching of first- and third-person information is involved as we have proposed in IRT.

6. Conclusion

Recent discoveries in the neuroscience of social understanding have opened a new window through which to evaluate theories of social understanding. In the present chapter we have primarily examined our own intentional relations theory (Barresi and Moore 1996) in light of these new discoveries. IRT has three important elements. First, it postulates a distinction between first- and third-person information pertaining to intentional relations, as well as a requirement that both forms of information be combined in order to generate representations of intentional action that are shared between, or equally applicable to, self and others. Second, it postulates that a distinction may be made between a level of social understanding at which first- and third-person information are integrated without being attributable to individual agents and more complex levels of social understanding at which integrated representations are recognized to be properties of individual agents. In human ontogeny (and possibly in phylogeny), the latter levels of social understanding are founded on the former level. Third, it postulates that under

certain conditions first- and third-person information about intentional relations may be processed separately so that the activities of self and other are represented independently. In humans, such a condition is seen in autism.

The neuroscience of social understanding shows that integrating first- and third-person information through matching these two types of information occurs in the understanding of action, emotion, as well as epistemic IRs of self and other. The fact that matching between first- and third-person aspects of IRs for self and other occurs immediately on-line for a variety of IRs is congruent with the notion that both aspects are necessary to fully extract the meaning of these activities. Such matching occurs early on in life, though this process of mind-sharing does not develop into understanding individual minds until later in development. On our account, it is only through processes that bring about shared psychological states between individuals early on, and provide the initial basis for social understanding, that later development of our usual understanding of individual minds becomes a possibility.

Although TT might account for some instances of theories of mind generated purely from behavior, it is only in autistic individuals where we see exaggerated “theories” of this type. However, in autistic individuals there is evidence of a failure in mapping first- and third-person information from very early on in life, which prevents shared mental activity in dyadic interactions. ST does better than TT in accounting for a variety of phenomena involving emotional empathy, and understanding epistemic states. But it cannot account, without special pleading, for matching phenomena involved in action understanding. Again, autistic individuals provide a window into the problem. They can generalize either first- or third-person representations separately from self to other or the reverse. However, because they did not initially engage in shared mental life with others, they have problems understanding the meaning of social activity when the integration of both first- and third-person information is involved. Without prior matching and integrating these two types of information in earlier shared mental activity associated with dyadic and triadic interactions, the concepts that they generate based either on behavior alone or internal states alone are diminished when compared to our usual understanding of IRs of self and other. Thus, we believe that matching theories like IRT provide the best account of how we come to understand our own as well as other minds.

References

- Appery, I.A., Samson D., Chiavarino C. and Humphreys G.W. 2004. "Frontal and left temporo-parietal contributions to theory of mind: Neuropsychological evidence from a false belief task with reduced language and executive demands." *Journal of Cognitive Neuroscience* 16: 1773–84.
- Avenanti, A., Bueti, D., Galati, G. and Aglioti, S.M. 2005. "Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain." *Nature Neuroscience* 8: 955–960.
- Barresi, J. 2001. "Extending self-consciousness into the future." In *The Self in Time: Developmental Perspectives*, C. Moore and K. Lemmon (eds.), 141–161. Mahwah, NJ: Lawrence Erlbaum Associates.
- Barresi, J. 2004. "Intentional relations and divergent perspectives in social understanding." In *Ipseity and Alterity: Interdisciplinary Approaches to Intersubjectivity*, S. Gallagher and S. Watson (eds.), 74–99. Rouen: Presses Universitaires de Rouen.
- Barresi, J. and Moore, C. 1996. "Intentional relations and social understanding." *Behavioral and Brain Sciences* 19: 107–154.
- Blanke, O. and Mohr, C. 2005. "Out-of-body experience, heautoscopy, and autoscopic hallucination of neurological origin Implications for neurocognitive mechanisms of corporeal awareness and self consciousness." *Brain Research Reviews* 50: 184–199.
- Carr, L., Iacoboni, M., Dubeau, M.C., Mazziotta, J.C. and Lenzi, G.L. 2003. "Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas." *Proceedings of the National Academy of Science, USA* 100: 5497–5502.
- Craig, A.D. 2003. "Interoception: The sense of the physiological condition of the body." *Current Opinion in Neurobiology* 13: 500–505.
- Damasio, A.R. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace.
- Dapretto, M., Davies, M.S., Pfeifer, J.H., Scott, A.A., Sigman, M., Bookheimer S.Y. and Iacoboni, M. 2006. "Understanding emotions in others: mirror neuron dysfunction in children with Autism Spectrum Disorder." *Nature Neuroscience* 9: 28–30.
- Decety, J. and Grezes, J. 2006. "The power of simulation: Imagining one's own and other's behavior." *Brain Research* 1079: 4–14.
- Decety, J. and Sommerville, J. 2003. "Shared representations between self and other: A social cognitive neuroscience view." *Trends in Cognitive Sciences* 7: 527–533.
- Frith, U. and de Vignemont, F. 2005. "Egocentrism, allocentrism, and Asperger syndrome." *Consciousness and Cognition* 14: 719–738.
- Gallagher, S. and Hutto, D.D. this volume. "Understanding others through primary interaction and narrative practice."
- Gallese, V., Keysers, C. and Rizzolatti, G. 2004. "A unifying view of the basis of social cognition." *Trends in Cognitive Science* 8: 396–403.
- Gallese, V., Fadiga, L., Fogassi, L. and Rizzolatti, G. 1996. "Action recognition in the premotor cortex." *Brain* 119: 593–609.
- Goldman, A. 1992. "In defense of the simulation theory." *Mind and Language* 7: 104–119.
- Gopnik, A. 1993. "How we know our minds: The illusion of first-person knowledge of intentionality." *Behavioral and Brain Sciences* 16: 1–14.
- Gordon, R. 1986. "Folk psychology as simulation." *Mind and Language* 1: 158–171.

