ARTICLE IN PRESS

Consciousness and Cognition xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

How from action-mirroring to intention-ascription? ☆

Pierre Jacob *

Institut Jean Nicod, UMR 8129, ENS/EHESS/CNRS, Ecole Normale Supérieure, Pavillon Jardin, 29, rue d'Ulm, 75005 Paris, France

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Mirror neurons
Mindreading
Prior intention
Motor intention
Embodied simulation

ABSTRACT

This paper is devoted to an assessment of the three-step model offered by Gallese and colleagues in support of the thesis that the function of the mirror mechanism is to mindread an agent's intention. The first step of the model is the acceptance of the direct-matching model of action understanding. The second step is the endorsement of a different model of mirror neuron activity, i.e. the model of chains of logically related mirror neurons (or motor chains) whose application to action-mirroring is supposed to show that the mirror mechanism enables an observer to predict the goal of the agent's forthcoming action. The third step is the endorsement of the 'deflationary' account of intention-ascription according to which to ascribe an intention to an agent is to predict the goal of the agent's forthcoming action. I argue that each step of the model faces insuperable objections.

1. Introduction

The discovery of mirror neurons by Giacomo Rizzolatti and his group in the early 1990's was a great discovery and a fascinating one. Mirror neurons were first discovered in the ventral premotor cortex and subsequently in the inferior parietal lobule of non-human primates by single-cell recording. They were reported to fire both when an animal performs a goal-directed action onto some target and also when the same animal observes another perform the same kind of transitive action. Evidence for the existence of mirror neuron activity in humans – i.e. a human 'mirror mechanism', as Gallese and Sinigaglia (2011) have recently called it –, has been reported on the basis of a variety of experimental techniques. Many studies have also reported that the human mirror mechanism can be elicited by a wider variety of actions than in non-human primates, including intransitive actions, not directed towards a specific physical target.

While the mirror mechanism was first documented in the brains of both non-human and human primates during the execution and the perception of actions, further evidence has been taken to show that the human mirror mechanism also underlies both the first-personal experiences of sensations and/or emotions and the third-personal recognition of (and perhaps also the second-personal response to) another's sensations and/or emotions. One plausible way to interpret the evidence for the activity of the mirror mechanism in the domains of sensations and/or emotions is simply to take notice of the fact that human actions encompass not only *goal-directed* but also *expressive* actions. While an agent's overt goal-directed action is a cue to the agent's goal, an agent's overt expressive action is a cue to the agent's sensation and/or emotion. In what follows I will focus on the activity of the mirror mechanism in the execution and the perception of goal-directed actions.[1]

☆ This article is part of a special issue of this journal on Mirror Neurons and Intentionality. I am grateful to Dr Antonella Corradini for both inviting me to the Workshop on Mirror Neurons and Intentionality in Milan on September 4, 2011, and for editing the special issue of Consciousness and Cognition, and to two anonymous referees for their comments on a previous version of this paper. This work was supported by a grant from the French Ministry of Research (ANR-BLAN SOCODEV).

* Fax: +33 146333933.
E-mail address: Jacob@ehess.fr

[1] Goldman (2008, 2009) has argued that the definition of mirroring should not be restricted to action-mirroring on the grounds that it must also apply to sensations and emotions. But if one draws the distinction between goal-directed and expressive actions, then one may economically restrict mirroring to action.

1053-8100/$ - see front matter
http://dx.doi.org/10.1016/j.concog.2013.02.005

Ever since the publication of Gallese and Goldman's (1998) seminal paper, the cognitive neuroscientists who discovered mirror neurons have interpreted their discovery in terms of the *simulation* approach to *mindreading*: the mirror mechanism is taken to enable an observer to ascribe an intention to the agent of an overt action via a process of mental simulation. As Goldman (2008, 2009) has pointed out, the mirror mechanism might contribute to mindreading another's intention in at least two ways: the mirror mechanism might either causally contribute to, or be constitutive of, mindreading. If the relation is causal, then the mirroring event and the mindreading event are distinct. If the former constitutes the latter, then there is a single event which is both a mirroring and a mindreading event. While Goldman (2008, 2009) has explicitly endorsed the former weaker view, Gallese, Rochat, Cossu, and Sinigaglia (2009, p. 108) have explicitly endorsed the stronger identity view: "The activation of the mirror neuron system is intrinsically constitutive of action and intention understanding, at least at the level of basic actions".

On the face of it, Gallese and colleagues' strong identity claim faces the following challenge. The mirror mechanism might perhaps enable an observer to *share* an agent's intention. But to mindread another's intention is to *ascribe* an intention to another. So the challenge for Gallese's strong identity claim is to close the gap between sharing an agent's intention and ascribing the intention to the agent. In recent years, Gallese has offered a three-step model of intention-ascription based on action mirroring that provides a tentative answer to this challenge.

- The first step is the acceptance of the direct-matching model of action understanding, whose application to action-mirroring is supposed to enable an observer to understand an agent's goal and to ascribe it to the agent.
- The second step is the endorsement of a different model of the mirror mechanism, i.e. the model of chains of logically related mirror neurons (or model of motor chains) whose application to action-mirroring is taken to show that the mirror mechanism enables an observer to predict the goal of the agent's forthcoming action.
- The third and final step is the endorsement of the so-called 'deflationary' account of intention-ascription according to which to ascribe an intention to an agent is to predict the goal of the agent's forthcoming action. In addition to endorsing this three-step model of intention-ascription based on action-mirroring, Gallese also embraces what he calls an *embodied* approach to mental simulation.

My twofold goal in this paper is to argue not only that the three-step model fails to establish Gallese's strong identity claim between action-mirroring and mindreading but also that there is a tension between features of the three-step model and the embodied approach to mental simulation. The paper is in six sections. In the second section, I argue that the uncritical acceptance of the mindreading interpretation of the mirror mechanism may have to some extent prematurely discarded a different interpretation of mirror neuron activity. In the third section, I examine the dilemma Csibra (2007) raised for the direct-matching model of action understanding and the solution favored by advocates of the three-step model of intention-ascription. In the fourth section, I argue that there is a tension between the solution to Csibra's dilemma accepted by advocates of the mindreading interpretation and the embodied approach to mental simulation. In the penultimate section, I turn to the model of chains of logically related mirror neurons and ask the question whether sharing an agent's chain of logically related mirror neurons could be sufficient for predicting the goal of the agent's next act. In the final section, I assess the deflationary model of intention-ascription.

