



**HAL**  
open science

## The Direct-Perception Model of Empathy: a Critique

Pierre Jacob

► **To cite this version:**

Pierre Jacob. The Direct-Perception Model of Empathy: a Critique. *Review of Philosophy and Psychology*, 2011, 2, pp.519-540. ijn\_00753488

**HAL Id: ijn\_00753488**

**[https://hal.science/ijn\\_00753488](https://hal.science/ijn_00753488)**

Submitted on 19 Nov 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## The Direct-Perception Model of Empathy: a Critique

Pierre Jacob

Published online: 6 August 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** This paper assesses the so-called “direct-perception” model of empathy. This model draws much of its inspiration from the Phenomenological tradition: it is offered as an account free from the assumption that most, if not all, of another’s psychological states and experiences are unobservable and that one’s understanding of another’s psychological states and experiences are based on inferential processes. Advocates of this model also reject the simulation-based approach to empathy. I first argue that most of their criticisms miss their target because they are directed against the simulation-based approach to mindreading. Advocates of this model further subscribe to an expressivist conception of human behavior and assume that some of an individual’s psychological states (e.g. her goals and emotions, not her beliefs) can be directly perceived in the individual’s expressive behavior. I argue that advocates of the direct-perception model face the following dilemma: either they embrace behaviorism or else they must recognize that one could not understand another’s goal or emotion from her behavior alone without making contextual assumptions. Finally, advocates of the direct-perception model endorse the narrative competency hypothesis, according to which the ability to ascribe beliefs to another is grounded in the ability to understand narratives. I argue that this hypothesis is hard to reconcile with recent results in developmental psychology showing that preverbal human infants seem able to ascribe false beliefs to others.

---

I am grateful to Gyury Gergely for inviting me to the International Conference on Intersubjectivity and the Self, Marie Curie DISCOS Project, in Budapest (June 2010), where a version of this paper was presented. I am grateful to Shaun Gallagher and Dan Zahavi for their oral and written responses to the paper and to Dan Hutto for inviting me to contribute a paper to the present special issue of RPP and for his own comments on the paper. Thanks also to Frédérique de Vignemont for numerous discussions on this topic. I gratefully acknowledge support of a grant from the French ministry of research (ANR-BLAN SOCODEV). I dedicate this paper to the memory of my friend Marc Jeannerod, who died on July 1st, 2011.

---

P. Jacob (✉)  
Institut Jean Nicod, ENS/EHESS/CNRS, Paris, France  
e-mail: piotrjacob@gmail.com

## 1 Introduction

This paper is about two related concepts, which have recently been at the center of a new embodied approach to human social cognition: *primary intersubjectivity* and *empathy*. While the expression ‘primary intersubjectivity’ was coined by the developmental psychologist Trevarthen (1979), the word ‘empathy’ was introduced into English at the beginning of the twentieth century as a translation of the German word *Einfühlung*, which means “feeling into” (another’s psychological experience).<sup>1</sup> Empathy is currently the focus of much work in both empirical cognitive social neuroscience and philosophy. Furthermore, much of this research is conducted within a simulation-based approach to mindreading. Arguably, the word is meant to apply paradigmatically to the experience of one individual (the *empathizer*) who comes to *share* another’s (the *target’s*) affective or emotional experience (e.g. pain) as a result of her awareness of the other’s experience. Given the emphasis on the link between empathizing and sharing, it is not surprising that the prevalent contemporary views of empathetic phenomena are simulation-based accounts according to which empathy is the output of a process of mental *simulation*.

Some advocates of the simulation-based approach to empathy hold the view that so-called processes of *mirroring* (or mimicry), exemplified by the activity of mirror neurons (MNs), underlie all instances of empathetic experiences (cf. Gallese 2001; Freedberg and Gallese 2007; Rizzolatti and Craighero 2005; Rizzolatti and Sinigaglia 2008). Others (e.g. Goldman 2011) accept a hybrid view according to which there are two routes to empathy: a lower-level route based on processes of mirroring and a higher-level route based on pretense or imagination. But all simulation-based approaches stress the role of the psychological and/or neurological similarity between an empathizer’s and her target’s psychological state (or experience) in empathy.

By stressing the *similarities* (at the expense of the dissimilarities) between the experiences of two individuals, one of which empathizes with the other, simulation-based approaches seem to generate puzzles of their own. Of course, two individuals could coincidentally both simultaneously feel pain in their respective left hand if both their left hands were simultaneously injured by two independent causes or by a common cause. But this would not make either experience of pain an empathetic experience. Only if you shared my experience of pain in my left hand as a result of your awareness of my pain could your experience of pain count as empathetic. But if so, then the question arises: how and to what extent could you empathize with, or share, my pain in my left hand unless we shared my left hand? Since we do not share my left hand, it is really puzzling how you could share my pain in my left hand. Towards evading or dissolving some of these puzzles, some philosophers have turned to an alternative view of empathy, which I shall label the *direct-perception* model of empathy.<sup>2</sup>

As the label suggests, the direct-perception model is a *perceptual* model: its core idea is that empathy is a kind of *perceptual experience* of another’s psychological

<sup>1</sup> For the historical background, cf. Stueber (2008).

<sup>2</sup> What I call “the direct-perception model of empathy” combines assumptions, some of which are held by Gallagher (2008a) and others by Zahavi (2007, 2008).

state. The direct-perception model of empathy is an extension of the direct-perception model of what I shall call here *primary intersubjectivity in the broad sense*.<sup>3</sup> Primary intersubjectivity in the broad sense, which has been advertised as the basis of all human interactions and social understanding, is supposed to rely on “certain embodied practices—practices that are emotional, sensory-motor, perceptual and non-conceptual” (Gallagher 2001: 85). From a developmental perspective, primary intersubjectivity in the broad sense encompasses what Trevarthen (1979) distinguished respectively as *primary* and *secondary intersubjectivity* in the narrow sense. As Gallagher (2001, 2004) sees it, primary intersubjectivity in the narrow developmental sense in turn encompasses both what Baron-Cohen (1995) called ID (for *intentionality* detector) and EDD (for *eye-direction* detector). Secondary intersubjectivity in the narrow developmental sense starts with what Baron-Cohen (1995) called SAM (for *shared-attention* mechanism). As Gallagher (2001: 91) puts it, “primary, embodied intersubjectivity is not primary simply in developmental terms. Rather it remains primary across all face-to-face intersubjective experiences, and it subtends the occasional and secondary intersubjective practices of explaining or predicting what other people believe, desire or intend in the practice of their own minds.”

Now, what I call here the direct-perception model of *empathy* derives quite naturally from the direct-perception model of *primary intersubjectivity* in the broad sense, on the further assumption that empathetic understanding just *is* the kind of social understanding delivered by primary intersubjectivity in the broad sense. This very assumption seems part of Zahavi’s (2008: 516) Schelerian view that “empathy has to do with a basic understanding of expressive others” or that “empathy, properly understood, is not a question of feelingly projecting oneself into the other, but rather an ability to experience behavior as expressive of mind, i.e. an ability to access the life of the mind of others in their expressive behavior and meaningful action” (Zahavi 2007: 37). Most of the paper is a critique of the major assumptions involved in the direct-perception model of primary intersubjectivity in the broad sense.<sup>4</sup>

There are two basic critical strands to the direct-perception model of primary intersubjectivity in the broad sense, one of which specifically targets the simulation-based approach to mindreading in general and to empathy in particular, and the other of which is directed to both the simulation-based and the so-called “theory-theory” approaches to mindreading. On the one hand, the direct-perception model is offered as an account of empathy free from the interpersonal similarity condition, which is arguably the major assumption of the simulation-based approach to empathy, and according to which not unless an empathizer’s experience stands in some suitable *similarity* relation to her target’s experience could the former be said to empathize with the latter. As I shall argue momentarily, however, it is not entirely clear that the direct-perception model can meet the challenge of offering an account

<sup>3</sup> Cf. Gallagher (2001, 2004, 2008a, b).

