From mirror neurons to joint actions

Elisabeth Pacherie, Jérôme Dokic

HAL Id: ijn_00352579

https://jeannicod.ccsd.cnrs.fr/ijn_00352579

Submitted on 13 Jan 2009

ELSEVIER

From mirror neurons to joint actions

Action editors: Luca Tummolini and Cristiano Castelfranchi

Elisabeth Pacherie *, Jérôme Dokic

Institut Jean Nicod, CNRS-EHESS-ENS, 1 bis, avenue de Lowendal, 75007 Paris, France

Abstract

The discovery of mirror neurons has given rise to a number of interpretations of their functions together with speculations on their potential role in the evolution of specifically human capacities. Thus, mirror neurons have been thought to ground many aspects of human social cognition, including the capacity to engage in cooperative collective actions and to understand them. We propose an evaluation of this latter claim. On the one hand, we will argue that mirror neurons do not by themselves provide a sufficient basis for the forms of agentive understanding and shared intentionality involved in cooperative collective actions. On the other hand, we will also argue that mirror neurons can nevertheless play an important role in an account of the production and understanding of joint action, insofar as they provide the basic constituents of implicit agent-neutral representations and are useful elements in a process of online mutual adjustment of participants' actions.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Mirror neurons; Understanding of action; Joint action; Social cognition

1. Introduction

In 1988, Giacomo Rizzolatti and members of his research team in Parma showed that neurons in area F5 in the monkey ventral premotor cortex do not code elementary movements but goal-related actions (Rizzolatti, Carmada, Gentilucci, Luppino, & Matelli, 1988). They also discovered among F5 neurons a class of visuomotor neurons that fire both when the monkey observes an action performed by another individual and when it executes the same or a similar action. These neurons became known as mirror neurons (MNs). There is also now strong indirect evidence in favor of the existence of mirror neuron systems in humans.

The discovery of mirror neurons has given rise to a number of interpretations of their functions together with speculations on their potential role in the evolution of specifically human capacities. Thus, MNs have been thought to ground many aspects of human social cognition, such as under-

standing of action and intentional agency (Gallese, 2003a; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fogassi, & Gallese, 2001), imitation (Gallese, 2003a; Rizzolatti et al., 2001; Wohlschlager & Bekkering, 2002), empathy (Gallese, 2001, 2003b), mind-reading (Gallese, 2003a; Gallese & Goldman, 1998) and the emergence of language (Arbib, 2002, 2003; Rizzolatti & Arbib, 1998).

In this paper we will focus on action understanding and its relevance to joint action. We will defend two main claims. Our first claim, against maximalist interpretations of what MNs are for, is that MNs do not provide by themselves a sufficient basis for an (implicit) understanding of intentional agency. Our second claim is that, by providing the basic constituents of implicit agent-neutral representations, MNs may nevertheless play an important role in the capacity to engage in cooperative collective actions and to understand them.

In Section 1, we briefly review the main empirical evidence for the existence of MNs in monkeys and humans. Section 2 examines two competing interpretations of these data and introduces two important distinctions relevant to

---

their assessment: the distinction between informational content and representational content and the distinction between self-reference and self-relativity. In Section 3, we argue that it is necessary to distinguish between at least *four levels of action understanding* (perceptual understanding, motor understanding, agentive understanding and meta-representational understanding). We also claim that mirror neurons underlie motor understanding but that further mechanisms and processes are required to explain agentive understanding. In Sections 4 and 5, we turn to joint action. We first examine some standard philosophical analyses of joint actions and the difficulties they are confronted with in accounting for their essential cooperative dimension. We argue that cooperative joint actions require from the participating agents the capacity to form *agent-neutral representations* of the instrumental (means-end) structure of these actions and to use them to control their execution. We further argue that the mirror system already provides the basic constituents of implicit agent-neutral representations and we sketch a model in which the essential cooperative features of joint actions are taken into account at the most basic level.

2. What are mirror neurons?

In a series of single neuron recording experiments on macaque monkeys, Rizzolatti and his colleagues investigated the functional properties of neurons in area F5, the rostralmost sector of the ventral premotor cortex that controls hand and mouth movements. A fundamental functional property of area F5 is that most of its neurons do not discharge in association with elementary movements but are active during purposeful object-oriented actions, such as grasping, tearing, holding or manipulating objects. Although the majority of neurons in F5 are purely motor neurons, area F5 also contains two classes of visuomotor neurons: canonical neurons and mirror neurons.

Canonical neurons are activated during the execution of goal-related movements and also discharge during object observation, typically showing congruence between the type of grip they motorically code and the size/shape of the object that visually drives them. Mirror neurons are activated both during the execution of purposeful, goal-related hand actions, and when the monkey observes similar hand actions performed by another agent. In a recent series of experiments, a proportion of mirror neurons were also found to respond not only when the monkey executes or observes an action but also when it hears that same action performed by another agent. These "audio-visual mirror neurons" do not only respond to the sound of actions, they also discriminate between the sounds of different actions (Keysers et al., 2003; Kohler et al., 2002). Neither the sight of the other agent alone nor the sight of the object alone are effective in evoking mirror neurons' responses. Similarly, mimicking the action without a target object or performing the action by using tools are much less effective in evoking responses.

In one third of the mirror neurons, the congruence between the visual and motor responses is strict, e.g. a neuron will only discharge when a precision grip is executed or observed. In two thirds of the neurons, the congruence is broad, e.g. the neuron's visual response will be elicited by observed actions similar but not identical to the executed actions it also codes for. Broadly congruent neurons appear to generalize across different ways of achieving the same goal. It should also be noted that in some cases, there is congruence in a logical or causal sense, rather than according to sameness of goal. Thus, Fogassi and Gallese (2002) give the example of a neuron that responded when the monkey observed the experimenter placing a piece of food on a tray and when the monkey grasped the same piece of food, where the two actions can be considered part of a logical sequence.

One important question concerns the source of the visual input to F5 mirror neurons. There is a striking resemblance between mirror neurons and neurons in the anterior part of the superior temporal sulcus (STSa) described by Perrett and co-workers, in particular a class of cells that are selectively responsive to hand-object interactions, such as reaching for, retrieving, manipulating, picking, tearing and holding (Perrett et al., 1989). These STS-neurons could provide a particularly well-suited source of visual input to F5 mirror neurons. One problem though is that area F5 does not receive direct projections from the STS region. However, visual information could be fed to area F5 from STSa through an intermediate step in the posterior parietal cortex, in particular area 7b or PF in the inferior parietal lobule. Indeed, researchers from the Parma group have recently found neurons with "mirror properties" in area 7b (Fogassi, Gallese, Fadiga, & Rizzolatti, 1998; Gallese, Ferrari, & Umiltà, 2002).

