

Action mirroring and action understanding: an alternative account

Gergely Csibra

Observed actions elicit covert motor activations in observers that, in case they were executed, would generate similar actions to the observed ones. I challenge the most popular explanation offered for these phenomena, according to which such action mirroring is generated by direct matching and serves the function of action understanding in terms of their goals. I propose that action mirroring is generated by action reconstruction via top-down emulation from action interpretation produced outside the motor system. Such action mirroring does not follow but anticipates ongoing actions and enables, beyond predictive tracking, action coordination with others. I argue that the available empirical evidence is more compatible with this alternative model than with the direct-matching account.

Action mirroring

Plenty of evidence shows that when human and non-human observers watch (or listen to) others' actions, their own motor system also tends to be activated. In addition, the motor activation induced by action observation¹ often roughly corresponds to the motor program that the observer would have to execute to perform the observed action. These phenomena, which I shall term collectively *action mirroring*, can be demonstrated in many ways. Behavioral methods reveal motor priming and motor interference effects (e.g. Brass *et al.*, 2001), neurophysiological measurements show covert muscle and motor neuron excitation (e.g. Fadiga *et al.*, 1995), and neuroimaging studies indicate automatic activation of motor and premotor areas of the cerebral cortex (e.g. Buccino *et al.*, 2001) upon action observation. But perhaps the clearest evidence of action mirroring comes from single-cell studies in monkeys demonstrating that a subset of premotor and parietal neurons, called *mirror neurons*, discharge both when the animal executes a certain motor act and when it perceives the same act performed by others (Rizzolatti and Craighero, 2004).

¹ Throughout this paper, I discuss only examples of *visual* observation of actions, but my argument can easily be extended to other modalities, such as audition.

To explain such a rich set of phenomena, one has to specify what functions they serve and what mechanisms these functions are achieved by. Action mirroring has been suggested to subserve imitation (Iacoboni *et al.*, 1999; Rizzolatti and Craighero, 2004; Iacoboni, 2005), intersubjectivity (Gallese, 2003), and empathy (Iacoboni, 2005). However, the more basic, and evolutionarily more ancient, function that action mirroring serves is thought to be simpler: it helps the observer to understand observed actions by extracting and representing the goal, or the meaning, of those actions (Rizzolatti *et al.*, 2001; Rizzolatti and Craighero, 2004; Blakemore and Frith, 2005). In fact, it seems to be “generally accepted that the fundamental role of mirror neurons is to allow the observing individual to understand the goal of the observed motor act” (Fogassi *et al.*, 2005, p. 665).

As for the mechanisms that allow action understanding by mirroring, the most popular proposal is known as the *direct-matching hypothesis* (Rizzolatti *et al.*, 2001). According to this hypothesis, “an action is understood when its observation causes the motor system of the observer to ‘resonate’” (Rizzolatti *et al.*, 2001, p. 661), and this resonance allows the observer to figure out the outcome, and ultimately the goal, of the action because “he know[s] its outcomes when he does it” (Gallese *et al.*, 2004, p. 396). In other words, action mirroring provides a simulation device for goal understanding by automatically and mandatorily duplicating the observed action in the observer’s motor system.

In this paper I challenge both of these claims. I shall propose that the primary function of action mirroring is not action understanding in terms of goals but *predictive action monitoring*. I shall also suggest that action mirroring in the observer is achieved not by direct matching but by *emulative action reconstruction*. Here is a brief outline of the argument I shall make in this paper. In the next section (‘Visuo-motor translation during action mirroring’) I shall argue that action mirroring cannot be direct but must be based on some kind of interpretation of the observed action. The section on ‘Action mirroring and goal understanding’ will review the empirical evidence regarding this interpretation and will conclude that it often includes assumptions about the goal of the observed action. This implies that action understanding may precede, rather than follow from, action mirroring. Then in ‘Goal understanding without simulation’ I shall illustrate by findings from human infants that this is possible: goals can be understood without simulation. Finally, I shall return to the question of what functions action mirroring serves. I shall suggest that by serving the primary function of predictive action monitoring, action mirroring will also support human-specific phenomena that are not less important than imitation and empathy: action coordination and engagement in joint actions.

Visuo-motor translation during action mirroring

Action mirroring requires activating the motor program that, if it were executed, would perform a similar action to the observed one. To achieve this, the brain has to perform a translation from the visual representation in which the observed action arrives at the neural system into motor code. In other words, just like the cortical mechanisms that transform the visual information about graspable objects into motor commands that guide grasping (Jeannerod *et al.*, 1995; Murata *et al.*, 1997), the mirror system also has to find a match for the visual representation of an action in terms of motor commands.

Direct matching versus reconstruction

The direct-matching hypothesis proposes that such a translation is performed by a resonance mechanism. This metaphor suggests that certain parts of the motor system, namely the mirror neurons (MNs), or more generally the mirror neuron system, are ‘tuned’ to visually represented information about actions, and become automatically activated whenever a match occurs. In other words, direct matching, which is defined as “mapping the visual representation of the observed action onto our motor representation of the same action” (Rizzolatti *et al.*, 2001, p. 661) refers to a cognitively unmediated matching process that *duplicates* the motor program of the observed action without interpreting it. The ‘directness’ of this process is also emphasized by the assertion that the result of the visual analysis of an action is “devoid of meaning” (*ibid.*), and so actions are interpreted and understood only after the visuo-motor transformation has been performed.

A full ‘direct-matching’ account of action-mirroring mechanisms will have to answer two questions about the nature of this matching process. First, one has to specify how the tuning of the MNs, which will determine the outcome of the direct-matching process, is established. This problem is analogous to the ‘correspondence problem’ known in theories of imitation (Heyes, 2001; Brass and Heyes, 2005), and it has been proposed that direct links between visual and motor representation of the same action are generated by general associative learning processes during ontogenesis (Keysers and Perrett, 2004; Brass and Heyes, 2005; Heyes and Bird, Chapter 21). The second question for a direct-matching process is to determine *what* exactly is to be reproduced from the observed act. Does the mirror system have to duplicate every minute detail of the observed act (including, for example, direction and speed of motion, angles between joints, etc.) in order to facilitate its understanding? If not, how to determine the appropriate level of visuo-motor transformation? I shall return to this question later, but first I introduce an alternative mechanism for action mirroring to be contrasted with direct matching.

