PREPRINT: Gallagher, S. 2012. Neurons, neonates and narrative: From embodied resonance to empathic understanding. In A. Foolen, U. Lüdtke, J. Zlatev and T. Racine (eds.), *Moving Ourselves, Moving Others* (167-96). Amsterdam: John Benjamins. (please cite original publication)

Neurons, neonates and narrative: From empathic resonance to empathic understanding[1]

Shaun Gallagher
Department of Philosophy
University of Memphis (USA)
School of Humanities
University of Hertfordshire

The problem I want to address is a classic one in the philosophy of mind. In that context it is called the problem of other minds, but it is a problem that is debated in and across a number of disciplines and approaches – the problem of intersubjectivity in phenomenology, empathy or understanding others in hermeneutics, social cognition or theory of mind (ToM) in cognitive sciences, psychology, and developmental studies, and most recently, under similar titles, we find discussions of motor resonance processes in the cognitive neurosciences. The basic question addressed under these different headings is: How are we able to understand other people – their intentions, their behaviors, their mental processes? All of these different titles for the problem, however, are themselves problematic and in some way beg the question. To cast the problem in terms of 'mind', 'inter-*subjectivity*', 'cognition', 'empathy', or 'motor resonance', already biases the way one is tempted to solve the problem. One strategy for balancing out, if not canceling out these different biases, is to take an interdisciplinary approach, and that is what I will do here. I will review several debates that are ongoing across these various disciplines, and, in contrast to certain standard views, I will map out an alternative position that will draw support from neuroscience, developmental psychology, phenomenology, and narrative theory.

1. Two debates about empathy

In the past several years a debate on the question of how we understand others has taken shape, motivated by the discovery of mirror neurons and what is more generally

[1] Earlier versions and elements of this paper were presented at various conferences and colloquia, including *Systèmes résonnants, empathie, intersubjectivité* at the Ecole Normale Supérieure, Paris (March 2005), *Beyond dichotomies, across the boundaries: Interdisciplinary investigations of dynamic interactions in biological and social sciences,* at the University of Minnesota (April 2005); an interdisciplinary colloquium at the University of Lund, Sweden (June 2005), and more recently at the *Vision Lab* at Harvard (April 2008), and the Mind and Culture seminar at Rutgers University (April 2008). I would like to thank all of the participants at these meetings who raised questions and offered comments, especially Vittorio Gallese, Evelyn Fox Keller, Jean-Luc Petit, Jean-Michel Roy, Giacomo Rizzolatti, Louis Sass, and Jordan Zlatev. I am also grateful to the University of Copenhagen's Body and Mind Research Priority Area, and the Center for Subjectivity Research for supporting my work during my stay as Visiting Professor.

referred to in the cognitive neurosciences as *resonance systems*. Briefly, mirror neurons are located in the premotor cortex and parietal areas and are activated in two conditions: (1) when the subject engages in intentional actions of a specific sort (e.g., actions that involve reaching and grasping) and (2) when the subject sees someone else engaging in the same kind of action.[2] More generally, it has been shown that there are overlapping neural areas (shared representations) in the brain that are activated when the subject intentionally *acts* in specific ways, *observes* the same kind of actions, or *imagines* such actions (Jeannerod 1997; Ruby and Decety 2001; Grezes and Decety 2001).

1.1 The early debate and some terminological strategies

The contemporary debate, which I will summarize shortly, was clearly prefigured by a similar one that took place at the beginning of the 20th century, based on behavioral and phenomenological observations which suggested that embodied, sensory-motor and action-related processes were important for explaining our understanding of others. Thus, Theodore Lipps, Max Scheler, Edmund Husserl, and others contested these issues in terms that involved concepts like analogy, projection, expression, and empathy. Lipps (1903), for example, discussed the concept of *Einfühlung*, which he equated with the Greek term *empatheia*, which in turn motivated Titchener's (1909) translation of the term as 'empathy'. Lipps attributed our capacity for empathy to a sensory-motor mirroring, an involuntary, "kinesthetic" inner imitation of the observed vital activity expressed by another person. Husserl, and other phenomenologists, including Scheler, and later, Heidegger, and Merleau-Ponty, developed phenomenological critiques of Lipps' account, contending that empathy is something more than these involuntary processes and that in some cases empathy happens as a solution or supplement to the breakdown or inadequacy of the more basic, automatic, perceptual understanding of others (see Zahavi 2001; 2005 for a good summary of these debates).

In light of this early discussion problems of terminology and definition become apparent. We noted that Lipps equated *Einfühlung* with *empatheia* which, in its use in late philosophical Greek, meant simply being in a state of passion, not necessarily a passion related to another person. *Hineinversetzen*, a term used by Dilthey (1926), which literally means putting oneself in the place of the other, has also been translated into English as 'empathy'. Dilthey associated this term with *Nacherleben* – to re-live something in our experience. Terminological and translation problems are also complicated by the fact that various authors make various differentiations among the concepts of empathy, sympathy, and compassion, and no two authors seem to agree on what is what, or if one is more basic than the other. Scheler (1923), for example, prefers the term 'sympathy', which means to have an accordance of feeling, and seems intrinsically intersubjective. This term is also used by Hume, who suggests a more psychological sense: "The minds of men are mirrors to one another" (1739/1978: 365). The idea that the self naturally mirrors the other goes back at least as far as Aristotle and his analysis of friendship. Hume, however, regards sympathy not as an immediate feeling, but as the result of an inference we make about the emotional state of the other (1739/1978: 576).

---

[2] These neurons were originally discovered by single cell neuron recordings in the macaque monkey (Rizzollati et al. 1999; 1996), and there is evidence that they exist in humans (Fadiga et al. 1995). See Dinstein et al. (2008) and Hickok (2009) for some critical limitations on claims about mirror neurons.

Faced with this terminological difficulty, there are several strategies that one could take. First, one could simply abandon the term 'empathy' and opt for some other word. Second, one could stipulate one's use of the word 'empathy' to signify just the basic resonance processes that are described by the research on mirror systems. Third, one could limit the use of the term to whatever is needed for understanding others that is more than just those basic resonance processes. I will pursue a fourth (and more Wittgensteinian) option designed to short-circuit the terminological and conceptual issues, and suggest that the term 'empathy' has many different meanings depending on the context in which we use it. In effect, there are many different kinds of empathy -- from those processes that involve basic resonance ("elementary empathy") to more sophisticated kinds that involve higher cognitive functions ("empathic understanding").[3] Rather than worry about how to define the term 'empathy', I want to look at what it takes, *in a full sense*, to understand others. What I mean by 'in a full sense' is what it takes for a mature adult to understand what another person intends. But to work out what it takes for a mature, adult human to understand another human, we cannot ignore how this ability develops and originates in early infancy, or what animal studies can tell us about the perception of intentional action.

1.2 The contemporary debate

The contemporary debate, motivated by the new data on mirror systems, shared representations and motor resonance, features proposals similar to Lipps' theory, associating empathy directly with motor resonance processes (e.g., Gallese 2001), and in contrast, proposals that suggest that empathy involves something more than motor resonance processes (e.g., Decety 2005). In light of these debates we want to ask: What is the relationship between resonance processes and empathy? And, what, if anything, is the "something extra" that is needed for empathy?

Neuroscience and developmental psychology inform the contemporary debate in ways that motivate a focus on two discussion areas. The first area concerns the operation of basic resonance processes, about which we have good science and good, although limited consensus; the second area concerns some kind of higher-order empathic understanding, about which we have no consensus. We can map out the terms of this debate by looking at representatives on either side of the major dividing line between these two areas. Vittorio Gallese (2001; 2003), for example, focuses his analysis of empathy on basic resonance processes. Jean Decety (2003, 2005) suggests that empathy involves a form of higher-order understanding.

Gallese is part of the neuroscientific team that discovered mirror neurons, and he builds his theory of empathy on the fact that

> when we observe goal-related behaviours … specific sectors of our pre-motor cortex become active. These cortical sectors are those same sectors that are active when we actually perform the same actions. In other words, when we observe actions performed by other individuals our motor system

[3] The distinction here follows Dilthey (1926), who suggested a distinction between *elementary understanding* and mature *empathic understanding*. As we'll see, regardless of the terminological dissensus, there is a building consensus over the idea that there are two processes at stake: an early, automatic process, and a more mature and controlled process.

