

Preprint: Gallagher, S. (in press -2006). *Logical and phenomenological arguments against simulation theory*. In D. Hutto and M. Ratcliffe (eds.), *Minding our Practice: Folk Psychology Re-assessed*. Springer Publishers.

## Logical and phenomenological arguments against simulation theory<sup>1</sup>

*Shaun Gallagher*  
Philosophy and Cognitive Sciences  
University of Central Florida  
gallaghr@mail.ucf.edu

Theory theorists conceive of social cognition as a theoretical and observational enterprise rather than a practical and interactive one. According to them, we do our best to explain other people's actions and mental experience by appealing to folk psychology as a kind of rule book that serves to guide our observations through our puzzling encounters with others. Seemingly, for them, most of our encounters count as puzzling, and other people are always in need of explanation. By contrast, simulation theorists do their best to avoid the theoretical stance by using their own experience as the measure of everyone else's. When it comes to explaining how we understand other people some of the very best contemporary philosophers, psychologists, and neuroscientists are simulationists. For example, Vittorio Gallese, Alvin Goldman, Robert Gordon, Jane Heal, Susan Hurley, and Marc Jeannerod. This short list of simulationists, however, already involves some problems. Not everyone on this list understands simulation in the same way. In effect, there are different simulation theories, and although it is important to distinguish them, and I will do so before I go much further, I will in the end argue against all of them. For several reasons I don't think that the concept of simulation explains our primary and pervasive way of understanding others, any more than theory theory does.

It should be obvious from what I've just said that being against simulation theory (ST) in this way does not mean that I favor its main competitor, theory theory (TT). The alternative to both of these approaches is not something that I will directly try to argue for here. I've done so elsewhere (Gallagher 2001, 2004, in press). But part of the argument that I develop against simulation here, feeds into and depends on that alternative. So a

---

<sup>1</sup>An earlier version of this paper was entitled "Against simulation," and presented at a Workshop with Marc Jeannerod, Jean-Michel Roy, and Roberto Poli. CPER Research program: Representationalism and the foundations of cognitive science: new problems, new perspectives. Ecole Normale Supérieure, Lyon (May 2005). Subsequently I found Rebecca Saxe's paper with the same title, so I changed mine. I presented a version of this paper at the University of Florida Philosophy Colloquium (September 2005). I thank members of these various meetings, as well as Dan Hutto and Matthew Radcliffe for constructive feedback. I wrote this when I was Visiting Professor at the University Research Priority Area: Body and Mind and Center for the Study of Subjectivity at the University of Copenhagen and I am grateful for their support.

quick summary is in order. Here it is. There is good evidence from developmental psychology, neuroscience, and phenomenology that we understand others, primarily and pervasively, in a non-mentalistic and perceptual way, which is enhanced around the age of four years by a developing narrative competency. This doesn't rule out the possibility that in rare cases we do take a theoretical stance, or that we sometimes use simulation routines to solve puzzling cases. But these are the rare cases. Ordinarily, in everyday encounters, and in the pragmatic and social contexts that characterize our everyday encounters with others, we perceive their movements, gestures, facial expressions, and speech acts as meaningful and intentional, without worrying about their minds or about how to explain or predict their actions. If, instead of directly interacting with others, we are called upon (or we call upon ourselves) to think more deeply about them, our tendency is not to theorize about them, or to automatically put ourselves in their place, in the manner of a simulation, but to generate a narrative framework that would facilitate our understanding of them (Hutto 2003, 2004; Gallagher, in press; Gallagher and Hutto, forthcoming).

Let me start the argument against simulation theory by outlining several different versions of simulation theory. Simulation theory, as an approach to theory of mind, has been developed in at least three different versions. An argument against one version will not necessarily work against all of them. So it is important to identify the specific targets in order to line the arguments up in the right way. The three different versions of ST are based on three different conceptions of what constitutes the simulation involved. In the first version, the simulation is something explicit, that is, an exercise of conscious imagination and deliberative inference. Goldman is a good representative of this position. In the second version, the simulation is also explicit, but does not involve a deliberative inference. This is Gordon's notion of radical simulation. In the third version, the simulation is something subpersonal, either a functional mechanism that is cashed out in terms of neuronal processes (e.g., Hurley 2005), or these neuronal processes themselves (e.g., Gallese 2001). There are, of course, hybrid theories that combine, within ST, explicit and implicit processes (e.g., Gallese and Goldman 1998), and hybrid theories that combine some version of ST with TT (see Currie and Ravenscroft 2002; Nichols and Stich 2003; Mitchell 2005; and Saxe 2005).

### **Explicit simulation and some arguments against it**

In the explicit version of ST, simulation involves conscious or introspective mental states in which I imagine myself in the other's situation and use the model (simulation) that is generated to predict the other's mental states. Goldman, for example, argues that simulation is explicit insofar as it involves a conscious introspective use of the imagination to conceptually manipulate propositional attitudes (beliefs, desires). "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 2005a). According to Goldman, simulation involves three steps.

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 2005b, 80-81.)