An alternative mechanism to direct matching in action mirroring is *action reconstruction*. This mechanism reproduces the observed action through interpreting it at some level of visual analysis (see below) and feeding the result of this interpretation into the observer’s own motor system. The idea of an action reconstruction mechanism comes from the recognition that visuo-motor translation is necessarily based on a visual analysis of the observed action, which must already have determined the decomposition of the action into relevant units, separated or merged its parts into smaller or larger segments, and disregarded some aspects of it (e.g. the position of the non-active hand in grasping) as irrelevant for reproduction. This kind of interpretation of actions is akin to the mid-level visual analysis of objects that determines the initial segmentation of a scene into relevant units to be analyzed further. Just like the mid-level scene analysis translates texture and motion information into objects (or object files, see Kahneman and Treisman, 1984), an analogous analysis can translate movements and body parts into actions to be mirrored. Visual analysis can go a long way in interpreting actions (and sequences of actions, see Subiaul *et al.*, 2004), as the activation of cells in the superior temporal sulcus (STS), which do not have motor properties, demonstrates (Perrett *et al.*, 1989; Pelphrey and Morris, 2006). Such interpretations may even include assumptions about the goal of the

observed action. If observed actions receive mid- or high-level interpretation within the visual system before becoming transformed into motor code, the result of this analysis will provide the input to the motor system to be reproduced. What mirroring can then achieve is the reconstruction of the motor command needed to perform that action (cf. the ‘ideomotor principle’, James, 1890; Hommel *et al.*, 2001). In the last section of this paper I shall return to the question of what function such an action reconstruction could serve.

Imitation versus emulation

The distinction between direct matching and action reconstruction is analogous to the imitation versus emulation distinction used in the social learning literature. In this context, *imitation* refers to the overt reproduction of an observed action, while *emulation* (or at least one type of emulation, called goal emulation or end-state emulation) is the reproduction of the *outcome* of an action by the observer’s own means (Tomasello, 1996; Whiten and Custance, 1996; Zentall, 1996; Custance *et al.*, 1999; Call *et al.*, 2005; Huang and Charman, 2005; Gergely and Csibra, 2006; Tennie *et al.*, 2006). While the distinction between these two types of overt action reproduction mechanisms can be conceptualized in many different ways, the crucial contrast between imitation and emulation lies in what information is copied over from the demonstrator to the observer: the means or the end (see also, Chaminade *et al.*, 2002). Suppose, for example, that you observe someone licking an envelope to seal it. When you want to seal an envelope next time, you may also lick it, imitating the action that you have observed. Alternatively, you may have learned from your observation that one has to apply moisture on the envelope in order to seal it. If you achieve this by a wet sponge, rather than by licking, you have attained the same end by alternative means, i.e. you have emulated the observed action.

Note that the reproduction of an observed action may be the same whether it is performed by imitation or emulation. Imitating an action will normally bring the action’s outcome with itself, as long as it is a goal-directed means action (as opposed, for example, to a gesture). Likewise, if the observer has similar effectors and biological constraints to the model, it is likely that she will emulate the outcome of the model’s behavior by the same means as it is achieved by the model, i.e. she will faithfully reconstruct the observed action. This is why, in studies of imitation, unusual or inefficient goal-directed actions are demonstrated to participants in order to test whether they tend to emulate the outcome by their own, more efficient way, or imitate the observed action faithfully (Meltzoff, 1988; Gergely *et al.*, 2002; Horner and Whiten, 2005). For example, to test whether infants are capable of deferred imitation, Meltzoff (1988) demonstrated an unusual action to them, in which the model switched on a box-light by pushing it with his forehead. If infants emulated the outcome, they would just use a simpler action to achieve the same goal: they could just push the box with their hands.

The distinction between imitation and emulation, or between action duplication and action reconstruction, however, is not absolute but relative. One way to characterize the relation between these concepts is to imagine imitation and emulation as two ends of a

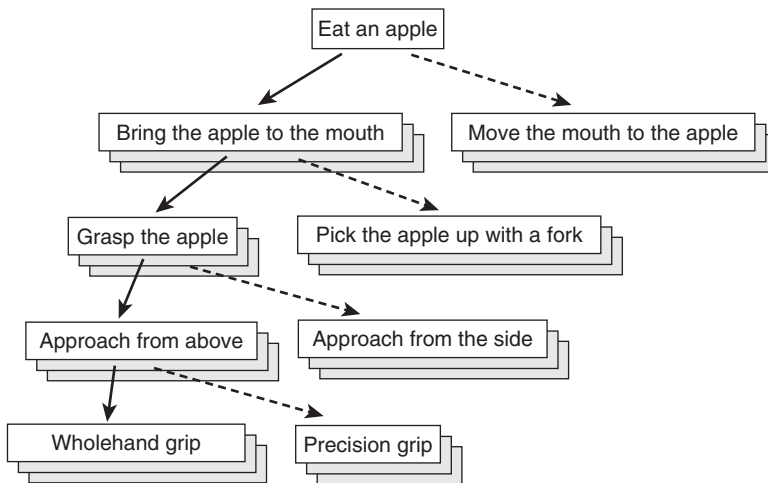


Figure 20.1 A simple action hierarchy.

continuum of decreasing fidelity of action reproduction (Whiten and Custance, 1996). Out of two copies of an action, the one that represents the less faithful reproduction would be considered an emulative response, while the other may be judged as an example of imitation. However, compared to a more accurate reproduction the latter response would be likely to be seen as emulation. Another, and perhaps better, way to characterize the relative nature of these concepts is describing them in relation to the action hierarchies of goal-directed actions. Specifically, whether a particular instance of action reproduction is considered to be imitation or emulation depends on what level the target action is defined in the action–goal hierarchy.

Action hierarchies

Actions are organized in a hierarchical manner (Jeannerod, 1994), which theories of imitation (Byrne and Russon, 1998) or action perception (Hamilton and Grafton, Chapter 18) cannot ignore. A simple action hierarchy is shown in Figure 20.1. In this scheme, the overarching goal of an action is the consumption of an apple, which requires a sequence of steps to achieve.² One of these steps could be to bring the apple to the mouth, which can be considered as a subgoal towards the higher-level goal of eating the apple. This subgoal, in turn, will also require a sequence of acts to complete; one of them will likely be ‘grasping the apple’. To achieve this sub-subgoal, one has to perform a series of movements, which, in turn, can also be analyzed into smaller units down to the level of individual muscle activations.

² From a theoretical point of view, it would be better to characterize goals and subgoals as states of affairs than as actions. Thus, it would be more appropriate to separate out the goal state (‘apple in the mouth’) from the action that achieves this (‘bringing the apple to the mouth’). For the sake of simplicity, however, I do not make this distinction in the text and in Figure 20.1.

Note, however, that often there are alternative means to achieve a certain goal or subgoal (see the right side of Figure 20.1). One can, for example, move one's mouth to the apple instead of the other way around (for many species this would be a more appropriate way of eating), or move the apple with a fork instead of by hand. These actions will also achieve the corresponding subgoals (apple in the mouth, getting hold of the apple to control its movement), but may involve different subactions and movements in the hierarchy below. When evaluating whether the reproduction of an action is an imitative or an emulative response, we have to judge whether it represents the same means as the model's action (compared to potential alternatives). If it does, it is considered to be an imitative response, if it does not (but achieves the same goal or subgoal), it is an emulative response. Note, however, that this evaluation requires us to specify which level of hierarchy is relevant for the reproduction of a given action.