> ‘resonates’ along with that of the observed agent. (Gallese 2001: 38)

For Gallese, our understanding of the other person's action relies on a neural mechanism that matches, in the same neuronal substrate, the observed behavior with a behavior that we (observers) could execute. This lived bodily motor equivalence between what we observe others doing, and the capabilities of our own motor system allows us to use our own system as a model for understanding the other's action. "Empathy is deeply grounded in the experience of our lived-body, and it is this experience that enables us to directly recognize others not as bodies endowed with a mind but as *persons* like us" (2001: 43). Thus Gallese uses action understanding as a framework to define empathy. In support of this he cites the work of Lipps: "When I am watching an acrobat walking on a suspended wire, Lipps (1903) notes, *I feel myself inside of him"* (2001: 43). Gallese's argument, however, is based on the neuroscience of mirror neurons. "I submit that the neural matching mechanism constituted by mirror neurons — or by equivalent neurons in humans — … is crucial to establish an empathic link between different individuals" (2001: 44). Gallese appeals to simulation theory to extend this model to include expressive aspects of movement that give us access to the emotional states of others (Gallese and Goldman 1998). I will return to the concept of simulation theory shortly.

Jean Decety (2002, 2003, 2004, 2005), in contrast, contends that empathy does not imply simply an action or emotion resonance initiated by the action or emotion state of the other. It also requires a minimal comprehension of the *mental states* of this person. He does not deny the importance of resonance systems, especially in early infancy, and he accepts that we have an innate capacity to feel that other people are *"like us."* But we also quickly develop the capacity to put ourselves *mentally* in the place of others. He also emphasizes that in this process difference is just as important as similarity. Empathy is founded on our capacity to recognize that others are similar to ourselves, but to do so without confusing ourselves with the other.

According to Decety (2005; Decety & Jackson 2004; Jackson, Meltzoff & Decety 2005) then, three fundamental components interact to create empathy:

1. a component of motor resonance (*resonance motrice*) whose release is generally automatic and nonintentional;
2. insight into the subjective mental perspective of the other which may be controlled and intentional;
3. the ability to differentiate between self and other.

The third component is satisfied at the same basic neuronal level of resonance, and specifically by Georgieff and Jeannerod's (1998) concept of a "who system." On this model, the shared representations (activated neuronal areas) for my action and the perception of the other person's action overlap and create the required resonance that enables the automatic recognition of the other's action as similar to action of which I am capable. But, as Georgieff and Jeannerod indicate, the overlap is not complete, so that different sets of neurons that are activated when I act are not activated when I see others act, and vice versa. This difference, they propose, allows for the differentiation between self and other (but see Legrand 2007, and note 6 below).

Both Gallese and Decety agree that basic resonance systems are in place, not only in early infancy, but also in non-human primates. The major difference between the positions represented by Gallese and Decety concerns the second component. For Gallese, this component is not something *more* than what the resonance systems already deliver, automatically; for Decety, this is the "something extra" that is needed for empathic understanding. For a precise understanding of the difference between these two positions, however, and to specify what the "something extra" is that is allegedly needed for empathic understanding, we need to briefly review a third debate that falls under the heading "theory of mind" (ToM).

2. The theory of mind debate

There are several strong candidate theories to be considered as solutions to the question of what more than resonance or mirror systems one needs in order to understand others in the full sense. Under the general title of theory of mind (ToM), the established contenders are "theory theory" (TT) and simulation theory (ST).

2.1 Theory theory

TT claims that one must take a theoretical stance toward the other in order to infer or "mindread" their beliefs, desires, or intentions (e.g., Baron-Cohen 1995; Carruthers and Smith 1996; Premack and Woodruff 1979). That is, the ability to understand others is based on having a certain kind of knowledge, a theory (or folk psychology) of how people behave in general. According to some theory theorists, the folk psychology may be learned via experience; other theorists contend that it is based on innate capacities that emerge developmentally. There is also disagreement about whether we use the theory explicitly (consciously) or implicitly (non-consciously). All theory theorists, however, embrace three suppositions, although sometimes only implicitly.

(1) that we understand others to be other *minds* that are perceptually inaccessible, and that "mindreading" involves an attempt to *explain* or *predict* the behavior of the other person on the basis of what the subject infers to be the other person's beliefs or desires;
(2) that in our encounters with others we are primarily observers;[4] and
(3) at least for many theory theorists, use of folk psychological theory is our primary and pervasive way of understanding others, once we reach the age of 4 years.

[4] Peter Carruthers denies this is the case. "In particular, it is simply false that theory-theorists must (or do) assume that mentalizing usually involves the adoption of a third-person, detached and observational, perspective on other people. On the contrary, theory theorists have always emphasized that the primary use of mindreading is in interaction with others (which Gallagher calls "second-personal") (2009b, 167). Yet the third-person observational stance is consistently implied throughout most of the literature on TT, and in most discussions of false-belief tasks, which are set up as third-person observational tasks. The observational stance is directly tied to the idea that the task in social cognition is to "explain and predict" the actions of others. To take up a task of theory-based explanation clearly implies a third-person process based on observations of the other person's actions. Even for Carruthers the task seems to be just this: "to provide fine-grained intentionalistic predictions and explanations" based on "inferences from observation" (1996). And he indicates that "we surely use our mind-reading system, for example, when processing a *description* of someone's state of mind as well as when observing their behavior" (Carruthers 2001); and he characterizes mindreading as something done by "a third-party observer" (2009a, 134).

This is the age that we begin to pass false-belief tests, and this is taken as evidence that we attain a theory of mind at this age (see below).

2.2 Simulation theory

In contrast to the knowledge-rich approach of TT, ST claims that in our attempt to understand others we do not need a folk psychology; rather, we employ our own mind as a model on which we simulate the other's mind by creating "as if" or pretend beliefs, desires, intentional states (e.g., Goldman 2006; Gordon 1996; Heal 1996). We then make inferences about the other person's mind on the basis of the simulation. As in TT, there is debate about whether our simulating ability is a product of experience or is an innate capacity activated by experience. Traditional ST suggests that simulation is explicit (conscious or introspective, involving imaginary enactments).[5] A more implicit version of ST suggests that simulation routines are non-conscious, and are performed automatically and sub-personally at the level of mirror neurons (e.g., Gallese and Goldman 1998; Goldman 2006). Similar to the suppositions that inform TT, many simulation theorists adopt the mentalizing supposition that we understand others as other minds that are perceptually inaccessible; that we primarily take the observational stance toward others; and that simulation is the primary and pervasive way in which we understand others.

Consider the claim about taking an observational stance. Both TT and ST are based on the idea that our attempts to understand others are always made in the mode of observation. Person A observes the behavior of person B and then resorts to either theory or simulation to predict or explain B's action. In TT this observational stance involves an explicit third-person perspective. As a third-person observer I consult the theory and infer what is in the other's mind and accordingly predict or explain the other's behavior. In ST this observational stance involves a first-person perspective in the sense that the simulation process is accomplished in a model of my own first-person thoughts, beliefs, desires, etc. I simulate from a first-person perspective, and then infer or project beliefs or desires to the other in order to predict or explain their behavior.

2.3 Interaction theory

In contrast to both of these approaches, I have argued elsewhere (Gallagher 2001a; 2004; 2005; 2007a&b) that our primary and pervasive way of encountering others is not characterized by observation, but by interaction. That is, the others we try to understand are usually people with whom we are interacting, engaged in some communicative act, or in some common task, or situated in some common setting. Our primary way of understanding others is worked out not via 3rd-person observation or 1st-person simulation, but via real (2nd-person) interaction in pragmatic and social contexts. As an alternative to TT and ST, I've proposed interaction theory (IT) that appeals to evidence from phenomenology and developmental psychology. IT challenges the ToMistic models of TT and ST on each supposition.

---

[5] Goldman, for example, describes it this way: "When a mindreader tries to predict or retrodict someone else's mental state by simulation, she uses pretense or imagination to put herself in the target's 'shoes' and generate the target state" (Goldman 2005; see Goldman 1989).

1. It (or IT) rejects the mentalizing supposition, that is, the Cartesian idea that other minds are hidden away and inaccessible, and cites evidence that in many cases knowing the other person's intentions, emotions, and dispositions is simply a matter of perceiving their embodied behavior in the situation. In most cases of everyday interaction no inference is necessary.
2. IT rejects the spectatorial supposition that we are primarily observers of others' behaviors. Our normal everyday stance toward the other person is not third-person, detached observation; it is second-person interaction. For the most part we are interacting with them on some project, or in some communicative practice, or in some pre-defined relation.
3. IT rejects the supposition of universality in regard to either theory or simulation. Rather, mindreading, as either TT or ST describes it, is at best a specialized activity that is relatively rarely employed. Our everyday understanding of others depends more on embodied and situated ways of perceiving and interacting with them, and is enhanced by narrative practices.

