22

Understanding Others: Embodied Social Cognition

Shaun Gallagher

Philosophy and Cognitive Sciences, University of Central Florida, Orlando, FL, USA and University of Hertfordshire Hatfield, GB

The standard and dominant approaches to social cognition rarely emphasize intersubjective interaction, and even when they do mention interaction they frame the problem in terms of two minds that have to communicate across the seemingly thin air of an unbridgeable gap. From this viewpoint, interaction is not a solution but simply an another way to state the problem of other minds. Consider, for example, the following formulation:

> … the study of social interaction … is concerned with the question of how two minds shape each other mutually through reciprocal interactions. To understand interactive minds we have to understand how thoughts, feelings, intentions, and beliefs can be transmitted from one mind to the other.
>
> (Singer et al., 2004, p. xvii)

On standard accounts of theory of mind (ToM) this gap between minds is bridged by some kind of cognitive processes in one mind providing the means to infer what is going on in the mind of the other, since the mind of the other is imperceptible. What one needs to bridge this gap is either theory (folk psychology), or simulation, or a combination of theory and simulation that will permit an inferential form of mind-reading or "mentalizing."

In this chapter, after reviewing some of the traditional ToM models of social cognition, I outline an alternative model on the basis of evidence from developmental psychology and phenomenology. In this alternative model, embodied, second-person interaction plays a central (although not an exclusive) role in our ability to understand other people. Finally, I discuss a recent

development of simulation theory (ST) that champions an embodied simulationist approach.

Traditional ToM accounts make little mention of how the body might fit into the process of understanding others. At best, we take an observational stance toward the other's body and treat it as the source of evidence for constructing an inference. Proponents of theory theory (TT) contend that inference formation happens as the result of a mental consultation with a theory or a set of folk-psychological rules that will allow one to deduce an explanation of the observed behavior in terms of beliefs and desires understood as the other's mental states. ST eschews theory and opts for simulation routines that are run on the mechanisms of one's own mind. Here, for example, is a clear statement of how an explicit simulation works:

> First, the attributor creates in herself pretend states intended to match those of the target. In other words, the attributor attempts to put herself in the target's 'mental shoes'. The second step is to feed these initial pretend states [e.g., beliefs] into some mechanism of the attributor's own psychology … and allow that mechanism to operate on the pretend states so as to generate one or more new states [e.g., decisions]. Third, the attributor assigns the output state to the target … [e.g., we infer or project the decision to the other's mind].
>
> (Goldman, 2005, pp. 80–81)

Both approaches share certain fundamental assumptions. First, they assume that the problem is best posed as one that involves lack of access to other minds. Minds are hidden away behind or beyond the behavior that may be manifested. The task, then, is to explain or predict the behavior in terms of mental states that can only be inferred.

A second assumption taken up by both TT and ST is that theory use or simulation use, respectively, constitute the primary and pervasive means for social cognition. Thus we find proponents of these ToM approaches making universalistic claims, of which the following are good examples.

> [H]umans everywhere interpret the behavior of others in … mentalistic terms because we all come equipped with a "theory of mind" module (ToMM) that is compelled to interpret others this way, with mentalistic terms as its natural language.
>
> (Tooby & Cosmides, 1995, p. xvii)

> It is hard for us to make sense of behavior in any other way than via the mentalistic framework.—'attribution of mental states is to humans as echolocation is to the bat. It is our natural way of understanding the social environment'.
>
> (Baron-Cohen, 1995, pp. 3–4; see also Leslie, 2000; Currie & Sterelny, 2000; Frith & Happé, 1999; Wellman, 1993; Karmiloff-Smith, 1992; Malle, 2002, for similar statements).

> The strongest form of ST would say that all cases of (third-person) mentalization employ simulation. A moderate version would say, for example, that simulation is the *default* method of mentalization … I am attracted to the moderate version … Simulation is the primitive, root form of interpersonal mentalization.
>
> (Goldman, 2002, pp. 7–8)

Third, they assume that our relations with others are always from an observational stance. Perception is characterized as a third-person process where one person is observing the behavior of the other person rather than interacting with him/her in a second-person fashion. This observational stance is very clear in almost all false-belief tests, which TT appeals to as scientific evidence about the development of our mentalizing ability. For example, a subject (often a child) is asked to observe the behavior of two other children (or sometimes puppets). Sally puts a marble in a basket and leaves the room; another child, Anne, moves the marble from the basket to a box. Sally comes back in the room and the subject is asked where Sally will look for the marble. Four year olds tend to answer correctly that Sally will look in the basket; 3 year olds tend to answer incorrectly that Sally will look in the box, where the marble actually is. This is taken as evidence that the 3-year-old subjects (and some autistic subjects) are unable to appreciate that having a different perspective could lead to Sally's false belief; 4-year-old children apparently have developed a ToM that can deal with false beliefs (Wimmer & Perner, 1983; Leslie & Frith, 1988). Such experiments are designed so that the subject is simply a third-person observer of events; the subject never participates in the events or interacts with Sally or Anne. Theory theorists also fail to point out that even the youngest of the non-autistic children tested interact with the experimenter, and tend to understand what the experimenter wants them to do.