Suppose, for example, that the model performed the apple-eating action and chose to perform it through the subgoals and subactions shown on the left side of Figure 20.1. In response, an imitator grasped an apple, brought it to her mouth and ate it. However, she took the apple by her thumb and index finger rather than using a whole-hand grip. Did she imitate the model's action? If action reproduction is evaluated at a low level where grip type is defined, she did not imitate but emulated the observed action. If, however, the model's action is defined at a higher level of the action hierarchy, the answer is affirmative; after all, she moved the apple to her mouth and not the other way around, and she grasped it by her hand instead of using a tool. At this level of analysis it is irrelevant whether minute details of her actions (like the grip type) matched to that of the model or not. For example, when infants imitate the head-touch action in variations of Meltzoff's (1988) experiment, they hardly ever copy exactly what the model has done. In fact, they frequently touch the light by their cheek, nose or chin, sometimes even kissing or biting the box (G. Gergely, personal communication). All these behaviors are considered to be imitative responses because the infants chose the same effector (their head) to perform the action as did the model. In other words, they imitated because they matched their action to the model's action at the appropriate level of description ('pushing the box by the head'). However, below that level they executed the action by their own means, i.e. they achieved this subgoal by emulation.

In fact, as we descend the action hierarchy, *any* action reproduction will at some point be seen as emulation rather than imitation because the differences between the imitator's and the model's body will not allow perfect matching in all movement parameters. This is, however, not a problem for imitation because, unless the imitator is a professional dancer or a mime artist, reproduction of perfect angles of joints, speed and acceleration of limb movement, etc. is irrelevant. My point here is not that imitation as such does not really exist, but rather that any instance of imitation is actually achieved by emulation. When someone imitates, she chooses a certain level of description of the observed action, and reproduces that level by reconstructing it in her own motor system.³

³ In this respect, imitation is always goal-directed (Wohlschläger *et al.*, 2003), although the 'goal-directed emulation' term would express better the nature of action reproduction.

From this perspective, whether something is imitation or emulation is not a well-formulated question. Instead, one should ask how the imitator decides what is the relevant level in the action hierarchy on which an action is to be reproduced (see Williamson and Markman, 2006).

Bottom-up versus top-down

The analysis that I have applied to overt action reproduction (i.e. imitation) also applies to covert action reproduction (i.e. mirroring). Just as the *mechanism* of imitation is always emulation at a lower level, the *mechanism* of motor mirroring is always reconstruction. There is no mysterious mirroring process that directly transforms action observation into motor code. Rather, the observed action is analyzed at some level of precision and the result of this analysis is mapped onto the observer's motor system. One can call this mapping process 'direct matching' (Rizzolatti *et al.*, 2001) and such mappings may be established by 'direct' associations (see Chapter 21), but what is mapped during mirroring is not an uninterpreted signal but a description of the observed action at some level of the action hierarchy. The fine details of the resulting motor activation in the observer do not directly originate from the observation but are reconstructed from this description.

In this context, the crucial question about action mirroring and its relation to action understanding concerns the *level* of action interpretation where the mapping from visual to motor code takes place. The intuition behind the idea of 'direct matching' and the 'resonance' metaphor probably is that visuo-motor translation during mirroring occurs at a relatively low level (see the left side of Figure 20.2). Hierarchical models of motor control assume that higher and lower level motor modules are reciprocally connected to each other (Wolpert *et al.*, 2003). Thus, the low-level motor activation generated by mirroring could propagate upwards in the observer's own hierarchically organized motor system to estimate what higher level subgoals and goals might have generated the observed action (Wolpert *et al.*, 2003). This bottom-up activation may be the key to how the motor system contributes to goal understanding by mirroring, and it is the central idea behind simulative understanding of observed actions.

Alternatively, mirroring, just like imitation, can also be achieved at a higher level of action interpretation (when this is available). If the action interpretation system can construe the observed behavior in terms of higher-level or further goals, these can be mapped onto the observer's own action control system, within which it can propagate downwards to generate the corresponding motor code by covert emulation (see the right side of Figure 20.2). This top-down activation is also a kind of simulation, but it is predictive in nature, generating motor actions for goal conjectures (cf. Gallese and Goldman, 1998) rather than the other way around. This account of action mirroring proposes that observed actions are interpreted to the highest possible level before they are passed on to the motor system for reconstruction.

The relation between the two kinds of simulative exploitation of the observer's own motor system by action mirroring is illustrated in Figure 20.2. The difference between the two models lies in two factors: the action interpretation level at which visuo-motor translation is performed, and the propagation direction of activation within the

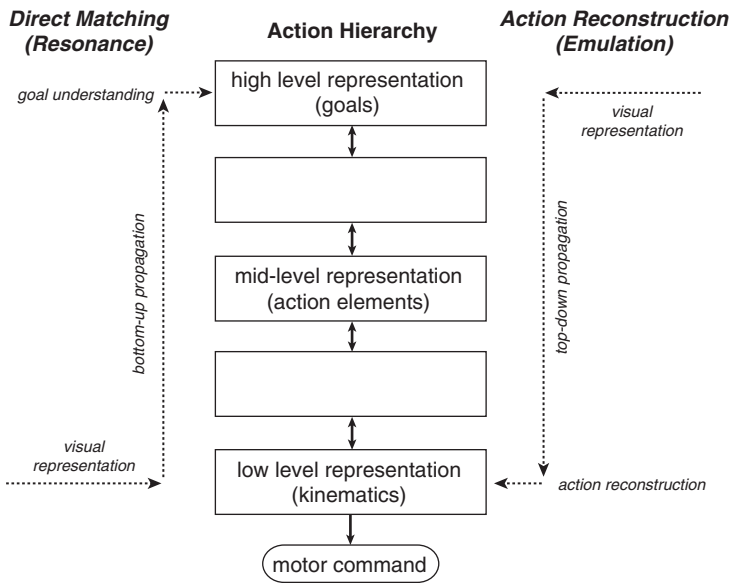


Figure 20.2 Contrasting ‘direct matching’ and ‘action reconstruction’ in a hierarchical action representation.

action control system following mirroring. Low-level mirroring (i.e. ‘direct matching’) and bottom-up propagation supports a simulation system that facilitates the understanding of the goals behind observed actions, while high-level mirroring and top-down propagation (i.e. ‘action reconstruction’) allows predictive emulation of observed actions on the basis of the high-level interpretation of the action achieved without low-level motor activation.

Action mirroring and goal understanding

In the previous section, I have provided a theoretical framework in which the relation between action mirroring, goal understanding, action hierarchy, simulation and emulation can be conceptualized. The question of what function action mirroring serves, however, is not a theoretical but an empirical question. In particular, the two alternative models of action mirroring (shown in Figure 20.2) are to be tested against the empirical findings accumulated in cognitive and neuroscientific research. Fortunately, the two models provide slightly different predictions for what kind of mirroring phenomena one would expect to find.