3. The science of social cognition

Let us consider the evidence for and against these different approaches. First, there are a number of phenomenological problems with the *explicit* versions of TT and ST (see Gallagher 2005 for a full discussion). To put it most simply, taking a theoretical stance or running a simulation routine is not the way it seems to happen in everyday interaction. If we take a close look at our experience as we encounter others, rather than consulting a theory or running a simulation, we seem rather to have a direct perception of how it is with others. In most cases we know what they mean or intend without consulting a theory or simulation model. The contrast to such everyday or ordinary encounters may be found in more difficult or puzzling cases, where we may not know the person, or where we find their behavior strange. In such cases we may indeed revert to the observational mode and attempt to draw on a theory or simulation routine in order to figure out what is going on. But this is relatively rare in the larger context of everyday interaction. Furthermore, in our normal second-person interactions, we do not stand back in order to draw up an explanation. Rather, we engage with others in pragmatic activities or social practices which may involve evaluative understandings of others (and of myself in light of how others view me).

To say that this is the way it seems, from a phenomenological perspective, however, does not necessarily rule out implicit versions of TT and ST. It may be the case that it does not seem that we appeal to theory, or that we run a simulation routine, but in fact we may be doing so non-consciously. Phenomenological evidence would not be able to show that this is or is not the case. Thus, we need to look at the empirical, scientific evidence offered in support of these implicit versions of TT and ST.

3.1 Theory theory and false-belief tasks

When theory theorists turn to science to find support they most frequently appeal to false-belief tests. For example, in a content change task, a child might be shown a box that appears from its packaging to contain sweets or candies. She is asked what she thinks is inside, and she naturally answers "candies." The box is then opened to reveal that something other than candies are inside, e.g., there may be pencils inside the box. The

child is then asked to think about another person, X, who may walk into the room. When shown the box, what will the other person say the box contains? It turns out that the four-year old will correctly predict that X will say that there are candies in the box. On average, however, three-year olds will reply that X will say that there are pencils in the box. This seemingly demonstrates that on average, children who are three and younger do not have a developed theory of mind since they are not able to see that another person could have false beliefs.

This kind of experiment, and other variations of false-belief tests, set up a situation in which the child is seemingly forced to rely on an abstract and theoretical approach. Often times there is no other real person X. Sometimes a puppet is used; other times the test is based on a story about some fictitious child named Sally or Maxi. In such cases the child is not involved in any kind of real second person interaction with X. For example, there are no movements or facial gestures that X might present; there is no meaningful encounter with X, and to the extent that is the case, the test is more about problem solving than it is about social cognition. Moreover, while three-year olds might have problems passing the false-belief test,[6] and explaining or predicting the action of a third person, they seemingly have no problems understanding what the experimenter is asking, or understanding the scenario that is presented to them. That is, they seem to have little or no difficulty engaging in the second-person interaction that defines their relationship with the experimenter. Regardless of what we conclude about such issues, these kinds of false-belief tests can tell us nothing about an *implicit* use of theory, since the task that is set for the child is set at a conscious, and even metarepresentational level. It is not a test for some implicit process, since the child is asked explicitly to respond explicitly, and nothing rules out the possibility that the four-year old may be using an explicit logic to arrive at the correct answer. As far as I know, there is no scientific evidence that our normal encounters with others are characterized by implicit appeals to theory. Even those studies that examine implicit brain processes in the context of theory of mind tests propose only that such activity is the substrate of an explicit process of reasoning about mental processes (e.g., Saxe et al. 2004). False-belief tasks may have much to tell us about when children come to develop a concept of belief, but they tell us very little about social cognition, unless we already assume that our understanding of others is by way of a theoretical inference about beliefs that are hidden away in the other's mind – suppositions that clearly guide the design of such experiments.

3.2 Simulation theory and the neuroscience of mirror neurons

In contrast to TT and its appeal to false belief tests, it may seem that ST has more significant scientific support in the form of the recent neuroscience of motor resonance processes. If, as in ST's formulations of an implicit simulation, the claim is that the simulation is sub-personal, instantiated in the workings of mirror neurons, or more general resonance systems, and is therefore automatic and nonconscious, then it seems that phenomenological objections have no force since the scope of phenomenology is limited to conscious processes. But let's take a close look at what is claimed about implicit neural simulation.

---

[6] This claim has now been complicated by recent studies that show that infants as young as 15 months are capable of passing properly designed (non-verbal) false belief tests (Onishi and Baillargeon 2005).

The general claim is that one's motor system reverberates or resonates in one's encounters with others. Mirror neurons in my motor system are activated when I perceive another person performing an intentional action, for example. One claim that can be made by explicit simulation theorists is that these processes underpin (or are the neural correlates) of explicit acts of simulation (Goldman 2006; Ruby and Decety 2001). For the implicit simulation theorists, however, these subpersonal processes themselves just are a simulation of the other's intentions. Gallese captures it clearly in his claim that activation of mirror neurons involves "automatic, implicit, and nonreflexive simulation mechanisms …" (Gallese 2005, 117; also see Gallese 2007). On this hypothesis, at the explicit, phenomenological level, one is not explicitly simulating; rather one is experiencing an empathic sense of the other person, and this is the result of a simulation process that happens on the subpersonal level.

Implicit neural ST understood in these or in similar terms is the growing consensus. Thus, for example, Marc Jeannerod and Elizabeth Pacherie write:

> As far as the understanding of action is concerned, we regard simulation as the default procedure …. We also believe that simulation is the root form of interpersonal mentalization and that it is best conceived as a hybrid of explicit and implicit processes, with subpersonal neural simulation serving as a basis for explicit mental simulation (Jeannerod and Pacherie 2004, p. 129; see Jeannerod 2001; 2003).

Goldman (2006) now distinguishes between simulation as a 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" (147). The claim is that mirror neuron activation is a simulation not only of the goal of the observed action but of the intention of the acting individual, and is therefore a form of mind-reading. Mirror neurons discriminate identical movements according to the intentional action and contexts in which these movements are embedded (Fogassi et al. 2005; Icoboni et al. 2005; Kaplan and Iacoboni 2007). Neural simulation has also been extended as an explanation of how we grasp emotions and pain in others (Avenanti and Aglioti 2006; Minio-Paluello, Avenanti and Aglioti 2007; Gallese, Eagle, Migone 2007). The idea that "simulator neurons" are responsible for understanding actions, thoughts, and emotions is taken up by Oberman and Ramachandran (2007) who provide evidence that the mirror neuron system as an internal simulation mechanism is dysfunctional in cases of autism.

In contrast to the claims of implicit neural ST, there are several reasons why mirror neuron activation should not be thought of as a form of simulation, and there is an alternative interpretation of the neuroscientific evidence about the mirror system that is more consistent with IT. First, let's consider the reasons for not considering mirror neuron activation an implicit simulation. The first reason concerns the meaning of 'simulation' as defined by ST. According to standard accounts of ST, (1) simulation involves pretense, and (2) has an instrumental character, i.e., it is characterized in terms of a mechanism or model that we manipulate or control in order to understand something to which we do not have direct access. These two aspects of simulation are ubiquitous in the ST literature, and are considered essential to the concept of simulation. Goldman

(2002, 7), for example, explains that simulation involves "pretend states" where, “by pretend state I mean some sort of surrogate state, which is *deliberately adopted* for the sake of the attributor's task … In simulating practical reasoning, the attributor *feeds* pretend desires and beliefs into her own practical reasoning system.” Both the instrumental and pretense character of simulation are reflected in this account. Dokic and Proust’s (2002, viii) description reflects the instrumental character: simulation means "*using* one's own evaluation and reasoning mechanisms as a model for theirs." Gordon (2004: 1) locates this instrumentalism at the neuronal level by suggesting that on the "cognitive-scientific" model, "one's own behavior control system is employed as a *manipulable model* of other such systems. (This is not to say that the "person" who is simulating is the model; rather, only that *one's brain can be manipulated to model other persons*)." Adams (2001, 384) indicates that “it is a central feature of ST that one takes perceptual inputs off-line,” that is, that simulation involves pretense. Bernier (2002, 34) also takes both instrumental and pretense aspects to be essential elements of simulation.

> According to ST, a simulator who runs a simulation of a target would use the resources of her own decision making mechanism, in an "off-line" mode, and then the mechanism would be fed with the mental states she would have if she was in the target's situation.