One might think that there would be problem with the very first step: "the attributor creates in herself pretend states intended to match those of the target." This suggests that the simulator already has some idea of what's going on with the other person. The question then is where does that knowledge come from and why isn't that already the very thing we are trying to explain. Hybrid theorists who combine TT and ST suggest that folk psychology provides, not a sense of what is going on with the other person, but some general rules about how people think and behave in a certain situations, and that this is what the simulationist can use to generate the pretend mental states needed for the simulation process (e.g., Currie and Ravenscroft 2003). In contrast, Goldman appeals to sub-personal mirror resonance processes (discussed below), although he then faces the problem of how to translate these processes into a conceptual grasp of propositional attitudes. He proposes to solve the latter problem by an appeal to phenomenal properties of propositional attitudes (2002). Goldman suggests that a belief feels different from a desire because it is generated by different sub-personal processes, which are themselves generated by differential activations induced by our perception of the other.

There are both logical and phenomenological arguments that have been made against this explicit version of ST. First, we should note that explicit ST is similar to an older theory called the argument by analogy. Gordon and Cruz (2003) recognize this forerunner as such. In the first half of the 20<sup>th</sup> century arguments were raised against this theory by a number of philosophers, including Max Scheler and Gilbert Ryle, and some of the same arguments work against ST. Ryle, for example, argued that the logic of simulation isn't correct because the idea of imputing to a variety of others what is true of my simulated action ignores the diversity of their actions. "[T]he observed appearances and actions of people differ very markedly, so the imputation to them of inner processes closely matching [one's own or] one another would be actually contrary to the evidence" (Ryle 1949: 54). A similar objection to the logic of simulation was raised by Scheler (1954). If I project the results of my own simulation on to the other, I understand only myself in that other's situation, but I don't understand the other.<sup>2</sup> Given the diversity of motives, beliefs, desires, and behaviors in the world, is not clear why such a simulation process would be at all reliable. Scheler also suggests another argument against ST. The explicit simulation process seems cognitively too complicated. Infants, and perhaps even animals, seem capable of understanding the intentions of others, but it would be difficult to attribute the complex cognitive processes involved in simulation to them.<sup>3</sup>

---

<sup>2</sup> Hurley points to this as one difference between TT and ST. "When I use practical reason off-line in mind-reading, I don't formulate normative laws from which I make inferences; rather, I activate my own normative and deliberative dispositions" (2005).

<sup>3</sup> There is now good scientific evidence to suggest that infants are indeed capable of understanding the intentions of others. Meltzoff (1995) shows that children at 18 months of age are capable of recognizing and completing another person's failed intention. The experimenter pretends to have great difficulty accomplishing a certain task with a toy, and presents an incomplete action to the child. The child, who

Here I'll propose one more argument, which I will call the simple phenomenological argument, against explicit ST. On the explicit version of ST, simulation is not only explicit but pervasive. That is, we use it all the time, or at least it is the default way of understanding others. Goldman (2002: 7-8) thinks this is a moderate claim.

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.<sup>4</sup>

If simulation is both explicit and pervasive, then one should have some awareness of the different steps that one goes through as one consciously simulates the other's mental states. But there is no phenomenological evidence for this. When I interact with or come to understand another person, there is no experiential evidence that I use such conscious (imaginative, introspective) simulation routines. That is, when we consult our own common experience of how we understand others, we don't find such processes. Of course, this is not to say that we never use simulations, but that in itself is telling. It may be the case that confronted with some strange or unaccountable behavior I do try to understand the other person by running a simulation routine (or by appealing to theory). I think this is the rare case, however. Moreover, it tends to stand out in its rarity. I can easily become aware that I am in fact taking this approach, and it is all the more apparent when I do this, simply because it tends to be the exception. But this tells against the idea that I employ simulation in the usual everyday circumstance. Most of our encounters are not third-person puzzles solved by first-person procedures. They are second-person interactions in which I easily have a sense of what is going on with the other person based on our common pragmatic or socially contextualized interactions, with no cognitive simulation required.

A possible defense of explicit ST is to make it a little implicit. Perhaps explicit simulation can be made so habitual that it becomes implicit, so that we do it without being aware that we do it, in the same way that we drive a car without being explicitly aware of all of our driving habits, or in the same way that an expert may employ cognitive strategies that become so habitual that the expert is no longer aware of how she does what she does. The simple phenomenological objection would be that if such implicit processes stay at the personal level, they would remain accessible to conscious reflection, or at least they would become apparent, as unworkable habits, in problematic situations when our habitual strategies break down. We can become aware of a habit that we are not usually aware of in such circumstances. This simply does not seem to be the case for the sort of simulation process described by explicit ST. Indeed, we may find ourselves initiating

---

takes the toy and completes the task with little effort, demonstrates that she understands what the experimenter desired to do. Even earlier, infants seem capable of parsing intentions (see Baird and Baldwin 2001; Baldwin and Baird 2001). Whether some animals are capable of understanding intentions is still a debated issue.

<sup>4</sup> Third-person mentalization signifies simply that one person is trying to understand another person, rather than trying to understand herself (which would be first-person). Some theory-of-minders contend that we use simulation (or theory) to understand our own minds.

simulation processes (or in some cases appealing to folk psychology) precisely in cases where our ordinary abilities to understand others break down.