Single neuron recording is an invasive technique not used in humans. Several studies using different methodologies have nevertheless yielded converging evidence for the existence of a similar mirror matching system in humans (for a review, see Rizzolatti, Craighero, & Fadiga, 2002). Brain imaging experiments in humans have shown an important overlap between brain areas activating during action execution and action observation. For instance, during hand action execution and observation, a cortical network involving sectors of Broca's region, premotor cortex, STS region and posterior parietal cortex is activated (Decety et al., 1994; Grafton, Arbib, Fadiga, & Rizzolatti, 1996; Grèzes & Decety, 2001; Rizzolatti et al., 1996; Stephan et al., 1995). A recent brain imaging study (Buccino et al., 2001) also showed that when we observe goal-related behaviors executed with different effectors such as the hand, the mouth or the foot, different sectors of our premotor cortex become active. These sectors are the same that are active when we perform the same actions. In addition, several studies indicate that, unlike in monkeys, the observation of not just transitive actions (i.e. object-directed actions) but also of intransitive actions may produce activation of the motor cortex (Buccino et al., 2001; Fadiga,

Fogassi, Pavesi, & Rizzolatti, 1995; Grèzes, Costes, & Decety, 1998). Thus, although evidence for mirror mechanisms in humans is only indirect, the important overlap between areas activated during action execution and during action observation strongly suggests that many frontal and parietal areas have mirror properties.[1]

3. What are mirror neurons for?

With these neurophysiological data in place, we can turn to the following questions: What is it exactly that mirror neurons encode? What purpose do they serve? In what sense is their function different from the function of STSa cells coding face and body movements as goal-directed actions? That the answers to these questions are not completely obvious can be brought out by comparing the two competing accounts given on the one hand by Gallese (2001, 2003b) and on the other by Knoblich and Jordan (2002).

Gallese endorses what we might call a "maximalist" interpretation of mirror neurons, which consists in the following three main claims:

(1) What mirror neurons code is "the relationship in motor terms between the *agent* and the *object* of the action" (2003b: p. 173). More specifically the relationship in question is a dynamic interaction leading to a certain end-state or goal.

(2) This interaction is represented independently of the self-other distinction; in other words the interaction is represented regardless of the identity of the agent, self or other. It is, to use Gallese's phrase, "we-centric".

(3) Mirror neurons underlie the capacity for understanding others as intentional agents pursuing goals.

In contrast, Knoblich and Jordan propose a much more deflationary interpretation, which we dub "minimalist". According to the minimalist interpretation, what is coded by mirror neurons is best described as perceivable action effects. In other words, the perceived effect the observed action exerts on the object is matched to a possible effect that could be also exerted by one of the observer's own actions. Knoblich and Jordan also note that "one implication of the action effect notion is that the kind of action understanding the mirror system provides is ego-centered and does not necessarily include an explicit representation of another agent. As a consequence, organisms endowed with a mirror system may have the ability to understand

that objects are affected in a way in which they could also affect them, but they may not understand that the peer who is producing the action is an agent like themselves" (Knoblich & Jordan, 2002). We can thus sum up the minimalist interpretation by the following claims:

(1′) What mirror neurons code is "the perceived effect the action exerts on the object" (2002: p. 116).
(2′) This perceived effect is encoded in "ego-centric" terms, as a type of effect I can produce by performing a certain motor act.
(3′) The main function of mirror neurons is to increase the tendency of the agent to reproduce the same action (response facilitation).

To clarify the nature of the disagreement between maximalist and minimalist interpretations, let us introduce a distinction between the informational content and the representational content of mirror neurons. MNs can be conceived as *natural signs* having an *informational content*. They carry information about whatever in the world covaries with their activation. Now, the kind of informational content MNs have is not an object of dispute. Their activation covaries with the presence of an action in the subject's environment. More precisely, their informational content is *that an agent is interacting with an object in a specific goal-directed manner*, where the agent can be the subject herself or another subject.[2]

The notion of information content should not be confused with that of *representational content*. A natural sign may not represent anything, and if it represents something, it typically does not represent or (as we shall also say) *articulate* all the aspects of its informational content (the complex event in the world its activation covaries with). The representational content of a state is at least partly a matter of what functions it serves in the system (Dretske, 1988). Unlike informational content, representational content is directly relevant to the issue of *understanding*. When we know what MNs represent, we have a better idea of what kind of action understanding they can afford.

In this paper, we shall suppose that there are at least two criteria of representation. The first criterion is the possibility of *misrepresentation*. A state is a representation only if it is produced by a system that is also capable of misrepresenting the world. The second criterion is the possibility of *contrast*. A state represents an aspect of the world only if it is produced by a system that is also capable of producing representations of contrasting aspects. Something is always represented, at least implicitly, *as opposed to* something else.

As an illustration of how these criteria apply to MNs (assuming of course that they have a representational

---

[1] It is important to note that mirror systems understood in a wide sense involve more than just mirror neurons. They may include further mechanisms that exploit mirror neuron resources together with information from other origins. In this paper however, we will be concerned only with mirror neurons, their contents and functions. We therefore distinguish between mirror systems (broad sense) and mirror neuron systems (narrow sense).

[2] Since mirror neurons will also fire when a monkey is preparing to act on a object but prevented to do it before a certain delay, it would perhaps be more correct to say that what they encode is *that an agent is interacting or preparing to interact with an object in a goal-directed manner*.

content), consider the question of whether the *time* of the action is represented in the MN system. MNs activation covaries with currently observed, prepared or executed actions, but are the times of actions articulated in the system? It is difficult to see how the MN system could *mis*represent the time of an action. Moreover, the system cannot represent an action *as* taking place at the present time unless it can represent an action *as* taking place at other, contrasting times, in the past or in the future. Now surely we do not want to say that the representational content of MNs is *tensed*. Time is thus not part of what is represented or articulated in the mirror system.

A more controversial issue is whether the *agent* is part of the representational content of MNs. Let us examine in more detail the opposition between the maximalist and the minimalist interpretations, explicitly conceived as a dispute about the representational content of mirror neurons. According to the maximalist interpretation, mirror neurons represent particular types of agent-object relational interactions. These interactions themselves can be further described as means-end relations, where the end-state is a different state of the agent (e.g. to take possession of a piece of food, to throw away an object, to break it, or to bring it to the mouth) and where the means are specified in a precise or a more abstract way depending on whether the mirror neuron in question is strictly or broadly congruent. Furthermore, Gallese's second claim suggests that although the agent parameter is part of the representational content of mirror neurons, the identity of the agent (self or other) is not articulated. Schematically then, in Gallese's picture, the representational content of mirror neurons can be specified in the following way:

> An *unspecified agent* is performing a *motor act of type M* on a *target object*.