2. Two basic options

As Rizzolatti, Fogassi, and Gallese (2004, p. 431) have characterized their initial discovery, "mirror neurons are a specific class of neurons that discharge both when the monkey performs an action and when it observes a similar action done by another monkey or the experimenter".[2] By advancing the cautious hypothesis that the mirror mechanism might be "a primitive version, or possibly a precursor in phylogeny, of a simulation heuristic that might underlie mindreading", Gallese and Goldman (1998, p. 498) have stressed the *synchronous interpersonal neural similarity* between the agent's and the observer's brains, which results from mirror neuron activity, thereby linking the mirror mechanism to the simulation approach to mindreading. Thus, the mindreading interpretation of the mirror mechanism is entirely consistent with the above definition of the mirror mechanism: it depends on the simulation approach to mindreading, which in turn is rooted in the claim that action-mirroring generates a synchronous interpersonal neural similarity between the agent's and the observer's brains.

However, Rizzolatti et al. (2004) do not define mirror neuron activity directly in terms of a synchronous interpersonal neural similarity across two distinct brains, but rather in terms of a *non-synchronous intrapersonal neural similarity* within single brains, at different times, in different tasks, i.e. the execution and the perception of action. In fact, synchronous interpersonal neural similarity between the agent's and the observer's brains could hardly be achieved unless there were a mechanism active in single brains, at different times, in different tasks (i.e. the execution and the perception of action). Two distinct brains could not stand in some appropriate similarity relation at a single time unless both brains were independently endowed with a mechanism (the mirror mechanism) active in two different tasks at different times. In other words,

[2] Because he thinks that the definition of mirroring should not be restricted to action but should also apply to emotions and sensations, Goldman (2008) has offered an original definition of a mirroring process that both considerably internalizes the definition of mirror neuron activity in an agent's brain and considerably liberalizes the definition of mirror neuron activity in an observer's brain. As I explained above, I think that the distinction between goal-directed and expressive actions makes this costly maneuver unnecessary, but I will not examine it any further in this paper.

Please cite this article in press as: Jacob, P. How from action-mirroring to intention-ascription? *Consciousness and Cognition* (2013), http://dx.doi.org/10.1016/j.concog.2013.02.005

synchronous interpersonal neural similarity across two distinct brains at a single time presupposes asynchronous intrapersonal neural similarity at different times in two different tasks. Presumably, natural selection operates on individuals' brains, not on pairs of individuals' brains. Only individual members of a species whose brains contained a mechanism active at different times in the execution and the perception of action could have been selected by evolution – not sets of pairs of individual members of a species, whose brains stood in some suitable similarity relation when one executed a grasping action and the other watched the former.

Admittedly then non-synchronous intrapersonal neural similarity is a more basic property of the mirror mechanism than synchronous interpersonal neural similarity. If so, then the latter must be a by-product of the former. This point was made by Hurley (2008, p. 758), as part of her criticism of Goldman's (2006) mental-simulation approach to mindreading, when she argued that "re-use", not similarity, is "the core generic sense of process-driven simulation". However, while Hurley (2008) emphasized the primary role of intrapersonal re-use, and de-emphasized the role of interpersonal similarity, she took it for granted, as Goldman (2006) did, that the primary function of action-mirroring is mindreading and that the correct approach to third-person mindreading must be mental simulation.[3]

But I want to consider an entirely different interpretation of action-mirroring based on the principle that there could not be synchronous interpersonal neural similarity between two different individuals unless there was a single mechanism in each individual's brain active at different times in different tasks. It looks as if a neural mechanism (the mirror mechanism) that is active in both the execution and the perception of instances of grasping is a mechanism whose function seems to be to abstract away from the many differences between executing and observing an act of grasping. For example, only an *agent* executing a task of grasping a target, not an observer perceiving the action, will both have efference copy information about his motor instruction and also haptic information about the target. Only the agent, not an observer, will be in a position to predict the sensory consequences of her action before executing it. Furthermore, among several studies, Kohler et al. (2002) and Keysers et al. (2003) have reported that single neurons in the monkey ventral premotor cortex selectively fire when the animal both executes and also sees, hears and both hears and sees such actions as peanut breaking, ring grasping or paper ripping. So not only does the mirror mechanism seem able to discard the many differences between executing and perceiving an action, but in perceptual tasks, it also seems able to abstract away from the further differences between vision and audition.

In a nutshell, the evidence shows that the mirror mechanism is active when an animal executes and perceives one and the same action (e.g. grasping). It also shows that it is able to achieve cross-modal integration in perceptual tasks. Now it seems as if the function of a mechanism able to deliver a representation of an action whose content brackets the differences between the motor, visual and auditory representations of one and the same action is to deliver a representation of this action with *conceptual* content. The view that the function of the mirror mechanism is to provide conceptual representations of actions makes sense of the variations in the statistical congruency between the motor and the perceptual properties of mirror neurons noted by Csibra (2005, 2007).[4]

This conceptualist interpretation of the mirror mechanism is in agreement with the view that mirror neuron activity underlies action recognition.[5] It is also in agreement with the view expressed by Rizzolatti et al. (2000, p. 542) and Craighero, Metta, Sandini, and Fadiga (2007) according to which area F5 of the monkey ventral premotor cortex (where mirror neurons were first discovered) is "a store of motor schemas or a *vocabulary* of actions", i.e. a *motor vocabulary*: "Neurons forming these vocabularies store both knowledge about an action and the description... of how this knowledge should be used". Thus, the motor vocabulary stored in F5 consists of mental schemas or symbols whose semantic role is to denote actions and such that the meanings of complex symbols depend systematically on the meanings of their constituents. If so, then mirror neuron activity does not directly underlie the attribution of a psychological state (e.g. an intention) to an agent. But it may contribute the abstract (amodal) content of the concept of the action (e.g. grasping) to determining the content of the agent's intention (to e.g. grasp a target). Finally, Gallese (2003, p. 1238) himself endorsed the conceptualist interpretation when he wrote: "If the different modes of presentation of events as intrinsically different as sounds, images, or willed effortful acts of the body are nevertheless bound together within a circumscribed, informational lighter level of semantic reference, what we have here is a mechanism instantiating conceptualization".