### **Radical simulation**

Robert Gordon's notion of simulation is radical in the sense that, on his view, the capacity for simulation is what allows us, in the first place, to recognize another person as someone who "mind-endowed" (Gordon 2004: 2). It is not simply that we understand the other to have beliefs and desires, and then use simulation to ascertain those mental states. Rather, simulation is the process that gives us to understand that the other is minded. At the same time, radical simulation is less mentalistic than the traditional explicit ST. I do not retreat introspectively to my own mind to run simulation routines by manipulating propositional attitudes like beliefs and desires. Rather, I put myself in the other person's perspective and look to see what she thinks is true about the world. This involves a transformation that takes place on the personal level. By using my imagination, I imagine/simulate what the other person must think in her situation. I do not imagine myself in her situation; I imagine *her* in her situation, by imaginatively occupying her situation. The transformation involves an egocentric shift, but does not involve either introspection of my own mental states, or inference making about the other's mental states. I'm not concerned with mental states at all. I imagine, in the first-person, how the other person sees the world.

The point I am making is that once a personal transformation has been accomplished, there is no remaining task of mentally transferring a state from one person to another, no question of comparing [the other person] to myself. For insofar as I have recentered my egocentric map on [the other person], I am not considering what [I] would do, think, want, and feel in the situation. (Gordon 1995a: 54)

How do we effect this transformation? How do we know what the other person sees or thinks?

My own view [...] is that the method we ordinarily use is limited to identifying states in the first person, but, thanks to our capacity for imaginatively transforming ourselves into other "first persons", it is not exclusively a one-person method. (58)

We use our imagination and we use ascent routines. When, for example, I am asked whether I believe that the car is red, I don't start an introspective routine in order to find a belief located some place in my mind; rather, I look at the car to see if it is red, or I remember the car's color. I answer the question about my belief by appealing to a more primary cognitive procedure – perception or memory (see Evans 1982). In the same way, when I am asked whether my neighbor believes the car is red, I look at the car from her first-person perspective (Gordon 1995a, 1995b). I settle the question of whether the other person believes

P by asking, within the framework of the simulation, whether it is the case that P.  
I concern myself with the world from the other person's perspective.

The description of such ascent routines, which radically discounts mentalism and the propositional attitudes, is clearly a more parsimonious phenomenological account of our epistemological activities. It is, however, more difficult to cash out simulative transformations; and since these transformations are meant to be as explicit as the simulation routines described by Goldman, one should be able to get a purchase on them using phenomenological currency. Thus, the simple phenomenological argument works against radical ST also. Although Gordon does away with the need for an extra step involving inference, because we are "already there" in the other's perspective, these transformations still require an "as if" component. Otherwise my own first-person perspective on the world would simply collapse into the first-person perspective of the other and the self/non-self distinction would disappear.

In ordinary circumstances, however, my experience of the other, even in the observational mode that seems to dominate theory-of-mind discussions, is not characterized by either an "as if" component or by a collapse of the self/non-self distinction. When I look out my window and see a man standing across the road, I need nothing more than my own perception of the environment to understand that his perspective is an opening onto just the same environment that I see, and I need not go further than that in most cases. That is, I don't have to transform myself into his perspective to know that he happens to see the road from an angle that differs from my view. I can see that this must be the case simply from the differences that define our positions vis a vis the road, and from the orientation and postural stance of his body. Ordinarily, it is enough to say that he sees the same road. And as for asking what might be going on inside the man's head, Gordon is right that I usually have no reason to ask. Even if the man suddenly starts jumping up and down in an excited fashion as he is looking and pointing up the road, I do not have to think about what propositional attitudes he might be entertaining, since I can see that he is excited about something. Nor, however, do I attempt (or need to attempt) to put myself in his place in order to find out what might be exciting him; I do not attempt an egocentric shift of my perspective. In fact, I'm not quite sure what I could learn by doing so. Rather, I move to a position from which I can look up the road to see what he is seeing. No simulation is required, and if there is a transformation involved, it is simply the transformation of my own physical movement. I come to an understanding of his action, and his perception, and his excitement, through an ascent routine that is focused on our mutually perceived world, rather than on what might be happening within his mind, or within his perspective. Of course this may fail. If I see nothing up the road that might be the occasion of his excitement, then indeed, I may have to use some simulative or theoretical routines to work out what the man is doing. This won't get me very far, however, unless I gain some further information. And without further information, simply by transforming my egocentric perspective into his I will remain puzzled. Perhaps, by simulation, I would hypothesize that he is playing a joke on me, or, by appeal to theory, that he is delusional. But I would still need more information about the man's character – I would need to know the man's story – to determine whether my simulative or theoretical supposition was correct. This kind of strategy, however, is motivated only in unusual circumstances

when the other person's behavior is puzzling. Most of my interactions with others are not like this, and are not observational in nature. Most are second-person interactions where pragmatic and social contexts are shared; and in cases where I know the person, I already know their character and possess a good narrative framework to allow understanding without the mental acrobatics required for simulation routines or transformations.

### **Implicit simulation**

ST can easily counter the simple phenomenological argument by moving to the more serious version of implicit simulation.<sup>5</sup> ST has gained more ground in recent years by appealing to good neuroscientific evidence involving sub-personal activation of mirror neurons, shared representations, or more generally, resonance systems. If simulation is sub-personal, and not something of which we would be aware, then phenomenology is not in a position to raise objections, since phenomenology gives us access only to conscious experience.