ST also takes observation to be the starting point, and inferential judgment to be the ending point of the intersubjective process. To put ourselves in the other's shoes, we need to first observe where those shoes are; that is, we need to observe the behavior of the other person before retreating into our own mind to run the simulation. The entire description of the simulation process is governed by the observational stance.

There are various debates within TT and ST, one of which concerns whether the processes involved are explicit (consciously controlled) or implicit. The strongest version of the implicit model is discussed below, but it should be noted that we could raise a simple phenomenological objection against explicit models that evoke conscious or introspective forms of theorizing or simulation. Simply put, if we carefully consult our everyday ordinary experiences of encountering others, we do not find ourselves taking observational stances in the third person; we do not find ourselves always trying to explain or predict their behavior, or attempting to get into their heads to ascertain what their beliefs or desires are. Most of our encounters are second person, interactive encounters, and most of what we need for understanding others is often readily available.

AN EMBODIED APPROACH

What we are calling an embodied or interactive approach involves a complex set of practices that can be found from infancy onward. From this viewpoint, much of what we call the mind is not something hidden away, but is something

that is more directly accessible. Consider the phenomenologist Max Scheler's characterization of intersubjective perception as a direct perception.

> For we certainly believe ourselves to be directly acquainted with another person's joy in his laughter, with his sorrow and pain in his tears, with his shame in his blushing, with his entreaty in his outstretched hands … And with the tenor of this thoughts in the sound of his words. If anyone tells me that this is not 'perception', for it cannot be so, in view of the fact that a perception is simply a 'complex of physical sensations … I would beg him to turn aside from such questionable theories and address himself to the phenomenological facts'.
>
> (Scheler, 1954, pp. 260–261)

The idea is that there is a good amount of information that we can pick up in an ongoing direct perception of the other person's body that will give us a sense of what is going on with them. We can perceive their feelings and intentions in their postures, movements, facial expressions, gestures, vocal intonations, and actions. Scheler is not alone on this issue. We can find in Wittgenstein a number of similar statements.

> Look into someone else's face, and see the consciousness in it, and a particular shade of consciousness. You see on it, in it, joy, indifference, interest, excitement, torpor, and so on. … Do you look into yourself in order to recognize the fury in *his* face?
>
> (Wittgenstein, 1967, p. 229)

> In general I do not surmise fear in him—I *see* it. I do not feel that I am deducing the probable existence of something inside from something outside; rather it is as if the human face were in a way translucent and that I were seeing it not in reflected light but rather in its own.
>
> (Wittgenstein, 1980, p. 170)

That we do not look into ourselves to see what the other person is experiencing suggests that it is not a simulation process. And to say that I am not surmising or deducing the other's experience means that it is not through a theoretical inference that I gain access to the other.

Although, from this viewpoint, access is not regarded as a problem, this is not to say that the other person is entirely transparent, or that the meaning of all behavior can be perceptually grasped; behavior is often ambiguous, people are not always revealing of their emotions and thoughts. The claim here is not that direct perception can penetrate to the soul of the other person and discover his/her innermost emotional states. Nor is the claim that we can never be misled by what we perceive. The claim is rather that for the most part, in most of our encounters in everyday life, direct perception delivers a significant amount of important information for understanding others. In addition, it would only be something that I discover through these means that would lead me to the idea that perhaps something more is going on with the other person.

Moreover, in ordinary everyday encounters with others, I am not taking an observational stance; I am not off to the side thinking or trying to figure out what they are doing. Rather, I am responding to them in an embodied way, and I am part of the situation. As we will see shortly, our own motor and emotional systems

are intricately involved in our perception of others, and we should think of perception here as an enactive rather than a passive process. What we call social cognition is first of all social interaction. What I perceive in these cases does not constitute something short of understanding. Rather my understanding of the other person is constituted within the perception–action loops that define the various things that I am doing with or in response to others.