<sup>4</sup> Whereas Zahavi (2001, 2007, 2008) seems to subscribe to the direct-perception model of both primary intersubjectivity in the broad sense and empathy, Gallagher (2001, 2004, 2007, 2008a, b) seems to subscribe only to the direct-perception model of primary intersubjectivity in the broad sense. Both stress the fact that primary intersubjectivity in the broad sense underlies social understanding based on *everyday* encounters (or interactions) with others.

of empathy entirely free from the interpersonal similarity condition. On the other hand, as the label suggests, the direct-perception model is also offered as an account of empathy free from the widely shared joint assumptions that most, if not all, of another's psychological states and experiences are unobservable and that one's representation and understanding of another's psychological states and experiences are based on inferential processes. As I will argue, it is not entirely clear how the direct-perception model can reject both assumptions without embracing behaviorism.

The positive account of empathy offered by advocates of the direct-perception model relies on two major ingredients. The first ingredient is the so-called "interaction" theory of primary intersubjectivity in the broad sense, according to which our basic everyday social understanding of another's experiences is based on *second-person interactions* with, not on third-person observations of, another's actions. This second-person interactionist view is meant to bridge the gap between the direct perception of *inanimate objects* and the direct *social* perception of conspecifics. The second ingredient is the *narrative competency* hypothesis, according to which the social understanding of another's psychological states and experiences that goes beyond the powers of second-person interactions is based on the ability to understand narratives.

The goal of this paper is to assess the direct-perception model of primary intersubjectivity in the broad sense and, by extension, of empathy. Applied to empathy, the model involves the four following assumptions, the first two of which constitute the second-personal interactionist alternative to the standard third-personal observational approach to mindreading:

- (1) An individual's basic goals and emotions are directly expressed in his expressive behavior.
- (2) One's basic understanding of another's goal and/or emotion is not inferential because (i) one directly perceives it in the other's expressive behavior and (ii) it is "pragmatically contextualized comprehension" (Gallagher and Hutto 2008).
- (3) Empathizing with another consists in one's basic non-inferential understanding of another's goal and/or emotion.<sup>5</sup>
- (4) Understanding of another's psychological life that falls beyond second-person interactions is based on narrative competency.

The paper falls into five sections. In the first section, I sketch a five-tiered simulation-based account of *empathy* and also offer some arguments against the identification between empathy and primary intersubjectivity in the broad sense (assumption (3)). In the second section, I examine the critiques of simulation-based approaches to empathy offered by advocates of the direct-perception model. In the third section, I scrutinize the extension of the direct-perception model of inanimate objects to social perception. In the fourth and penultimate section, I turn to the second-personal interaction-based model of the understanding another's basic goals

<sup>5</sup> Acceptance of assumption (3) is characteristic of the direct-perception model of *empathy*. Acceptance of assumption (3) depends in part on the terminological decision to identify empathetic understanding with understanding generated by basic intersubjectivity. As I mentioned above, Zahavi (2007, 2008) explicitly endorses assumption (3). Deprived of assumption (3), the direct-perception model applies to *basic intersubjectivity*.

and emotions. Finally, I turn to the narrative competency account of the understanding of another's psychological states that are beyond the second-person interactionist account.

## 2 A Five-Tiered Model of Empathy<sup>6</sup>

Arguably, any acceptable account of empathetic understanding should be able to discriminate an empathizer's own experience from four related though distinct psychological processes. First of all, it must specify what makes an *empathizer's* psychological experience distinct from her *target's*. Secondly, it must specify what distinguishes an empathetic experience from an experience generated by a process of affective or emotional *contagion*. Thirdly, it must specify what makes an empathetic response to another's experience distinct from a *sympathetic* response. Finally, it must also specify what makes an empathetic experience distinct from a task of standard *mindreading*.

In order to capture the distinction between an empathizer's experience and her target's, we can turn to the notion of a *vicarious* experience. For example, whereas a target's experience of *standard* pain is caused by some bodily injury or other, an empathetic experience of pain is caused by the empathizer's awareness of her target's standard pain. In fact, the notion of a vicarious experience applies both to empathetic responses and to responses generated by a process of affective or emotional contagion. For example, whereas a standard experience of fear or disgust is caused by its intentional object, one can also be caused to experience a vicarious experience of either fear or disgust by a contagious process whereby one catches another's fear or disgust.

I will now state five conditions on empathy, on the basis of which one can discriminate empathetic experiences from contagious responses, sympathetic responses and tasks of standard mindreading:

- (i) *Affectivity condition*: both target and empathizer must experience some affective state or other.
- (ii) *Interpersonal similarity relation condition*: the target's experience  $s$  and the empathizer's experience  $s^*$  must stand in some similarity relation (e.g. both must experience some kind of pain or fear).
- (iii) *Causal path condition*: the empathizer's being in affective state  $s^*$  is caused by the target's being in affective state  $s$ .
- (iv) *Ascription condition*: there could not be empathetic understanding unless the empathizer ascribed the appropriate affective state to the target.
- (v) *Caring condition*: the empathizer must care about the target's affective life.

Satisfaction of the affectivity condition (i) is what distinguishes one's empathetic response to another's experience of e.g. pain from an ascription of pain to another, generated by a standard process of mindreading: in a standard process of mindreading another's pain, one forms the belief that another is in pain. Believing

<sup>6</sup> This section summarizes the view of empathetic experiences, which is presented in detail in de Vignemont and Jacob (submitted), as it applies to empathetic pain.

that another is in pain is different from experiencing empathetic pain. Satisfaction of the interpersonal similarity condition (ii) distinguishes an empathetic from a sympathetic response to another's experience of e.g. pain. One may sympathize with either another's pain or another's jealousy by feeling *sorry* for the other's pain or jealousy. If so, then a sympathetic response satisfies condition (i), but not condition (ii). Condition (iii) is meant to distinguish empathetic experiences from cases in which two unrelated individuals happen to satisfy conditions (i) and (ii) by mere coincidence. Satisfaction of the ascription condition (iv) distinguishes an empathetic response from a vicarious experience generated by a contagious process. For example, one may catch another's fear (as in crowd panic) or another's anxiety without being aware that one's vicarious experience of either fear or anxiety is being caused by another's. Finally, much neuroscientific evidence about empathy for pain shows that like sympathetic responses, but unlike vicarious experiences generated by contagion, empathetic experiences are subject to top down modulation by several contextual factors. For example, in a well-known fMRI experiment, conducted by Singer et al. (2006), participants first played an economic game with confederates who played either fairly or unfairly. Singer et al. (2006) report reduced activity in the affective component of the pain-system in the brains of male participants (not female participants), when they became aware that unfair (not fair) players received painful stimulation.<sup>7</sup> Thus, the caring condition (v) is meant to capture the fact that empathy is not the default response to one's awareness of another's experience of e.g. pain.<sup>8</sup>

Arguably, one distinctive feature of this five-tiered model of empathy is the interpersonal similarity condition (ii), which makes it a simulation-based model of empathy. As I said above, the direct-perception model of empathy is predicated on the rejection of the interpersonal similarity condition (ii). Now consider the following example offered by an advocate of the direct-perception model presumably as a counter-example to the interpersonal similarity condition. Suppose you empathize with my grief or sorrow caused by the loss of a child. Suppose further that my grief or sorrow represents the loss of my child and your empathetic response represents my grief or sorrow. If so, then our respective experiences do not have the same intentional object: whereas the intentional object of my grief or sorrow is the loss of my child, the intentional object of your empathetic experience is my grief or sorrow (cf. Zahavi 2008: 516–517). If they do not have the same intentional objects, then clearly my experience of grief or sorrow and your empathetic experience are distinct types of experience. But showing that they are distinct experiences is compatible with the claim that your empathetic experience satisfies the interpersonal similarity condition (ii). Arguably, on the assumption that the intentional object of your empathetic experience is my experience of grief or sorrow, the content of the latter may be said to be *part* of the content of the former. If so, then this semantic mereological relation might well warrant a similarity relation between the two contents and your empathetic experience would then satisfy condition (ii).