In contrast, according to the minimalist interpretation, mirror neurons represent perceivable action effects, do not include a representation of the other agent, and are ego-centered. There are different things this notion of ego-centeredness could refer to.

One first possibility is that the content of mirror-neurons is ego-centered simply in the sense that the actions they can represent must be part of the agent's own motor repertoire. This reading of ego-centeredness is not especially contentious and, presumably, Gallese would not take issue with it.

As a second possibility, it may be that MNs are ego-centered in the sense that their function is not to understand the actions of others but to steer the agent's own behavior. This is a claim explicitly made by Knoblich and Jordan. Now, they do not distinguish between two explanations of the fact that mirror neurons serve this function. On one explanation the agent parameter is part of the representational content of mirror neurons and its value is always the self:

> There is an action effect of type E which I [the self] could achieve by performing a motor act of type M.

However, given our criteria for what counts as a representation, the claim that the self is represented in the contents of MNs is implausible. It could only be sustained if it were possible, at that level of representation, to *contrast* either self with non-self or self with others. In other words, MNs cannot represent the self as such unless they are also capable of producing representations with a content of the form "This (other) agent is doing F", or "Non-self is doing F". Knoblich and Jordan themselves acknowledge that another agent cannot be represented as such within the MN system, and they say nothing about the availability of a self/non-self contrast.

Another more plausible explanation is available. For a representation to steer the behavior of an agent, it is not required that this representation include an explicit reference to her. To borrow a distinction from Perry (1993), the representation might *concern* the agent without being *about* her. When the representation merely concerns the agent, it is self-relative without having any component referring to the self. Perry claims that this is the case with perceptual representations:

> What each of us gets from perception may be regarded as information concerning ourselves, to explain connections between perception and action. There is no need for a self-referring component of our belief, no need for an idea or representation of ourselves. When a ball comes at me, I duck; when a milk shake is put in front of me, I advance. The eyes that see and the torso or legs that move are parts of the same more or less integrated body. And this fact, external to the belief, supplies the needed coordination (1993: p. 219).

A perceptual representation with the content "This ball is approaching" might cause the agent to duck even though it is not self-referential. The representation concerns the agent in virtue of facts about her functional architecture, in particular the fact that local perception is causally wired to local action. Similarly, the representation "It's raining" is typically not about the local weather, but merely concerns it. What makes it concern the local weather is the fact that the agent only uses the representation to act on the local weather, for instance by opening her umbrella. One reason to consider these representations as self-relative rather than self-referential has to do with the contrast criterion. At the relevant levels of representation, the subject and the place of the represented rain are not represented as opposed to other, real or possible subjects and places.

If we are right, the representational content of MNs is best specified in an agent-neutral way, for instance as follows:

> There is an *effect of type E* achieved by performing a *motor act of type M*.

Mirror neurons with this content can be self-relative but they need not be. Whether they are self-relative or not depends on what they are used for. They are certainly self-relative when they are used to steer the agent's own behavior. Thus, a representation can be ego-centered in the sense that it steers the behavior of the agent to which it belongs (self-relativity) without being ego-centered in the more demanding sense of involving an explicit representation of the agent as self (self-reference). We shall call representations that do not involve an explicit reference to an agent "a-centered". A-centered representations, whether self-relative or not, are representationally indifferent to the distinction between self and others.

According to Gallese (2001, pp. 40-2), the MN system may originally have developed to achieve a better control of action performance. Canonical neurons code a motor program for acting on a target object, and a copy of this motor program signal is fed to mirror neurons. The motor activity of mirror neurons acts as a sort of "simulator" of the programmed action, used to predict its consequences and thus achieve more precise control of the motor performance. Then, the coupling of the MN system with visual input may have allowed it to be used for different purposes, namely to represent other individuals' actions. Gallese seems to be committed to the claim that the representational content of MNs was different when they were first used as efference copies – we presume that they could not be we-centric at that stage. In contrast, we claim that the representational content of MNs is the same at both stages, namely agent-neutral or a-centered.

In a nutshell, a representation can be indifferent to the self/other distinction without being we-centric in Gallese's sense. At first blush, indifference to the self/other distinction may be more economically explained in terms of a-centric representations. Gallese therefore needs further reasons to maintain that the representational content of MNs is we-centric. We suspect that his main reason for this claim is his view that the MN system grounds our understanding of others as agents pursuing goals. For instance, Gallese (2001, p. 35) claims that purely visual representations of a behavior (representations in the STS) do not allow for its coding/representing as an *intended*, mind-driven behavior. We agree with this claim concerning visual representations, but, as we shall argue in Section 4.2, we think that the same goes for motor representations in the mirror system.

4. Four levels of action understanding

In the light of the above discussion and of further elements to be introduced shortly, it appears necessary to distinguish at least four levels of action understanding. Let us note at the outset that understanding an action in the sense of having a neural representation of an action should be distinguished from understanding an action in the sense of having a conscious, personal-level representation of an action. The former kind of understanding can

be called implicit understanding and the latter explicit understanding. Here we will mostly be interested in implicit forms and levels of action understanding, although the fourth level of action understanding we consider is arguably one that takes place at the personal level. The four levels we want to consider are visual understanding, motor understanding, agentive understanding and meta-representational understanding. We now examine each level in turn.

4.1. Visual understanding

As shown by Perrett and his co-workers (for a review see Perrett, 1999), several types of neurons in the STS region of the visual system of monkeys are relevant to understanding action. One type of cell encodes the visual appearance of the face and body while they are static or in motion. A second type of cell in the STS is selectively responsive to particular types of body movement but remains inactive to static images of the body. A third type of cell responds specifically to face and body movements as goal directed actions. As Perrett (1999) points out, the coding achieved by the first two types of cells can be understood with sole reference to the visual appearance and movement of a body or face within the visual image. Actions, however, are more than just movements in that they involve goal-directed behavior. One way visual information can be used to specify the goal-directed nature of movement is by relating body movements to other aspects of the visual environment. Movement coding thus becomes action coding. This type of coding is precisely what cells of the third type achieve.