Now, the cognitive neuroscientists who discovered mirror neurons have often construed the output of action-mirroring from the standpoint of what they call *motor intentionality*, by which they mean to refer to the *motor representation* of an agent's action (or goal of her action), which they take to be shared by the agent and the observer (for example, cf. Rizzolatti & Sinigaglia, 2007; Rizzolatti & Sinigaglia, 2010). Building on Jeannerod's (1994) seminal investigation of *motor imagery* in humans, several authors (including Jeannerod, 2006 and Pacherie, 2000) have argued that while the content of an agent's *prior* intention to perform a goal-directed action is conceptual (or propositional), the non-conceptual content and format of the agent's *motor* intention may suitably be linked to the content of motor imagery, i.e. the content of the output of

[3] She also wrongly, I think, inverted the respective functions of simulation and mindreading, when she wrote that "according to simulation theory, mindreading aims at simulation of, and hence, matching the target's mental states" (Hurley, 2008, p. 757). Matching the target's mental state is the function of mental simulation, which is a step in the process of third-person mindreading, whose function is to ascribe the mental state to another. Ascription requires exiting the simulation stage. By contrast, Goldman's (2006) own account is a simulation-plus-projection account.

[4] Gallese and Sinigaglia (2011, p. 513) recognize that "interpersonal similarity between a simulator's and a target's mental state or process does not qualify as mental simulation unless it arises from intrapersonal reuse of the simulator's own mental state or process". But they do not further take intrapersonal neural similarity as a step towards a conceptualist interpretation of the mirror mechanism (cf. Section 3).

[5] Following Sperber (2005), I have argued for such an interpretation of mirror neuron activity in Jacob (2009). Meini and Paternoster (2012) also argue for a similar interpretation, but they hold a different view of concept-possession.

Please cite this article in press as: Jacob, P. How from action-mirroring to intention-ascription? *Consciousness and Cognition* (2013), http://dx.doi.org/10.1016/j.concog.2013.02.005

the process whereby the agent *imagines* performing the action without executing it. The reason why the content of the agent's prior intention (to e.g. turn the light on) must be conceptual is that it must be able to interact with the conceptual content of the agent's beliefs and desires.

Recently, Butterfill and Sinigaglia (2012) have further argued that a proper account of an agent's goal-directed action requires that the agent be able to coordinate two distinct representations of the outcome of her action: the outcome of the agent's action must be represented by both her intention and the motor representation of her own action. They take the outcome of the agent's action to be conceptually represented by the agent's intention, on the grounds that the content of the agent's intention must interact appropriately with the conceptual contents of her beliefs and desires. But they argue that the outcome of the agent's action must be non-conceptually represented by the agent's motor representation of her action. On their account, the content and format of the agent's motor representation of the outcome of her action is on a par with the non-conceptual content and format of motor imagery. Since the outcome of the agent's action is both represented by the agent's intention and her motor representation, the two kinds of representation must differ in format. They further offer a solution to the "interface" problem of how the two distinct representations of the outcome of the agent's action can match each other. They argue that the agent's intention conceptually represents the outcome by means of a demonstrative action concept that "defers" to the agent's motor representation of the outcome.

I would like to make three short points about Butterfill and Sinigaglia's (2012) interesting line of thinking about the distinction between the contents and formats of respectively an agent's intention and the motor representation of her action. First, mirror neuron activity was discovered in the brains of non-human primates and what is known about the content and format of motor imagery is based on the investigation of humans, not non-human primates. As a result, it is not entirely clear to what extent human motor imagery can shed light on the content of the output of action-mirroring. Secondly, it is not entirely clear whether Butterfill and Sinigaglia's distinction between the agent's intention and the agent's motor representation of her action leaves any room for the agent's *motor intention*, which shares the world-to-mind direction of fit of the agent's prior intention and the non-conceptual content of the agent's motor representation. Thirdly, the distinctive ways the outcome of an agent's action is being represented by respectively the agent's intention and her motor representation can be captured along several different lines. For example, the agent's intention (whether her prior intention or her motor intention) has a world-to-mind direction of fit, not a mind-to-world direction of fit. So it represents the outcome in a prescriptive, not a descriptive, way. Furthermore, the intention is a cause of the achievement of the outcome. By contrast, the motor representation of the agent's action has a mind-to-world direction of fit and so it represents the outcome in a descriptive, not a prescriptive, way. This distinction is consistent with the hypothesis that the content of the output of action-mirroring must be sufficiently abstract to be commonly entertained by an agent who performs an act and by an individual who sees, hears and both sees and hears the act being performed by another.

3. Answering Csibra's dilemma

Not only is the motor vocabulary or conceptualist interpretation of the mirror mechanism an alternative to the mindreading interpretation, but the latter asymmetrically depends on the former: there could not be synchronous interpersonal neural similarity between two distinct brains unless there were non-synchronous intrapersonal neural similarity within a single brain in different tasks (but of course the converse does not hold). Nonetheless, the simulation-based approach to mindreading has become the orthodox interpretation of the mirror mechanism among the cognitive neuroscientists who discovered mirror neurons. On Gallese's three-step model, the main function of the mirror mechanism is to mindread another's intention. The first step of Gallese's model is the acceptance of the *direct-matching model of action understanding*, whose application to action-mirroring is supposed to enable an observer to understand an agent's goal and to ascribe it to the agent.

The direct-matching model of action understanding itself involves the three following steps. First, perceiving an agent's goal-directed action causes the observer to covertly rehearse (or imitate) the agent's movements by mapping them onto the observer's motor repertoire (without executing them). Secondly, by covertly rehearsing the agent's movements, the observer comes to share the agent's goal. Thirdly, by sharing the agent's goal, the observer comes to understand it and somehow to ascribe it to the agent.[6]

Csibra (2007) has raised the following dilemma for the direct-matching model of action-understanding, the first horn of which is that action-mirroring consists in mapping the agent's executed *movements* onto the observer's motor repertoire in accordance with the first step of the direct-matching model. The second horn is that action-mirroring enables the observer to share and understand the agent's *goal* in accordance with the second and third steps of the direct-matching model. The dilemma between the two horns arises from the fact that an agent's movements do not stand in a one-to-one relation to her goal: not only can an agent recruit different movements in the service of a single goal, but she can also recruit one and the same movement in the service of different goals. So, if what is being mapped onto the observer's motor repertoire by action-mirroring is the agent's overt motor act (or bodily movements), then it is unlikely that mirroring could deliver a representation (or understanding) of the agent's goal. Conversely, if the output of mirroring is a representation (or understanding) of the agent's goal, then it is unlikely to be generated by the mapping of the agent's observed movements onto the observer's motor repertoire.

[6] While 'intention' refers to an agent's psychological state, 'goal' may refer either to a possible non-mental state of affairs in the environment, i.e. a goal-state, which the agent aims to achieve, or to a mental representation of such a goal-state (cf. Gergely & Csibra, 2003).