The aspect of pretense is one of the things that distinguishes simulation from a theoretical model or a simple practice of reasoning (see Fisher 2006). This characterization also purportedly applies to the sub-personal processes of the motor system in iST: "our motor system becomes active *as if* we were executing that very same action that we are observing" (Gallese 2001: 37). The neurons that respond when I see your intentional action, respond "*as if* I were carrying out the behavior …" (Gordon 2005: 96). For ST, in all of its various forms, the concept of simulation clearly needs to meet these two conditions: it is a process that I control in an instrumental way (in the explicit version it is "deliberately adopted"), and it involves pretense (I put myself "as if" in the other person's shoes).

It seems clear, however, that neither of these conditions is met by mirror neurons. First, in regard to the instrumental aspect, if simulation is characterized as a process that I (or my brain) instrumentally use(s), manipulate(s), or control(s), then it seems clear that what is happening in the implicit processes of motor resonance is not simulation. At the personal level, I do not manipulate or control the activated brain areas -- in fact, I have no instrumental access to neuronal activation. Nor does it make sense to say that at the subpersonal level the brain activates a model or methodology in order to generate an understanding of something else. Indeed, in precisely the intersubjective circumstances that we are considering, these neuronal systems do not take the initiative; they do not activate themselves. Rather, they are activated by the other person's action. The other person *has an effect on us* and *elicits* this activation. It is not us (or our brain) *initiating* a simulation; it’s the other who does this to us. This is a case of perceptual elicitation rather than executive control.

Second, in regard to pretense, in sub-personal mirror processes there can be no pretense. This is obviously the case if we understand neurons as vehicles or mechanisms: neurons either fire or don’t fire; they don’t pretend to fire. More to the point, however,

and to adopt the standard terminology, in regard to their representational function, what these neurons represent or register cannot involve pretense in the way required by ST. Since mirror neurons are activated both when I engage in intentional action and when I see you engage in intentional action, the mirror system is neutral with respect to the agent; no first- or third-person specification is involved (deVignemont 2004; Gallese 2005; Hurley 2005; Jeannerod and Pacherie 2004). In that case, it is not possible for them to register *my* intentions as pretending to be *your* intentions; there is no "as if" of the sort required by ST because there is no 'I' or 'you' represented.[7]

One could go against the standard characterization of simulation and argue for a more minimal conception. Goldman (2006; Goldman & Sripada 2005), for example, in reference to neural simulation, acknowledges a discrepancy between the ST definition of simulation and the working of subpersonal mirror processes. "Does [the neural simulation] model really fit the pattern of ST? Since the model posits unmediated resonance, it does not fit the usual examples of simulation in which pretend states are created and then operated upon by the attributor's own cognitive equipment (e.g. a decision-making mechanism), yielding an output that gets attributed to the target. …" To address this discrepancy Goldman and Sripida propose a generic definition of simulation:

> However, we do not regard the creation of pretend states, or the deployment of cognitive equipment to process such states, as essential to the generic idea of simulation. The general idea of simulation is that the simulating process should be similar, in relevant respects, to the simulated process. Applied to mindreading, a minimally necessary condition is that the state ascribed to the target is ascribed as a result of the attributor's instantiating, undergoing, or experiencing, that very state. In the case of successful simulation, the experienced state matches that of the target. This minimal condition for simulation is satisfied [in the neural model] (Goldman and Sripada 2005, 208).

There is good reason to think, however, that matching, as a minimal condition for simulation, cannot be the pervasive or default way of attaining an understanding of others. There are many cases of encountering others in which we simply do not adopt, or find ourselves in, a matching state. Furthermore, with respect to implicit neural ST, if simulation were as automatic as mirror neurons firing, then it would seem that we would not be able to attribute a state different from our own to someone else. But we often do this in cases where we see someone acting in a way that actually motivates the opposite reaction in us, for example, if I see someone enjoying acting in a way that for me is disgusting (Gallagher 2007a). In such cases, neither my neural states, nor my motor actions (I may be retreating with gestures of disgust just as the other person is advancing

[7] I think this is the case even if there is some aspect of mirror neuron activation that differentiates between my action and the other's action, e.g., the frequency of the spiking activity, as Legrand (2007) suggests, or possibly the timing of activation relative to other sensory-motor processes (Gallagher 2005). This last point is especially important, as Legrand points out (personal communication) , given that the brain functions in terms of neuronal assemblies which are "built" according to temporal constraints (synchrony of action potentials). The "same" neuronal activation slightly delayed may end up forming an entirely different assembly, and having a different function. Such differences would subtend a simple agent discrimination rather than a simulation.

with gestures of enthusiasm), nor my feelings/cognitions match his. Yet I understand his actions and emotions (which are completely different from mine), and I do this without even meeting the minimal necessary condition for simulation, that is, matching my state to his at any level.

Consider, in addition, the difficulties involved if we were interacting with more than one other person, or trying to understand others who are interacting with each other. Is it possible to enter into the same, or what are likely different states, and thereby simulate the neural/motor/mental/emotional states of more than one person at the same time? Or can we alternate quickly enough, going back and forth from one person to the other, if in fact our simulations must be such that we instantiate, undergo, or experience, the states in question? How complicated does it get if there is a small crowd in the room? Would there not be an impossible amount of cognitive work, or subpersonal matching required to predict or to understand the interactions of several people if the task involves simulating their mental states, especially if in such interpersonal interactions the actions and intentions of each person are affected by the actions and intentions of the others (Morton 1995 makes a similar point).

Finally it should be noted that the scientific research on mirror neurons suggests good reasons to think mirror neuron activation does not involve a precise match between motor system execution and observed action. Between 21 and 45% of neurons identified as mirror neurons are sensitive to multiple types of action; of those activated by a single type of observed action, that action is not necessarily the same action defined by the motor properties of the neuron; furthermore, approximately 60% of mirror neurons are *broadly* congruent, which means there may be some relation between the observed action(s) and their associated executed action, but not an exact match. Only about one-third of mirror neurons show a one-to-one congruence (Csibra 2005). Activation of the broadly congruent mirror neurons, therefore, may represent a complementary action rather than a similar action (Newman-Norlund et al. 2007, 55). In that case they could not be simulations.

In denying that mirror neurons are simulating or matching in such cases, I am not denying that mirror neurons may be involved in our interactions with others. Indeed, it is likely that they do contribute to our ability to understand others or to keep track of ongoing intersubjective relations. What I am denying is that they constitute simulations in any acceptable use of that term. There is a much more parsimonious interpretation of mirror neuron activation which is consistent with the IT approach to social cognition.

3.3 Interaction theory, intersubjective enactive perception, and evidence from developmental psychology

The alternative interpretation of the mirror neuron data suggests that rather than simulation, mirror neuron activation is part of the neuronal processes that underlie a form of intersubjective enactive *perception*. That is, the articulated neuronal processes that include activation of mirror neurons or shared representations may underpin a non-articulated immediate perception of the other person's intentional actions, rather than a distinct process of simulating their intentions. On this view, perception is a temporally dynamic and enactive process.

We know that mirror neurons fire 30-100 milleseconds after appropriate visual stimulation. This short time scale motivates the question of precisely where to draw the

line between perceptual processes and something that would count as a sub-personal simulation. Distinctions at the neuronal level between activation of the visual cortex and activation of the premotor cortex, do not constitute a distinction between processes that are perceptual *simpliciter* and processes that involve something more than perception (see Gallagher 2008). Even if neuronal processes that involve information-flow from sensory cortex to pre-motor cortex take some time (as much as 100 milleseconds) it is not clear that we should identify this dynamic flow as constituting a two-step process (perception plus simulation) rather than a temporally extended and enactive perceptual process. If we think of perception as an enactive process (Varela, Thompson, and Rosch 1991), as involving sensory-motor skills rather than as just sensory input/processing – as an active, skillful, embodied engagement with the world rather than as the passive reception of information from the environment – then it may be more 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. Mirror activation, on this interpretation, is not the initiation of simulation; it's part of an enactive intersubjective perception of what the other is doing.