For instance, cells in the ventral bank of the STS are selectively responsive to hand actions, with different sub-populations of these cells showing selectivity for specific hand-object interactions, such as reaching for, retrieving, manipulating, holding or tearing (Perrett et al., 1989). These cells respond when both the appropriate hand movements and the appropriate movements of the object acted upon are visible. Furthermore, the hand movements and the object movements must be appropriately related. For instance cells responsive to object manipulation do not respond to the sight of the hand movements combined with object movements when the hand and object are spatially separated by a few centimeters (Perrett et al., 1989). According to Perrett, this indicates that the cells are sensitive to the causal relation (in the form of spatial contiguity) between effector and object acted upon. One further important property of these cells is that they generalize across several different instances of the same action, including the sight of the action from different perspectives, at different distances or speeds, performed on different objects, performed by the monkey's own hand as well as by other hands.

This makes it possible to claim that actions and not just movements are represented and classified by cell populations in the STS. Importantly, goal-directedness, the

distinguishing feature of actions as opposed to mere movements, is represented in terms of specific spatial relations between biological movements (of an agent) and other aspects of the visual environment (an object, another individual or a place in the environment). As emphasized by Perrett (1999), the kind of action understanding made possible by the visual coding of actions achieved by the STS is behavioral rather than mentalistic. It makes possible the formation of learned associations between visually coded actions and their visual consequences. Exploitation of these learned associations makes it possible to predict the visual consequences of the seen action.

4.2. Motor understanding

The main difference between the kind of visual understanding instantiated by STS-neurons and the motor understanding found at the level of MNs is that the action (movement + goal) that was encoded in STS in visuo-spatial terms is now motorically encoded; i.e. encoded in terms of the motor commands necessary to achieve the goal. For instance, whereas hand movements are represented in STS in terms of spatial configurations and spatial transitions (rotations, translations) among them, they are represented by MNs in terms of specific motor programs.

What are the benefits of a motor understanding of actions? As we have seen, according to the hypothetical scenario proposed by Gallese (Fogassi & Gallese, 2002; Gallese, 2001), the motor activity of mirror neurons could be interpreted as the result of an efference copy signal of the motor program encoded by canonical neurons. This system may originally have developed to achieve better control of action performance, and its later coupling with visual input may have allowed it to be used to represent others' actions. If the system initially developed to predict the consequences of the agent's own actions, one immediate benefit of its being coupled with visual input would be to allow it to predict the consequences of observed actions on the basis of its "knowledge" of the consequences of its own actions.

As suggested by Knoblich and Jordan, another benefit of the motor understanding of observed actions and their consequences would be to help the observer react appropriately. In some instances, the appropriate reaction would be to perform a similar action. For instance prepare to grasp for a piece of food when observing another agent doing the same thing. The activation of mirror neurons during action observation would therefore serve a motor function. Rizzolatti et al. (2001) dismiss this interpretation as unlikely, pointing that "when the monkey observes another monkey grasping a piece of food, the obvious action to take would be for instance to approach the other monkey, but certainly not to repeat the observed action" (2001, p. 667). Of course, if "repeating the observed action" is taken in a very narrow sense, performing the very same action on the very same object, they are right. But if it is taken in a broader sense, producing an action of a similar type, it could make perfect sense at least in ecological conditions. Peanuts may

become available only one at a time in a laboratory setting, but when monkeys are foraging in the wild, it would make sense for a monkey observing another monkey picking a berry to prepare for the same action as presumably other berries are to be found in the vicinity. Similarly, the kind of broad congruence of MNs that was described in Section 2 as logical or causal congruence suggests that the motor understanding of observed actions may have as one of its functions to prepare an appropriate response. In the example given by Fogassi and Gallese (2002), to prepare to grasp a piece of food, when observing an experimenter placing it on a tray.

4.3. Agentive understanding

To illustrate the contrast between agentive and both perceptual and motor understanding, consider the following passage by Tomasello (1999):

Nonhuman primates see a conspecific moving toward food and may infer, based on past experience, what is likely to happen next, and they may even use intelligent and insightful strategies to affect what happens next. But human beings see something different. They see a conspecific as trying to obtain the food as a goal, and they can attempt to affect this and other intentional and mental states, not just behaviour. (1999, p. 21)

Visual understanding of action can be described as the understanding of specific spatio-temporal relations among external events, where the antecedent event is one of biological motion. Motor understanding of actions is understanding of actions as motor procedures that produce certain effects. Visual and motor representations of actions may thus at best underlie an understanding of agents as self-propelled beings that make things happen. But an agentive understanding of action is something different. It involves understanding the more complex relationships between an agent with a goal, the instrumental means used and the effects produced. In effect, it involves moving from intentionality-in-the-world to agent-intentionality, from goals as relational structures in the world to goals as intentional relations between agents and the world.

We claim that this kind of understanding requires mechanisms and representational skills beyond those found at the level of the MN system. One line of evidence for this claim is the fact that, as the majority of primatologists now seem to agree, agentive understanding of actions and understanding of others as intentional agents is unique to humans (Povinelli, 2001; Povinelli, Perilloux, Reaux, & Beirschwale, 1998; Premack, 1986; Tomasello, 1999). For instance, chimpanzees do not appear able to discriminate between intentional and non-intentional actions. Since, however, MN systems are not unique to humans but are also found in apes and monkeys, it is difficult to maintain that they provide a sufficient basis for agentive understanding.

According to Tomasello (1999), what non-human primates lack is the ability to understand goals and behavioral means as separable in the actions of others. In the absence of this ability, they "focus on the changes of state (including changes of spatial position) of the objects involved during the demonstration [of actions involving tool use], with the actions of the demonstrator being in effect just other physical motions. The intentional state of the demonstrator, and thus her behavioral methods as distinct behavioral entities, are simply not part of their experience" (1999, pp. 30–31). Indeed, the way MNs encode actions is by representing the goal in terms of the motoric means used to achieve it. Thus, at this level goals and means cannot be represented separately.

A genuine agentive understanding of action requires a representational redescription of the action, where goals and means are represented separately, thus making it possible to represent agents as directed towards goals independently of the means employed and to understand the intentional relation of the agent to the goal as the mediating force that explains the means-end sequence. Once the agent parameter is articulated as a separate component of the representation of the action, further attribution mechanisms can come into play to determine the value of the agent parameter (self vs. non-self or self vs. other).[3]

4.4. Meta-representational understanding

By meta-representational understanding, we mean the kind of understanding of an action that relates it not just to the immediate intention or goal the agent is striving to achieve, but to his or her reasons for pursuing that goal. This kind of understanding of action involves attributing to agents beliefs and desires that rationalize their intentions and actions. It therefore requires the possession of concepts of beliefs, desires and other mental states. These meta-representational skills and the kind of understanding of action they make possible are unique to humans.