If action mirroring is achieved by visuo-motor translation at a low-level of action interpretation, we should find that low-level mirroring is always present and is matched well to the kinematics of the observed action. In contrast, if motor mirroring is produced by action reconstruction, we would expect to find that motor activation may not necessarily accompany action observation and, since it is produced by emulation, it may be

different from the observed action. This section reviews some studies that are relevant for judging which of these predictions are confirmed by empirical findings.

Some aspects of basic mirroring phenomena are highly consistent with low-level action interpretation for mirroring. For example, observers seem to automatically imitate simple transitive and intransitive hand movements, measured by a facilitatory effect on performing compatible, and an inhibitory effect on executing incompatible, actions (Brass *et al.*, 2001; Press *et al.*, 2005; Bertenthal *et al.*, 2006). Observing someone making arm movements also interferes with executing different arm movements (Kilner *et al.*, 2003). These findings indicate a low-level mirroring process because the kinematic aspects of the observed action seem to have had an effect on the observer's kinematically similar or different movements. However, these phenomena are not relevant for judging between the two models of action mirroring because when higher-level action interpretation is not available, the emulator model also predicts low-level mirroring.

Low-level congruency between observed and mirrored actions also occurs in mirroring goal-directed actions. For example, a sizeable proportion (19–41%) of MNs in the monkey's ventral premotor and parietal cortex are classified as 'strictly congruent' (di Pellegrino *et al.*, 1992; Gallese *et al.*, 1996, 2002; Ferrari *et al.*, 2003). These neurons respond only to observed actions that are performed the same way (e.g. using the same grip type) as characterized by the motor properties of the same cell. Similarly, observation of goal-directed grasping actions tends to activate the same muscles that the observer should use for executing the same action (e.g. Fadiga *et al.*, 1995). These findings are indicative of low-level action mirroring but, since they involve observation of prototypical actions, they do not exclude the possibility that such mirroring phenomena are generated by emulative action reconstruction. Further findings, however, are more compatible with the predictions drawn from the emulative than from the direct-matching model of action mirroring.

Congruency in mirroring goal-directed actions

Monkeys' MNs do not respond to mimicked actions, for example, when they observe the experimenter pretending to grasp something in the absence of any objects (Gallese *et al.*, 1996). This finding is puzzling if action mirroring is performed by low-level direct matching because the low-level kinematics of a mimicked action is presumably similar to that of an object-directed action, and is available for mirroring. Similarly puzzling is the fact that observing a reaching act for an occluded target object elicits MN activation whereas the same movement does not trigger a MN response when the monkey knows that there is no food behind the occluder (Umiltà *et al.*, 2001). After all, the visual input is the same in the two conditions and low-level action mirroring is possible. These findings are often cited in support of the claim that MNs play a role in goal understanding, but what they really indicate is that MNs *reflect* action understanding rather than contribute to it. Consistent with the predictions drawn from the emulation model of action mirroring, MN activation in these studies seems to be conditional on action understanding and not the other way around.

Similar phenomena were also demonstrated in human neuroimaging studies. Action mirroring in humans is not restricted to transitive, object-direct actions (e.g. Bertenthal

et al., 2006), and the human ventral premotor cortex is sometimes activated by the observation of intransitive actions that do not involve any objects (Iacoboni *et al.*, 1999). Nevertheless, this activation is often higher for goal-directed actions than for actions whose goal is not evident (e.g. Nishitani and Hari, 2000; Chaminade *et al.*, 2001; Koski *et al.*, 2002; Johnson-Frey *et al.*, 2003).

It is not possible to measure the exact match of observation-induced motor activation in neuroimaging studies (though there is a good correspondence on the level of effectors, see Buccino *et al.*, 2001). However, such data are available from MN studies with monkeys. Although many MNs are ‘strictly congruent’, i.e. represent a good one-to-one correspondence between observed and executed actions, the majority of them (~60%, see Fogassi and Gallese, 2002) fall into the ‘broadly congruent’ category. Some of these neurons respond to two or even three types of observed actions (Gallese *et al.*, 1996), which are often related to each other in meaningful ways. A MN that is active during grasping by hand may also be activated by the observation of ‘grasping with the mouth’ (Gallese *et al.*, 1996). It is difficult to see how low-level motor mirroring could produce such a mismatch, but this kind of mirroring fits perfectly with the emulation model of mirroring. If the monkey has ‘understood’ the immediate goal of the action outside the motor system, from which the motor activation reconstructs the observed action, we would expect exactly this kind of correspondence between observation and execution: matching actions at a higher level with occasional mismatches at a lower level along the action hierarchy.

Even when a MN is activated only by the observation of a single action, it is not necessarily the same action as defined by the motor properties of the neuron. For example, di Pellegrino *et al.* (1992) reported that in many MNs, the effective observed and effective executed actions were *logically* related. “For example, the effective observed action was placing an object on the table, whereas the effective executed action was bringing food to the mouth or grasping the object” (di Pellegrino *et al.*, 1992, p. 179). It is hard to see how low-level action mirroring could result in such MN activation, while emulative mirroring, and especially predictive emulation, can easily explain such a phenomenon (see ‘Action mirroring is anticipatory’ below).

Further goals and intentions

The activation of MNs depends not only on the presence of target objects, but may also be tailored to the further, or higher level, goals of the observed individual. In a recent study, Fogassi *et al.* (2005) trained monkeys to perform two actions: grasping an object and putting it into their mouth (i.e. eating it), or grasping an object and putting it into a container (placing). Although the first part of these actions (grasping) was kinematically similar to each other, the researchers found separate sets of MNs in the inferior parietal lobule, which were preferentially activated before and during grasping according to the subsequent, to-be-executed action. In other words, some MNs showed higher activation when the monkey grasped the object to eat it, while others were more active when the monkey was about to place the object into the container. Crucially, these neurons responded similarly when the monkey observed the same actions performed by an experimenter. When the experimenter was about to eat the object, the ‘grasping to eat’

neurons were selectively activated; when he was about to place the object into the container, the other set of MNs fired. This is a clear demonstration that MNs take into account the further goal, and not just the perceived action, when responding to observed actions.

A neuroimaging study has recently concluded that the human inferior frontal cortex behaves the same way during action observation. Iacoboni *et al.* (2005) presented observers with actions (grasping a cup) either out of context, or in contexts that indicated one of two underlying intentions ('drinking tea' or 'cleaning up'). BOLD activation in the right inferior frontal cortex was higher when the context indicated the intention, i.e. the further goal, than when the action was presented alone, and the authors also found evidence for differential activation for the two intentions within the same region.

It is not clear how low-level mirroring would explain the differential motor activations in these studies, because these experiments were specifically designed to make it impossible for the observer to figure out the intention of the actor from observation of the action alone. The researchers in both studies made sure that, whatever the intention behind the observed action would have been, the perceptual and motor properties of the initial action (grasping) were as similar as possible. Thus, low-level action mirroring and propagating such activation upwards in the observer's action control system could not contribute to the understanding of distinct further goals or intentions for the observed actions.