This interpretation of mirror neuron activation provides a tight fit with the interaction theory of social cognition, which can be further supported by developmental studies. On this account the capacities for human interaction and intersubjective understanding are already operative in infancy in embodied practices that are emotional, sensory-motor, perceptual, and nonconceptual. Evidence from developmental psychology suggests that infants much younger than three-years are able to perceive the intentions and feelings of others in their movements, gestures, and actions. Our access to others is based on certain innate or early developing capacities manifested at the level of *perceptual* experience. This is the notion of *primary intersubjectivity* (Trevarthen 1979), a set of capacities that allow us to see, in the other person's bodily movements, facial gestures, eye direction, etc. what they intend and what they feel. Neonates less than an hour old, for example, are capable of imitating the facial gestures of another human (Meltzoff and Moore 1977, 1983; Gallagher and Meltzoff 1996). Although there is a debate about how precisely to characterize this behavior (see, e.g., the papers in Meltzoff and Prinz 2002) there is growing consensus that it involves the mirror resonance systems, even if they are not fully developed in the infant (see, e.g., Gallagher 2001b; Hurley 2005; Williams et al. 2001). Primary intersubjectivity also includes capacities for eye tracking, and for parsing various movements of the head, the mouth, the hands, and more general body movements as meaningful or goal-directed. Such perceptions are important for a non-mentalistic (pre-theoretical, nonconceptual) understanding of the intentions and dispositions of other persons, and they are operative by the end of the first year (Baldwin and Baird 2001; Baldwin et al 2001; Johnson 2000; Johnson *et al*. 1998). This is not a form of "mindreading" in the sense of discerning mental states hiding behind observed behavior; rather, seeing the actions and expressive movements of the other person is already to see their meaning. No inference to a hidden set of mental states (beliefs, desires, etc.) is necessary.

The infant already has a pre-reflective sense of itself as an experiencing subject; it has a perceptual sense that certain kinds of entities (but not others) in the environment are indeed such subjects; it has a sense that in some way these entities are similar to and in other ways different from itself. This is a non-mentalising understanding of the

intentions and dispositions of other persons, a perceptual grasp of emotional, embodied, enactive meaning. Moreover, it is part of what is primarily a second-person *interaction* rather than a third-person observation. One can see this in the timing and emotional mirroring of infants' behavior (Hobson 2002). Infants "vocalize and gesture in a way that seems [affectively and temporally] 'tuned' to the vocalizations and gestures of the other person" (Gopnik and Meltzoff 1997, 131). In this regard, it is the interaction itself that contributes something that is not reducible to the actions of the individuals involved.

> [T]he intentionality in the mother-infant interaction does not reside in any individual mind; it emerges as a product of their social interaction. Thus, what is intentional about the mother-infant interaction cannot be explained simply in terms of the mother's and infant's intentions with respect to each other. (Gibbs 2001)

On average, around the age of 9 mos. to 1 year, when the capacity for joint attention begins, the infant goes beyond person-to-person immediacy and enters into the contexts of shared attention, interacting with others in a way that allows for learning about the surrounding world, what things mean and what they are for. This is the beginning of secondary intersubjectivity (Trevarthen and Hubley 1978).

> The defining feature of secondary intersubjectivity is that an object or event can become a focus *between* people. Objects and events can be communicated about. … the infant's interactions with another person begin to have reference to the things that surround them (Hobson 2002: 62).

Merleau-Ponty (1962: 353) put it this way: "No sooner has my gaze fallen upon a living body in the process of acting than the objects surrounding it immediately take on a fresh layer of significance." At 18 months infants can re-enact to completion the goal-directed behavior that an observed subject does not complete, showing that they recognize the unfulfilled intentions of others (Herrmann et al. 2007; Meltzoff 1995). Secondary intersubjectivity gives us the capacity for socially and pragmatically contextualized understanding, a more developed understanding of others in context. Through all of this the infant or young child is not trying to discover mental states in the other person's head; they are trying to discover meaning in the other person's world, which is the same world that they share and in which they interact with the other.

The evidence provided by these developmental studies is not ignored by theory theorists, but rather interpreted as indicative of some "precursors" to fully developed ToM (e.g., Baron-Cohen 1995; Currie 2008). Baron-Cohen identifies three basic mechanisms that contribute to this development: the intentionality detector (ID), the eye direction detector (EDD), and the shared attention mechanism (SAM), but, on his account, these mechanisms are inadequate for explaining the more mature ToM abilities that come online at around the age of four. Gopnik and Meltzoff (1998) cite much of the same evidence mentioned here, but they interpret this as already a form theorizing in practice. The infant is honing its theoretical skills, constantly making inferences about the behavior of others, and testing them out in quasi-experimental fashion. Infants are

small scientists gradually building a folk psychology that will come to rule our more mature interpretations of others.

The capabilities and practices of primary and secondary intersubjectivity, however, are not stages that we pass through, and are not replaced by more sophisticated theory-governed interpretations, and in that sense cannot count as precursors to use of folk psychological theory. These capabilities do not disappear in adulthood; as studies of perception of emotional expression using simple point light displays demonstrate, they mature and become more subtle (Dittrich et al. 1996).[8] As adults, for example, when we see a smiling face (and other facial gestures) we automatically, involuntarily, and non-consciously attune to it with an enactive, mimetic, response (Schilbach et al., 2008). Face perception includes an enactive element through which we engage with and respond to stimuli instead of a mere passive perception of face-based cues (Schilbach et al. 2008). Even as adults we frequently need to go no further than what is already the 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, through a perception that is already informed by my interaction with the other person, as well as by my previous situated experiences, my habitual ways of understanding, and by cultural norms and established practices, 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 her gestures and style of movement), then in our normal ordinary engagements the work of understanding is already sufficiently accomplished for most practical purposes, 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 gaining some understanding of her is there in her action and in our shared world.

4. Empathy and narrative competency

Primary and secondary intersubjective capacities do not rule out the possibility of misunderstanding, unresolved ambiguity, or that the other person may in some circumstances be a real puzzle. I may not have enough information perceptually or contextually (or otherwise) to make sense out of what the other person is doing. But in a broad range of normal circumstances enough meaning for our everyday intersubjective interactions is already available in the perception of movements, gestures, facial expressions, and so on, as well as in cues provided by pragmatic and social contexts. This, however, cannot be the complete story. Our mature ability for understanding (or misunderstanding) others, even if it does not leave primary and secondary intersubjective capabilities behind, is enhanced by a different kind of practice. In opposition to TT and ST, however, I want to argue that this enhancement is not a matter of theorizing or simulating; it involves communicative and narrative competency.[9]

---

[8] As Merleau-Ponty notes, through these early developing capabilities the child can appropriate objects and "learn to use them as others do, because the [motor resonance of the] body schema ensures the immediate correspondence of what he sees done and what he himself does … and the unsophisticated thinking of our earliest years remains as an indispensable acquisition underlying that of maturity" (Merleau-Ponty 1962: 354).

[9] It's important to note that a complete explanation of our intersubjective capabilities cannot be captured by any one of these aspects. The thought that primary intersubjectivity is the full explanation, for example, the

4.1 The development of communicative and narrative competence

There is good evidence that sometime around the age of two years, a number of things happen that lead to a capacity for empathic understanding. Decety and Jackson (2004) note:

> It is around the 2nd year that empathy may be manifested in prosocial behaviors (e.g., helping, sharing, or comforting) indicative of concern for others. Studies of children in the 2nd year of life indicate that they have the requisite cognitive, affective, and behavioral capacities to display integrated patterns of concern for others in distress (Bretherton, Fritz, Zahn-Waxler, & Ridgeway 1986). During this period of development, children increasingly experience emotional concern "on behalf of the victim," comprehend others' difficulties, and act constructively by providing comfort and help (Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman 1992).

What does it take for this kind of empathy (empathic understanding) to emerge? We can point to a number of important developments in the child around this age. At 12-18 months we see the development of secondary intersubjectivity in which children start to see things in pragmatic contexts: objects start to get their meaning from the way people interact with them. Children begin to make sense of the world through their interaction with others – a process that De Jaegher and Di Paolo call "participatory sense-making" (2008). Just around the same time the ability for mirror self-recognition emerges, and this provides the child with a more objective sense of self, in contrast to an earlier, proprioceptively-based sense of self (Gallagher 2005). In addition, sometime between 15-24 months, children start to speak, or as Merleau-Ponty might put it, language starts to acquire them and advances their communicative capacities. Finally, between 18-24 months, children start to manifest an ability for autobiographical memory.

> By 18-24 months of age infants have a concept of themselves that is sufficiently viable to serve as a referent around which personally experienced events can be organized in memory…. The self at 18-24 months of age achieves whatever 'critical mass' is necessary to serve as an organizer and regulator of experience.… This achievement in self-awareness (recognition) is followed shortly by the onset of autobiographical memory… (Howe 2000: 91-92).