- Halligan, P.W., Fink, G.R., Marshall, J.C. and Vallar, G. 2003. "Spatial cognition: Evidence from visual neglect." *Trends in Cognitive Sciences* 7: 125–133.
- Harris, P. 1989. *Children and Emotion*. Oxford: Basil Blackwell.
- Hobson R.P. 1991. "Against the theory of 'theory of mind.'" *British Journal of Developmental Psychology* 9: 33–51.
- Hobson, R.P. 1998. "The intersubjective foundations of thought." In *Intersubjective Communication and Emotion in Early Ontogeny*, S. Braten (ed.), 283–296. Cambridge: Cambridge University Press.
- Hobson, R.P. 2002. *The Cradle of Thought. Exploring the Origins of Thinking*. London: Macmillan.
- Hobson, R.P. and Hobson, J. this volume. "Engaging, sharing, knowing: Some lessons from research in autism."
- Humphrey, N. 1984. *Consciousness Regained*. Oxford: Oxford University Press.
- Hutchison, W.D., Davis, K.D., Lozano, A.M., Tasker, R.R. and Dostrovsky, J.O. 1999. "Pain-related neurons in the human cingulate cortex." *Nature Neuroscience* 2: 403–405.
- Iacoboni, M. 2005. "Understanding others: Imitation, language, empathy." In *Perspectives on Imitation: From Neuroscience to Social Science*, S. Hurley, and N. Chater (eds.), 77–99. Cambridge, MA: MIT Press.
- Iacoboni, M., Kaplan, J. and Wilson, S. in press. "A neural architecture for imitation and intentional relations." In *Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions*, C. Nehaniv and K. Dautenhahn, (eds.). Cambridge, UK: Cambridge University Press.
- Jackson, P.L., Brunet, E., Meltzoff, A.N. and Decety, J. 2006. "Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain." *Neuropsychologia* 44: 752–761.
- Keysers, C. and Perrett, D.I. 2004. "The neural correlates of social perception: A Hebbian network perspective." *Trends in Cognitive Sciences* 8: 501–507.
- Keysers, C., Wicker, B., Gazzola, V., Anton, J., Fogassi, L. and Gallese, V. 2004. "A touching sight: SII/PV activation during the observation and experience of touch." *Neuron* 42: 335–346.
- Knoblich, G. and Jordan, J.S. 2002. "The mirror system and joint action." In *Mirror Neurons and the Evolution of Brain and Language*, M.I. Stamenov and V. Gallese (eds.), 115–124. Amsterdam: John Benjamins.
- Leslie, A.M. 1987. "Pretense and representation: The origins of 'theory of mind.'" *Psychological Review* 94: 412–426.
- Leslie, A.M. 1994. "ToMM, ToBy, and Agency: Core architecture and domain specificity." In *Mapping the Mind: Domain Specificity in Cognition and Culture*, L.A. Hirschfeld and S.A. Gelman (eds.), 119–148. New York: Cambridge University Press.
- Lou, H.C., Luber, B., Crupain, M., Keenan, J. P., Nowak, M., Kjaer, T.W., Sackeim, H.A. and Lisanby S.H. 2004. "Parietal cortex and representation of the mental Self." *Proceedings of the National Academy of Science, USA* 101: 6827–6832.
- Matsuzawa, T. 2007. Comparative cognitive development. *Developmental Science* 10: 97–103.
- Moll, H. and Tomasello, M. 2005. "12- and 18-month-old infants follow gaze to spaces behind barriers." *Developmental Science* 7: F1–F9.
- Moore, C. 1999. "Intentional relations and triadic interaction." In *Developing Theories of Intention*, P. D. Zelazo, J. W. Astington and D. R. Olson (eds.), 43–62. Mahwah, NJ: Lawrence Erlbaum Associates.