We call this level "meta-representational" not because we think that the level of agentive understanding does not involve representations of mental states, namely intentions, but because the fourth level involves representations that are more complex in at least two ways. First, it may be argued that the mental states that are represented at the third level are relational, i.e. they involve relations of an agent to a goal in the world, whereas the meta-representational level is concerned with propositional attitudes that cannot always be construed as relations to the world. Second, at the meta-representational level of understanding, the focus is on the relations between intentions and other mental states that rationalize them. Indeed, much recent developmental literature suggests that the child masters an agentive understanding of actions much before he develops a full-blown

meta-representational understanding of them (Bartsch & Wellmann, 1995; Malle, Moses, & Baldwin, 2001).

5. Joint action: classical analyses and their limitations

In the broadest sense, the label "collective actions" can be used to cover all cases where a certain global effect is the result of the actions of several individuals. Creating a traffic jam is a collective action in this sense, as one single motorist does not suffice to create a traffic jam. Note, though, that this is a collective action in the weakest possible sense, since it is not even required that the participating agents share a goal, much less that they cooperate to achieve that goal. At the other end of the spectrum of collective actions are joint cooperative actions. An instance of joint cooperative action would be jazz musicians involved in a jam session. Here, the participating agents share the same goal, creating a new version of a jazz standard; they act towards that shared goal by playing their parts and they cooperate by adjusting their individual performance to the performance of other members of the band to achieve the common goal. There are various intermediate cases between collective action in the weak sense described above and joint cooperative actions. Here we will focus on the latter kind of actions, joint actions for short.

To help bring to the fore what is distinctive of joint actions, let us consider the following example by Searle:

> Imagine that a group of people are sitting on the grass in various places in a park. Imagine that it suddenly starts to rain and they all get up and run to a common, centrally located shelter. Each person has the intention expressed by the sentence "I am running to the shelter". But for each person, we may suppose that his or her intention is entirely independent of the intentions and behaviour of others. In this case there is no collective behaviour; there is just a sequence of individual acts that happens to converge on a common goal. Now imagine a case where a group of people in a park converge on a common point as a piece of collective behaviour. Imagine that they are part of an outdoor ballet where the choreography calls for the entire corps de ballet to converge on a common point. We can imagine that the external bodily movements are indistinguishable in the two cases; the people running to the shelter make the same types of bodily movements as the ballet dancers. Externally observed, the two cases are indistinguishable, but they are clearly internally different. (Searle, 1990, pp. 402–403)

The problem then is to specify what this internal difference consists in. One important clue is that in the first situation the convergence on a common goal is a mere accident. The intention and behavior of each individual are entirely independent of the intentions and behavior of others. Moreover, as Searle remarks, this is so even if each person knows that the other people intend to run to the

---

[3] For discussions of the possible forms attribution mechanisms could take, see Georgieff and Jeannerod (1998) and Jeannerod and Pacherie (2004).

shelter and knows that the other people know that he or she intends to run to the shelter. Searle claims that in the second situation, by contrast, the individual "I intentions" are derivative from "we-intentions".

A second reason for thinking that collective intentional behavior is not the same as the summation of individual intentional behavior is that "often the derived form of an individual intention will have a different content from the collective intention from which it is derived" (Searle, 1990, p. 403). The example Searle gives is that of a football team trying to execute a pass play. No individual member of the team can have "we are executing a pass play" as the entire content of his intention, for no one can execute a pass play by himself. Each player must make a specific contribution to the overall goal.

Saying that what is distinctive of joint actions is that they involve we-intentions does little more than give the problem a name, as long as one does not specify what we-intentions consist in. We will here consider two analyses of we-intentions, proposed respectively by Searle (1990) and Bratman (1992) and point out some of their shortcomings.

According to Searle, we-intentions cannot be analyzed into sets of I-intentions, even supplemented with beliefs, including mutual beliefs about the intentions of other members of a group. Analyses that attempt such a reduction fail to account for the cooperative and coordinated character of joint actions. Searle shows through counter-examples that the existence of mutual beliefs among members of a group is not sufficient to ensure cooperation. Thus, business school graduates who have been exposed to Adam Smith's theory of the hidden hand may come to believe that the best way for somebody to help humanity is by pursuing his own selfish interests. Each may form a separate intention to thus help humanity by pursuing his own selfish interests and not cooperating with anybody and they may all have mutual beliefs to the effect that each has such an intention. In such a case, despite all the businessmen having the same goal as well as mutual beliefs about their respective intentions, there is no cooperation and no collective action. What they lack is an intention to cooperate mutually. Mutual beliefs among members of a group do not ensure the presence of such an intention.

According to Searle, this cooperative dimension of collective actions can be captured only if it is accepted that the intentions attributable to the individuals that take part in collective actions are different in type from the intentions attributable to those same individuals when they engage in individual actions. The idea then is that to account for cooperation we have to introduce a specific type of mental states: we-intentions. What needs to be spelled out is the sense in which we-intentions are special and, relatedly, the sense in which they can be said to imply cooperation. Prima facie, there are three possibilities. The first is that what makes we-intentions special has to do with the type of entities they can be attributed to. The second is that what makes them special are features of their contents

and thus that the dimension of cooperation is linked to specific features of these contents. Finally the third possibility is that rather than the contents or the possible bearers of we-intentions, it is the psychological mode itself – i.e., the fact that the psychological mode is that of we-intending instead of I-intending – that implies the notion of cooperation. Before we examine each possibility in turn, let us note that they are not mutually incompatible: we-intentions could in principle be special in all three ways.

Searle rejects the first option on the ground that it would force one to admit the existence of some forms of primitive collective entities, an ontological commitment he sees as unreasonable. He also rejects the second option, claiming that the content of we-intentions is of a form already present in some complex cases of singular intentions, the content of which encompasses a by-means-of relation. The idea is that in the case of singular intention of, e.g., firing a gun by pulling the trigger, there is only one intention and one action, with the relation of the means-intention to the overall intention being only part-whole. Similarly, for Searle, in the case of collective actions, there is only one complex: the singular intentions of the participating agents are related to the collective intention as means to ends and this relation is simply part-whole. To borrow Searle's example of two cooks, say Paul and Gilbert, preparing a hollandaise sauce together, the content of Paul's we-intention would be something like "that we make the sauce by means of me steering" and the content of Gilbert's intention could be rendered as "that we make the sauce by means of me pouring". It is important to note that there is nothing in the by-means-of relation per se that implies cooperation. For instance, I can intend that we go to the police station by means of me dragging you, and clearly in such a case no cooperation need be involved. Thus, there is nothing in the analysis Searle offers of the form of the content of collective intentions that makes it necessary that the dimension of cooperation essential to collective intentions be reflected in their contents.