In contrast, these results fit perfectly with the emulation model of action mirroring. Although the actions themselves did not carry information about the further goal of the actor in the studies cited above, the context did. For example, the monkey in the Fogassi *et al.* (2005) study could figure out the further goal of the observed action from the kind of object (food or non-food) involved, and whether or not a container was present. Such a goal attribution allowed them to emulate the action needed to achieve the goal by their own motor system, and this explains the differential activation of MNs during the observation of the initial action, which itself was not different between experimental conditions.

Mirroring non-executable actions

Action mirroring, almost by definition, requires the same or similar effectors and biological constraints between actor and observer. This may be the reason why mirroring phenomena cannot normally be elicited by nonhuman, robotic actions (Castiello *et al.*, 2002; Kilner *et al.*, 2003; Tai *et al.*, 2004; but see Press *et al.*, 2005), and this was the original explanation for why MNs did not appear to respond to actions performed by a tool, like grasping with a pincer (Gallese *et al.*, 1996). A recent study, however, casts doubt on the existence of such a constraint in action mirroring.

Ferrari *et al.* (2005) subjected monkeys to a long visual training of tool-using actions, like picking up food items by sticks. The monkeys passively observed as the experimenters manipulated these tools, and never learned themselves how to use them. In fact, when tested, one of the monkeys did not even make an attempt to use a stick to reach a desirable piece of food outside its cage. In spite of this, some MNs in their premotor cortex started to respond to the observation of tool use after 2 months of observational training.

When they saw the experimenter using a stick to pick up a piece of food, some MNs that were active while the monkeys themselves grasped food by hand, discharged. This is a clear example of mirroring activation for an action that the observer is unable to perform, which is incompatible with the idea of low-level motor mirroring. Nevertheless, the mirroring process was not random. As Ferrari *et al.* (2005) observed, “most tool-responding mirror neurons ... show a very good similarity between *the goal* of the observed and executed effective actions” (p. 216, italics added). In other words, MNs responded to the sight of a non-executable action with a different action that the monkey could have used to achieve the same goal. This is exactly what the emulation model of action mirroring predicts for observed actions whose goals are interpreted outside the motor system and then fed into the observer’s action control system for reconstruction.

Action mirroring can even occur in response to biologically impossible actions. Costantini *et al.* (2005) measured observers’ brain activation by functional magnetic resonance imaging while they were presented with finger movements that were within or outside the normal range of such actions. The impossible action depicted a hand with the little finger moving laterally for 90°. The results showed that the human ventral premotor cortex (part of the human ‘mirror neuron system’) was activated equally by the possible and impossible actions. Costantini *et al.* (2005) concluded that “the premotor system does not take into account the biomechanical constraints the observed movements would involve if they were actually executed” (p. 1765). Is it possible that such ‘mirroring’ of an impossible action is produced by ‘direct matching’? As there is no matching action in the observer’s repertoire, this is unlikely. However, it is conceivable that the visual system can provide an appropriate description of the end-state of such an action (‘the little finger is perpendicular to the others’), which then the motor system attempts to approximate, albeit unsuccessfully, using the available motor programs. In other words, mirroring the observed action can be attempted by driving the motor system top-down from a mid-level interpretation of it.

Level of mirroring and goal understanding

While most findings that demonstrate action-mirroring phenomena fit both accounts of action mirroring, the ones I reviewed above appear to be incompatible with the direct-matching model. These findings show that low-level motor ‘resonance’ (i) is not mandatory (pace Gallese, 2006) but may depend on the interpretation of the observed action as goal-directed action, (ii) takes into account extra-motor, contextual information relevant to potential goals, and (iii) is evoked by non-executable actions when their goals can be estimated from visual information. These aspects of action mirroring suggest that goal understanding is not the output but the input of the mirroring process, and covert reproduction of the observed actions is generated by top-down emulation rather than by bottom-up propagation of activation from low-level motor resonance.

All these findings reflect a tension between two conflicting claims about action mirroring implied by the direct-matching hypothesis: the claim that action mirroring reflects

low-level resonance mechanisms, and the claim that it reflects high-level action understanding. The tension arises from the fact that the more it seems that mirroring is nothing else but faithful duplication of observed actions, the less evidence it provides for action understanding; and the more mirroring represents high-level interpretation of the observed actions, the less evidence it provides that this interpretation is generated by low-level motor duplication.

Let me illustrate this point on the study of how MN activation reflects ‘intention understanding’ in monkeys (Fogassi *et al.*, 2005), discussed above in ‘Further goals and intentions’. Marc Jeannerod (personal communication) suggested that the slight kinematic variation between the monkeys’ ‘grasping to eat’ and ‘grasping to place’ actions might explain the activation difference across MNs. Fogassi *et al.* (2005) did not report the kinematic profile of the actions that the monkeys observed from the experimenter. If we assume, as Jeannerod suggested, that the observed actions included the same kinematic differences as in the monkeys’ actions, and the monkeys’ parietal MNs were sensitive to these parameters, then their activation represents a low-level mirroring phenomenon (in fact, it represents a lower-level mirroring than any of the earlier studies had demonstrated). However, nothing in this study would then suggest that the monkeys would have understood the ‘intention’ behind the observed actions. In contrast, if we accept Fogassi *et al.*’s (2005) argument that the selectivity of MNs was independent of the kinematic parameters and reflected ‘intention’ understanding based on contextual cues, then nothing in this study provided evidence that such an understanding is based on low-level mirroring (i.e. motor resonance). One cannot have one’s cake and eat it too: the discharge of a set of MNs cannot represent the activation of the observer’s motor system at low and high levels at the same time.

The emulation account of action mirroring, according to which mirroring is generated by action reconstruction from the highest available level of action interpretation, avoids this pitfall because it explains *both* why action mirroring may reflect the inferred goal of an action, *and* why it appears to provide a partial motor duplication for observed actions. However, such an account entails that the goal of an observed action can be estimated without the involvement of the low-level motor system.

Goal understanding without simulation

Understanding the goal of an observed action involves figuring out the content of the intention that generated the action.⁴ Since there are always an infinite number of different intentions that may have produced any particular action (Jacob and Jeannerod, 2005), the extraction of the goal from an action is an *inverse problem* (Csibra and Gergely, 2007). Inverse problems (which usually attempt to infer causes from effects) do not have

⁴ We have argued elsewhere that extracting the content of an intention (the goal) does not necessarily imply representing it *as* the content of a mental state (Csibra and Gergely, 1998; Gergely and Csibra, 2003). As this distinction is not relevant in the context of the present paper, I will not discuss it further.

analytical solutions, but their solution can be estimated by statistical methods. One way to describe the task of goal attribution is as a Bayesian problem (Baker *et al.*, 2006 see also Chapter 16). Under this description, the probability that a certain goal explains a certain action is estimated as

$$p(\text{goal} | \text{action}) = \frac{p(\text{action} | \text{goal}) \times p(\text{goal})}{p(\text{action})}$$

This equation says that this probability is proportional to the probability that this action is generated to achieve that particular goal state and to the probability of the goal state, and inversely proportional to the probability of the action. These last two terms [$p(\text{goal})$ and $p(\text{action})$] can be estimated by the accumulated experience with the frequencies of actions and outcomes (interpreted as goals). But where can the $p(\text{action} | \text{goal})$ term come from?