<sup>7</sup> For further evidence about modulation of empathetic responses, cf. Decety et al. 2011; Lamm et al. 2007, 2009.

<sup>8</sup> In order to meet the five conditions above, de Vignemont and Jacob (submitted) endorse an E-imagination based account of empathetic pain, which contrasts with both mirroring-based and hybrid accounts.

### 3 Real and Not-So-Real Troubles with Simulation Theory

I share some of the reservations of advocates of the direct-perception model of empathy towards at least two features of the simulation-based (S-B) approach to *mindreading*.<sup>9</sup> First, consider Goldman's (2011) claims that "the term 'empathize' [is] roughly equivalent to 'simulate' (in an inter-subjective fashion)" and that "empathy is a key to mindreading [...], the most common form of mindreading." On this basis, the S-B approach to mindreading seems to turn empathy into a *default* response to one's awareness of another's experience of e.g. pain. But as I noted above in defense of the caring condition (v) of the five-tiered model of empathy, the empirical evidence shows that empathy is not the default response to one's awareness of another's experience of pain.

Secondly, acceptance of a S-B model of *empathy* does not amount to acceptance of a S-B model of *mindreading*. On the five-tiered model of empathy sketched in the previous section, the interpersonal similarity condition (ii) is a necessary condition on *empathy*. But as conditions (iv) and (v) of the five-tiered model show, it cannot be a sufficient condition. Whether the interpersonal similarity condition is meant to be a *necessary* or merely an *enabling* condition on standard *mindreading*, it cannot be a sufficient condition either for two related reasons. It is one thing to intend to do something (or to experience pain); it is something else to believe that someone else intends to do so and so (or experiences pain). Furthermore, whereas one does not need mastery of the concept of *intention* in order to form an intention, arguably one needs mastery of the concept of *intention* in order to ascribe an intention to another, i.e. to form the belief that another intends to do so and so. Arguably, this is why Goldman's (2006) influential version of the S-B approach to mindreading is really a *simulation-and-projection* approach, whereby the mindreader could not complete her mindreading task unless she projected onto her target the psychological state generated by the simulation routine, in accordance with the interpersonal similarity condition.

The full version of Goldman's (2006) S-B approach to mindreading also involves a further distinction between *low-level* and *high-level* tasks of mindreading achieved by two distinct kinds of processes of mental simulation: whereas so-called processes of "mirroring" (exemplified by the activity of mirror neurons) underlie tasks of low-level mindreading, so-called processes of "Enactment-imagination" (E-imagination for short or pretense) underlie tasks of high-level mindreading.

On behalf of the direct-perception model of primary intersubjectivity in the broad sense, Gallagher (2007) has offered specific arguments against S-B approaches to both high-level and low-level tasks of mindreading, which he calls respectively *explicit* and *implicit* uses of mental simulation. As Gallagher (2007: 355) characterizes it, what makes a use of mental simulation explicit is that it "is a conscious or introspective process in which I imagine myself in the other's situation and use the model (the simulation) that is generated to predict the other's mental states." What makes it implicit is that it is a sub-personal process (*Ibid.*: 356).

<sup>9</sup> Furthermore, on the pretense-theoretic version of the simulation account of the prediction of another's decision, a mindreader uses her own decision-mechanism to run the simulation routine by giving it pretend inputs (i.e. representations of the other's beliefs, desires and intentions). As a result, the mindreader's decision mechanism generates a pretend output (i.e. a pretend decision). But if the problem of the pretend inputs can be solved, then why is the simulation routine needed after all?

Gallagher (2007) offers two complementary arguments against explicit S-B approaches to high-level tasks of mindreading, one of which is what he calls “the Ryle-Scheler argument”, and the other of which he calls “the simple phenomenological argument.” According to the Ryle-Scheler argument, “if I project the result of my own simulation onto the other, I understand only myself in that other’s situation.” As Goldman (2006) emphasizes, on the S-B approach, the mindreader provides her own cognitive mechanism (e.g. decision- mechanism) and feeds it so-called “pretend inputs” (e.g. pretend goals and pretend beliefs). Only if she manages to quarantine her own beliefs and desires from the pretend inputs will she accurately simulate the right output and thereby project onto, and ascribe to, her target the right psychological state (e.g. decision). The Ryle-Scheler argument takes failures of quarantine as evidence against the S-B approach. But as Goldman (2006) suggests, failures of quarantine might be taken as empirical evidence that mindreaders do in fact use a pretense-based simulation routine and fail to bracket their own beliefs and desires in the process.

According to the simple phenomenological argument,

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. There is no experiential evidence that I use such conscious (imaginative, introspective) simulation routines when I interact with or come to understand another person (Gallagher 2007: 356).

I am not overly impressed by the simple phenomenological argument because I think that it betrays a misleading use of the implicit/explicit distinction. Arguably, the implicit/explicit distinction applies to cognitive tasks (of e.g. mindreading), which can be achieved either explicitly (via the use of language) or implicitly (without using language). Now mental simulation is the name of a cognitive heuristic or mechanism, whose steps, operations and computations are meant to be beyond the scope of conscious awareness. The appeal to phenomenological evidence here covertly reflects the acceptance of the first-person authority of introspection on cognitive scientific theorizing.

I now turn to Gallagher’s (2007) argument against implicit S-B approaches to low-level tasks of mindreading, which are supposed to be achieved by mirroring or mirror neuron (MN) activity. I will call it “the control and pretense argument”:

implicit simulation theory interprets neural resonance as simulation [...] The concept of simulation, as defined by simulation theory [...] clearly needs two conditions: it is a process that I can control in an instrumental way [...] and it involves pretense (I put myself in the other’s shoes) [...]. Thus, according to simulation theory, simulation involves the instrumental use of the first-person model to form third-person ‘as-if’ or ‘pretend’ mental states (*Ibid.*: 360).

I will reconstruct the control and pretense (CP) argument by means of the following three premises (CP1-CP3):

- (CP1) The best alleged example of implicit mental simulation is MN activity (or mirroring).
- (CP2) Process P counts as a simulation process only if (i) P is subject to control and (ii) P involves pretense, i.e. the creation of pretend states.

(CP3) Mirroring is (i) neither subject to control (ii) nor does it involve the creation of pretend states.

Conclusion: Therefore, mirroring is not genuine simulation.

As premise (CP2) clearly shows, Gallagher's control and pretense argument puts the pressure on the assumption, widely accepted since Gallese and Goldman's (1998) classical paper, that mirroring is a process of mental simulation, whose output meets the interpersonal similarity condition and which underlies tasks of low-level mindreading. Elsewhere I have expressed my own reservations about the exact contribution of mirroring to tasks of (low-level) mindreading (cf. Jacob 2008a, 2009). However, I am inclined to accept (CP1), the view that mirroring is a process of mental simulation whose output satisfies the interpersonal similarity condition. Furthermore, accepting (CP2) for the sake of argument, it might be possible to resist the conclusion of Gallagher's control and pretense argument by rejecting (CP3) on the following grounds.

First, consider (CP3i), the claim that mirroring is not subject to control. According to Rizzolatti et al.'s (2001) direct-matching model of mirroring, MN activity is supposed to underlie an observer's understanding of the agent's action by mapping (or matching) the agent's act onto the observer's own motor repertoire. On the direct-matching model, MN activity in an observer's brain enables the observer to form the same motor representation of the agent's act as the agent's own representation. By enabling the observer to map the agent's act onto her own motor repertoire, MN activity grounds the observer's understanding of the observed action into her own ability to perform the act. If so, then contrary to (CP3i), mirroring is precisely what anchors an observer's understanding into motor control.