The option Searle favors is the third one, namely, that what makes we-intentions special is the psychological mode itself, not the possible subjects of we-intentions nor their contents. More specifically, what is special about we-intentions is that they are mental states that "make reference to collectives where the reference to the collective lies outside the bracket that specifies the propositional content of the intentional state" (Searle, 1990, p. 408).

One problematic feature in Searle's analysis has to do with the dimension of cooperation that he says is essential to joint action. Nothing in the structure of the content of we-intentions as laid out by Searle seems to capture the notion of cooperation. As we have already noted, there is nothing in the by-means-of relation per se that implies cooperation. According to Searle, in order to account for the cooperative character of we-intentions, we must appeal to Background capacities. What collective intentionality presupposes is "a Background sense of the other as a candidate for cooperative agency; that is, it presupposes a

sense of others as more than mere conscious agents, indeed as actual or potential members of a cooperative activity'' (Searle, 1990, p. 414). Background capacities, according to Searle, are not themselves representational. Rather, they are a set of non-intentional or preintentional capacities that enable intentional states to function. In other words, they are biological or neurophysiological phenomena rather than intentional phenomena. Yet, by drawing a line between the realm of the intentional and the Background and by considering that the dimension of cooperation is part of the Background, Searle acknowledges that he cannot account for it in intentional terms.

While there is no denying that all our cognitive capacities rest on an underlying basis of biological capacities, it seems both unduly hasty to sweep cooperation under the rug of Background presuppositions and controversial to claim that the required background capacities are non-representational. Although a capacity for cooperation may depend in part on certain background biological capacities, we shall try to show later that these background capacities can be spelled out in part in representational terms.

But first let us consider Bratman's attempt at capturing the cooperative dimension of joint action. Bratman agrees with Searle that collective intentional behaviour is not analyzable as just the summation of individual intentional behavior even supplemented with mutual beliefs or mutual knowledge. Yet, contrary to Searle, he thinks that collective intentionality is not a primitive phenomenon and that a non-circular reductive analysis may be possible. Bratman (1992) identifies three features of joint actions (which he calls shared cooperative activities) that such an analysis would have to account for. The first feature is a commitment to a joint activity (e.g. we paint the house together). The second is the mutual responsiveness of each participating agent to the intentions and actions of the others. The third is a commitment to mutual support by which each agent is committed to supporting the efforts of the other to play her role in the joint activity. None of these features is by itself sufficient to make an activity a joint action, but, according to Bratman, taken together these three features are characteristic of joint actions.

Bratman is careful to ensure that his analysis is non-circular. In this regard the notion of a commitment to a joint activity may seem suspect. However, Bratman makes it clear that in his analysis the notion of a joint activity should be understood in a cooperatively neutral way. Second, one may wonder whether it makes sense to appeal to my intention that we do something together, insofar as it may be thought that one can only intend one's own actions. Here, Bratman introduces a distinction between intending and attempting and defends a planning conception of intentions that emphasizes the role of future-directed intentions as elements of partial plans. This conception of intentions allows him to be more liberal about what can be intended than about what can be attempted since references to things other than our own actions can function appropriately in our plans.

Since Bratman construes commitment to a joint activity in a cooperatively neutral way, this commitment does not suffice to ensure that the activity that follows is a shared cooperative activity or a collective action in Searle's sense. The originality of Bratman's analysis comes from the way in which he construes the two further features of mutual responsiveness and commitment to mutual support. These are analyzed in terms of meshing subplans and interdependent intentions. For an activity to be a shared cooperative activity, it must be the case that each agent intends that the group performs this joint action in accordance with subplans that mesh, where for subplans to mesh it is not necessary that there be full agreement in the agents' subplans, but merely that there be some way the action can be done that would involve the successful execution of those individual subplans. This meshing of subplans in turn implies that the intentions of the participating agents must be interlocking. Each agent should bring into the content of his intention the efficacy of the other participants' intentions as well as the efficacy of his own intentions. By thus requiring that the intentions of the participating agents be interlocking, Bratman moves away from the classical reductive analyses of collective action according to which the crucial link among the attitudes of the participating agents is simply cognitive. Mutual belief or mutual knowledge are not sufficient to ensure that an intention is shared or collective; but neither is it necessary to posit collective intentionality as a primitive form of intentionality. What is crucial is the specific form of interdependence of the individual intentions. Finally, insofar as the intention of each participant includes his intending that the relevant intentions of the others be successfully executed, commitment to mutual support is also involved in shared cooperative activities. As Bratman phrases it, the intentions of the participants should be minimally cooperatively stable. In other words, there must be at least some circumstances in which each participant would be prepared to help the others do their part in the joint activity.

Bratman's analysis may be judged more satisfactory than Searle's. It does not lay open to the criticism Searle addresses to traditional reductive analyses of collective intentionality since it does not maintain that the crucial link among the attitudes of agents involved in collective behaviour is a purely cognitive link. It brings to the fore three essential features of shared intentions that are not taken into account by traditional analyses. Finally, contrary to Searle, it tries to capture what is distinctive of shared intentions in terms of a special kind of interdependence among the individual intentions of the participants, rather than by postulating a mysterious form of primitive we-intentionality supposed to be in the head of individual agents. Thus, according to Bratman, the contents of the intentions of individuals involved in collective action are special insofar as they make reference to the intentions of the other participants – each agent must have intentions in favour of the efficacy of the intentions of the others – and are self-reflexive as well – each agent must have intentions concerning the efficacy of his or her own intentions.

Despite its illuminating character, Bratman's analysis is not totally unproblematic. The characterization he proposes seems tailored to forms of joint actions that are already rather sophisticated. His analysis focuses on future-directed intentions that involve rational deliberation and conscious planning. It also requires that the participating agents have the concept of an intention, since the contents of their intentions make reference to the intentions of the other participants. One may therefore wonder how more elementary forms of joint action, that may not involve prior rational deliberation and conscious planning, and that may indeed be performed by creatures that do not possess meta-representational capacities and do not have the concept of an intention, could be accounted for.