Evidence for this can be found in many developmental studies, and generally falls under what the developmental psychologist Colwyn Trevarthen calls "primary intersubjectivity" (Trevarthen, 1979). We do not arrive in the world as a *tabula rasa*—and our slate starts to fill up very quickly. Developmental studies consistently tell us that neonate perception is already relatively smart. The newborn infant can pick out a human face from the crowd of objects in its environment, with sufficient detail that will enable it to imitate the gesture it sees on that face (Meltzoff & Moore, 1977, 1994). There is an increasing evidence that infants automatically attune to smiles (and other facial gestures) with an enactive, mimetic, response (Schilbach et al., 2008). The young infant is visually attracted to movement and in specific ways to biological movement, and auditorily attracted to certain kinds of sounds, such as its mother's voice. Infants "vocalize and gesture in a way that seems [affectively and temporally] 'tuned' to the vocalizations and gestures of the other person" (Gopnik & Meltzoff, 1997, p. 131). Human infants show a wide range of facial expressions, such as complex emotional, gestural, prosodic, and tactile face-to-face interaction patterns which are absent or rare in non-human primates (Falk, 2004; Herrmann et al., 2007), but notably without the intervention of theory or simulation. Moreover and in a non-mentalizing way, they are able to see bodily movement as expressive of emotion, goal-directed intentional movement, and they are able to perceive other persons as agents. This does not require advanced cognitive abilities, inference, or simulation skills; rather, it is a perceptual capacity that is "fast, automatic, irresistible and highly stimulus-driven" (Scholl & Tremoulet, 2000, p. 299).

Infants are able to detect correspondences between visual and auditory information that specify the expression of emotions as early as 5–7 months (Walker, 1982; also Hobson, 1993, 2002). At 9 months, infants follow the other person's eyes (Senju, Johnson & Csibra, 2006), and start to perceive various movements of the head, the mouth, the hands, and more general body movements as meaningful, goal-directed movements. Baldwin and colleagues, for example, have shown that infants at 10–11 months are able to parse some kinds of continuous action according to intentional boundaries (Baldwin & Baird, 2001; Baird & Baldwin, 2001). Such perceptions give the infant, by the end of the first year of life, a non-mentalistic, perceptually based embodied understanding of the intentions and dispositions of other persons (Baldwin, 1993; Johnson et al., 1998; Allison et al., 2000; Johnson, 2000). These capabilities do not disappear in adulthood but they mature and become more sophisticated (see Dittrich et al., 1996). This can be clearly shown in a micro-analysis of the postures, movements, gestures, gazes, and facial expressions of people as they engage in a novel task and where communication

among them is intrinsic to the actions that they take (see Niedenthal et al., 2005; Lindblom, 2007).

This initial set of direct perceptual practices does not give us the full account of social cognition, and the information we pick up directly from the other person's embodied comportments is far from sufficient for the often rich and nuanced understanding that we can have of the other person. This primary intersubjectivity, however, is immediately supplemented and enhanced by a secondary intersubjectivity (Trevarthen & Hubley, 1978). Expressions, intonations, gestures, and movements, along with the bodies that manifest them, do not float freely in the air; we find them in the world, and infants soon start to notice how others interact with the world. Infants begin to tie actions to pragmatic contexts around the age of 1 year; they enter into *contexts* of shared attention—shared situations—in which they learn what things mean and what they are for. Behavior representative of joint attention begins to develop around 9–14 months (Phillips et al., 1992). In such interactions, the child looks to the body and the expressive movement of the other to discern the intention of the person or to find the meaning of some object. The child can understand that the other person *wants* food or *intends* to open the door; that the other can *see* him (the child) or is *looking at* the door.[1] They begin to see that another's movements and expressions often depend on meaningful and pragmatic contexts and are mediated by the surrounding world. Others are not given (and never were given) primarily as objects that we encounter cognitively, or in need of explanation. We perceive them as agents whose actions are framed in pragmatic and socially defined contexts. It follows that there is not one uniform way in which we relate to others, but that our relations are mediated through the various pragmatic (and ultimately, institutional) circumstances of our encounters. Indeed, we are caught up in such pragmatic circumstances, and are already existing in reference to others, from the very beginning (consider for example the infant's dependency on others for nourishment), even if it takes some time to sort out which agents provide sustenance, and which ones are engaged in other kinds of activities.