Secondly, consider (CP3ii), the claim that mirroring is not pretense because it does not involve the creation of pretend states. However, MN activity is supposed to enable an observer to form a motor representation of an agent's action similar to the agent's own representation. In other words, MN activity in an observer's brain is supposed to be shared by the observer's and the agent's brain: it can be said to underlie the preparation of an action, whether the action is being performed or not. If so, then contrary to (CP3ii), MN activity would seem to qualify as a paradigm instance of a mechanism that underlies the pretense of action.<sup>10</sup>

#### 4 From Direct Object-Perception to Direct Social Perception

There are two core ideas in the direct-perception model of primary intersubjectivity, one of which is that our everyday basic social understanding of other people is a kind of social perceptual experience, and the other of which is that the social perceptual experience of another's goal and emotion is as direct as our perceptual experience of inanimate objects. As Gallagher (2008a: 537) recognizes, the claim that social perception is direct could not be assessed unless we knew exactly what is

<sup>10</sup> Although I do not agree either with his defense of Goldman's simulation-based approach to mindreading or with each of his criticisms of Gallagher's arguments against simulation-based approaches to mindreading, I have learnt a great deal from Barlassina's (2010) paper presented at Institut Jean Nicod in the Spring of 2010 and subsequent discussions.

meant by ‘direct’. So the question arises: what does it take for the perception of inanimate objects to be *direct*? In response, Gallagher (2008a) offers a clarification of the directness of object-perception in terms of the two complementary concepts of *smart* and *enactive* perception.

Here is how Gallagher (2008a: 536) characterizes the smartness of object-perception:

Let us distinguish between a smart and a not-so-smart perception [...] I open my eyes and I see a certain unrecognized red mass with a specific shape just in front of me. My eyes are working fine, thank you. My visual cortex is processing all the preliminary visual information, and what vision delivers is the meaningless red mass. Let’s call this a not-so-smart perception. In contrast, in the very same situation, when I open my eyes I see my car [...] I see the shape and color [...] as aspects of something that is amazingly recognizable as my car [...] I simply see my car [...] Let’s call this a perception with some degree of smarts.

On this account, object-perception qualifies as smart only if it enables Shaun to recognize his car or to see his car *as* his car, upon opening his eyes in its presence. But clearly eye opening could not confer to perception the required smartness unless further conditions were satisfied. First of all, it is highly unlikely that Shaun would see the shape and color of his car as “aspects of something that is amazingly recognizable as [his] car” if he faced it either at night or at a distance of 300 m or both. Nor would opening his eyes enable Shaun to see his car as his car, at the right distance and in good day light, if, unbeknownst to him, his car had been replaced overnight by a perceptually indistinguishable car parked at the same place. Finally, suppose that I know neither Shaun’s car nor that he owns a car. If so, then opening my eyes alone would certainly not enable me to see Shaun’s car as his car when I am facing it at the right distance and in good day light. Thus, what makes object-perception smart lies in one’s knowledge of the right contextual cues, including previous knowledge of perceived objects and knowledge of the background conditions that are best for perceptual object-identification. But if so, then it is far from clear that smartness can support the direct character of object-perception.

Nor is it clear that appeal to the enactivist account of the content of object-perception can really help to clarify the directness of object-perception. According to the full enactivist account (of e.g. O’Reagan and Noë 2001 and Noë 2004), the content of one’s perceptual experience of an object is constituted by one’s tacit knowledge of the sensory consequences of one’s actions onto the object. As a move in the direction of the enactive account of object-perception, Gallagher (2008a: 257) says that he does not see his car just as some object among others, but instead as an object that he can *use* (climb into and drive). In short, he sees it as *drivable*. Now, the appeal to the enactivist account of object-perception here raises at least two related problems.

First of all, there is a gap between the mere Gibsonian point that one can see one’s car as *drivable* and a genuine enactivist account of the content of one’s perceptual experience of an object (e.g. one’s car). No doubt, one can see one’s car as drivable, but one can also see it as *non-drivable* (if e.g. the engine has been disassembled or after a bad accident). One can also see it as on sale, as a member of one’s collection of cars or as an instance of the brand of Mustang cars and so on. Furthermore, one

could only see one's car as drivable if one had already identified one's car, i.e. sorted it out from the background and from distracting cars in a visual scene. So the question arises whether, and if so, how seeing one's car as drivable could enable one to see one's car as e.g. non-drivable or on sale. More generally, any enactivist account faces the question: how could tacit knowledge of the sensory consequences of one's own actions onto an object enable one to bind the color, shape, texture of the body of my unmoving car together with its spatial position relative to me when seen from the outside of the car? Unless one were antecedently able to sort it out from neighboring distractors and from the background, how could one plan to act on some target? For someone to know (implicitly or explicitly) how to act on an object—for anyone to know the sensory consequences of one's own actions upon some object—one must be able to independently track the object or distinguish the putative target of one's action from other surrounding objects and from the background. On the full enactive account, one's visual experience of an object is constituted by one's knowledge of which sensory consequences one's actions would have upon the perceived object. But now if one's knowledge of the sensory consequences of one's actions upon the object requires some independent and prior tracking or sensory detection of the object, then the enactive account seems plainly threatened by circularity.<sup>11</sup> If so, then it is unlikely that appeal to the enactive account could help clarify the direct theory of object-perception.

I now turn to the extension of the direct-perception model to *social* perception. Advocates of the direct-perception model reject the interpretation of mirroring according to which MN activity in an observer's brain is taken to be a process of mental simulation underlying tasks of low-level mindreading. Instead, they interpret MN activity as an enactive process extended from object-perception to social perception:

If we think of perception as an enactive process, as involving sensory-motor skills, rather than just sensory input/processing—as an active, skillful, embodied engagement with the world, rather than 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. MN activation is not the initiation of simulation; it subtends a direct intersubjective perception of what the other is doing. (Gallagher 2007: 541–542)

As I see it, the problem is that the direct-perception model of social perception faces a dilemma: if what underlies social perception is MN activity, then social perception cannot be both resonant and enactive. As Gallagher characterizes it, social perception would be enactive if it were “an active, skillful, embodied engagement with the world, rather than the passive reception of information from the environment.” If social perception is enactive in this sense, then it is not a resonance process, but either a cooperative or a competitive *contingent* response: seeing another stumble might, depending on the context, cause one either to try and catch him or to make sure that he falls down. Conversely, if social perception is a process of resonance, then arguably it is the passive mental duplication (or rehearsal) of

<sup>11</sup> For further discussion, see Jacob (2008b).

another's perceived action (not followed by execution). But if so, then it cannot be an enactive process in the relevant sense.

## 5 Can One Directly Perceive Another's Goal or Emotion?

Advocates of the direct-perception model of empathy reject the standard neo-Cartesian asymmetry between the presupposed direct non-inferential first-personal access to the contents of one's own mind and the non-direct inferential third-personal access to the minds of others based on the observation of others' overt behavior. As Zahavi (2008: 518) insightfully puts it, the standard neo-Cartesian asymmetry jointly “underestimates the difficulties involved in self-experience and overestimates the difficulties involved in the experience of others.” The second-personal interactive model is meant as an alternative to the third-personal inferential and observational model of mindreading or, as Hutto (2004) puts it, it is meant as a way to overcome “the limits of spectatorial folk psychology.”<sup>12</sup> In accordance with assumptions (1) and (2) of the direct-perception model of primary intersubjectivity in the broad sense, the second-personal interactive model in turn jointly relies on the expressivist conception of behavior and the application of the direct model of social perception to others' expressive behavior.