First, let's review the recent neuroscience. The basic finding in this regard is that one's motor system reverberates or resonates in one's encounters with others. My motor system is activated when I perceive another person performing an intentional action, for example. Mirror neurons in the pre-motor cortex and in Broca's area 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). Also, specific overlapping neural areas (shared representations), in parts of the frontal and parietal cortexes, are activated under the following conditions: (1) when I engage in intentional actions; (2) when I observe some other person engage in that action; and (3) when I imagine myself or another person engage in that action (e.g., Grezes and Decety 2001). These subpersonal mechanisms are said to constitute a simulation of the other's intentions (Gallese 2001; Gallese and Goldstein 1998). Gallese captures it clearly in his claim 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 'resonates' along with that of the observed agent .... action understanding heavily relies on a neural mechanism that matches, in the same neuronal substrate, the observed behaviour with the one [the observer could execute] ... (2001: 38-39)

---

<sup>5</sup> It is interesting to note, however, that the implicit version of ST is actually an argument against the explicit version of ST. That is, if our understanding of others is in fact mediated by an implicit and automatic simulation process, then we have little need for the more explicit version. Indeed, to the extent that an implicit ST would explain the phenomenological scarcity of explicit simulation, it would support the simple phenomenological argument against explicit simulation. Along this line Gallese states: "Whenever we face situations in which exposure to others' behavior require a response by us, be it active or simply attentive, we seldom engage ourselves in an explicit, deliberate interpretive act. Our understanding of a situation most of the time is immediate, automatic, and almost reflex like" (2005:102).

The processes themselves are very real, and there is good neuroscientific evidence to support this. But is it appropriate to characterize these processes as simulations, as Gallese goes on to do?

According to this hypothesis, 'understanding' is achieved by modeling [simulating] a *behaviour* as an *action* with the help of a motor equivalence between what the others do and what the observer does. (39)

This is a subpersonal process generated by "automatic, implicit, and nonreflexive simulation mechanisms ..." (Gallese 2005: 117). He refers to his model as the "shared manifold hypothesis" and distinguishes between three levels (Gallese 2001: 45):

- The *phenomenological level* is the one responsible for the sense of similarity ... that we experience anytime we confront ourselves with other human beings. It could be defined also as the *empathic level* ....
- The *functional level* can be characterized in terms of simulation routines, *as if* processes enabling models of others to be created.
- The *subpersonal level* is instantiated as the result of the activity of a series of mirror matching neural circuits.

As we saw, one possible objection to ST is that since it employs a model that is first-person, or at least confined to my own system (a simulation in my own mind or motor system) nothing justifies inferring anything about what must be going on in the other person. As Dan Hutto (in press) suggests, "what is needed is a reliable method of indicating the *other's* perspective on events, not a projection of our own on or into theirs." Does a sub-personal simulation lock us up within our own 1st-person system?

Defenders of the implicit version of ST have an answer to this. Mirror neurons (and shared representations) are, it is claimed, *neutral* -- neither first nor third person -- they are activated both for my own action and for observation of the other's action: activation of the system simulates the intentional action but not the agent (deVignemont 2004; Gallese 2005; Hurley 2005; Jeannerod and Pacherie 2004). In this case, the sub-personal simulation process, like its explicit cousin, involves a multi-step process. First, we perceive the other's behavior; this is followed immediately by activation of shared representations -- in neutral mode (registering "naked intentions"); and this is followed by a determination of agency (i.e., a specification of who did the action -- me or the other person (Jeannerod and Pacherie 2004). This final step is accomplished by what Georgieff and Jeannerod (1998) have called the "Who" mechanism.

It will pay to stop and consider Jeannerod and Pacherie's claims about naked intentions. They assume that an articulation at the level of the neural activations between those responsible for (1) registering in the perceiving system the ("naked") intention in an action, and (2) registering the agency for the action, means that there is an articulation in experience between the perception of intention and the experience of agency. "We can be aware of an intention, without by the same token being aware of whose intention it is. ... something more than the sole awareness of a naked intention is needed to determine its

author." If in fact the brain can process information about intentions without assigning agency to the intentions, is it legitimate to say that our experience is similarly articulated? Jeannerod and Pacherie suggest that it is.

When the naked intention one is aware of yields an overt action, the extra information needed to establish authorship may be found in the outside world. The question 'Is this intention mine?' would then be answered by answering the question: 'Is this my body performing the corresponding action?' (140).