First, it can be estimated on purely statistical basis: the more times a certain outcome has been associated with a certain action, the more likely that they will go together again (Hommel *et al.*, 2001). A clear example of such associative learning is the study by Ferrari *et al.* (2005), in which monkeys learnt to associate goals with unfamiliar tool-assisted actions performed by experimenters (see ‘Mirroring nonexecutable functions’ above). This finding nicely demonstrates that such associative learning does not necessarily involve motor simulation. Second, the probability that a certain action is performed to achieve a certain goal can be estimated by motor simulation. The observer can feed the goal into her own motor system (switched to ‘pretend’ mode), and generate the action that she would be most likely to perform to achieve it (Gallese and Goldman, 1998). This is a powerful method for estimating the $p(\text{action} | \text{goal})$ term, with the obvious limitation that it can only be used for understanding known goals of individuals with similar motor constraints to the observer (Csibra and Gergely, 2007).

Finally, the likelihood that an action is performed to achieve a certain goal can also be estimated by *teleological reasoning* (Csibra and Gergely, 1998; Gergely and Csibra, 2003). Such reasoning⁵ assumes that agents tend to conserve energy, and achieve their goals in the most efficient way available to them. In other words, teleological reasoning assumes that $p(\text{action} | \text{goal}) \approx$ (efficiency of action towards goal), and takes into account the situational constraints when evaluating efficiency. It has been demonstrated in several studies that human infants apply such teleological reasoning to interpret actions they observe. Here I shall briefly describe three studies showing that infants attribute goals (i) to non-human agents, (ii) to impossible actions, and (iii) to pretence actions. None of these goal attributions is based on motor simulation.

⁵ The term ‘reasoning’ here does not imply conscious or deliberate cognitive processes. Just as certain operations within the visual (Scholl, 2005) and the motor (see Chapter 16) systems can be described as performing Bayesian inferences, the Bayesian reasoning process that applies teleological assumptions to infer goals can also operate automatically and without awareness.

Understanding goals of nonhuman agents

To investigate whether infants attribute goals to nonhuman agents, we presented 12-month-olds with computer animations, in which a circle repeatedly approached another one by jumping over an obstacle (Gergely *et al.*, 1995). Having habituated to this action, infants were then confronted with a modified situation, in which we removed the obstacle that had separated the agent from its goal, and we showed them two different events. One of them depicted the agent performing the same ‘jumping’ action as before, which was unnecessary and inefficient in the absence of the obstacle, and in the other event the agent approached its goal via the most efficient straight pathway. Looking-time measures indicated that, although it was perceptually more similar to the previously seen action, the infants found the ‘jumping’ action unexpected in this situation, suggesting that they predicted an efficient goal approach. A control condition, in which the infants were habituated to a non-necessary ‘jumping’ action, confirmed that this expectation was based on the goal attribution they had made when watching the efficient goal approach in the experimental condition.

Since the original report, this finding has been extended to younger infants (Csibra *et al.*, 1999; G. Csibra, unpublished data), different kinds of computer animations (Csibra *et al.*, 2003; Wagner and Carey, 2005), and different kinds of agents (Sodian *et al.*, 2004; Kamewari *et al.*, 2005). Studies with other paradigms also suggest that infants do not refrain from attributing goals (or, at least, preferences) to nonhuman objects (Luo and Baillargeon, 2005; Bíró and Leslie, 2007). None of these findings can be explained by motor simulation or resonance, because the goal-directed agents the infants observed in these studies did not have bodies or motor systems similar to the infants. Even the study that replicated the original results with human agents in 6-month-old infants (Kamewari *et al.*, 2005) involved motor actions (e.g. walking) that the participants were too young to be able to perform.

Understanding the goal of an impossible action

In a recent study, we tested whether 6–8-month-old infants were willing to extend goal attribution to a biologically impossible action (V. Southgate *et al.*, unpublished data). We reasoned that if young infants have not yet accumulated sufficient knowledge about the biological constraints of human bodies (Slaughter *et al.*, 2002), but expect that goal-directed actions take the most efficient course, they may erroneously predict a biologically impossible action if it were physically more efficient than a biologically possible goal approach. Briefly, we familiarized infants to a video clip in which a hand reached toward, grasped, and retrieved a ball from behind an obstacle. We then tested them with video clips showing a different situation, in which a second, new obstacle was introduced. Since the two obstacles together made it impossible for the hand to simply reach for the ball, we offered two new actions as solutions to the infants. In one of them, the hand pushed the second obstacle away and then retrieved the object the same way as during the familiarization, while the second action involved the hand snaking around the two obstacles forcing the forearm to take an S shape. While this second action is clearly

impossible to perform by a human arm, were it biologically possible, it would represent a more efficient goal-directed action than the first one because it would get to the target object in a single step.

Our looking-time measures indicated that the infants found the efficient but impossible action more compatible with the original goal approach than the possible but less efficient action. In the control condition, in which the infants were familiarized to a nonefficient goal approach, we did not find such differential responses. This finding confirms that young human infants are willing to accept an objectively impossible action as a well-formed, goal-directed action if it appears to be more efficient than its alternatives. Note that this result suggests not only that infants, unlike adults, were not surprised to see a biologically impossible action, but also that they positively thought that this action was better in the given situation than the possible one. It is hard to see how a simulation-based goal understanding could explain such a finding.

Understanding pretence actions

Onishi *et al.* (2007) presented 15-month-old infants with the following sequence of actions. The actor turned two upside-down cups upright while demonstrating that they were empty. She then took an empty bottle, turned it over one of the cups as if she was pouring something into the cup, but no liquid left the bottle. She put down the bottle, and then took one of the cups and pretended to drink from it. Infants looked much longer at the 'drinking' action when it was performed with the cup that was not involved in the mimed pouring than when it was performed with the cup that should have contained some liquid, had the pouring been real (see Onishi *et al.*, 2007, for the control conditions). This demonstrates that infants expect that sequentially organized actions are directed towards a specific goal state: a pouring action is anticipated to result in an outcome state that provides the enabling condition (liquid in the cup) for a subsequent drinking action from the same cup. Crucially, infants recognized the causal and teleological relatedness of these actions even if no liquid was present and therefore neither the subgoal nor the final goal was ever achieved.

This finding is interesting in the context of this paper because infants at this age rarely drink from cups, are unable to pour from a bottle, and do not produce pretend actions themselves. Thus, even if they mirrored the observed action sequences faithfully, it would not have allowed them to understand these actions. In contrast, teleological reasoning, and their background knowledge about cups and bottles, did enable them to figure out the pretended goal of the pouring action and its relatedness to the subsequent drinking action. Goal understanding is possible without motor simulation.