On the direct-perception model, “empathy is a basic, irreducible, form of intentionality [...] directed towards the experience of others” (Zahavi 2008: 517). One's everyday empathetic understanding of another is the experience of another's mind. But experiencing another's mind is not going through the same experience as the other. For example, in empathizing with another's sorrow caused by the other's loss of a child, one can directly experience the other's sorrow. But the empathetic experience of the other's sorrow is not itself an experience of sorrow—the two experiences have distinct intentional objects. Now what is supposed to make one's empathetic understanding of another's affective state or emotion a perceptual experience of the other's mind is the *expressivist* conception of behavior. As Zahavi (2008: 518) puts it,

affective and emotional states are not simply qualities of subjective experience, rather they are given *in* expressive phenomena, i.e. they are expressed in bodily gestures and actions, and they thereby become visible to others [...]. If the realm of expressive phenomena is accepted as the primary datum or primitive stratum of perception, the access to the mind of others will no longer present the same kind of problem. Indeed [...] expressive phenomena—in particular facial expressions and gestures, but also verbal expressions—can present us with a direct and non-inferential access to the experiential life of others.

It is uncontroversial that an individual's goal-directed or intentional behavior betrays the goal or intention that caused the agent's executed movements. An individual's expressive behavior can also be said to betray the emotion or affective state that caused the agent's overt movements. In both cases, it is uncontroversial

<sup>12</sup> Bogdan (1997: 104), whom Hutto credits for coining the expression, defines the “spectatorial” stance as portraying “the subject as a remote object of observation and prediction.”

that one's understanding of another's goal or emotion should be "informed and influenced" by the other's goal-directed or expressive bodily behavior. This much will be accepted by advocates of the standard approach, according to which mindreading another's goal or emotion results from an inferential process based on the observation of the other's behavior. But of course on the standard approach, another's expressive behavior should be sharply distinguished from her psychological states, including her emotional and affective experiences: the former cannot constitute the latter. Nor could one observe or perceive another's emotional or affective experiences: one can only perceive (or observe) another's expressive behavior. From the standpoint of the standard approach, it seems as if the second-personal interactive model of empathy faces the following two dilemmas, the first of which shows that this model is surprisingly vulnerable to the charge of behaviorism.<sup>13</sup>

The first dilemma rests on whether or not another's expressive behavior could constitute her affective experience: either another's overt bodily expressions *do* constitute her emotional or affective experiences or they do not. If they do not, then clearly by perceiving another's expressive behavior, one does not thereby perceive the other's emotional or affective experience. But if they do, then it is quite unclear how the second-personal interactive model can distance itself from behaviorism.

The second dilemma turns on the scope and limits of the expressivist conception of behavior. Arguably, the expressivist conception of behavior is meant to apply to some, not all, of an individual's psychological states, e.g. her emotions and goals, not her beliefs. But consider an agent's goal-directed behavior consisting of her grasping a glass and drinking a sip from it.<sup>14</sup> An advocate of the expressive conception of behavior would be likely to argue that one could directly perceive the agent's goal in her goal-directed behavior. However, the agent would probably not execute the same goal-directed behavior if she *believed* that what is in the glass is poisonous. So the agent's goal-directed behavior is the product of the agent's goal *and* belief. If so, then the agent's goal-directed behavior does not merely express her goal, it also expresses her belief. So the direct-perception theorist now faces a dilemma: if the reason why an agent's goal can be directly perceived in her goal-directed behavior is that an agent's goal-directed behavior itself is "soaked with the meaning of the mind", as Zahavi (2008: 520) puts it, then what is the further principle that enables the phenomenologist to assert that only an agent's goal, not her relevant belief, can be directly perceived?

In Section 4, I argued that what makes object-perception smart is the perceiver's background knowledge of contextual cues. Similarly, I am now going to argue that one's understanding of another's goal or emotion depends on one's knowledge of the relevant contextual cues. I start with understanding another's emotion. Wicker et al. (2003) have reported fMRI evidence showing that the brain areas (in particular the anterior insula) that are active when one experiences disgust as a result of inhaling a disgusting odorant are also active when one recognizes another's experience of disgust from observing the other's facial expression of disgust. Both Wicker et al.

<sup>13</sup> It may look awkward that a phenomenological approach to understanding others is open to the charge of behaviorism. But it is.

<sup>14</sup> Arguably, there are some differences between an agent's goal-directed and an agent's expressive behavior. But the argument could easily be run in terms of an agent's expressive behavior.

(2003) and Gallese et al. (2004) have offered a simulation-based interpretation of this finding as showing the existence of “a common mechanism for understanding the emotion in others and feeling the same emotions in ourselves.” Now, advocates of the direct-perception model of empathy might argue that this finding is evidence in favor of the second-personal interactionist view of social perception whereby one directly perceives another’s disgust *in* his facial expression, in accordance with the Scheler (1954) dictum that we are “directly acquainted” with another’s emotion or affective state on the grounds that one can perceive it *in* the other’s expressive behavior.

However, as shown by a quick inspection of Wicker et al.’s (2003) video-clips displaying the facial expression of an individual who experiences disgust as a result of inhaling an odorant emanating from a glass, the visual perception of another’s facial expression of disgust might not be a sufficient basis for ascribing disgust to another. The facial expression of another’s disgust might be difficult to discriminate from the facial expression of other affective states. For example, there is evidence reported by Dailey et al. (2002) that participants find it very hard to discriminate between the facial expressions of disgust and anger. Furthermore, Meeren et al. (2005) and Aviezer et al. (2008) have provided evidence that in case of incongruent bodily and facial expressions (e.g. a facial expression of disgust and a bodily expression of anger), the former may win over the latter (e.g. participants judge that the face expresses anger, not disgust). Arguably, unlike other affective experiences, an experience of disgust has an environmental intentional object (i.e. the substance contained in the glass) that is also the cause of the experience. Clearly, the glass present in Wicker et al.’s (2003) video-clips is not *in* the individual’s facial expression; instead, it is part of the *context* of the individual’s facial expression. Thus, only by representing relevant contextual cues might one be able to disambiguate between the facial expression of another’s experience of respectively pain and disgust. If so, then one cannot perceive the other’s disgust *in* the other’s facial expression alone, independently of the perception of contextual cues which are not part of the other’s expression.

I now turn to understanding another’s goal (or intention). In a series of elegant experiments, Fogassi et al. (2005), Iacoboni et al. (2005) and Cattaneo et al. (2007) have reported evidence showing that MN activity recorded during either the execution or the perception of an act of grasping a target can be modulated by the most likely act that the agent *will* execute, e.g. of either bringing the target to the mouth (and eating it) or placing it into a container.<sup>15</sup> On the *mirroring-first* view of understanding another’s intention offered by Iacoboni et al. (2005: 539), these findings are taken to show that “premotor mirror neuron areas—areas active during the execution and the observation of an action—previously thought to be involved only in action recognition are actually also involved in understanding the intentions of others.” But of course, there is a relevant asymmetry between MN activity in the brain of respectively an *agent* of an act of grasping and an *observer* of an act of grasping performed by another. In normal cases, the agent of an act of grasping

<sup>15</sup> Fogassi et al. (2005) recorded the activity of single cells in area IPL of macaque monkeys in both executive and observational tasks of grasping a target. Iacoboni et al. (2005) showed human participants video-clips of actions of grasping a cup either within or outside a context. Cattaneo et al. (2007) recorded and compared the electromyographic (EMG) activity respectively in typically developing children and in children with autism, in both executive and observational tasks of grasping a target.

knows whether the next most likely act she will perform is bringing the target to her mouth or placing it into a container because she knows her prior intention. But an observer of an act of grasping performed by another can only guess from contextual cues and/or psychological background assumptions about the agent whether the agent will next bring the target to her mouth or place it into a container.