Phenomenologically (experientially), however, intentions in almost all cases come already clothed in agency. The "who" question hardly ever comes up at the level of experience, because the neural systems have already decided the issue -- one way or the other -- i.e., even if I'm wrong about who is acting, I am still experiencing or perceiving the intention as already determined in respect to agency. The wonderful thing about the "Who system" is that it is completely neurological and sub-personal -- and the results of its activation are hardly ever experientially manifested as "making a decision about who did the action." Rather, the results of its activation are experientially manifested as "X's action" where X is either you or me. Indeed, our direct perception is highly reliable in regard to discriminating between self and non-self. Pathologies and oddly arranged experiments may reveal "who" problems, but in normal ecological behavior it is generally clear whose intention/action it is. As we know from philosophers like Wittgenstein, Shoemaker, and Evans, the identification question -- "Someone is intending to pick up the apple, is it me?" -- just doesn't come up. There is, in effect, no necessary isomorphism between the phenomenological level and the neuronal level. So if the neuronal processes can be defined as involving a step-wise process, this does not mean that a step-wise process needs to show up in phenomenology.<sup>6</sup>

This sets the stage for the first argument against implicit ST. The claim I want to make is that the neural resonance processes of which we have been speaking, in fact, are part of the processes that underlie intersubjective *perception* rather than *simulation*. That is, these processes have to do with a direct perception of the other person's intentions, rather than a distinct mental process of simulating their intentions. This claim requires that we conceive of perception as a temporal phenomenon, and as enactive, and thus as involving motor processes.

First, mirror neurons fire 30-100 ms after appropriate visual stimulation (Gallese, private correspondence). What is, even in neurological terms, a short amount of time between activation of the visual cortex and activation of the pre-motor cortex, raises the question of where precisely to draw the line between the act of perception and something that would count as a simulation. Even if it is possible to draw a line between activation of the visual cortex and activation of the pre-motor cortex, this does not mean that this line distinguishes, on either a functional or phenomenological level between perception and simulation as a step-wise process.

---

<sup>6</sup> On the question about isomorphism between subpersonal and personal levels, see Gallagher 1997; Hurley 2005; Varela 1996.

Simulation, according to various versions of ST, involves a step-wise process that begins with perception and ends with some form of understanding. We first see an action that we need to understand; we then simulate it in our own mind or motor system; we then attribute agency for the action, or infer or understand something about the other's experience. But if neuronal processes that send information from sensory cortex to pre-motor cortex take some time (as much as 100 ms), it is not clear that we should identify two functionally distinct steps rather than a temporally extended and enactive perceptual process. That is, at least in terms of temporal parameters, the fact that at the neurological level A (sensory processing) is followed by B (activation of mirror neurons) does not mean that at a functional level one should distinguish between perception and simulation, or that at a phenomenological level there is a similar step-wise process happening.

If we think of perception as an enactive process (e.g., Noë 2004; Hurley 19XX) – as sensory-motor, and not just as sensory input – then it may be more appropriate to think of the resonant processes as part of the structure of the perceptual process when perception is of the action of conspecifics or members of close species. Mirror activation is not the initiation of simulation; it's part of a direct perception of what the other is doing. At the phenomenological level, when I see the other's action or gesture, I see (*directly perceive*) the meaning in the action or gesture. I see the joy or I see the anger, or I see the intention in the face or in the posture or in the gesture or action of the other. I see it. I don't have to simulate it. And I immediately see that it is *their* action, gesture, emotion, or intention, and it is extremely rare that I would be in a position to confuse it with my own. Although Jeannerod and Pacherie (2004) defend a version of ST, they nicely express the phenomenological alternative: "Perception and action are closely integrated and when we visually perceive actions, we seem to be immediately sensitive to the distinctive properties of intentional behavior" (p. 139).

Of course the simulationist can accept the phenomenology ("Yes indeed, that is what *seems* to happen") and still hold to the interpretation that these specific subpersonal processes involve simulation. In fact, however, if we use Gallese's distinctions, this interpretation goes from the functionalist level to the neurological level, but with no clear justification. That is, what happens on the neurological level is simply a complex sequence of neuronal activations. If we look at those processes from a functionalist perspective already framed by ST, then we tend to read those processes as involving simulation. If, in contrast, we look at those processes from a phenomenological level that suggests a direct perception of the other's intentions, then we tend to read those processes as perceptual without simulation. Can the simulationist offer any convincing evidence that the activation of resonance processes is in fact a simulation?

This brings us to a second argument against the implicit version of ST. What theorists of implicit simulation (Gallese, Jeannerod, Pacherie), and even critics of implicit ST, like Saxe, call 'simulation', is not simulation in any genuine sense of the word. Consider, first, two definitions of 'simulation' offered by the OED. (1) Simulation is an imitation, in the sense of something *not real* – counterfeit; to simulate means to feign, to pretend. We can find this use of the term in Wittgenstein. "Why can't a dog simulate pain? Is he too honest? Could we teach a dog to simulate pain? Perhaps it is possible to teach him to

howl on particular occasions as if he were in pain, even when he is not" (Wittgenstein 1958: § 250). (2) Simulation in the sense of a simulator: a model (a thing) that we can *use* or do things with so we can understand the real thing. We can find both senses of the term in the literature of ST. Consider the following characterizations (italics are mine). Simulation means "*using* one's own evaluation and reasoning mechanisms as a model for theirs ..." (Dokic and Proust 2002: viii). 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. (Goldman 2002: 7).

The surrogation or pretense however, is of a precise kind. Bernier (2002) makes this explicit as an essential element found in ST.