Along with a developing communicative competence, autobiographical memory, and a more objective sense of self, comes the capacity for self-narrative. It may be that 2-year olds work more from scripts than from full-fledged narratives; their autobiographical

---

basis for Carruthers (2009b) criticism. He focuses on what Bruckner et al. (2009) happily call "a weakly integrated swarm of first-order [sensory-motor] mechanisms," i.e., those aspects that constitute primary intersubjectivity, and he claims that "Appealing just to sensorymotor skills (as Gallagher does) is plainly inadequate to account for the flexibility of the ways in which adults and infants can interact with others" (p. 167). IT does *not* limit the explanation of intersubjectivity, however, "just to" the sensory-motor processes found in primary intersubjectivity; rather, it in addition consistently points to the capacities involved in secondary intersubjectivity and narrative competency.

memories have to be elicited by questions and prompts (Howe 2000; Nelson 2003; 2009). But from 2-4 years, children fine-tune their narrative abilities by means of a further development of language ability, autobiographical memory, and the growing stability of their sense of self.

Through narratives we also learn from others and engage more fully in participatory sense-making. Katherine Nelson (2003: 31) suggests that, in Jerome Bruner's (1986) phrase, the "landscape of action" narrative emerges in 2-year olds, "with respect to the child's own experience, which is forecast and rehearsed with him or her by parents." Self-narrative requires building on our experiences of others and their narratives. Thus, "children of 2-4 years often 'appropriate' someone else's story as their own" (Nelson 2003: 31). As Dan Hutto (2008) has pointed out, the fact that in most cultures children grow up surrounded by stories that transmit cultural meanings and values initiates them into practices of understanding reasons for action. The pragmatic and social contexts of secondary intersubjectivity become semantically enriched with the development of this kind of narrative competency.

4.2 Narrative competency and empathic understanding

I want to argue that this development of communicative and narrative competency is a necessary component of empathic understanding. I don't mean that empathic understanding requires an occurrent or explicit story telling: but it does require the ability to frame the other person in a detailed pragmatic or social context, and to understand action in that context in a narrative way. My own action, and the actions of others have intelligibility and begin to make sense when we can place them in a narrative framework (see McIntyre 1981). This kind of narrative scaffolding is an extension of secondary intersubjectivity and an enhancement of participatory sense-making. Our understanding of others and their situations, and hence the possibility of empathizing with them, is not based on attempts to get into their heads in a mentalising fashion, since we already have access to their embodied actions and the rich worldly contexts within which they act – contexts that can be translated into narratives that operate to widen or make more specific the meaning/significance of actions and expressive movements.

Through narrative competency the more primary form of empathy based on the activation of resonance systems is brought to a more conceptual level. If, for example, I see someone crying, I may immediately, on the basis of resonance processes, empathize with him on a very concrete, but still ambiguous (non-valenced) level of concern. Only when I find out his story will I be able to move to a level of empathic understanding. If, however, his story is that he is crying because he lost the gun with which he was going to kill me, then it is unlikely that any sort of positive empathic understanding will result, although I may still understand his intentions, his actions, and maybe even his motives. The story, the narrative, helps to fill in the circumstances, and for understanding of the empathic sort, one needs to understand the circumstances. Dilthey puts us on the right track.

> It is necessary to distinguish the state of mind which produced the action by which it is expressed from the circumstances of life by which it is conditioned. … [In some cases] action separates itself from the background of the context of life and, unless accompanied by an explanation of how circumstances,

> purposes, means and context of life are linked together in it, allows no comprehensive account of the inner life from which it arose. (Dilthey 1926/1988: 153).

Dilthey's account, however, remains too mentalistic; it is not the inner life or the mental life that we attempt to access, but simply the other's life in its worldly/situational contexts, and that's what narrative can capture.

As deWaal (1996) points out, differentiation between self and other is important for distinguishing empathy from emotional contagion, which involves a complete identification with the other. This is also emphasized by Reed (1994: 288): "When one empathizes, one perceives a situation from another's point of view without losing track of one's own point of view." These different perspectives are worked out and stabilized through communicative and narrative practices. To occupy a position within a narrative, and to distinguish it from another, requires more than a minimal (pre-reflective, non-conceptual, proprioceptive/kinaesthetic) self-awareness -- it requires a conceptual, objective, narrative self that is aware of itself as having a point of view that is different from others.

4.3 Narrative and ToM

One might think that this ability to distinguish different points of view, or to have this narrative competency depends on already having a theory of mind. Janet Astington (1990) argues in this way. She cites the distinction between the *landscape of action* (a narrative of simple actions) and the *landscape of consciousness* (a folk-psychological narrative which expresses "what those involved in the action know, think, or feel, or do not know, think, or feel" – Bruner 1986: 14). To understand narrative, and by extension, to empathize, she argues, we need access not only to the character's actions but also *to their minds*. We gain the latter either through folk psychological theory or simulation. Astington therefore suggests that children younger than 4 years prefer descriptive accounts of actions (the landscape of action) to folk-psychological narratives (the landscape of consciousness). Children at 4 years (when they acquire ToM) start to prefer narrative stories that include mental terms. In folk-psychological narratives we find verbs signifying mental states (thinking, remembering, desiring, believing, etc.) and attribute them to characters in the narrative: The character *believes* X; the character *desires* Y; or the character *intends* to do Z. Once we can see things in this way, Astington proposes, then we can understand the characters and their different points of view.

We can find evidence against this prioritizing of ToM and folk-psychology from experiments conducted by Bruner himself. He offers good experimental evidence against the importance of mental or folk-psychological terms (and by implication, ToM) for understanding narratives. In a study of narrative comprehension in adults (Feldman, Bruner et al. 1990), two different versions of the same story are presented. One version has a rich language of consciousness; the characters are construed as having specific mental states. A second version of the same story is stripped of all such language and is reduced to a pure language of action. Different subjects are asked to read one of the versions and then to tell the gist the story; they are asked to recount the facts of the story, and to do so in the order they occurred in the story. The results showed no significant

differences between landscape-of-action narratives and landscape-of-consciousness narratives, (1) when providing the gist; (2) in recounting the facts of the story; (3) in recounting the order of events; or even (4) in the use of reader-related mental verbs when they recount the landscape-of-consciousness narrative.

4.4 Narrative and resonance processes

While the presence of mental terms, or a folk psychological vocabulary, and by implication, ToM, may not make a difference for narrative understanding, the presence or absence of resonance processes, especially in the affective order, do seem to make a difference. Within the context of a narrative, affective resonance (as represented in expressive movements and gestures) needs to be consistent if empathic understanding is to emerge. This has been shown by Decety and Chaminade (2003). Subjects were presented with a series of video clips showing actors telling sad and neutral stories, as if they had personally experienced them. The stories were told with either congruent or incongruent motor expression of emotion. As a measure of empathy the subjects were then asked to rate the mood of the actor and how likable they found that person. When the subjects were exposed to sad stories (eliciting an empathic understanding) versus neutral stories, there was increased activity in emotion processing-related structures (including the amygdala and parieto-frontal areas) predominantly in the right hemisphere. But when the story-tellers showed incongruent facial expressions (happy gestures while telling a sad story, for example) these areas were not activated, indicating an absence of empathy. These areas of neural activation respond not simply to perceived features of action and expression (and the subjectivity of the other person) but also to the larger story, the represented scene, the narrative circumstances of the other person, and how features of action and expression match or fail to match those circumstances. The affective resonance that comes along with expressive movements and gestures, and the pragmatic sense of the person's instrumental actions are not without relevance for empathic understanding. We have argued, however, that they are not enough; one needs to see these elements in the larger situation or in the context of the larger story.

5. Conclusion

If we return to the contemporary debate about empathy, we are now in a position to get a clearer idea of the different positions on this question. First, on an account that is consistent with the neural or implicit simulation theory, empathy is automatically generated in the activation of the mirror system. Gallese, for example, equates empathy with the motor resonance processes that he also equates with an implicit simulation. Thus, in his "shared manifold hypothesis" he distinguishes three levels of analysis (see Gallese 2001: 45).

- The *phenomenological level:* the sense of similarity with persons like us -- the *empathic* level involving actions, emotions and sensations.
- The *functional level:* simulation routines, *as if* processes enabling models of others to be created.
- The *subpersonal level:* mirror matching neural circuits -- resonating body schemas.

Accordingly he claims, "… sensations, pains and emotions displayed by others can be empathized, and therefore understood, through a mirror matching mechanism" (2001: 45).

In contrast to this, Decety suggests that basic resonance processes are not sufficient for empathy. For empathy one needs to have in place a more sophisticated theory of mind (perhaps in the form of an explicit simulation ability) in addition to the resonance processes. That is, something more than basic resonance activity is required for empathy.