Now advocates of the direct-perception model of social perception reject the simulation-based interpretation of mirroring. However, both the mirroring-first view and the direct-perception model of social perception seem to accept the assumption that MN activity in an observer's brain alone can take as input the perception of an agent's act of grasping a target and generate as output a representation of the agent's goal, where the latter can be thought of as the agent's prior intention to either eat the target or place it into a container (in Fogassi et al. 2005 and Cattaneo et al. 2007) or alternatively as the agent's intention to either drink a cup of tea or clean the cup (in Iacoboni et al. 2005). There is, however, a rival interpretation of these findings according to which MN activity in the observer's brain does not generate a representation of the agent's goal or prior intention from the perception of the agent's act. Instead, MN activity *results from* an independent representation of the agent's goal or prior intention, which is based on the perceptual processing of contextual cues (cf. Csibra 2007 and Jacob 2008a). Unlike the view commonly accepted by the mirroring-first view and the direct-perception model of social perception, this alternative interpretation of MN activity stresses the contribution of contextual cues to understanding another's goal or intention.

Now according to the second-personal interactive model of direct social perception, "we do not stand to the side as third-person observers" (cf. Hutto 2004). Instead advocates of this model stress the role of "pragmatically contextualized comprehension" (Gallagher and Hutto 2008) in our everyday basic understanding of others' goals and emotions:

the pragmatic situations, social practices, and socially defined roles through which our interactions with others happen turn out to do much of the work of shaping our understanding of others. We understand what they are doing and what their intentions are because their actions are set within environments that help to define meaningful actions. Specific contexts make sense out of their actions, gestures, and expressions (Gallagher 2007: 354).

By 'pragmatic situation' one may mean the context of another's action in general or more specifically the context of another's verbal and/or non-verbal communicative behavior. In either case, what enables an observer to understand another's action are not merely *situations* and *contexts*, but rather our capacities to *cognitively represent* situations and contexts, and to attribute an intention to the agent. Furthermore, if the agent is a speaker, then not unless the addressee was able to represent the speaker's communicative intention could she understand the agent's action.

## 6 The Narrative Competency Hypothesis and the Concept of *Belief*

Advocates of the direct-perception model of primary intersubjectivity in the broad sense (Gallagher 2007; Gallagher and Hutto 2008; Hutto 2008a, b) are willing to

concede that the second-personal interactive account of the experiential access to another's mind is limited to our everyday understanding of another's basic goals and emotions. Arguably, an agent's expressive or intentional behavior does not betray her (true or false) beliefs the way it betrays her basic goals and emotions. As a way of extending the kind of social understanding generated by primary intersubjectivity in the broad sense, advocates of the direct-perception model stress the role of children's *narrative competency*. As Hutto (2008a, b) in particular has argued, children's ability to frame others' actions in terms of a narrative structure helps them make sense of an agent's *reasons* for action, where an agent's reasons include her propositional attitudes.

Arguably, there is a weaker and a stronger version of the narrative competency hypothesis. On the weaker version, children's ability to frame and understand people's actions in terms of a narrative is supposed to strengthen or enhance their capacity to interpret human actions in terms of an agent's reasons for actions. The weaker version would seem to be quite compatible with either a simulation-based or a theory-theory based approach to understanding other minds beyond the limits of the second-personal interactive approach. But on the stronger version, the appeal to the narrative competency hypothesis is supposed to be a genuine alternative to both the simulation-based and theory-theory approaches to mindreading. It is quite clear that advocates of the direct-perception model of primary intersubjectivity in the broad sense endorse the stronger of the two versions. For example, on behalf of the narrative practice hypothesis, Hutto (2008a: 53) claims that "children normally achieve [folk psychological] understanding by engaging in story-telling practices with the support of others" and Gallagher (2007: 354) hypothesizes that "this kind of narrative practice is not based on ToM [theory of mind] or folk psychology as understood by ST or TT, but may in fact be the basis for our use of folk psychological concepts (like beliefs, desires, reasons) in those rare cases when we encounter puzzling behavior and attempt to explain it." It is worth mentioning that only on the strong, not on the weak version, is the narrative competency hypothesis really inconsistent with either the simulation-based or the theory-theory approach to mindreading.<sup>16</sup>

So two related questions naturally arise for the narrative competency hypothesis: what is distinctive of the cognitive ability to frame and understand a narrative and how could it be the basis of one's understanding other minds beyond the limits of the second-personal interactive approach? From an epistemic point of view, a historical narrative (a story) is a linguistic *description* of a temporally or chronologically

<sup>16</sup> For a good instance of the oscillation between the weaker and the stronger version of the narrative competency hypothesis, cf. Hutto (2008b: 178), who argues that it is incompatible with either simulation-based or theory-theory approaches to mindreading and nonetheless recognizes that "narratives do crucially important but nonetheless limited work. They are not responsible for introducing an understanding of mental concepts, such as desire and belief for the first time, rather, being complex linguistic representations of particular events, they put on show how these attitudes can integrate with one another (and also how they fit with other mental states and stand with respect to other contextual factors) [...] kids already have a practical grasp on what it is to have a desire or belief before learning how to integrate their discrete understanding of these concepts in making sense of actions in terms of reasons [...] Through shared encounters with [FP] narratives children become familiar with the forms and norms of giving and asking for reasons, knowing how and when these apply. In this way they develop the capacity to give and receive reason explanations, as and when required."

ordered sequence of (causally or not causally) related *events*. Arguably, by describing event  $E_n$  as the last of an ordered sequence of events from  $E_1$  to  $E_{n-1}$ , a narrative provides a historical explanation and understanding of the occurrence of event  $E_n$ . The explanation and understanding of the occurrence of an event based on a historical narrative are different from the kind of explanation and understanding based on law-like or nomological generalizations provided by science. If the event to be explained is an individual's behavior or a change in some of her psychological states, then, as advocates of the direct-perception model emphasize, a narrative explanation is likely to differ from the kind of explanation based on tacit knowledge of a psychological theory comprised of law-like psychological generalizations in the theory-theory conception of mindreading. Furthermore, unlike scientific explanations based on law-like generalizations, and like the ability to mindread, narrative explanations seem distributed in all human cultures. From a psychological point of view, the ability to generate, understand and transmit stories and narratives must depend on language-understanding and working memory.

The basic challenge for the strong version of the narrative competency hypothesis is: why and how should narrative competency be, in Gallagher's (2007: 354) terms, "the basis for our use of folk psychological concepts like beliefs, desires, reasons"? In particular, how could it be the basis of human children's ability to represent others' *false beliefs*? What could it be about the memorable chronological sequence of events (encoded in a narrative) that could anchor the basic (and arguably unique)<sup>17</sup> human ability to represent another's false belief? Given what it takes to understand and frame a narrative, the idea of grounding a human child's ability to understand and represent another's false belief into her narrative competency would only make sense for children who have been shown to successfully pass the standard false belief task at around age 4. As Gallagher and Hutto (2008) have stressed, "it is notable that many false-belief tests are presented in the form of a narrative and could be interpreted as tests for a certain level of narrative competency."<sup>18</sup>

Surely, this is true of *standard* false belief tasks of the Ann-Sally kind, in which the child knows the true location of an object and is explicitly asked where a character with a false belief will look for an object. But as many psychologists (e.g. Bloom and German 2000; Leslie 2000, 2005; Baillargeon et al. 2010) have emphasized, there is more to successfully passing a standard false belief *task* than the ability to represent another's false belief. For example, in addition to the ability to understand the natural language in which the question is asked and access the representation of the agent's false belief, the child must also have the ability to *inhibit* any prepotent tendency to answer the test question based on her own true belief or knowledge about the location of the object. Clearly, standard false belief tasks are, in Baillargeon et al.'s (2010: 110) terminology, *elicited-response* tasks "in which children are asked a direct question about an agent's false belief."