Please cite this article in press as: Jacob, P. How from action-mirroring to intention-ascription? *Consciousness and Cognition* (2013), http://dx.doi.org/10.1016/j.concog.2013.02.005

In response, Rizzolatti and Sinigaglia (2010) have opted for the second horn of Csibra's dilemma: on the basis of experimental results reported by Umiltà et al. (2008), they have argued that in the monkey brain, mirror neuron activity encodes *goals*. Umiltà et al. (2008) trained monkeys to grasp objects using both normal pliers and so-called 'reverse' pliers: grasping an object with normal pliers involves closing the fingers, but grasping with reverse pliers involves opening the fingers. Single cell recordings in area F5 in executive and perceptual tasks by Umiltà et al. (2008) show that F5 neurons discharged during the same phase of grasping 'regardless of whether this involved opening or closing of the hand' (Rizzolatti & Sinigaglia, 2010, p. 266). In other words, what seems to matter to the firing of F5 neurons is the agent's goal (grasping) irrespective of the difference between the agent's closing and opening his fingers.

4. What is so special about embodied simulation?[7]

Since they published their joint (1998) paper together in support of the mindreading interpretation of the mirror mechanism, Gallese's and Goldman's respective views have diverged to some extent. While Goldman (2006, 2008, 2009) has argued for a merely causal link between the mirror mechanism and mindreading another's intention, Gallese defends a stronger constitutive link: "The activation of the mirror neuron system is intrinsically constitutive of action and intention understanding, at least at the level of basic actions" (Gallese et al., 2009, p. 108). Gallese further advocates an *embodied* approach to mental simulation. I will examine Gallese's strong claim in favor of the constitutive link between mirroring and mindreading another's intention in the next section, when I turn to his deflationary view of intention-ascription. In the present section, I want to suggest that there is a deep tension (or a dilemma) between Rizzolatti and colleagues' solution to Csibra's dilemma and the endorsement of an embodied approach to mental simulation.

This tension shows up quite clearly in a recent paper by Gallese and Sinigaglia (2011, p. 516), in which they construe embodied simulation as "the reuse of mental representations that are bodily in format" and they further argue that "there is substantial evidence that the mirror mechanism is selective for motor goals and motor intentions, regardless of the body effectors and kinematic features enabling their accomplishments". The dilemma for this view is the following: on the one hand, in response to Csibra's dilemma, Gallese and Sinigaglia (2011) clearly take it that the mirror mechanism codes an agent's goals (and intentions), not the agent's movements. But on the other hand, they also characterize the mirror mechanism as a process of mental simulation whereby the goal is coded in a bodily format. So the question is: how could the mirror mechanism code an agent's goal in a bodily format "regardless of the body effectors and kinematic features enabling its accomplishments"?

So the next question arises: What makes an approach to mental simulation embodied? What makes the question fairly complicated is that many, but not all, advocates of the *embodied cognition* program mean to reject what Wheeler (2005, p. 81) calls the *neurocentric* assumption "at work within orthodox cognitive science", according to which "the causal factors that explain the adaptive richness and flexibility of naturally occurring intelligent behavior are located neither in the agent's non-neural body nor in her environment, but pretty much exclusively in her brain". As Clark (2008a, 2008b) and Jacob (2012b) have further argued, there are currently two distinct ways of rejecting the *neurocentric* assumption: one is Shapiro's (2011) *body-centric* view according to which an individual's mind (or cognition) is constituted by both the individual's brain and the brain's non-neural bodily environment, which makes an individual's mind a *two*-place relation between the individual's brain and its non-neural bodily environment. The other one is Clark and Chalmers' (1998) and Clark's (2008b) *extended mind* thesis according to which an individual's mind (or cognition) is constituted by the individual's brain, the brain's non-neural bodily environment and the individual's non-bodily environment, which makes an individual's mind a *three*-place relation between an individual's brain, its non-neural bodily environment and its non-bodily environment.

Not all advocates of *embodied* cognition, however, are willing to reject the neurocentric assumption. Instead, some advocates of embodied cognition (including Gallese) seem happy to recognize that an individual's brain is part of her body and want to emphasize instead the contribution of the human motor and/or perceptual systems to higher human cognitive capacities (e.g. mindreading). Gallese himself has offered three answers to the question what makes mental simulation embodied, all of which seem consistent with neurocentrism.

First, Gallese (2009a, p. 493) argues that mental simulation is embodied if it is "a mandatory, pre-rational, non-introspectionist functional mechanism". Secondly, Gallese (2009b, p. 520) argues that what makes mental simulation embodied is that it is "a crucial functional mechanism of intersubjectivity by means of which the actions, emotions, and sensations of others are mapped by the same neural mechanisms that are normally activated when we act or experience similar emotions and sensations". While Gallese's second condition amounts to defining the process of mental simulation as interpersonal neural similarity, both answers are clearly consistent with a neurocentric approach to mental simulation. Finally, Gallese and Sinigaglia (2011, p. 513) have recently adopted Goldman and de Vignemont's (2009) proposal that what makes a representation embodied is its *bodily format*: "the idea is that just as a map and a series of sentences might represent the same route with a different format, so mental representations might have partly overlapping contents (e.g. a motor goal, an emotion or sensation) while differing from one another in their format (e.g. bodily instead of propositional)".

Building on Jeannerod's (1994, 2006) insight (echoed by Pacherie, 2000), Butterfill and Sinigaglia have argued that, unlike thinking about an action of grasping a cup, executing the action, perceiving the action executed by another and imagining

[7] This is the title of a Gallese and Sinigaglia's (2011) paper.

Please cite this article in press as: Jacob, P. How from action-mirroring to intention-ascription? *Consciousness and Cognition* (2013), http://dx.doi.org/10.1016/j.concog.2013.02.005

performing the action without executing it, all three may be said to recruit motor representations with the same "performance profile", to the extent that these representations may interfere with one another. In line with Gallese and Sinigaglia (2011), they conclude that what makes embodied simulation *embodied* is not the content, but the bodily format, of the motor representation of the (goal of the) action.