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*. (Bernier 2002: 34)

For ST, a simulation is not simply a model that we use to understand the other person -- theoretical models would suffice if this were all that is required. Even the fact that the model is constituted in our own mechanisms is not sufficient. Rather, I must use the model "as if" I were in the other person's situation. As Gallese puts it, "our motor system becomes active *as if* we were executing that very same action that we are observing" (2001: 37). Gordon puts this right at the neuronal level: the neurons that respond when I see your intentional action, respond "*as if* I were carrying out the behavior ..." (2005: 96). If we call this the "pretense condition," then simulation has these two characteristics: it is a process that I control (in the explicit version it is "deliberately adopted"), and it involves a pretense condition (I put myself "as if" in the other person's shoes).

We find these two characteristics in almost every description of simulation in ST. Mental simulation is a cognitive "*ability or heuristic or methodology*" (Jacob 2002, who cites Gordon for the latter term) -- by which I "*engage* in pretense," *put myself* in someone else's shoes, *compare* my experience to their experience, and *predict* their mental state, emotion, or behavior. We *use* ourselves as a model ... I *create* in myself some pretend beliefs ... and so forth. This is the way simulation is characterized not only by theorists of explicit simulation, but also by theorists of implicit simulation. The pretense condition mentioned by Gallese is accomplished in a simulation considered to be "an interactive *model* of what cannot be known in itself" (2003). At the subpersonal level, the brain in a stepwise fashion is *modelling* the intentional action of others. Gordon (2004: 1) suggests 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)."

If simulation is characterized as a process that I (or my brain) uses or controls, if this is what simulation is, then it seems clear that what is happening in the implicit processes of motor resonance is not simulation. We, at the personal level, do not *do* anything with the activated brain areas -- in fact, we have no access to neuronal activation, and we can't use it as a model. Nor does it make sense to say that at the subpersonal level the brain is *using* a model or methodology, or *comparing* one experience with another, or *creating* pretend states, or that one set of neurons makes use of another set of neurons as a model.<sup>7</sup> The neuronal systems do not take the initiative; they do not activate themselves, but are activated by the other person's action. The perception of the other person's action automatically activates in our brain the same areas that are activated when we engage in similar action. The other person *has an effect on us*. The other *elicits* this activation. This is not a simulation, but a perceptual elicitation. It is not us (or our brain) *doing* it, but the other who does this to us.<sup>8</sup>

But wait. Even if resonance processes described as the activation of mirror neurons and shared representations are not initiated and controlled, there is another way to think of simulation on the subpersonal level that does involve controlled comparison. In the motor control literature and in discussions of agency, theorists do talk about comparators and models that the brain uses. The term 'simulation' is also used. Efference copy sent through forward control mechanisms, for example, are said to constitute a simulation of an intended movement in order to compare it with an ongoing movement, to predict its success. The brain uses this simulation to make fast non-conscious corrections that keep the action on track. This use of the term seems closer to its general meaning even though these processes are automatic. Some theorists have appealed to these processes as possible mechanisms involved in the simulation of another's action (e.g., Gallese 2001; Hurley 2005: 181-188; see Iacoboni, cited in Millikan 2005: 188n2). Our own motor system comparators are activated to simulate and thereby anticipate the other's action. The brain could be said to predict the other person's actions in this way. On this account the perception of the other's action is automatically informed by a sub-personal simulation; perception of action involves a loop through the motor control comparator. Can ST adopt this model of simulation?

The problem in this case is that the pretense condition is not met; there is no "*as if* it were I" involved, and in that regard it fails to be the kind of simulation required by ST. If indeed the subpersonal simulation is neutral in regard to who's action is at stake (Gallese 2005; Jeannerod and Pacherie 2004; Hurley 2005), then it can be only a representation of an intentional action in my motor system, but not a representation of my own motor action as if it were the other's. Even if a "Who" mechanism adds a specification of

---

<sup>7</sup> I'm not sure the problem is solved by attributing such things to the simulation itself, as Hurley (2005) does: "Simulation *uses* certain processes to generate related information, rather than *theorizing about* them in separate meta-processes."

<sup>8</sup> It may seem contradictory to claim in the previous argument that perception is enactive, or as Noë says, "perception is action," and in this argument to claim that the activation of the resonance system is the result of a passive elicitation, so that the motor aspect of perception does not involve our action, but is a case of us being affected by the other. I think that a fuller account of enactive perception has to be able to accommodate this passive, affective aspect of perception (see Gallagher 2005).

agency, differentiating between self and non-self, it is not at all clear that, as Gordon (2005: 96) suggests, the neurons respond "*as if* I were carrying out the [other's] behavior" in any sense in which the "as if" registers sub-personally.<sup>9</sup> A specification in my motor system that the action belongs to another is not equivalent to the specification "*as if* I were carrying out the action." If this is a simulation of intentional action, it is nonetheless not the kind of simulation that ST needs; it may be nothing more than motor priming or emulation, or what Hurley calls mirroring (2005: 184).

It is therefore not clear why we should think of the activation of resonance systems as a simulation process of the sort required by ST. This is not to deny that there are resonance processes at work in our perception of the other person. Moreover, the nature of the resonance processes involved in such encounters makes our perception of other conspecifics different from our perception of objects and instruments. But it doesn't make social cognition the result of an implicit simulation.