<sup>17</sup> Cf. Call and Tomasello (2008).

<sup>18</sup> Sometimes, as noted by Herschbach (2008), advocates of the narrative competency hypothesis seem to take the view that the assumption that an agent's psychological state is "hidden" (an assumption jointly accepted by simulation-based and theory-theory based approaches to mindreading) is appropriate "in those rare cases when we encounter puzzling behavior" such as an agent's action based on a false belief (cf. Gallagher 2007: 354 and Zahavi 2008: 515). But human actions based on false beliefs are far from unusual.

Furthermore, a number of studies conducted in the past 5 years or so and based on *spontaneous response* tasks suggest that the ability to represent another's false belief is present in preverbal human infants. Some of these spontaneous response tasks are based on the violation-of-expectation (VOE) paradigm; others on the anticipatory-looking (AL) paradigm.<sup>19</sup> For example, in the first familiarization trial of Onishi and Baillargeon's (2005) VOE study, 15-month-olds saw a human female agent grasp and briefly play with a toy located in between a yellow and a green box before hiding it inside the green box and pause with her hand inside the green box. In the second and third familiarization trials, the infants saw the agent reach inside the green box (as if she was retrieving the toy) and then pause. Next, infants were exposed to four different belief-induction trials, in which infants saw the toy move from one box to the other in either the presence or the absence of the agent. Finally, during four test trials, infants saw the agent retrieve the toy from either the yellow or the green box on the basis of either a true or a false belief. Onishi and Baillargeon (2005) report that infants looked reliably longer either when they saw the agent retrieve the toy in the wrong location while she had a true belief about the toy's location or when they saw the agent retrieve the toy in the right location while she had a false belief.<sup>20</sup>

In the familiarization trials of Southgate et al.'s (2007) AL study, 25-month-olds saw a bear puppet hide a ball in one of two boxes while a female human agent was looking with her head above a panel with two small doors, one above each box. After the bear hid the toy in one of the boxes, the two doors lit up, a chime sounded simultaneously, and after a short delay, the agent retrieved the toy through one of the doors. In the test trials, the agent saw the bear hide the toy in one of the two boxes until a phone rang and the agent turned her head away until the phone stopped ringing. While the agent was not looking, the bear retrieved the toy and left the scene with it. When the phone stopped ringing, the agent turned her head towards the boxes, the doors lit up and the chime sounded. Using an eye-tracker to measure infants' looking behavior, Southgate et al. (2007) coded the location of the first saccade following the illumination of the windows. They report that infants reliably looked at the door above the box where the agent falsely believed the toy had been hidden by the bear puppet, revealing thereby their ability to anticipate an agent's action based on her false belief.

As I mentioned above, advocates of the narrative competency hypothesis are pulled between a stronger and a weaker version of their view. If they accept the challenge raised by children's ability to represent and reason about others' false beliefs, then, as Gallagher (*in press*) recognizes, they must restrict the applicability of their hypothesis to elicited-response tasks in which children are asked a direct question about an agent's false belief. But if so, then they apparently leave unexplained the experimental results showing the ability of preverbal human

<sup>19</sup> For a recent extensive review, cf. Caron (2009).

<sup>20</sup> In another VOE study, Surian et al. (2007) reported that 13-month-olds looked longer when a caterpillar approached the true location of its preferred food, when the food has been hidden there in the absence of the caterpillar. In still another VOE study, He and Baillargeon (2010) report that 11-month-olds looked reliably longer when an agent, who had *not* seen the shortening of the length of a toy, retrieved the toy from the shorter of a pair of boxes when only the other box would have been large enough to contain the toy before it was shortened in the agent's absence.

infants to represent and reason about others' false beliefs in either VOE or AL spontaneous response tasks. Furthermore, endorsing this move, as Gallagher (*in press*) seems tempted to, would create a complete gap between the ability to represent another's false belief according to whether it is tested by a spontaneous or an elicited-response task.

On the face of it, a child's ability to understand and enjoy the story of *Little Red Riding Hood* must involve the ability to ascribe to the main character in the story the false belief that the wolf is her grand-mother. But to claim that the child's ability to represent another's false belief derives from her ability to enjoy the story seems like putting the cart before the horse. Arguably, so-called "narrative competency" could be one of the necessary conditions for successfully passing an *elicited-response* false belief task without being a necessary (let alone a sufficient) condition for being able to represent and reason about another's false belief. As Gallagher (*in press*) argues, "narrative competency does seem to be involved in many of the false belief tests given to 3–5-year-olds, and [...] narrative competency may be required to understand the question." But only if one *denied* that there is *more* to successfully passing an elicited-response false belief task than the ability to represent and reason about another's false belief could one conclude that narrative competency is either a necessary or a sufficient condition for the ability to represent and reason about another's false belief. On my view, the fact that preverbal human infants do seem to represent another's false belief in spontaneous false belief tasks is strong evidence that there is more to passing an elicited-response false belief task than representing another's false belief. Furthermore, acceptance of the distinction between the ability to represent another's false belief and the various further abilities required to successfully pass an elicited-response false belief task is consistent with the assumption that the former is already in place in preverbal human infants.

## 7 Conclusions

No doubt, there is a certain amount of stipulation in the way one decides to use the word 'empathy.' One of the explicit purposes of the direct-perception model is to by-pass the inter-personal similarity condition on empathy, which is at the core of the simulation-based approach. But as I have pointed out in Section 2, it is really unclear that the direct-perception model can really free itself from the inter-personal similarity condition. The direct-perception model of empathy is an extension of the direct-perception model of primary intersubjectivity in the broad sense. The latter is meant as a rejection of the assumptions shared by both simulation-based and theory-theory-based approaches to mindreading that others' psychological states are "hidden away and are therefore not accessible to perception" (Gallagher 2008a, b: 536), that they "remain concealed and hidden" and that "all that I have experiential access to is their behavior" (Zahavi 2008: 519).

The direct-perception model involves two basic ingredients: a second-personal interactionist account of one's everyday understanding of another's basic goals and emotions and a narrative competency account of another's psychological states that are beyond the second-personal interactionist account. On the second-personal interactionist account, one can directly perceive another's basic goals and emotions

in his or her expressive behavior. But as I have argued in Sections 4 and 5, contrary to the direct-perception model, one's ability to ascribe a basic goal or emotion to another on the basis of the perception of her intentional or expressive behavior crucially relies on background assumptions and the representation of relevant contextual cues.

The narrative competency hypothesis is supposed to support the direct-perception model by extending one's understanding of another's psychological life beyond the limits of the second-personal interactionist account. But as I have argued in Section 6, neither the strong nor the weak version of the narrative competency hypothesis seem to have the adequate resources to account for the ability to represent and reason about another's false beliefs. If I am right, then neither of the two basic components of the direct-perception model of primary intersubjectivity in the broad sense is a viable option. If so, then the direct-perception model of primary intersubjectivity in the broad sense can be a rival to neither the simulation-based or the theory-theory approaches to either mindreading in general or empathy in particular.