As we noted, children do not simply observe others; they are not passive observers. Rather they interact with others and in doing so they develop further capabilities in the contexts of those interactions. If the capacities of primary intersubjectivity, like the detection of intentions in expressive movement and eye direction, are sufficient to enable the child to recognize dyadic relations between the other and the self, or between the other and the world, something more is added to this in secondary intersubjectivity. As noted, in joint attention, beginning around 9–14 months, the child alternates between monitoring the gaze of the other and what the other is gazing at, checking to verify that they are continuing to look at the same thing. Indeed, the child also learns to point at approximately this same time. At 18 months, children comprehend what another person intends

[1]This is not taking an intentional stance, that is, treating the other *as if* they had desires or beliefs hidden away in their minds; rather, the intentionality is perceived in the embodied actions of others.

to do with an instrument in a specific context. They are able to re-enact to completion the goal-directed behavior that someone else fails to complete. Thus, the child, on seeing an adult who tries to manipulate a toy and who appears frustrated about being unable to do so, quite readily picks up the toy and shows the adult how to do it (Meltzoff, 1995; Meltzoff & Brooks, 2001; see also Herrmann et al., 2007).[2]

Our understanding of the actions of others occurs on the highest, most appropriate pragmatic level possible. That is, we understand actions at the most relevant pragmatic (intentional, goal-oriented) level, ignoring possible subpersonal or lower level descriptions, but in most cases also ignoring interpretations in terms of beliefs, desires, or hidden mental states. Rather than *making an inference* to what the other person is intending by starting with bodily movements, and moving from there to the level of mental events, we see actions as meaningful in the context of the physical and intersubjective environment. If, in the vicinity of a locked door, I see you reach for a set of keys, I would know your intentions as much from the door and the keys, your bodily posture and expression as from anything that I postulate in your mind. We interpret the actions of others in terms of their goals and intentions set in contextualized situations, rather than abstractly in terms of either their muscular performance or their beliefs. The environment, which is not only a physical location, but also a pragmatic context and a social situation, is never perceived neutrally (without meaning), either in regard to our own possible actions, or in regard to the actions and possibilities of others. In this regard, the world itself does much of the work involved in social cognition. As Gibson's theory of affordances (Gibson, 1979) suggests, we see things in relation to their possible uses, and therefore never as a disembodied observer. Likewise, our perception of the other person, as another agent, is never of an entity existing outside of a situation, but rather of an agent in a pragmatic context that throws light on the intentions (or possible intentions) of that agent.

There is much more to say about the role of socially defined situations and the roles that people play in them. As children develop, and precisely because they have the embodied capabilities defined by primary and secondary intersubjectivity,

[2]Onishi and Baillargeon (2005) have recently shown that infants at 15 months apparently mentalize the false beliefs of others. The data from their experiments suggest that infants see what the other person intends to do and is surprised (or at least notices) when the behavior of the other violates what the infant knows about the context (specifically about who has seen or not seen certain events). Although Onishi and Baillargeon interpret the data entirely in a ToM framework of mentalizing the other's beliefs, an alternative interpretation in terms of perceived meaningful (contextualized) behavior, actions, and intentions is clearly available. See Woodward & Sommerville (2000): "[…] 12-month-old infants interpreted action in context in two senses: They used both the other actions performed by the actor and the causal constraints in the situation to interpret an ambiguous action …. infants as young as 6 months construe grasping as goal directed, infants under 12 months may be able to interpret the goal of an action on the basis of sequences [of actions in context]" (pp. 76–77). Appeal to hidden beliefs or mental states is not required. See also Király et al. (2003) and Biro et al. (2007).

they easily learn what to expect of other people in such situations, and these expectations define the default cultural framework for understanding others. When I enter a classroom or a grocery store, I can immediately see who the teacher is or who the cashier is, and I can intuitively understand what they are doing, and for my particular purposes that may be sufficient for my interactions. We have no need for theories or simulations; most of our social understanding is shaped by scripts and short narratives that we learn as children (Hutto, 2007). We do not ordinarily need to go further than the already rich and complex comprehension that we gain through the perception of a situated agent—that is, of an agent who is situated in an environment which also tells us something about what that person is doing and thinking. If I see the situation and what the agent is doing in it, and how the agent is doing it, and what the agent is expressing (e.g., through his/her gestures and style of movement), and this perception is already informed by my own interaction with them and others, as well as by my previous situated experiences, my habitual ways of understanding, and by cultural norms and established practices, and so forth, then in cases which we encounter in our normal ordinary engagements the work of understanding is already sufficiently accomplished and I do not have to go any further. I do not have to start thinking about what might be going on in the other person's mind since everything I need for understanding him/her is there in his/her action and in our shared world.

Again, there is more to be said about the role of narratives in fine-tuning our social understandings. We gain narrative competency as young children, and along with it comes the ability to employ a folk-psychological practice in those rare cases where we may be entirely puzzled about someone's actions (Hutto, 2007; Gallagher & Hutto, 2008). If the cashier is dancing on the counter, or the teacher starts to throw water balloons in class, then we may adopt an observational stance (as we duck) and start to theorize or simulate about what the state of his/her mind might be. This kind of practice, however, is the exception rather than the primary or pervasive way by which we come to understand others.