Within these debates there is a growing consensus around the idea that infants are capable of very basic or elementary empathic behavior, although disputes remain whether we should explain this behavior in terms of a pre-curser to theory of mind (Baron-Cohen 1995, Gopnik and Meltzoff 1998; Meltzoff 2002), something that is already simulation (Gallese 2001), or something closer to intersubjective enactive perception, as in interaction theory (De Jaegher and Di Paolo 2007; Gallagher 2001a; 2004; 2008a&b; Reddy 2008; Rochat 2004; Zahavi 2008). There is also some agreement that something more subtle and sophisticated happens as part of human social maturity. The following table (Table 1) summarizes a variety of positions on this point.

Table 1: Different views of empathic behavior

Basic processes starting in infancy	**"Something more"**	
Elementary understanding	Empathic understanding	Dilthey (1926)
Percursor processes (ID, EDD, SAM)	Theory of mind (TT)	Baron-Cohen (1995)
Low-level simulation	High-level simulation	Goldman (2006)
Basic empathy (implicit simulation)	Reenactive empathy (explicit simulation)	Steuber (2006)
Primary and secondary intersubjectivity	Communicative and narrative competency	Gallagher (2006); Gallagher and Hutto (2008)

The consensus is that there are at least two parts to this story. There is, of course, no consensus on what constitutes either of the parts – elementary or empathic understanding – or what the "something more" is that constitutes the more advanced capacity. In most cases it is acknowledged that what starts in infancy does not end in infancy but continues and is perhaps transformed by the more developed processes. I have argued against theory and simulation accounts, and in place of ToM capabilities, I have suggested that the development of communicative and narrative competency provides the "something more" needed for empathic (or even non-empathic) understanding.[10]

[10] Here I remind the reader that we've adopted a Wittgensteinian strategy in regard to the terminological problem with the term 'empathy'. I'm using this term in a broad sense to encompass a variety of phenomena. Notwithstanding this strategy let me point out that uses of terms like 'empathic understanding' (Dilthey) or 'reenactive empathy' (Steuber) suggest that our normal and everyday understanding is always empathic, and that may be tied to the specific theory of understanding to which these theorists hold. For example, Steuber equates empathy with simulation and considers ST the "default

In part, what I have attempted to map out here is a story about how we understand others, consistent with both phenomenological and scientific evidence. With respect to the basic capacities for understanding others, ToMistic approaches invest in either theory (folk psychology) or in simulation. These approaches have recently been trying to account for capacities that clearly develop in children younger than 4 years of age. ST has been helped by the recent advances in the neuroscience of resonance systems that may offer some explanation for these earlier capacities for intersubjectivity. IT takes as its starting point just such capacities for primary and secondary intersubjectivity, and argues, in contrast to TT and ST, that these are embodied, sensory-motor capacities of enactive perception. But clearly this is not sufficient to explain our more developed capacity for empathic understanding.

One option would be to say that IT provides a good account of intersubjectivity up until the fourth year of life, at which time the child acquires a theory of mind – an ability to use folk psychology or simulation routines to make inferences about other people's mental states. The use of mentalising inferences, however, seems to be more the exception than the rule, and would make empathic understanding more a matter of observational logic than of being moved by the other's situation. Rather than pursuing some form of TT or ST, or some hybrid version of ToM, I've tried to make the case for the importance of communicative and narrative competency to address this issue. The capacities of primary and secondary intersubjectivity, which characterize our human interactions in early and late infancy, are not replaced by a cold theoretical logic, or a self-controlled simulation. They are extended through language and autobiographical memory into a narrative competency that allows us to recognize the other person's circumstance and to construct an appropriately nuanced narrative understanding.

References

Adams, F. 2001. Empathy, neural imaging, and the theory versus simulation debate. *Mind and Language* 16 (4): 368-92.

Astington, J. 1990. Narrative and the child's theory of mind. In B. K. Britton and A. D. Pellegrini (eds.), *Narrative Thought and Narrative Language* (151-171). Hillsdale, NJ: Lawrence Erlbaum Associates.

Avenanti, A. and Aglioti S. M. 2006. The sensorimotor side of empathy for pain. In M. Mancia (ed.), *Psychoanalysis and neuroscience* (235-256). Milan: Springer.

Baldwin, D. A. and Baird, J. A. 2001. Discerning intentions in dynamic human action. *Trends in Cognitive Science,* 5 (4): 171-78.

Baldwin, D. A., Baird, J. A., Saylor, M. M., and Clark, M. A. 2001. Infants parse dynamic action. *Child Development*, 72 (3): 708-17.

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

---

position for our ability to understand rational agents within the folk-psychological context" (2006: 167). In that sense empathy is not a special way of understanding others; it is simply the way that we always understand others. Although I want to suggest that there is a variety of phenomena that might be called empathic, I want to stop short of equating empathy with any and all forms of intersubjective understanding.

Bernier, P. 2002. From simulation to theory. In J. Dokic and J. Proust (eds.), *Simulation and Knowledge of Action* (33-48). Amsterdam: John Benjamins.

Bretherton, I., Fritz, J., Zahn-Waxler, C., & Ridgeway, D. 1986. The acquisition and development of emotion language: A functionalist perspective. *Child Development 57*: 529-548.

Buckner, C., Shriver, A. Crowley, S. and Allen, C. 2009. How "weak" mindreaders inherited the earth. *Behavioral and Brain Sciences*, 32: 140-41.

Bruner, J. 1986. *Actual Minds, Possible Worlds*. Cambridge, MA: Harvard University Press.

Carruthers, P. 1996. Simulation and self-knowledge: a defense of theory-theory. In P.Carruthers and P.K. Smith (eds.), *Theories of Theories of Mind* (22-38). Cambridge: Cambridge University Press.

Carruthers, P. 2002. The cognitive functions of language. *Behavioral and Brain Sciences* 25 (6): 657-674.

Carruthers, P. 2009a. How we know our own minds: the relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, 32 (2009), 121-138.

Carruthers, P. 2009b. Mindreading underlies metacognition. *Behavioral and Brain Sciences*, 32: 164-176.

Carruthers, P. and Smith, P. K. (eds.) 1996. *Theories of Theories of Mind.* Cambridge: Cambridge University Press.

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

Currie, G. 2008. Some ways to understand people. *Philosophical Explorations* 11 (3): 211 – 218.

Decety, J. 2002. Naturaliser l'empathie [Empathy naturalized]. *L'Encéphale*, *28*, 9-20.

Decety, J. 2003. L'empathie ou l'émotion partagée [Empathy: Sharing emotions]. *Pour La Science*, 309: 46-51.

Decety, J. 2004. Empathie et mentalisation a la lumiere des neurosciences sociales. *Neuropsychiatrie: Tendances et Debats*, 23: 25-35.

Decety, J. 2005. Une anatomie de l'empathie. *Psychiatrie, Sciences Humaines, Neurosciences*. 3 (11): 16-24.

Decety, J., and Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, 41: 127-128.

Decety, J. & Jackson, P. L. 2004. The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*3 (2): 71-100

De Jaegher, H., & Di Paolo, E. 2007. Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciencee* 6 (4): 485-507.

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

de Waal, F. 1996. *Good Natured: The Origins of Right and Wrong in Humans and Other Animals.* Cambridge, MA: Harvard University Press.

Dilthey, W. 1926. *Gesammelte Schriften.* Vol. 7. Göttingen-Stuttgart: Vandenhoeck & Ruprecht. Partially translated in K. Mueller-Vollmer (ed.), *The Hermeneutics Reader* (152-64). New York: Continuum, 1988.

Dinstein, I., Thomas, C., Behrmann, M. and D.J. Heege. 2008. A mirror up to nature. *Current Biology* 8 (1): R13-R18.

Dittrich, W. H., Troscianko, T., Lea, S.E.G, and Morgan, D. 1996. Perception of emotion from dynamic point-light displays represented in dance. *Perception* 25: 727-38.

Dokic, J. and Proust, J. 2002. Introduction. In J. Dokic and J. Proust (eds.), *Simulation and Knowledge of Action* (vii-xxi). Amsterdam: John Benjamins.

Fadiga, L., Fogassi, L., Pavesi, G. and Rizzolatti, G. 1995. Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology* 73: 2608–11.

Feldman, C. F., Bruner, J., Renderer, B. & Spitzer, S. 1990. Narrative comprehension. In In B. K. Britton and A. D. Pellegrini (eds.), *Narrative Thought and Narrative Language* (1-78). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fisher, J. C. 2006. Does simulation theory really involve simulation? *Philosophical Psychology* 19 (4): 417–432

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.