I claimed above that there is a tension between two strands in Gallese and Sinigaglia's (2011) approach to embodied simulation: on the one hand, the solution to Csibra's dilemma requires that detailed information about bodily effectors and the kinematics of the agent's movements be disregarded from the content of the representation of the agent's goal generated by action-mirroring. On the other hand, the stress on embodiment pulls towards including detailed information about the agent's bodily effectors and the kinematics of her movements as part of the content of the representation of the agent's goal generated by action-mirroring. Could the tension be resolved by appeal to the distinction between the content and the format of a representation of the agent's goal?[8]

Given Goldman and de Vignemont's (2009) approach to embodied cognition further elaborated by Goldman's (2012), I don't think that appeal to the content/format distinction can resolve the tension. Their approach to embodied cognition is two-tiered. On the one hand, as Goldman (2012, pp. 73–74) makes clear, a representation derives its bodily format from its bodily content: a representation qualifies as embodied if it is generated by a mechanism (e.g. the somatosensory system or the motor system) whose primary function is to represent one's own bodily parts and states (e.g. pain, tickle, temperature, itch, muscular and visceral sensations, sensual touch, and other feelings from, and about, the body). For example, on Goldman's (2012, p. 79) account, action-mirroring counts as an embodied prescriptive representation of how the agent is to move her own *effectors* (in e.g. a task of executing an act of grasping). On the other hand, Goldman (2012) further assumes that an agent's representation of her own body with a bodily format can be *re-deployed* (or re-used) in a novel cognitive task, for representing other things than her own body parts. According to Goldman, the neuroscientific findings show that the embodied prescriptive representation of how to move the agent's effectors generated in an agent's brain by mirror neuron activity is *re-deployed* in action observation.

Thus, acceptance of Goldman's two-tiered conception of embodied cognition seems to mandate a solution to Csibra's dilemma which runs against the option favored by Rizzolatti, Gallese and Sinigaglia. On Goldman's conception, the output of action-mirroring represents an agent's motor command to one of the agent's effectors, not a representation of the agent's goal irrespective of particular effectors and the kinematics of the effectors' movements. In a nutshell, it is really unclear how to reconcile an embodied version of the simulation approach to mindreading with a solution to Csibra's dilemma.

5. The motor chain model of mirror neuron activity

So far, I have only discussed the first assumption of the three-step model of intention-ascription based on action-mirroring, i.e. the direct matching model of action-understanding. The three-step model of intention-ascription also involves a different model of action-mirroring (or mirror neuron activity): the model of "chains of logically related mirror neurons". The second assumption of the three-step model is that by sharing chains of logically related mirror neurons with the agent of a goal-directed action, an observer can predict the goal of the agent's forthcoming action. In the present section, I turn to this second assumption.

The model of chains of logically related mirror neurons, which was first considered in an early paper by Di Pellegrino, Fadiga, Fogassi, Gallese, and Rizzolatti (1992), was re-introduced in order to account for intriguing findings based on single cell-recordings in the monkey by Fogassi et al. (2005) and on brain-imaging in humans by Iacoboni et al. (2005). Fogassi et al. (2005) recorded single mirror neurons in the monkey inferior parietal lobule (IPL) while the animal executed an act of grasping a target which was embedded into two distinct more complex actions, one of which was eating (by bringing the target to the mouth) and the other of which was placing the target into a container. They also recorded single mirror neurons while the animal was watching another agent perform the same action of grasping embedded into one or the other of the two more complex actions. They found that different mirror neurons fire when the animal either performed or observed the same action of grasping, according to whether the following act was bringing to the mouth or placing into a container. These findings are currently interpreted as evidence that mirror neurons in the monkey IPL form motor chains or are organized along chains of intentional motor acts.

Iacoboni et al. (2005) used functional magnetic resonance imaging to study the brain activations of participants while they watched video-clips displaying three possible conditions. In the Context condition, they saw a teapot, a cup and a plate with cookies either orderly as if before tea or disorderly as if after tea. In the Action condition, they saw a human hand grasp a cup either by full prehension or precision grip with no contextual elements. In the Intention condition, they saw a human hand grasp a cup by either full prehension or precision grip in one of the two contexts described above. Iacoboni et al. (2005) found a significantly stronger signal in the posterior part of the inferior frontal gyrus and the adjacent sector of the ventral premotor cortex where hand actions are represented for the observation of the Intention condition than for the other two conditions.

Both Fogassi et al. (2005) and Iacoboni et al. (2005) take their findings to support the model of chains of logically related mirror neurons (or motor chains). On this model, mirror neurons are taken to form chains: one chain might link cells firing

[8] I am grateful to an anonymous reviewer for forcing me to be more explicit on this point.

Please cite this article in press as: Jacob, P. How from action-mirroring to intention-ascription? *Consciousness and Cognition* (2013), http://dx.doi.org/10.1016/j.concog.2013.02.005

for grasping to cells firing for bringing to the mouth. Another chain might link cells firing for grasping to cells firing for placing into a container. In accordance with the second assumption of the three-step model, the authors assume that sharing an agent's chain of logically related mirror neurons enables an observer to predict the goal of the agent's next act.[9]

Now, the question is whether sharing the agent's motor chain of logically related mirror neurons could be a *sufficient* condition for predicting the goal of the agent's next act. In the experiments by either Fogassi et al. (2005) or Iacoboni et al. (2005), the inference of the goal of the agent's next act depends on two dissociable parameters: the observation of the agent's executed act of grasping and contextual cues. What the data show is that the activation of chain CH1 of logically related mirror neurons in an observer's brain is triggered by the perception of the act of grasping in context C1, while the activation of chain CH2 of logically related mirror neurons in the brain of the same observer is triggered by the perception of the act of grasping in context C2. Given that the perceived act of grasping is the same in both cases, the data do not rule out the possibility that what enables the observer to infer the two distinct goals of the agent's two distinct next acts is the difference between context C1 and context C2. As the findings by Iacoboni et al. (2005) in the Context condition clearly show, the perception of the context alone fails to trigger mirror neuron activity. If so, then sharing chains of logically related mirror neurons with the agent will not be sufficient for predicting the goal of the agent's next act because knowledge of the context (which is not achieved by mirror neuron activity at all) is also necessary.

According to the model of motor chains, mirror neurons are organized into chains of motor acts. But clearly if mirror neurons are organized into motor chains, mirror neuron activity itself cannot select the right motor chain in the observers' brain: the right motor chain cannot be self-selected in the observer's brain; only the processing of contextual cues, which cannot be achieved by mirror neuron activity, may enable the observer to select the right motor chain.

6. The deflationary account of intention-ascription

The third assumption of the model of intention-ascription is what Gallese (2007, p. 662) has labeled the *deflationary* account of intention-ascription, according to which determining an agent's intention (or ascribing it to the agent) amounts to determining why a given act (e.g. grasping a cup) was performed, which in turn is equivalent to detecting the goal of the agent's "still not executed and impending subsequent act (e.g. bringing the cup to the mouth").[10] As a result of accepting both the view that sharing another's chain of logically related mirror neurons is a sufficient basis for predicting the goal of her forthcoming act and the deflationary account of intention-ascription, one can conclude that sharing an agent's chain of logically related mirror neurons enables an observer to ascribe an intention to the agent.