IMPLICIT SIMULATION OR EMBODIED PRACTICES

The embodied practices of primary and secondary intersubjectivity, involving direct perception and pragmatic contextualizations, clearly contrast with the claims made by theory theorists and simulation theorists who conceive of social cognition as a purely mentalistic or cognitive process. Recently, however, ST has appealed to the neuroscience of resonance systems and mirror neurons (MNs) as offering scientific evidence for a form of implicit simulation. This, of course, depends on a specific interpretation of the scientific data.

We know that the perceiver's motor system is activated when he/she perceives another person performing an intentional action. The same or overlapping neural areas in parts of the frontal and parietal cortices, and specifically, MNs in the

pre-motor cortex, in Broca's area, and in the parietal cortex of the human brain are activated both when the subject engages in specific instrumental actions, and when the subject observes someone else engage in those actions (Rizzolatti et al., 1996, 2000; Grèzes & Decety, 2001). Some simulation theorists claim that these processes underpin (or are the neural correlates) of explicit acts of simulation (Jeannerod & Pacherie, 2004, p. 129). *Implicit* simulation theorists, however, contend that these subpersonal processes themselves just are a simulation of the other's intentions. Vittorio Gallese, for example, claims that activation of MNs involves "automatic, implicit, and nonreflexive simulation mechanisms …" (2005, p. 117; see also Gallese, 2007). According to Gallese, one's empathic experience of the other person at the phenomenological level is underpinned by the activity of "mirror matching neural circuits" at the brain level, which he interprets as "simulation routines, *as if* processes enabling models of others to be created" at the functional level (2001, p. 45). On this hypothesis, at the explicit, phenomenological level, one is not explicitly (consciously) simulating; rather the simulation process remains entirely at the subpersonal level.

There is a growing consensus forming around this implicit simulation idea. Decety & Grèzes (2006, p. 6) summarize Rizzolatti's position in this way:

> By automatically matching the agent's observed action onto its own motor repertoire without executing it, the firing of mirror neurons in the observer brain simulates the agent's observed action and thereby contributes to the understanding of the perceived action

Goldman (2006) distinguishes between simulation as a form of high-level (explicit) mind-reading and simulation as a low-level (implicit) mind-reading where the latter is "simple, primitive, automatic, and largely below the level of consciousness" (p. 113), and the prototype for which is "the mirroring type of simulation process" (p. 147). Research suggests that MN activation is a simulation not only of the goal of the observed action but of the intention of the acting individual, and therefore a form of mind-reading. MNs discriminate identical movements according to the intentional action and contexts in which these movements are embedded (Fogassi et al., 2005; Iacoboni et al., 2005; Kaplan & Iacoboni, 2006). Neural simulation has also been extended as an explanation of how we grasp emotions and pain in others (Avenanti & Aglioti, 2006; Minio-Paluello et al., 2006; Gallese et al., 2007). Oberman & Ramachandran (2007), who amass evidence that the MN system as an internal simulation mechanism is dysfunctional in cases of autism, reinforce the idea that "simulator neurons" are responsible for understanding actions, thoughts, and emotions.

There are, however, several conceptual problems involved in calling subpersonal mirror resonance processes "simulations" (Gallagher, 2007a, b). There are good reasons to think that subpersonal processes, such as MN activation, fail to meet the definition of simulation as it is developed in ST. In that definition, simulation involves two essential aspects: first, simulation involves instrumental control of a model as we use it to understand something that we cannot understand

directly. Second, simulation involves pretense—the idea that we use our own mental states "as if" they were the mental states of others. In contrast, however, subpersonal mirroring processes do not have an instrumental character, nor are they under our control. Rather, they are automatic and, indeed, they are elicited by the actions of others. The perceiver does not launch an MN activation as a means for making sense of the other's action; rather, the process is one of perceptual elicitation where the perceived action calls forth the activation of these neurons. Furthermore, because MNs are activated both when I act and when I see someone else act, they are neutral with respect to who the agent is (deVignemont, 2004; Jeannerod & Pacherie, 2004; Gallese, 2005; Hurley, 2005). As a result, MNs do not involve pretense, which requires distinguishing one agent (me) from another (you). There is no I or you registered in MNs, per se (see Georgieff & Jeannerod, 1998).