## References

- Aviezer, H., R.R. Hassin, J. Ryan, C. Grady, J. Susskind, A. Anderson, M. Moscovitch, and S. Bentin. 2008. Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science* 19(7): 724–32.
- Baillargeon, R., R.M. Scott, and Z. He. 2010. False-belief understanding in infants. *Trends in Cognitive Sciences* 14(3): 110–118.
- Baron-Cohen, S. 1995. *Mindblindness, an essay on autism and theory of mind*. Cambridge: MIT.
- Barlassina, L. 2010. Sticking to simulation or why Gallagher's arguments don't refute Goldman's theory of mindreading. Paper presented at Institut Jean Nicod, Spring 2010.
- Bloom, P., and T. German. 2000. Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* 77: B25–B31.
- Bogdan, R. 1997. *Interpreting minds: The evolution of a practice*. Cambridge: MIT.
- Call, J., and M. Tomasello. 2008. Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences* 12(5): 187–192.
- Caron, A.J. 2009. Comprehension of the representational mind in infancy. *Developmental Review* 29: 69–95.
- Cattaneo, L., M. Fabbri-Destro, S. Boria, C. Pieraccini, A. Monti, G. Cossu, and G. Rizzolatti. 2007. Impairment of actions chains in autism and its possible role in intention understanding. *Proceedings of the National Academy of Science* 104(45): 17825–17830.
- Csibra, G. 2007. Action mirroring and action understanding. In *Sensorimotor foundations of higher cognition. Attention and performance XXII*, ed. P. Haggard, Y. Rossetti, and M. Kawato. Oxford: Oxford University Press.
- Dailey, M.N., G.W. Cotrell, C. Padgett, and R. Adolphs. 2002. EMPATH: a neural network that categorizes facial expressions. *Journal of Cognitive Neuroscience* 14(8): 1158–1173.
- Decety, J., Echols, S. and Correll, J. 2011. The Blame Game: The Effect of Responsibility and Social Stigma on Empathy for Pain. *Journal of Cognitive Neuroscience*.
- de Vignemont, F. and Jacob, P. submitted. What is it like to feel another's pain?
- Fogassi, L., P.F. Ferrari, B. Gesierich, S. Rozzi, F. Chersi, and G. Rizzolatti. 2005. Parietal lobe: from action organization to intention understanding. *Science* 308: 662–667.
- Freedberg, D., and V. Gallese. 2007. Motion, emotion and empathy in esthetic experience. *Trends in Cognitive Sciences* 11: 197–203.
- Gallagher, S. 2001. The practice of mind, theory, simulation or primary interaction? *Journal of Consciousness Studies* 8(5–7): 83–108.
- Gallagher, S. 2004. Hermeneutics and the cognitive sciences. *Journal of Consciousness Studies* 11: 10–11.
- Gallagher, S. 2007. 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. Understanding others: embodied social cognition. In *Handbook of cognitive science: An embodied approach*, eds. P. Calvo, and A. Gomila. Academic Press.
- Gallagher, S. in press. Empathy, simulation, and narrative.
- Gallagher, S., and D. Hutto. 2008. Understanding others through primary interaction and narrative practice. In *The shared mind: Perspectives on intersubjectivity*, ed. J. Zlatev, T.P. Racine, C. Sinha, and E. Itkonen, 17–38. Amsterdam: John Benjamins.
- Gallese, V. 2001. The ‘shared manifold’ hypothesis, from mirror neurons to empathy. *Journal of Consciousness Studies* 8(5–7): 33–50.
- Gallese, V., and A.I. Goldman. 1998. Mirror neurons and the simulation theory of mindreading. *Trends in Cognitive Sciences* 12: 493–501.
- Gallese, V., C. Keysers, and G. Rizzolatti. 2004. A unifying view of the basis of social cognition. *Trends in Cognitive Sciences* 8(9): 396–403.
- Goldman, A.I. 2006. *Simulating minds, the philosophy, psychology and neuroscience of mindreading*. New York: Oxford University Press.
- Goldman, A.I. 2011. Two routes to empathy, insights from cognitive neuroscience. In *Empathy: Philosophical and psychological perspectives*, ed. A. Coplan and P. Goldie. Oxford: Oxford University Press.
- He, Z., and R. Baillargeon. 2010. False-belief understanding in 11-month-old infants. Manuscript in preparation.
- Herschbach, M. 2008. False-belief understanding and the Phenomenological critics of folk psychology. *Journal of Consciousness Studies* 15(12): 33–56.
- Hutto, D.D. 2004. The limits of spectatorial folk psychology. *Mind and Language* 19(5): 548–573.
- Hutto, D.D. 2008a. *Folk psychological narratives: The sociocultural basis of understanding reasons*. Cambridge: MIT.
- Hutto, D.D. 2008b. The narrative practice hypothesis: clarifications and implications. *Philosophical Explorations* 11(3): 175–192.
- Iacoboni, M., I. Molnar-Szakacs, V. Gallese, G. Buccino, J.C. Mazziotta, and G. Rizzolatti. 2005. Grasping the intentions of others with one’s own mirror neuron system. *PLoS Biology* 3: 529–535.
- Jacob, P. 2008a. What do mirror neurons contribute to human social cognition? *Mind and Language* 23(2): 190–223.
- Jacob, P. 2008b. The scope and limits of enactive approaches to visual experience. *Perception* 37(3): 446–461.
- Jacob, P. 2009. The tuning-fork model of social cognition: a critique. *Consciousness and Cognition* 18(1): 229–243.
- Lamm, C., H.C. Nusbaum, A.N. Meltzoff, and J. Decety. 2007. What are you feeling? Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *PLoS One*. 212: e1292.
- Lamm, C., A.N. Meltzoff, and J. Decety. 2009. How do we empathize with someone who is not like us? A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience* 22(2): 362–376.
- Leslie, A. 2000. “Theory of Mind” as a mechanism of selective attention. In *The new cognitive neurosciences*, ed. M.S. Gazzaniga. Cambridge: MIT.
- Leslie, A. 2005. Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences* 9(10): 459–462.
- Meeren, H.K., C.C. Heijnsbergen, and B. de Gelder. 2005. Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences, USA* 102: 16518–16523.
- Noë, A. 2004. *Action in perception*. Cambridge: MIT.
- Onishi, K.H., and R. Baillargeon. 2005. Do 15-month-old infants understand false beliefs? *Science* 308 (5719): 255–258.
- O’Reagan, K., and A. Noë. 2001. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 24: 939–1031.
- Rizzolatti, G., and L. Craighero. 2005. Mirror neuron: a neurological approach to empathy. In *Neurobiology of human value*, ed. J.-P. Changeux et al. Berlin: Springer-Verlag.
- Rizzolatti, G., and C. Sinigaglia. 2008. *Mirrors in the brain: How our minds share actions and emotions*. Oxford: Oxford University Press.
- Rizzolatti, G., L. Fogassi, and V. Gallese. 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Review Neuroscience* 2: 661–670.
- Scheler, M. 1954. *The nature of sympathy*. Trans. P. Heath. London: Routledge and Kegan Paul.
- Singer, T., B. Seymour, J. O’Doherty, K.E. Stephan, R.J. Dolan, and C.D. Frith. 2006. Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439: 466–469.

- Southgate, V., A. Senju, and G. Csibra. 2007. Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science* 18(7): 587–592.
- Stueber, K. 2008. *Empathy*. *Stanford Encyclopaedia of Philosophy* <http://plato.stanford.edu/entries/empathy>.
- Surian, L., S. Caldi, and D. Sperber. 2007. Attribution of beliefs by 13-month-old infants. *Psychological Science* 18: 580–586.
- Trevarthen, C.B. 1979. Communication and cooperation in early infancy: a description of primary intersubjectivity. In *Before speech*, ed. M. Bullowa, 321–347. Cambridge: Cambridge University Press.
- Wicker, B., C. Keysers, J. Plailly, J.-P. Royet, V. Gallese, and G. Rizzolatti. 2003. Both of us disgusted in *my insula*: the common neural basis of seeing and feeling disgust. *Neuron* 40: 655–664.
- Zahavi, D. 2001. Beyond empathy, phenomenological approaches to intersubjectivity. *Journal of Consciousness Studies* 8(5–7): 151–67.
- Zahavi, D. 2007. Expression and empathy. In *Folk psychology re-assessed*, ed. D.D. Hutto and M. Ratcliffe, 25–40. Dordrecht: Springer.
- Zahavi, D. 2008. Simulation, projection and empathy. *Consciousness and Cognition* 17: 514–522.