In accordance with the deflationary account of intention-ascription, Iacoboni et al. (2005, p. 0533) have argued that their own findings "strongly suggest that coding the intention associated with the actions of others is based on the activation of a neuronal chain formed by mirror neurons coding the observed motor act and by 'logically related' mirror neurons coding the motor acts that are most likely to follow the observed one, in a given context". The application of the deflationary account of intention-ascription here is open to the two following critical comments.

First, when applying the deflationary account of intention ascription, Iacoboni et al. (2005, p. 0530) entertain the hypothesis that the mirror neuron system might code "the global intention associated with the observed action". What is meant by "the agent's global intention"? In the context of their study, Iacoboni et al. (2005) must mean that the activity of a given chain of logically related mirror neurons codes the agent's *prior* intention to e.g. drink and that by sharing the agent's chain of logically related mirror neurons, an observer can ascribe to the agent the *prior* intention to e.g. drink. If so, then it is surprising that Rizzolatti and Sinigaglia (2010, p. 271) backtrack from this strong claim and instead endorse the weaker claim that "the studies reviewed above [by which they refer to Iacoboni et al. (2005) study] indicate that the parieto-frontal mirror network may subserve the understanding of the *motor* intention underlying the actions of others" [my emphasis].

In the Action and Intention conditions of Iacoboni et al.'s (2005) study, the agent has the *prior* intention to either drink or clean and the *motor* intention to grasp the cup using either full prehension or precision grip. One must choose between two strategies: one can argue that mirror neuron activity in participants' brains codes the agent's *motor* intention to grasp the cup with either full prehension or precision grip. Alternatively, one can argue that mirror neuron activity in participants' brains codes the agent's *prior* intention to either drink or clean. But one cannot argue at once that mirror neuron activity in the observer's brain codes both the agent's prior intention and her motor intention. In particular, one cannot do so in the context of Iacoboni et al.'s (2005) study because what constitutes decisive evidence in favor of ascribing to the agent one prior intention (drinking) over the other (cleaning) are the contextual elements (before tea vs. after tea), not the distinctive ways in which the agent grasps the cup (by full prehension vs. by precision grip). As it happens, in the Intention condition, as in the Action condition, participants see alternations between instances of grasping the cup with full prehension and with precision grip. However, only by arguing that mirror neuron activity in the observer's brains codes the agent's *motor* intention (*not* her prior intention) could one further claim that the content of mirror neuron activity is coded in a *specific bodily format* and therefore that the process of mental simulation involved in mirror neuron activity is *embodied*.

[9] I accept Barlassina's (2011) well-taken reply to my earlier criticisms of the model of chains of logically related mirror neurons in Jacob (2008).

[10] As Iacoboni et al. (2005, p. 0533) put it, "to ascribe an intention is to infer a forthcoming new goal, and this is an operation that the motor system does automatically". For a very interesting assessment of the deflationary account of intention-ascription, cf. Spaulding (in press).

Please cite this article in press as: Jacob, P. How from action-mirroring to intention-ascription? *Consciousness and Cognition* (2013), http://dx.doi.org/10.1016/j.concog.2013.02.005

Secondly, referring to mirror neuron activity, as mirror neuron theorists do, as ''coding'' an agent's intention is ambiguous between two distinct interpretations: ''coding'' an agent's intention might either mean *sharing* the agent's intention or *ascribing* the intention to the agent. Sharing the agent's intention is to have the same intention as the agent, namely to intend to do something or other (cf. Jacob, 2012a). But ascribing an intention to another is not to intend to do something or other: to ascribe an intention to an agent is to judge or believe that another intends to do so and so. While having the concept INTENTION is not necessary to intend to do something and therefore to share another's intention, it is necessary for believing that another intends to so something, and therefore to ascribe an intention to another. Nor (I believe) is intending to do something a necessary condition for believing that another intends to do something and therefore for ascribing the intention to do something to another. I am inclined to say that by sharing the agent's chain of logically related mirror neurons, an observer may share the agent's ability to predict the goal of the agent's next act. But it would, I think, be a mistake to identify the agent's prior intention to drink with her ability to predict the goal of her next act (bringing to the mouth) while she is performing the act of grasping the cup. Similarly, it would, I think, be a mistake to identify the observer's ability to ascribe to the agent the prior intention to drink with her ability to predict the goal of the agent's next act (bringing to the mouth) upon observing the agent grasp the cup.

7. Conclusions

In this paper, I have assessed each step of the impressive three-step model offered by Vittorio Gallese and his colleagues, who discovered mirror neurons, in support of the thesis that the function of mirror neuron activity is to enable an observer to ascribe an intention to an agent who performs a goal-directed action. Distancing myself from the detailed criticisms of each step of the three-step model of intention-ascription, I would now like to make explicit one of the fundamental bases of my disagreement with Gallese and colleagues' mindreading interpretation of the mirror mechanism in how one construes what it takes to achieve a task of mindreading.

What I think Premack and Woodruff's (1978) seminal paper and the subsequent discussion of this paper by the three philosophers Bennett (1978), Dennett (1978) and Harman (1978) showed is that one can predict an agent's likely behavior by *sharing* her motivations (goals and desires) and epistemic states (beliefs about the way the world is). But predicting another's likely behavior by sharing her epistemic and motivational states falls short of providing evidence that one can read another's mind for the simple reason that to share another's belief and desire about some states of affairs in the world just is to represent the world in the same way as the other individual does.

What would decisively show that one can read another's mind (or possesses a 'theory-of-mind') is that one could ascribe to an agent motivations and/or epistemic states that *one does not share*. For example, it is one thing to share an agent's intention; it is another thing to ascribe an intention to the agent. To share an agent's intention is to have an intention. To ascribe an intention to an agent is to believe or judge that another intends to do so and so. (Furthermore, one could not come to judge or believe that another intends to do so and so unless one possessed the concept INTENTION.) To judge or believe that an agent intends to do so and so, not to share her intention, is to mindread her intention. The significance of the famous standard elicited false belief task (designed by Wimmer and Perner, 1983) lies in the fact that while the participant has a true belief about some object's location, her task is to predict where an agent with a false belief will look for the object. Only if she can ascribe to the agent a belief whose content the participant does *not* share can she pass the standard elicited false belief task. While there are reasons to believe that there is more to passing the standard elicited false belief task than ascribing false beliefs to another, passing the standard false belief task is clearly a sufficient condition for displaying the ability to ascribe to another a psychological state distinct from one's own (cf. Baillargeon, Scott, & He, 2010; Bloom & German, 2000). Thus it is that the ability to ascribe to others psychological states that one does not share has been taken to be a hallmark of the ability to read others' minds.