These kinds of issues motivate a weakened or minimal definition of simulation which jettisons the instrumental and pretense aspects and defines simulation as simply a form of matching (Goldman & Sripada, 2005; Goldman, 2006). This strategy, however, fails to explain how we understand others who are engaged in very different activities from us, or who are experiencing very different emotions. For example, I may see someone acting in a certain way (picking up an insect, for instance) and clearly enjoying it, while at the same time I feel disgust about that very action and make a pushing away gesture. Neither my emotional state nor my motoric state matches up with the relevant states of the other person, yet I clearly understand his/her emotional and motor states—they are in fact motivating my own. Furthermore, there is neuroscientific evidence that shows that MN activation does not necessarily involve a precise match between motor system execution and observed action, but may be involved in "logically related" actions (e.g., complementary actions) or in anticipating future action (Csibra, 2005; Iacoboni et al., 2005). All of this goes against the idea that MNs are simulating anything.

To deny that mirror resonance processes constitute simulations, however, is not to deny that MNs may play an important role in our interactions with others, possibly contributing to our ability to understand others, or to keep track of ongoing intersubjective relations. Rather, the alternative and more parsimonious interpretation of MN activation is that it constitutes part of the neuronal correlates of direct intersubjective perception. That is, the articulated neuronal processes that include activation in various sensory areas, but also resonating activation of MNs in the motor system, are part of what underpins a non-articulated immediate perception of the other person's intentional actions, rather than a distinct process of simulating their intentions (Gallagher, 2007a, b, 2008).[3] On this view, we need to think of perception as an enactive process (Hurley, 1998; Noë, 2004; Varela et al., 1991), as involving sensory–motor skills rather than as just sensory input/processing, as

[3]Note that MN activation is only part of the story and likely not sufficient for social perception of intentions. MNs, for example, were first discovered in monkeys, but this does not mean that monkeys are capable of social perception in the same way that humans are.

an active, skillful, embodied engagement with the world rather than as the passive reception of information from the environment. In the context of social cognition, it seems appropriate to think of mirror resonance processes as part of the structure of the perceptual process when it is a perception of another person's actions. Accordingly, mirror activation is not the initiation of simulation; it subtends a direct intersubjective perception of what the other is doing. On this interpretation, MN activation fits properly with the direct perception account of intersubjective understanding and interaction, and helps to explain such capacities already operative in infancy in certain embodied practices—practices that are emotional, sensory–motor, non-conceptual, and directly perceptual—practices that involve a perceptual sense of others and that constitute a common bodily intentionality shared by both the perceiving subject and the perceived other (Gallagher, 2001, 2005).

CONCLUSION

On the embodied view of social cognition, the mind of the other person is not something that is hidden away and inaccessible. In perceiving the actions and expressive movements of the other person in the interactive contexts of the surrounding world, one already grasps their meaning; no inference to a hidden set of mental states (beliefs, desires, etc.) is necessary. When I see the other's action or gesture, I see (I *immediately perceive*) the meaning in the action or gesture; and when I am in a process of interacting with the other, my own actions and reactions help to constitute that meaning. I not only see, but I resonate with (or against), and react to the joy or the anger, or the intention that is in the face or in the posture or in the gesture or action of the other.

The alternative, non-simulationist interpretation of the neuroscience of MNs coheres with the larger non-ToMistic, interaction view of social cognition. This view, supported by evidence from developmental and neuroscientific studies, suggests that before we are in a position to theorize, simulate, explain, or predict mental states in others, we are already in a position to interact with and to understand others in terms of their contextualized expressions, gestures, and purposive movements, reflecting their intentions and emotions. We already have specific perception-based understandings about what others feel, whether they are attending to us or not, how they are acting toward us and others, whether their intentions are friendly or not, and so forth; and in most cases, we have this without the need for personal-level theorizing or simulating about what the other person believes or desires. Moreover, we understand this without the benefit of anything that on the subpersonal level could be considered an extra cognitive step, a simulation, or inference.

REFERENCES

Allison, T., Puce, Q., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Science*, *4*(7), 267–278.

Avenanti, A. & Aglioti, S. M. (2006). The sensorimotor side of empathy for pain. In M. Mancia, (Ed.), *Psychoanalysis and Neuroscience* (pp. 235–256). Milan: Springer.

Baird, J. A. & Baldwin, D. A. (2001). Making sense of human behavior: Action parsing and intentional inference. In B. F. Malle, L. J. Moses & D. A. Baldwin (Eds.), *Intentions and Intentionality: Foundations of Social Cognition* (pp. 193–206). Cambridge, MA: MIT Press.

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, *20*, 395–418.

Baldwin, D. A. & Baird, J. A. (2001). Discerning intentions in dynamic human action. *Trends in Cognitive Science*, *5*(4), 171–178.

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development*, *72*, 708–717.