## Conclusion

I have brought a number of arguments against various versions of ST. What I've called the simple phenomenological argument can challenge both explicit and radical versions, especially if we take these versions to describe our primary, pervasive, and perhaps default mode of understanding others. In such cases, ST should be able to call on phenomenological evidence to verify the simulation model. In fact, phenomenological reflection on our ordinary experience speaks against ST. In the same way that in our everyday encounters we find a scarcity of theoretical musing about others, we find a scarcity of simulation routines at work. One could argue, of course, that phenomenology is not always correct, even in regard to the most explicit experiences. But that argument would have to be extended too far if the claim were to be made that in fact we use simulation routines all the time, but phenomenology constantly and consistently misses this fact.

In addition to what I've called the simple phenomenological argument, phenomenologists and others, like Ryle, have pointed to logical and conceptual problems with ST. One of these objections challenges all versions of ST, even though it was originally brought

---

<sup>9</sup> It is possible to use brain imaging to identify neuronal areas differentially activated for first-person simulated action (the subject imagines herself doing the action) and third-person simulated action (the subject imagines that she is the other person doing the action) (Ruby and Decety 2001). Results show that there is a good deal of overlap of neuronal areas responsible for these different activities, but also some non-overlapping areas that supposedly accounts for the differences in agency and perspective. Similar studies show overlapping areas for first-person action and first-person simulated (imagined) action. Thus, although it seems reasonable that one might be able to point to a certain area of neuronal activity associated with pretense (imaginative enactment or simulation) in these cases, such activation is generated by a personal level, consciously controlled pretense. What has not been shown, as far as I know, is an automatic pretense condition that takes place without conscious control, on the subpersonal level. Indeed, the concept of pretense, the "as if," may simply be a personal level category that cannot be applied to neuronal activity *per se*. Although one might say that neurons respond "as if" X were the case (e.g., *as if* there were an oasis, even though it is a mirage) it is not clear that the "as if" aspect is neuronal rather than an interpretation of that activity.

against the earlier theory based on inference from analogy. Given the large diversity of motives, beliefs, desires, and behaviors in the world, it is not clear how a simulation process based on my own relatively narrow experience (or relatively unique circumstance) can give me a reliable sense of what's going on in the other person's mind, or in their behavior.

Implicit ST may be in a better position to answer this particular charge. Perhaps the common tendency to anthropomorphize (which isn't necessarily bad if the other we are concerned about is another human), or to use the intentional stance, is based on just those motor resonance systems that work on a basic inter-bodily level. Moreover, the test of ST shouldn't be framed in terms of reliability. After all, we often misunderstand others, and whatever mechanism accounts for intersubjectivity, it will necessarily have some degree of unreliability built in.

The problem with implicit ST is that it is only one, and not necessarily the best interpretation of the significance of motor resonance systems. I've argued that implicit resonance processes are not simulations in any sense that is useful for ST. Furthermore, if implicit ST does give an account of our primary and pervasive ability to understand others, it would count as an argument against explicit ST, since explicit simulations would be redundant in this case. Likewise, however, if our default mode of understanding others were based on explicit simulation, as Goldman contends, then the claims of implicit ST about the adequacy of motor resonance processes would be wrong. Goldman's view of implicit motor resonance processes is that they do not constitute simulations of a sort that would be sufficient to do the full job, but do generate some background information that is useful to initiate the explicit simulation process. I've argued, in contrast to both explicit and implicit ST, that implicit motor resonance processes are important enactive processes that contribute to the constitution of the perceptual access that we have to the intentions of others.

I do not claim that we get a full account of human intersubjectivity in the idea that we have *perceptual* access to the intentions of others. Perceptual access to the other person's bodily movements, gestures, facial expressions, and so forth does give us a sense of what is going on with them, what they mean and what they feel, and this, together with our interactions with them in pragmatic and social contexts, gives us a relatively reliable, but still minimal understanding of them. There is much more to say about the role of language and narrative competency in a fuller account of intersubjectivity (see Hutto 2003, 2004, and forthcoming; Gallagher in press, Gallagher & Hutto forthcoming). Even in that larger story, however, the theory of mind approaches that emphasize either simulation, or the role of folk psychology as background theory, have a minimal role to play in our normal and everyday interactions.