My claim that sharing another's intention falls short of mindreading her intention for the reasons offered in the above paragraph seems open to two criticisms. On the one hand, while the evidence (reviewed by Baillargeon et al., 2010) based on spontaneous non-elicited tasks is taken by many (but not all) developmental psychologists to suggest that preverbal human infants can make sense of an agent's instrumental action by taking into account the contents of both her motivations and epistemic states (including her false belief), it sounds unlikely (if not outrageous) to credit infants with the ability to ascribe to an agent a belief that the infants themselves take to be false.[11] On the assumption that preverbal human infants achieve tasks of mindreading, it would seem that ascribing an intention to another (as opposed to sharing another's intention) might be too strong a requirement on mindreading. On the other hand, the findings reported by Kovacs, Téglas, and Endress (2010) show that human adults come to share automatically another's perspective. On the assumption that this counts as mindreading another's perspective, these findings too show that ascribing a mental state to another would seem unnecessary as a condition on mindreading.[12]

I will consider the first objection first. In most (if not all) so-called spontaneous non-elicited ''false belief tasks'', what is being tested is participants' ability to make sense of an agent's false belief by measuring participants' surprise when the agent with a false belief reaches for the object at its actual location or by measuring participants' ability to look towards

[11] See e.g. Perner and Ruffman (2005) for skepticism.

[12] I am grateful to an anonymous referee for pressing me to be more explicit on these points.

the empty location where an agent with a false belief is going to reach for the object in anticipation of the agent's action. If and when an infant displays surprise upon observing an agent with a false belief reach for an object at its actual location or looks at the empty location in anticipation of the action of an agent with a false belief, it makes sense to assume that the infant is able to take the agent's epistemic perspective, represent the content of the agent's false belief and, therefore, to ascribe to the agent a belief that happens to be false. What the infants' surprise or anticipatory looking behavior does *not* require is that the infant draws an explicit comparison between the content of her own true belief and the content of the agent's false belief. It is not required that the infant recognizes that, while her own belief is true, the agent's belief is false. In other words, it is not required that she assesses the agent's belief *as false* or that she predicates the property of falsehood of the agent's belief.

I now turn to the findings by Kovacs et al. (2010): in their experiment, adult participants see a video displaying a blue smurf who places a ball on a table on which there is also an occluder. After the ball has been placed on the table, its own self-propelled motion takes it behind the occluder and it either remains there or not. The participants' psychophysical task is to press a button as fast as possible upon detecting the ball after the removal of the occluder. While Kovacs et al. (2010) report that adults are faster to press the button when they expect the ball to be behind the occluder than when they do not, they also found that adults are faster when they do not expect the ball to be there, but the smurf falsely expects it to be there. This shows that adults automatically take into account the smurf's false expectation, in spite of the irrelevance of the smurf's expectation to the psychophysical task. In Kovacs and colleagues' experiment, participants are not presented with an instrumental action performed by an agent as a means towards satisfying her desire or intention in light of her belief. Their task is to press a button as fast as possible, not to make sense of an agent's belief and desire. Although they see the blue smurf bring the ball and place it on the table, it is questionable whether participants are representing the content of the smurf's motivation at all. Thus, Kovacs et al.'s findings are surprising for two related reasons: one is that in spite of its complete irrelevance to the task, the smurf's false expectation influences participants' performance. The other is that participants seem to automatically take the smurf's epistemic perspective without ascribing any motivation to the smurf. While this is evidence that participants take the smurf's epistemic perspective into account, it is still an open question whether participants specifically tie the content of the smurf's false expectation to the smurf or whether they use this content as a way of mitigating their confidence in the truth of their own conflicting expectation.

In some of their writings, Gallese and his colleagues seem aware of the fact that they must face the challenge of offering an alternative conception of mindreading. Thus, Gallese, Keysers, and Rizzolatti (2004, p. 396) have described the mirror mechanism as "the fundamental mechanism that allows us a direct experiential grasp of the mind of others is not conceptual reasoning but direct simulation of the observed events". Presumably, by "a direct experiential grasp of the mind of others" what Gallese et al. (2004) mean is *sharing* another's mental state (e.g. an intention). Gallese and Sinigaglia (2011, p. 518) explicitly recognize that the mirror mechanism can only play a causal, not a constitutive, role in mindreading, "when mindreading needs a representational attribution, that is, when attribution concerns propositional attitudes such as beliefs and desires that can be construed as putative reasons for action". But they maintain that the mirror mechanism constitutes mindreading "when mindreading involves functional attribution of mental states or processes (such as a motor goal or a motor intention...) having a bodily format".

So now the question arises what are the grounds for the distinction between *functional* and *representational attributions* of psychological states to others? While the purpose of representational attribution is to construe others' mental states as putative *reasons* for their actions, what else could be accomplished by functional attributions? As Gallese and Sinigaglia (2011, p. 517) put it, "functionally, an attribution is a representation of a goal, intention or belief which plays some role in enabling one to deal with an agent by virtue of its being appropriately related to that agent's goal, intention or belief". As I understand it, a an instance of a functional (non-representational) attribution of a goal and/or intention to another would consist in forming a joint goal (i.e. moving a piece of furniture together) and taking the steps towards fulfilling it by engaging in joint action with another. Thus, sharing a joint goal in this sense does seem to "enable one to deal with an agent by virtue of being appropriately related to that agent's goal". Now the ability to share a joint goal and engage in joint action with another has been highlighted by Tomasello, Carpenter, Call, Behne, and Moll (2005) as a uniquely human cognitive and motivational capacity, not shared by non-human primates. If so, then mirror neuron activity would not seem sufficient to support sharing a joint goal and engaging in joint action. In fact, it seems to me that establishing a joint goal and fulfilling it by successfully executing a joint action require for each agent not only to track and mindread the other's motivational and epistemic states and to construe them as putative reasons for her action, but also to engage in communicative action. If this is correct, then far from being achieved by mirror neuron activity, functional attribution (exemplified by sharing a joint goal) seems to rely on, and require, representational attributions of mental states to others, i.e. construing others' mental states as reasons for their actions.

References

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*(3), 110–118.

Barlassina, L. (2011). After all, it's still replication: A reply to Jacob on simulation and mirror neurons. *Res Cogitans, 8*(1), 92–111.