- Moore, C. 2006. "Representing intentional relations and acting intentionally in infancy: Current insights and open questions." In *Human Body Perception from the Inside Out*, G. Knoblich, I. Thornton, M. Grosjean and M. Shiffrar (eds.), 427–442. New York: Oxford University Press.
- Moore, C. 2007. "Understanding self and other in the second year." In *Transitions in Early Socioemotional Development: The Toddler Years*, C.A. Brownell and C.B. Kopp (eds.), 43–65. New York: Guilford Press.
- Northoff, G. and Bermpohl, F. 2004. "Cortical midline structures and the self." *Trends in Cognitive Sciences* 8: 102–7.
- Prinz, W. 1997. "Perception and action planning." *European Journal of Cognitive Psychology* 9: 129–154.
- Rammani, N., and Miall, C.R. 2004. "A system in the human brain for predicting the actions of others." *Nature Neuroscience* 7: 85–90.
- Repacholi, B.M. and Gopnik, A. 1997. "Early reasoning about desires: Evidence from 14- and 18-month-olds." *Developmental Psychology* 33: 12–21.
- Rizzolatti, G., Fadiga, L., Gallese, V. and Fogassi, L. 1996. "Premotor cortex and the recognition of motor actions." *Cognitive Brain Research* 3: 131–141.
- Samson D., Apperly I.A., Chiavarino C. and Humphreys G.W. 2004. "Left temporoparietal junction is necessary for representing someone else's belief." *Nature Neuroscience* 7: 499–500.
- Samson, D., Apperly, I. A., Kathirgamanathan, U., and Humphreys, G.W. 2005. "Seeing it my way: A case of a selective deficit in inhibiting self-perspective." *Brain* 128: 1102–1111.
- Sebanz, N. and Frith, C. 2004. "Beyond simulation? Neural mechanisms for predicting the actions of others." *Nature Neuroscience* 7: 5–6.
- Singer, T. and Frith, C. 2005. "The painful side of empathy." *Nature Neuroscience* 8: 845–846.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. and Frith, C. 2004. "Empathy for pain involves the affective but not sensory components of pain." *Science* 303: 1157–1162.
- Singer, T., Seymour, B., O'Doherty, J.P., Stephan, K.E., Dolan, R.J. and Frith, C.D. 2006. "Empathic neural responses are modulated by the perceived fairness of others." *Nature* 439: 466–469.
- Sommerville, J.A., Woodward, A.L. and Needham, A. 2005. "Action experience alters 3-month-old infants' perception of others' actions." *Cognition* 96: B1–B11.
- Swanson, D., Apperly, I.A., Kathirgamanathan, U. and Humphreys G.W. 2005. "Seeing it my way: A case of a selective deficit in inhibiting self-perspective." *Brain* 128: 1102–1111.
- Tomasello, M. 1992. *First Verbs: A Case Study of Early Grammatical Development*. New York: Cambridge University Press.
- Tomasello, M. 1999. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28: 675–691.
- Trevarthen, C. & Hubley, P. 1978. Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. In *Action, Gesture, and Symbol: The Emergence of Language*. A. Lock (ed.), 183–229. New York: Academic Press.
- Wager, T.D., Rilling, J.K., Smith, E.E., Sokolik, A., Casey, K.L., Davidson, R.J., Kosslyn, S.M., Rose, R.M. and Cohen, J.D. 2004. "Placebo-Induced Changes in fMRI in the Anticipation and Experience of Pain." *Science* 303: 1162–1167.

- Wicker, B., Keysers, C., Plailly, J., Royet, J.-P., Gallese, V. and Rizzolatti, G. 2003. "Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust." *Neuron* 40: 655–664.
- Wilson, M. and Knoblich, G. 2005. "The case for motor involvement in perceiving conspecifics." *Psychological Bulletin* 131: 460–473.
- Woodward, A.L. 1998. "Infants selectively encode the goal object of an actor's reach." *Cognition* 69: 1–34.
- Woodward, A.L. 2005. "The infant origins of intentional understanding." *Advances in Child Development and Behavior* 33: 229–262.
- Zlatev, J. this volume. "The co-evolution of intersubjectivity and bodily mimesis."

Uncorrected proofs - © John Benjamins Publishing Company