## References

- Baird, J. A. and Baldwin, D. A. 2001. Making sense of human behavior: Action parsing and intentional inference. In *Intentions and Intentionality: Foundations of Social Cognition*, ed. B. F. Malle, L. J. Moses, and D. A. Baldwin (pp. 193-206). Cambridge, MA: MIT Press.
- Baldwin, D. A. and Baird, J. A. 2001. Discerning intentions in dynamic human action. *Trends in Cognitive Science* 5 (4): 171-78.
- Bernier, P. 2002. From simulation to theory. In J. Dokic and J. Proust (eds.), *Simulation and Knowledge of Action* (33-48). Amsterdam: John Benjamins.
- Currie, G. and Ravenscroft, I. 2002. *Recreative Minds*. Oxford: Oxford University Press.
- deVignemont, F. 2004. The co-consciousness hypothesis. *Phenomenology and the Cognitive Sciences* 3(1): 97-114.
- Dokic, J. and Proust, J. 2002. Introduction. In J. Dokic and J. Proust (eds.), *Simulation and Knowledge of Action* (vii-xxi). Amsterdam: John Benjamins.
- Evans, G. 1982. *The Varieties of Reference*, ed. J. McDowell. Oxford: Oxford University Press.
- Fadiga, L., et al. (1995), 'Motor facilitation during action observation: a magnetic stimulation study', *Journal of Neurophysiology*, 73, 2608-2611.
- Gallagher 1997. Mutual enlightenment: Recent phenomenology in cognitive science. *Journal of Consciousness Studies* 4 (3): 195-214
- Gallagher, S. 2001. The practice of mind: Theory, simulation, or interaction? *Journal of Consciousness Studies* 8 (5-7): 83-107.
- 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. (in press). The narrative alternative to theory of mind. *Consciousness and Emotion*.
- Gallagher, S. 2005. Book review: A new movement in perception: Review of Alva Noë's *Action in Perception*. *Times Literary Supplement* (London).
- Gallagher, S. and Hutto, D. (forthcoming). Primary interaction and narrative practice: From empathic resonance to folk psychology. In: Zlatev, Racine, Sinha and Itkonen (eds). *The Shared Mind: Perspectives on Intersubjectivity*. Amsterdam: John Benjamins.
- 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., & Goldman, A., 1998. Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences* 2: 493-501.

- Georgieff, N. and 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. 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. 2005a. Mirror Systems, Social Understanding and Social Cognition. *Interdisciplines*. (<http://www.interdisciplines.org/mirror/papers/3>).
- Goldman, A. 2005b. Imitation, mind reading, and simulation. In Hurley and Chater (eds.) *Perspectives on Imitation II* (80-81). Cambridge, MA: MIT Press.
- Gordon, R. M. 2004. Folk psychology as mental simulation. In N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. (URL = <http://plato.stanford.edu/archives/fall2004/entries/folkpsych-simulation/>).
- Gordon, R. M. 1995a. Folk Psychology as Simulation. In: M. Davis & T. Stone (Eds.). *Folk Psychology*. Oxford: Blackwell. 60–73.
- Gordon, R. M. 1995b. Developing commonsense psychology: Experimental data and philosophical data. APA Eastern Division Symposium on Children's Theory of Mind, December 27, 1995. ([http://www.umsl.edu/~philo/Mind\\_Seminar/New%20Pages/papers/Gordon/apakids9.htm](http://www.umsl.edu/~philo/Mind_Seminar/New%20Pages/papers/Gordon/apakids9.htm)).
- Gordon, R. and Cruz, J. 2003. Simulation theory. *Encyclopedia of Cognitive Science* London: The Nature Publishing Group Macmillan Reference Ltd.
- Grèzes, J. and Decety, J. (2001), Functional anatomy of execution, mental simulation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, 12, 1-19.
- Hurley, S. L. 2005. Active perception and perceiving action: The shared circuits model. In T. Gendler and J. Hawthorne (eds.), *Perceptual Experience*. New York: Oxford University Press.
- Hurley, S. L. 1998. *Consciousness in Action*. Cambridge, MA: Harvard University Press.
- Hutto, D. 2003. Folk-psychological narratives and the case of autism. *Philosophical Papers* 33 (3): 345-61.
- Hutto, D. 2004. The limits of spectatorial folk-psychology. *Mind and Language* 19: 548-73.
- Hutto, D. (in press). Unprincipled engagements: Emotional experience, expression and response.
- Hutto, D. (forthcoming). *Folk Psychological Narratives*. Cambridge, MA: MIT Press.
- Jacob, P. 2002. The scope and limits of mental simulation. In J. Dokic and J. Proust (eds.), *Simulation and Knowledge of Action* (87-109). Amsterdam: John Benjamins.
- Jeannerod and Pacherie 2004. Agency, simulation, and self-identification. *Mind and Language* 19 (2) 113-46.
- Meltzoff, A.N. 1995. Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology* 31: 838-850.
- Millikan, R. G. 2005. Some reflections on the theory theory-simulation theory debate. In Hurley and Chater (eds.) *Perspectives on Imitation II* (182-88). Cambridge, MA: MIT Press.
- Mitchell, J. P. 2005. The false dichotomy between simulation and theory-theory: the argument's error. *Trends in Cognitive Sciences* 9 (8): 363.
- Nichols, S. and Stich, S. 2003. *Mindreading*. Oxford: Oxford University Press.

- Noë, A. 2004. *Action in Perception*. Cambridge, MA: MIT Press.
- Rizzolatti, G., Fogassi, L. and 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. 539-52). Cambridge, MA: MIT Press.
- Rizzolatti, G., Fadiga, L., Gallese V. and Fogassi, L. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3: 131-141.
- 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.
- Ryle, G. 1949. *The Concept of Mind*. New York: Barnes and Noble.
- 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.
- Saxe, R. 2005. Hybrid vigour: Reply to Mitchell. *Trends in Cognitive Sciences* 9 (8): 364.
- Varela, F. J. 1996. Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, 3 (4): 330-49.
- Wittgenstein, L. 1958. *Philosophical Investigations*. Trans. G. E. M. Anscombe. Third edition. New York: Macmillan.