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

Gallagher, S. 2001b. Emotion and intersubjective perception: A speculative account. In *Emotions, Qualia and Consciousness* (95-100). London: World Scientific Publishers, and Naples: Instituto Italiano per gli Studi Filosofici.

Gallagher, S. 2004. Understanding interpersonal problems in autism: Interaction theory as an alternative to theory of mind. *Philosophy, Psychiatry, and Psychology* 11 (3): 199-217.

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

Gallagher, S. 2006. The narrative alternative to theory of mind. In R. Menary (ed.), *Radical Enactivism: Intentionality, Phenomenology, and Narrative* (223-29). Amsterdam: John Benjamins.

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

Gallagher, S. 2007b. Simulation trouble. *Social Neuroscience.* 2 (3-4): 353-65.

Gallagher, S. 2008a. Direct perception in the intersubjective context. *Consciousness and Cognition* 17: 535–543

Gallagher, S. 2008b. Inference or interaction: Social cognition without precursors. *Philosophical Explorations* 11 (3): 163-174.

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

Gallagher, S., & Meltzoff, A. N. 1996. The earliest sense of self and others: Merleau-Ponty and recent developmental studies. *Philosophical Psychology* 9: 211-233.

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 2005. 'Being like me': Self-other identity, mirror neurons and empathy, In Hurley, S. and Chater, N. (eds.), *Perspectives on Imitation* I (101-118). Cambridge, MA: MIT Press.

Gallese, V. 2003. The roots of empathy: The shared manifold hypothesis and the neural

basis of intersubjectivity. *Psychopathology* 36:171–180

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

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.

Gallese, V. & Goldman, A. 1998. Mirror neurons and the simulation theory of mind-reading', *Trends in Cognitive Sciences* 12: 493–501.

Georgieff, N. & Jeannerod, M. 1998. Beyond consciousness of external events: A 'Who' system for consciousness of action and self-consciousness. *Consciousness and Cognition*, 7: 465-77.

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

Goldman, A. 2005. Mirror Systems, Social Understanding and Social Cognition. *Interdisciplines*. (http://www.interdisciplines.org/mirror/papers/3)

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

Goldman, A. 1989. Interpretation psychologized. *Mind and Language* 4: 161-185.

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

Gopnik, A. & Meltzoff, A. 1998. *Words, Thoughts, and Theories*. Cambridge, MA: MIT Press.

Gordon, R. M. 1996. 'Radical' simulationism. In P. Carruthers and P. K. Smith (eds.), *Theories of Theories of Mind* (11-21). Cambridge: Cambridge University Press.

Gordon, R. M. 2004. Folk psychology as mental simulation. In N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. (http://plato.stanford.edu/archives/fall2004/entries/folkpsych-simulation/).

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

Heal, J. 1996. Simulation, theory and content. In P. Carruthers and P. K. Smith (eds.), *Theories of Theories of Mind* (75-89). Cambridge: Cambridge University Press.

Herrmann, E., Call, J., Hare, B., Tomasello, M. 2007. Humans evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* 317 (5843): 1360-66

Hickok, G. 2009. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience* 21 (7): 1229-1243.

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

Howe, M. L. 2000. *The Fate of Early Memories: Developmental Science and the Retention of Childhood Experiences*. Washington: American Psychological Association.

Hume, D. 1739/1978. *A Treatise of Human Nature*. Oxford: Clarendon Press.

Hurley, S. L. 2005. The shared circuits model. How control, mirroring, and simulation can enable imitation and mind reading. *Interdisciplines*. (http://www.interdisciplines.org/mirror/papers/5).

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): 529-35.

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

Johnson, S. *et al.* 1998. Whose gaze will infants follow? The elicitation of gaze-following in 12-month-old infants. *Developmental Science* 1: 233-38.

Jackson, P. L. Meltzoff, A. N. and Decety, J. 2005. How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage* 24: 771– 779

Jeannerod, M. 2003. The mechanism of self-recognition in humans. *Behavioural Brain Research* 142: 1–15

Jeannerod, M. 2001. Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage* 14: 103-109.

Jeannerod, M. 1997. *The Cognitive Neuroscience of Action.* Oxford: Blackwell Publishers.

Jeannerod, M. and Pacherie, E. 2005. Agency, simulation, and self-identification. *Mind and Language* 19 (2): 113-46.

Kaplan, J.T. and M. Iacoboni. 2006. Getting a grip on other minds: Mirror neurons, intention understanding, and cognitive empathy. *Social neuroscience* 1 (3-4): 175-83.

Legrand, D. 2007. Naturalizing the acting self: Subjective vs. anonymous agency. *Philosophical Psychology* 20 (4): 457-478

Lipps, T. 1903. Einfulung, innere nachahmung und organenempfindung, *Archiv. F. die Ges. Psy.* vol I, part 2 (Leipzig: W. Engelmann).

McIntyre, A. 1981. *After Virtue.* South Bend: University of Notre Dame Press.

Meltzoff, A. N. 2002. Imitation as a mechanism of social cognition: Origins of empathy, theory of mind, and the representation of action. In U. Goswami (Ed.), *Blackwell Handbook of Childhood Cognitive Development* (6-25). Oxford: Blackwell.

Meltzoff, A.N. 1995. Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, **31**: 838–50.

Meltzoff, A.N. and Moore, M.K. 1977. Imitation of facial and manual gestures by human neonates. *Science* 198: 75–8.

Meltzoff, A. N. and Moore, M.K. 1983. Newborn infants imitate adult facial gestures. *Child Development,* 54: 702-709.

Meltzoff, A. N. and Prinz, W. 2002. *The Imitative Mind: Development, Evolution, and Brain Bases.* Cambridge: Cambridge University Press.

Merleau Ponty, M. 1962. *Phenomenology of Perception.* Trans. C. Smith. London: Routledge and Kegan Paul.

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

Morton, A. 1996. Folk psychology is not a predictive device. *Mind* 105: 119-37.

Nelson, K. 2003. Narrative and the emergence of a consciousness of self. In G. D. Fireman, T. E. McVay, Jr. and O. J. Flanagan (eds.), *Narrative and Consciousness: Literature, Psychology, and the Brain* (17-36). Oxford: Oxford University Press.

Newman-Norlund, RD. Noordzij, ML. Meulenbroek, RGJ, Bekkering, H. 2007. Exploring the brain basis of joint attention: Co-ordination of actions, goals and intentions. *Social Neuroscience* 2 (1): 48-65.

Oberman, L. M. and 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. and Baillargeon, R. 2005. Do 15-month-old infants understand false beliefs? *Science* 308 (5719): 255-58.

Premack, D. and Woodruff, G. 1979. Does a chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 4: 515-26.

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

Reed, E. S. 1994. Perception is to self as memory is to selves. In U. Neisser and R. Fivush. *The Remembering Self: Construction and Accuracy in the Self-narrative* (278-92). Cambridge: Cambridge University Press.

Rochat, P. 2004. Emerging co-awareness. In G. Bremner and A. Slater (eds.), *Theories of Infant Development* (258-282). New York: Wiley-Blackwell.

Rizzolatti, G., Fadiga, L., Gallese, V. and Fogassi, L. 1996. Premotor cortex and the recognition of motor actions. *Cognition and Brain Research* **3**: 131–41.

Rizzolatti, G., Fadiga, L., Fogassi, L. and Gallese, V. 1999. Resonance behaviors and mirror neurons. *Arch. It. Biologie* **137**: 83–99.

Ruby, P. and Decety, J. 2001. Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nature Neuroscience* 4 (5): 546-50.

Saxe, R. Carey, S. and Kanwisher, N. 2004. Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology* 55: 87–124.

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. and Vogeley, K. 2008. What's in a smile? Neural correlates of facial embodiment during social interaction. *Social Neuroscience* 3 (1):37-50

Steuber 2006

Titchener, E. B. 1909. *Lectures on the Experimental Psychology of Thought-Processes*. New York: Macmillan.

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

Trevarthen, C. and 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* (183-229). London: Academic Press.

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

Williams, J. H.G., Whiten, A., Suddendorf, T. and Perrett, D. I. 2001. Imitation, mirror neurons and autism. *Neuroscience and Biobehavioural Review* 25: 287-295.

Zahavi, D. 2001. Beyond empathy: Phenomenological approaches to intersubjectivity. *Journal of Consciousness Studies* 8 (5-7): 151-67.

Zahavi, D. 2005. *Subjectivity and Selfhood*. Cambridge, MA: MIT Press.

Zahavi, D. 2008. Simulation, projection and empathy. *Consciousness and Cognition* 17: 514-522.

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. 1992. Development of concern for others. *Developmental Psychology* 28: 126-136.