The function of action mirroring

So far, I have argued that action mirroring must be based on some level of interpretation of the observed action, and tried to show that, when the action can be interpreted in terms of goals, such interpretation precedes, rather than follows from, action mirroring. Evidence from human infants demonstrates that goals of observed actions can be

understood without the involvement of the motor system, and suggests that action mirroring can be achieved by top-down emulation. This conclusion, however, raises a question about the function of action mirroring: If actions can be understood without motor simulation, then why does the brain go on and reproduce them in the motor system?

One possible answer to this question is that the observer's motor system does not reproduce, but *pre*reproduces the observed action. In other words, the motor activation in the observer may not mirror but anticipate, may not shadow but foreshadow, what the other is doing. The perception of dynamic events, whether or not they involve social stimuli, is always predictive in nature (Wilson and Knoblich, 2005), and there are many benefits of being able to anticipate the immediate future in social interactions (see 'Why anticipate actions').

Action mirroring is anticipatory

Mirror neurons do not passively reflect observed actions but seem to anticipate them (see figures in Gallese *et al.*, 1996). 'Grasping mirror neurons', for example, start to discharge hundreds of milliseconds before the observed hand touches the target object, as if they mirrored the future. If the function of MNs is to anticipate impending actions, the existence of 'logically related' MNs (see 'Congruency in mirroring goal-directed actions' above) also makes sense. Logically related MNs (38% of all MNs in di Pellegrino *et al.*, 1992) respond to observed actions with a different action, but with one that could 'logically' follow the observed one. While it is not easy to see how 'direct matching' could result in such mirroring, this kind of motor response is expected if MNs anticipate both the course of the current action and the potential subsequent actions. Similar neurons have also been found in the parietal cortex in the study that reported intention-sensitive MNs (Fogassi *et al.*, 2005).

The temporal resolution of neuroimaging studies based on haemodynamic responses may not allow accurate measurement of timing relation between observation and motor activation, but other methods suggest that humans also anticipate the next move of an observed individual. This is evident in their eye movement patterns (Flanagan and Johansson, 2003; Falck-Ytter *et al.*, 2006), as well as in motor (Kilner *et al.*, 2004) and ventral premotor cortex (Nishitani and Hari, 2000; Chaminade *et al.*, 2001; Ramnani and Miall, 2004) activation.⁶ In all these cases, the observer's own motor activation was triggered by their understanding of the goal, or immediate subgoal, of the observed action, and they seem to have simulated predictively what the other should do to achieve that goal. But goal understanding is not a necessary requirement for action prediction. If action mirroring generally serves a predictive function, it should use any information, and not just assumed goal states, that allows action anticipation.

⁶ In fact, the premotor cortex seems to implement predictive procedures even in domains outside action perception (Chapter 7).

One such information is cyclic repetition of movement. Borroni *et al.* (2005) measured the modulation of muscle reflexes of people who watched an experimenter repeatedly performing wrist flexion and extension. They found that the excitability of the corresponding muscles in the observers was modulated with the same rate at which the action was performed. When the frequency of the demonstration was 1 Hz, the reflex modulation followed a 1 Hz cycle; when the demonstration was faster at 1.6 Hz, the excitability of the muscles also appeared to be modulated at 1.6 Hz. This is a clear demonstration of motor mirroring. However, the phase of the excitability cycle was advanced compared to that of the demonstration, i.e. the reflex modulation *preceded* rather than followed the demonstration by about 160 ms. Thus, motor mirroring in this study anticipated the movement of the observed hand, which was made possible not by goal understanding but by the cyclic repetition of the observed action.

Goal understanding and action prediction

Is the proposal that action mirroring serves a predictive function different from the one it is meant to replace, i.e. that it leads to goal understanding? One can say, after all, that the purpose of goal understanding is itself predictive: it tells the observer the likely outcome of an observed action. It is true that goal attribution enables two kinds of prediction to be made (Csibra and Gergely, 2007). The first one is the goal itself; goal attribution normally predicts the goal to be achieved. The present proposal, however, assigns an additional predictive function to action mirroring. While goal attribution itself allows us to jump ahead in time and predict a hypothesized future state, it also enables us to fill up the intervening time by action anticipation. In this kind of prediction, the hypothesized goal of the action is not the output, but serves as the input, of the prediction process (cf. Gallese and Goldman, 1998).

Others collected good arguments for why simulating an action is not sufficient for recovering the goal or the intention behind it (Jacob and Jeannerod, 2005; Saxe, 2005). However, there are equally good arguments for the role of simulation in action prediction (e.g. Wilson and Knoblich, 2005; Prinz, 2006; Csibra and Gergely, 2007). In fact, if the observer has a good guess about what the actor is trying to achieve (i.e. what the 'goal' is), and the actor is conspecific with similar motor constraints to the observer, the most effective way to anticipate the actor's unfolding behavior is motor simulation by emulation. In this kind of predictive simulation, the observer feeds the hypothesized goal state into her own motor system, generates (but does not execute) the appropriate motor command, uses the corresponding forward model (Wolpert and Ghahramani, 2000; see also Chapter 16) to predict the visual consequences of this action, and attaches this prediction to the actor.

Thus, the present proposal is similar to other hypotheses (e.g. Gallese *et al.*, 2004) in that it interprets action mirroring as a simulation process. This simulation process, however, is not retrodictive, does not recover the intention that generated the action, but predictive, emulating the action needed to achieve a hypothesized goal.

Why anticipate actions?

What is the use of such online action anticipation of ongoing actions? First, it enables the observer to verify and revise goal attribution. In fact, in the first proposal for how action motor mirroring could support the understanding of intentions (Gallese and Goldman, 1998), the mirroring process was suggested to be exploited in this way (later to be abandoned for a ‘direct-matching’ approach). The only computational model that I am aware of and that attempts to model the understanding of intentions by MNs (Oztop *et al.*, 2005) follows the same strategy. In that model, the observer model ‘knows’ in advance the set of possible goals that the actor may try to achieve. When it observes a particular movement, it predictively simulates actions for these goal states sequentially until it finds one that matches the observed movement. The role of ‘mirror neurons’ here is to produce the expected sensory (visual) effects of the predicted action, which then can be compared to the observation. Note that what drives MNs in this model is not a ‘direct-matching’ mechanism but the goal conjectures generated outside the motor system. Yet, mirroring is a useful process to verify if these conjectures are valid.

Second, whether or not action anticipation is generated from goal conjectures, it is always beneficial to be ahead of events. For example, action anticipation gives the perceiver the opportunity in competitive situations to intervene in time if it becomes necessary, and allows for the quick recruitment of resources to deal with unexpected events. Perhaps this explains why MNs do not appear to respond to actions presented to the monkeys on a television screen (Ferrari *et al.*, 2003; Keysers and Perrett, 2004), and stop responding to actions performed on non-food objects “after a few or even the first presentation” (Gallese *et al.*, 1996). In these situations, the observed actions are irrelevant for the monkeys (they would not be able to get anything out of them), and hence they are not worth monitoring because intervention is either impossible (in case of actions seen on the television) or would be unrewarded (in case of nonfood objects).