Our purpose in the last section will therefore be to explore the contrast between the sophisticated joint actions that are the target of Bratman's analysis and more elementary forms of joint actions that may not require the complex cognitive machinery implied by his analysis. We will try to characterize the role MN representations may play in an account of what makes elementary or sophisticated joint actions possible. We think, contrary to Searle, that the Background underlying joint action can be characterized at least in part in representational terms. We will try to show that the cooperative features of shared intentions have a partial analogue at the subpersonal level.

6. What MNs can and cannot explain about joint action

Various species of social mammals engage in types of joint behaviour that involve some form of cooperation and coordination among participants. A well-known example is group-hunting in lions, wolves or chimpanzees. Although human beings are clearly capable of more sophisticated forms of joint activity, they also engage in unreflective joint actions that may share many of the features of other mammals' collaborative endeavors.

It therefore seems useful to draw a distinction between elementary or thin forms of joint action common to humans and other social mammals and sophisticated or thick forms of joint action, perhaps unique to the human species. We suggest that this distinction can be understood as a distinction between forms of joint action that require only a capacity for a perceptual and a motor understanding of action and forms of joint actions that also require a capacity for an agentive and a meta-representational understanding of action. We also want to suggest that actions of the former type involve cooperation without representation, in the sense that the participating agents need not represent the activity they take part in as a cooperative venture, hence need not represent the other participants' intentions and the relations these intentions bear to their own intentions. In contrast, joint actions of the latter type require the interlocking complexes of intentions described by Bratman. Finally, we think the cooperative features exhibited by thin joint actions may be explained in terms of subpersonal mechanisms of coordination in a way that
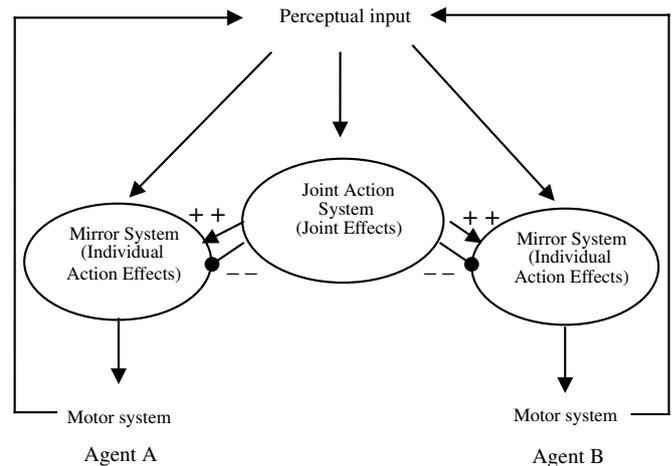

Fig. 1. The model of action coordination proposed by Knoblich and Jordan (2002).

gives MN representations a central role, while the role MN representations play in thick forms of joint actions may be more marginal.[4]

To illustrate the way in which action coordination may be achieved through the use of simple perceptual and motor representations, let us consider a model proposed by Knoblich and Jordan (2002). They argue that although the simple kind of perception-action match provided by MNs will not allow just by itself an individual to successfully coordinate his or her actions with those of others, it need only be supplemented with a representation of the joint action effects and a simple coordination mechanism to achieve this purpose. This additional machinery will make it possible for the individual to modulate his or her own actions in response to perceiving the effects of the others' actions. Knoblich and Jordan's model requires separate representations of joint action effects and individual action effects. As shown in Fig. 1, the joint and individual action effects are coupled by excitatory or inhibitory connections. These connections ensure that if there is an action of one's own which can increase the joint action effect, it will be triggered irrespective of who is responsible for the joint effect. Conversely, one's own action is inhibited if it hinders the production of an intended joint effect.

In this model, joint intentions at the level of the joint action system are represented in purely perceptual terms

---

[4] Note that the majority of human actions are probably hybrids of thin and thick actions with some aspects of the cooperative endeavor being explicitly represented – generally, the more global ones –, and others supported by subpersonal mechanisms of coordination. Another important characteristic of human joint actions is their heavy reliance on conventions and coordination devices, i.e. artefacts whose main function is to facilitate coordination. One clear example would be the set of conventions known as the Highway Code and the various artefacts (traffic lights, road signs, road markings, etc.) that help drivers coordinate their actions, while limiting the cognitive load involved in this coordination game. It is presumably simpler to stop when the traffic light is red than to try and figure out what the intentions of the drivers coming from other directions are.

and do not intrinsically exploit the MN system. The MN system is used to represent the individual actions of the participating agents. Therefore only perceptual and motor understanding of action are involved. Cooperation, in the form of action coordination, is achieved without being represented or intended as such. There is no commitment to a joint activity in the sense of a shared intention that we, say, catch the prey together, but simply a goal or intended action effect that agents happen to share, say, that the prey be caught. The mutual responsiveness and mutual support characteristic of thick joint action here simply translate as the modulation of one's own actions in response to perceiving the effects of a peer's action.

Such a model has serious limitations and cannot be a model of thick joint actions. First, the fact that there is a joint action is not essential to the model. For all it says, the joint effects might be the result of the agent's action (e.g. sailing) together with non-agentive forces (e.g. the wind). Second, the essential cooperative dimension of thick joint action is not present. As the join action unfolds, what each participant agent does is simply assess the situation – joint action effects – and decide what to do next. Third, the complementarity of the individual actions of participating agents reduces to either continue carrying out one's action if this increases the desired joint effect (excitatory connection) or stop carrying out one's action if it decreases the probability of the desired joint effect (inhibitory connection). In contrast to thin joint actions, thick or sophisticated joint actions require representations not just of actions and their effects, but also of agents and their intentions. Their three main features – commitment to a joint activity, mutual responsiveness and mutual support – require that subjects form representations that make explicit reference to themselves as well as to other participating agents. Thus, we must both intend that *we* J together, that *we* J by means of *me* A-ing and *you* B-ing, that *I* help *you* B-ing and *you* help *me* A-ing if necessary. The reflexivity and interlocking character of shared intentions require that each agent bring into the content of his or her own intentions the efficacy of the other participants' intentions.