Bennett, J. (1978). Some remarks about concepts. *Behavioral and Brain Sciences, 4*, 557–560.

Bloom & German, T. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*, B25–B31.

Butterfill, S. & Sinigaglia, C. (2012). Intention and motor representation in purposive action. *Philosophy and Phenomenological Research*.

Clark, A. (2008a). Pressing the flesh: a tension in the study of the embodied, embedded mind? *Philosophy and Phenomenological Research, 76*(1), 37–59.

Clark, A. (2008b). *Supersizing the mind, embodiment, action, and cognitive extension*. Oxford: Oxford University Press.
Clark and Chalmers (1998). The extended mind. *Analysis, 58*(1), 7–19.
Craighero, L., Metta, G, Sandini, G. & Fadiga, L. (2007). The mirror-neurons system: data and models. In von Hofsten & Rosander (Eds.), *Progress in Brain Research*, Vol. 164.
Csibra, G. (2005). Mirror neurons or emulator neurons? <http://mirrorneurons.free.fr>.
Csibra, G. (2007). Action mirroring and action understanding. In P. Haggard, Y. Rossetti & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition*. Attention and Performance XXII. Oxford: Oxford University Press.
Dennett, D. C. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences, 4*, 568–570.
Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91*, 176–180.
Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organisation to intention understanding. *Science, 308*, 662–667.
Gallese, V. (2003). A neuroscientific grasp of concepts: from control to representation. *Philosophical Transactions of the Royal Society, London B, 358*, 1231–1240.
Gallese, V. (2007). Before and below theory of mind: Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society, 362*(1480), 659–669.
Gallese, V. (2009b). Mirror neurons, embodied simulation, and the neural basis of social identification. *Psychoanalytic Dialogues, 19*(5), 519–536.
Gallese, V., & Goldman, A. I. (1998). Mirror neurons and the simulation theory of mindreading. *Trends in Cognitive Sciences, 2*, 493–501.
Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of social cognition. *Trends in Cognitive Sciences, 8*, 396–403.
Gallese, V. (2009a). Mirror neurons and the neural exploitation hypothesis: From embodied simulation to social cognition. In J. A. Pineda (Ed.), *Mirror neuron systems* (pp. 163–190). New York: Humana.
Gallese, V., Rochat, M., Cossu, G., & Sinigaglia, C. (2009). Motor cognition and its role in the phylogeny and ontogeny of action understanding. *Developmental Psychology, 45*(1), 103–113.
Gallese, V., & Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends in Cognitive Sciences, 11*, 512–519.
Gergely, G., & Csibra (2003). Teleological reasoning about actions: The naïve theory of rational actions. *Trends in Cognitive Sciences, 7*, 287–292.
Goldman, A. I. (2006). Simulating Minds the Philosophy, Psychology, and Neuroscience of Mindreading. Oxford: Oxford University Press.
Goldman, A. I. (2008). Mirroring, mindreading, and simulation. In J. A. Pineda (Ed.), *Mirror neuron systems* (pp. 311–330). New York: Humana.
Goldman, A. I. (2009). Mirroring, simulating and mindreading. *Mind and Language, 24*(2), 235–252.
Goldman, A. I. (2012). A moderate approach to embodied cognitive science. *Review of Philosophy and Psychology, 3*, 71–88.
Goldman, A. I., & de Vignemont, F. (2009). Is social cognition embodied? *Trends in Cognitive Sciences, 13*(4), 154–159.
Harman, G. (1978). Studying the chimpanzee's theory of mind. *Behavioral and Brain Sciences, 4*, 576–577.
Hurley, S. (2008). Understanding simulation. *Philosophy and Phenomenological Research, 77*(3), 755–774.
Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with One's own mirror neuron system. *Plos Biology, 3*, 529–535.
Jacob, P. (2008). What do mirror neurons contribute to human social cognition? *Mind and Language, 23*(2), 190–223.
Jacob, P. (2009). A philosopher's reflections on the discovery of mirror neurons. *Topics in Cognitive Science, 1*(3), 570–595.
Jacob, P. (2012a). Sharing and ascribing goals. *Mind and Language, 27*(2), 202–229.
Jacob, P. (2012b). Embodying the mind by extending it. *Review of Philosophy and Psychology, 3*(1), 33–51.
Jeannerod, M. (1994). The representing brain: neural correlates of motor intention and imagery. *Behavioral and Brain Sciences, 17*(2), 187–245.
Jeannerod, M. (2006). *Motor cognition*. Oxford: Oxford University Press.
Keysers, C., Kohler, E., Umiltà, M. A., Fogassi, L., Rizzolatti, G., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research, 153*, 628–636.
Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions; action representation in mirror neurons. *Science, 297*, 846–848.
Kovacs, A. M., Téglas, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*, 1830–1834.
Meini, C., & Paternoster, A. (2012). Mirror neurons as a conceptual mechanism? *Mind and Society, 11*(2), 183–201.
Pacherie, E. (2000). The content of intentions. *Mind and Language, 15*(4), 400–432.
Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science, 308*, 214–216.
Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 4*, 515–526.
Rizzolatti, G., Fogassi, L., & Gallese, V. (2000). Cortical mechanisms subserving object grasping and action recognition: A new view on the cortical motor functions. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge, Mass: MIT Press.
Rizzolatti, G., Fogassi, L., & Gallese, V. (2004). Cortical mechanisms subserving object grasping, action understanding, and imitation. In M. Gazzaniga (Ed.), *The cognitive neurosciences III* (pp. 427–440). Cambridge, MA: MIT Press.
Rizzolatti, G., & Sinigaglia, C. (2007). Mirror neurons and motor intentionality. *Functional Neurology, 22*(4), 205–210.
Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews. Neuroscience, 11*(4), 264–274.
Shapiro, L. (2011). *Embodied cognition*. London: Routledge.
Spaulding, S. (in press). Mirror Neurons and Social Cognition. *Mind and Language.*
Sperber, D. (2005). Mirror neurons or concept neurons? <http://www.interdisciplines.org/mirror/papers/1>.
Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675–691.
Umiltà, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., et al (2008). When pliers become fingers in the monkey motor system. *Proceedings of the National Academy of Sciences, 105*(6), 2209–2213.
Wheeler, M. (2005). *Reconstructing the cognitive world: The next step*. Cambridge MA: MIT Press.

Please cite this article in press as: Jacob, P. How from action-mirroring to intention-ascription? *Consciousness and Cognition* (2013), http://dx.doi.org/10.1016/j.concog.2013.02.005