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Biro, S., Csibra, G., & Gergely, G. (2007). The role of behavioral cues in understanding goal-directed actions in infancy. *Progress in Brain Research*, *164*, 303–322.

Csibra, G. (2005). Mirror neurons and action observation. Is simulation involved? ESF Interdisciplines. http://www.interdisciplines.org/mirror/papers/.

Currie, G. & Sterelny, K. (2000). How to think about the modularity of mind-reading. *The Philosophical Quarterly*, *50*(199), 145–160.

Decety, J. & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*, 4–14.

deVignemont, F. (2004). The co-consciousness hypothesis. *Phenomenology and the Cognitive Sciences*, *3*(1), 97–114.

Dittrich, W. H., Troscianko, T., Lea, S. E. G., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, *25*, 727–738.

Falk, D. (2004). Prelinguistic evolution in early hominids: Whence motherese? *Behavioral and Brain Sciences*, *27*(4), 491–503.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science*, *308*, 662–667.

Frith, U. & Happé, F. (1999). Theory of mind and self-consciousness: What is it like to be autistic? *Mind and Language*, *14*(1), 1–22.

Gallagher, S. (2001). The practice of mind: Theory, simulation, or interaction? *Journal of Consciousness Studies*, *8*(5–7), 83–107.

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

Gallagher, S. (2007a). Logical and phenomenological arguments against simulation theory. In D. Hutto, & M. Ratcliffe, (Eds.), *Folk Psychology Re-assessed* (pp. 63–78). Dordrecht: Springer Publishers.

Gallagher, S. (2007b). Simulation trouble. *Social Neuroscience*, *2*(3–4), 353–365.

Gallagher, S. (2008). Direct perception in the social context. *Consciousness and Cognition, 17*, 535–543.

Gallagher, S. & Hutto, D. (2008). Primary interaction and narrative practice. In: Zlatev, Racine, Sinha and Itkonen (Eds). *The Shared Mind: Perspectives on Intersubjectivity* (pp. 17–38). Amsterdam: John Benjamins.

Gallese, V. (2001). The shared manifold' hypothesis: From mirror neurons to empathy. *Journal of Consciousness Studies*, *8*, 33–50.

Gallese, V. (2005). Being like me: Self-other identity, mirror neurons and empathy. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation I* (pp. 101–118). Cambridge, MA: MIT Press.

Gallese, V. (2007). Before and below "theory of mind": Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society*, *B-Biological Sciences*, *362*(1480), 659–669.

Gallese, V., Eagle, M. N. & Migone, P. (2007). Intentional attunement: Mirror neurons and the neural underpinnings of interpersonal relations. *Journal of the American Psychoanalytic Association*, *55*(1), 131–176.

Georgieff, N. & Jeannerod, M. (1998). Beyond consciousness of external events: A "Who" system for consciousness of action and self-consciousness. *Consciousness and Cognition*, *7*, 465–477.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton-Mifflin.

Goldman, A. (2006). *Simulating minds: The philosophy, psychology and neuroscience of mindreading*. Oxford, England: Oxford University Press.

Goldman, A. (2005). Imitation, mind reading, and simulation. In Hurley & Chater (Eds.), *Perspectives on Imitation II* (pp. 79–93). Cambridge, MA: MIT Press.

Goldman, A. I. (2002). Simulation theory and mental concepts. In J. Dokic & J. Proust (Eds.), *Simulation and Knowledge of Action* (pp. 1–19). Amsterdam: John Benjamins.

Goldman, A. I. & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition*, *94*, 193–213.

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

Grèzes, J. & Decety, J. (2001). Functional anatomy of execution, mental simulation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, *12*, 1–19.

Herrmann, E., Call, J., Hare, B., & Tomasello, M. (2007). Humans evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, *317*(5843), 1360–1366.

Hobson, P. (1993). The emotional origins of social understanding. *Philosophical Psychology*, *6*, 227–249.

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

Hurley, S. L. (1998). *Consciousness in Action*. Cambridge, MA: Harvard University Press.

Hurley, S. L. (2005). Active perception and perceiving action: The shared circuits model. In T. Gendler, & J. Hawthorne, (Eds.), *Perceptual Experience*. New York: Oxford University Press.

Hutto, D. (2007). *Folk Psychological Narratives*. Cambridge, MA: MIT Press.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C. & Giacomo Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, *3*(3), 529–535.

Jeannerod & Pacherie, (2004). Agency, simulation, and self-identification. *Mind and Language*, *19*(2), 113–146.

Johnson, S. et al. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-old infants. *Developmental Science*, *1*, 233–238.

Johnson, S. C. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Science*, *4*, 22–28.