Obviously, MN representations are not up to such a representational challenge. As we argued in Section 3, they do not involve an explicit agent-parameter – they are a-centered – and they do not represent intentions but motor acts and their effects. But to say that MN representations do not provide a sufficient basis for a capacity for thick joint actions is not to say that they are useless. The a-centeredness of MN representations, hence their agent-neutrality, may be a limitation but is may also be considered a strength in some respects. First, the agent-neutrality of MN representations makes it possible to use them as the building blocks of the representation of the joint action sequence and thus to test its coherence while abstracting away from the division of labor among participants. Second, the coherence of the whole instrumental sequence is not represented once and for all; it is actively maintained in the course of the joint action. As Pierre Livet (1994)

has pointed out, to engage in a joint action is to be prepared to correct mistakes made by other participants, hence to allow for modifications of our own actions in response to others' errors or failed actions. A system producing agent-neutral motor representations, and producing them online, might be useful at this point. Each agent can represent the instrumental sequence in a way that allows for an adjustment both of the sequence itself and of the division of labor among participants. The two functions Gallese assigns to MNs – better control of one's own actions and understanding of the actions of others – can give rise to a third function: joint action control, where each agent adjusts his own actions as a function of the common goal and of the predicted consequences of the actions of other participants.

7. Conclusion

In this paper, we defended two main claims. Our first claim is that a maximalist interpretation of the representational content of mirror neurons is not warranted. Mirror neurons do not represent agents as such and cannot directly ground an agentive understanding of action. Our second, more positive, claim is that MNs can nevertheless play an important role in the explanation of joint actions. First, MN representations can help explain how creatures lacking both an agentive and a meta-representational understanding of actions may nevertheless engage in thin forms of joint actions. Second, a minimalist interpretation of the representational content of mirror neurons, far from limiting their usefulness, can yield a better account of the production and understanding of joint action. In particular, the agent-neutrality or a-centeredness of mirror neurons representations makes them plausible constituents of the representation of a joint action sequence and useful elements in a process of online mutual adjustment of participants' actions.

Acknowledgements

References

Arbib, M. A. (2002). The mirror system, imitation, and the evolution of language. In C. Nehaniv & K. Dautenhahn (Eds.), *Imitation in animals and artifacts*. Cambridge, MA: MIT Press.

Arbib, M. A. (2003). The evolving mirror system: a neural basis for language readiness. In M. H. Christiansen & S. Kirby (Eds.), *Language evolution: The states of the art*. Oxford: Oxford University Press.

Bartsch, K., & Wellmann, H. M. (1995). *Children talk about the mind*. Oxford: Oxford University Press.

Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review, 101*(2), 327–341.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience, 13*, 400–404.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., et al. (1994). Mapping motor representations with PET. *Nature, 371*, 600–602.

Dretske, F. (1988). *Explaining behavior*. Cambridge, MA: MIT Press.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology, 73*, 2608–2611.

Fogassi, L., & Gallese, V. (2002). The neural correlates of action understanding in non-human primates. In M. I. Stamenov & V. Gallese (Eds.), *Mirror neurons and the evolution of brain and language* (pp. 13–35). Amsterdam: John Benjamins.

Fogassi, L., Gallese, V., Fadiga, L., & Rizzolatti, G. (1998). Neurons responding to the sight of goal directed hand/arm actions in the parietal area PF (7b) of the macaque monkey. *Society of Neuroscience Abstracts, 24*, 257.

Gallese, V. (2001). The "Shared Manifold" Hypothesis: from mirror neurons to empathy. *Journal of Consciousness Studies, 8*(5–7), 33–50.

Gallese, V. (2003a). The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical Transactions of the Royal Society of London B, 358*, 517–528.

Gallese, V. (2003b). The roots of empathy: The shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology, 36*(4), 171–180.

Gallese, V., Ferrari, P. F., & Umiltà, M. A. (2002). The mirror matching system: a shared manifold for intersubjectivity. *Behavioral Brain Sciences, 25*, 35–36.

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

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593–609.

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

Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Experimental Brain Research, 112*, 103–111.

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

Grèzes, J., Costes, N., & Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: a PET investigation. *Cognitive Neuropsychology, 15*, 553–582.

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

Keysers, C., Kohler, E., Umiltà, M. A., Fogassi, L., Nanetti, L., & Gallese, V. (2003). Audio-visual mirror neurones and action recognition. *Experimental Brain Research, 153*, 628–636.

Knoblich, G., & Jordan, J. S. (2002). The mirror system and joint action. In M. I. Stamenov & V. Gallese (Eds.), *Mirror neurons and the evolution of brain and language* (pp. 115–124). Amsterdam: John Benjamins.

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.

Livet, P. (1994). La communauté virtuelle. Combas: Editions de l'Eclat.

Malle, B. F., Moses, L. J., & Baldwin, D. A. (Eds.). (2001). *Intentions and Intentionality: Foundations of social cognition*. Cambridge, MA: MIT Press.

Perrett, D. I. (1999). A cellular basis for reading minds from faces and actions. In M. Hauser & M. Konishi (Eds.), *Neural mechanisms of communication*. Cambridge, MA: MIT Press.

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., et al. (1989). Frameworks of analysis for the neural representations of animate objects and actions. *Journal of Experimental Biology, 146*, 87–114.

Perry, J. (1993). *The problem of the essential indexical*. Oxford: Oxford University Press.

Povinelli, D. J. (2001). On the possibilities of detecting intentions prior to understanding them. In B. F. Malle, L. J. Mosas, & D. A. Baldwin (Eds.), *Intentions and intentionality. Foundations of social cognition* (pp. 225–248). Cambridge, MA: MIT Press.

Povinelli, D. J., Perilloux, H. K., Reaux, J. E., & Beirschwale, D. T. (1998). Young and juvenile chimpanzees (*Pan troglodytes*) reactions to intentional versus accidental and inadvertent actions. *Behavioral Processes, 42*, 205–218.

Premack, D. (1986). *Gavagai!* Cambridge, MA: MIT Press.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences, 21*, 188–194.

Rizzolatti, G., Carmada, R., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of Area 6 in the macaque monkey. II Area F5 and the control of distal movements. *Experimental Brain Research, 71*, 491–507.

Rizzolatti, G., Craighero, L., & Fadiga, L. (2002). The mirror system in humans. In M. I. Stamenov & V. Gallese (Eds.), *Mirror neurons and the evolution of brain and language* (pp. 37–59). Amsterdam: John Benjamins.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., et al. (1996). Localization of grasp representations in humans by PET. 1. Observation versus execution. *Experimental Brain Research, 111*, 246–252.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2*, 661–670.

Searle, J. R. (1990). Collective intentions and actions. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication*. Cambridge, MA: MIT Press.

Stephan, K. M., Fink, G. R., Passingham, R. E., Silbersweig, D., Ceballos-Baumann, A. O., Frith, C. D., et al. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *Journal of Neurophysiology, 73*, 373–386.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Wohlschlager, A., & Bekkering, H. (2002). Is human imitation based on a mirror-neuron system? Some behavioural evidence. *Experimental Brain Research, 143*(3), 335–341.