Kaplan, J. T. & Iacoboni, M. (2006). Getting a grip on other minds: Mirror neurons, intention understanding, and cognitive empathy. *Social Neuroscience*, *1*(3–4), 175–183.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.

Király, I., Jovanovic, B., Prinz, W., Aschersleben, G. & Gergely, G. (2003). The early origins of goal attribution in infancy. *Consciousness and Cognition*, *12*(4), 752–769.

Leslie, A. (2000). Theory of mind as a mechanism of selective attention. In M. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (pp. 1235–1247). Cambridge, MA: MIT Press.

Leslie, A. & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology*, *6*, 315–324.

Lindblom, J. (2007). *Minding the Body: Interacting Socially through Embodied Action*. Linköping: Linköping Studies in Science and Technology, Dissertation No. 1112.

Malle, B. F. (2002). The relation between language and theory of mind in development and evolution. In T. Givón & B. F. Malle (Eds.), *The Evolution of Language out of Pre-Language* (pp. 265–284). Amsterdam: John Benjamins.

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

Meltzoff, A. N. & Brooks, R. (2001). Like me as a building block for understanding other minds: Bodily acts, attention, and intention. In B. F. Malle, L. J. Moses & D. A. Baldwin (Eds.), *Intentions and Intentionality: Foundations of Social Cognition* (pp. 171–191). Cambridge, MA: MIT Press.

Meltzoff, A. & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75–78.

Meltzoff, A. & Moore, M. K. (1994). Imitation, memory, and the representation of persons. *Infant Behavior and Development*, *17*, 83–99.

Minio-Paluello, I., Avenanti, A., & Aglioti, S. M. (2006). *Social Neuroscience*, *1*(3–4), 320–333.

Niedenthal, P. M., Barsalou, L. M., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review*, *9*(3), 184–211.

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

Oberman, L. M. & Ramachandran, V. S. (2007). The simulating social mind: The role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin*, *133*(2), 310–327.

Onishi, K. H. & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258.

Phillips, W., Baron-Cohen, S., & Rutter, M. (1992). The role of eye-contact in the detection of goals: Evidence from normal toddlers, and children with autism or mental handicap. *Development and Psychopathology*, *4*, 375–383.

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

Rizzolatti, G., Fogassi, L., & Gallese, V., (2000). Cortical mechanisms subserving object grasping and action recognition: A new view on the cortical motor functions. In M. S. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (pp. 539–552). Cambridge, MA: MIT Press.

Scheler, M. (1954). *The Nature of Sympathy*. Trans. P. Heath. London: Routledge and Kegan Paul. Original: *Wesen und Formen der Sympathie*. Bonn: Verlag Friedrich Cohen, 1923.

Schilbach, L., Eickhoff, S. B., Mojzisch, A., & Vogeley, K. (2008). What's in a smile? Neural correlates of facial embodiment during social interaction. *Social Neuroscience, 3*(1), 37–50.

Scholl, B. J. & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299–309.

Senju, A., Johnson, M. H., & Csibra, G. (2006). The development and neural basis of referential gaze perception. *Social Neuroscience*, *1*(3–4), 220–234.

Singer, W., Wolpert, D., & Frith, C. (2004). Introduction: The study of social interactions. In C. Frith & D. Wolpert (Eds.), *The Neuroscience of Social Interaction* (pp. xii–xxvii). Oxford: Oxford University Press.

Tooby, J. & Cosmides, L. (1995). Foreword to S. Baron-Cohen, *Mindblindness: An Essay on Autism and Theory of Mind* (pp. xi–xviii). Cambridge, MA: MIT Press.

Trevarthen, C. B. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before Speech* (pp. 321–347). Cambridge: Cambridge University Press.

Trevarthen, C. & Hubley, P. (1978). Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. In A. Lock (Ed.), *Action*, *Gesture and Symbol: The Emergence of Language* (pp. 183–229). London: Academic Press.

Varela, F. J., Thompson, E. & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge: MIT Press.

Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology*, *33*, 514–535.

Wellman, H. M. (1993). Early understanding of mind: The normal case. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding Other Minds: Perspectives from Autism* (pp. 10–39). Oxford: Oxford University Press.

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

Wittgenstein, L. (1967). *Zettel*. Eds. G. E. M. Anscombe & G. H. von Wright, trans. G. E. M. Anscombe. Berkeley: University of California Press.

Wittgenstein, L. (1980). *Remarks on the Philosophy of Psychology*, Vol. II. Eds. G. H. von Wright and H. Nyman, trans. C. G. Luckhardt & M. A. E. Aue. Oxford: Blackwell.

Woodward, A. L. & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science*, *11*, 73–77.