Third, predictive tracking may also support learning about the physical environment in which the observed action takes place. For example, the motor system is involved in the weight estimation of objects from observation of lifting actions (Hamilton *et al.*, 2004; Pobric and Hamilton, 2007) though probably not through low-level mirroring of action kinematics (Hamilton *et al.*, 2005). If the kinematics of a perceived action differ considerably from what is anticipated by emulation, this allows us to revise the estimated weight or other invisible physical parameters of objects involved in the action.

Finally, action coordination between individuals is virtually impossible without action anticipation. Even the simplest task, like taking a walk with someone or handing over an object to someone, requires precise adjustment of the timing of movements to the other party. Humans, unlike other animals, are frequently engaged in cooperative and joint actions (Knoblich and Jordan, 2002; Pacherie and Dokic, 2006; Sebanz *et al.*, 2006). Perhaps this provides an explanation for why predictive action mirroring is so ubiquitous in humans.

Conclusions

This paper is concerned with the relation between mirroring and understanding of actions. The term ‘understanding’ has many different meanings. For example, Gallese (2006) assigns a role of mirroring in *experiential* understanding of others, which is “direct grasping of the *sense* of the actions performed by others” (p. 16). My analysis, however, was restricted to the question of whether action mirroring is involved in extracting the immediate or further goal (i.e. the potential content of the underlying intentions) of an action. The popular conception of the causal role of mirroring in understanding the ‘meaning’ of actions involves a direct, unmediated, automatic, mandatory, resonance-like transfer mechanism, which miraculously generates a copy of the motor command responsible for the observed action, and forms the basis of bottom-up identification of the goals (or intentions) that have guided that action. I showed that there is a theoretical alternative to this scenario, in which the motor activation during mirroring is generated in a top-down manner within the action control system from the level where the observer has been able to interpret the other’s action. Beyond the fact that this alternative account has some theoretical advantages (e.g. it does not involve reversing the normal information flow in the motor system, as suggested by Blakemore and Decety, 2001), it is also more compatible with the empirical evidence.

On the other hand, the action–reconstruction account, just like ‘direct matching’, also asserts that action mirroring implements a simulation function in the motor system. This is, however, predictive, rather than retrodictive simulation mechanism. Such a simulation mechanism will, no doubt, also contribute to the ‘understanding’ of the observed action in a sense, though it may not supply the goal or the intention behind the action for free. Ironically, the function that I propose for such a simulation routine portrays the motor involvement in action observation as reflecting a perception-for-action procedure (allowing, for example, intervention or action coordination), while the direct-matching approach views it as a mechanism with a primarily epistemic purpose (i.e. ‘understanding’). It is not action simulation that makes action understanding possible, but the other way around: it is action understanding that makes action emulation efficient.

This conclusion implies that the term ‘mirroring’ may be misleading because it does not capture the true nature of motor activation during action observation. Inevitably, many people will see my account of MN function as deflationary, but it is not necessarily so. If MNs are, in fact, *emulator neurons*, and they do not inform us about goals but enable us to be engaged in joint actions with others, they may play a more important role in social interactions than what is usually ascribed to them—a function that would be difficult to achieve without them.

Acknowledgments

This work was supported by the UK Medical Research Council (G9715587). I thank to Luciano Fadiga, György Gergely, Teodora Gliga, Antonia Hamilton, Kazuo Hiraki, Pierre Jacob, Mark Johnson, Magda Marton, Dan Sperber, Victoria Southgate, John S. Watson, Daniel Wolpert, and Jennifer Yoon for their valuable comments on earlier versions of this paper.

References

- Baker, C. L., Tenenbaum, J. B. and Saxe, R. R. (2006) Bayesian models of human action understanding. *Advances in Neural Information Processing Systems*, **18**.
- Berenthal, B. I., Longo, M. R. and Kosobud, A. (2006) Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance*, **32**, 210–225.
- Bíró, S. and Leslie, A. M. (2007) Infants' perception of goal-directed actions: development through cue-based bootstrapping. *Developmental Science*, **10**, 379–398.
- Blakemore, S.-J. and Decety, J. (2001) From perception of action to the understanding of intention. *Nature Reviews Neuroscience*, **2**, 561–567.
- Blakemore, S.-J. and Frith, C. (2005) The role of motor contagion in the prediction of action. *Neuropsychologia*, **43**, 260–267.
- Borroni, P., Montagna, M., Cerri, G. and Baldissera, F. (2005) Cyclic time course of motor excitability modulation during the observation of cyclic hand movement. *Brain Research*, **1065**, 115–124.
- Brass, M. and Heyes, C. (2005) Imitation; is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, **9**, 489–495.
- Brass, M., Bekkering, H. and Prinz, W. (2001) Movement observation affects movement execution in a simple response task. *Acta Psychologica*, **106**, 3–22.
- Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G. and Freund, H.-J. (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, **13**, 400–404.
- Byrne, R. W. and Russon, A. E. (1998) Learning by imitation: a hierarchical approach. *Behavioral and Brain Sciences*, **21**, 667–721.
- Call, J., Carpenter, M. and Tomasello, M. (2005) Copying results and copying actions in the process of social learning: chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Animal Cognition*, **8**, 151–163.
- Castiello, U., Lusher, D., Mari, M., Edwards, M. and Humphreys, G. W. (2002) Observing a human or a robotic hand grasping an object: differential motor priming effects. In Prinz, W. and Hommel, B. (eds), *Common Mechanism in Perception and Action. Attention and Performance XIX*, pp. 315–333. Oxford University Press, Oxford.
- Chaminade, T., Meary, D., Orliaguette, J.-P. and Decety, J. (2001) Is perceptual anticipation a motor simulation? A PET study. *Neuroreport*, **12**, 3669–3674.
- Chaminade, T., Meltzoff, A. N. and Decety, J. (2002) Does the end justify the means? A PET exploration of the mechanisms involved in human imitation. *Neuroimage* **15**, 318–328.
- Costantini, M., Galati, G., Ferretti, A., Caulo, M., Tartaro, A., Romani, G.L. and Aglioti, S. M. (2005) Neural systems underlying observation of humanly impossible movements: an fMRI study. *Cerebral Cortex*, **15**, 1761–1767.
- Csibra, G. (unpublished) Goal attribution to inanimate agents by 6.5-month-old infants.
- Csibra, G. and Gergely, G. (1998) The teleological origins of mentalistic action explanations: a developmental hypothesis. *Developmental Science*, **1**, 255–259.
- Csibra, G. and Gergely, G. (2007) 'Obsessed with goals': functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, **124**, 60–78.
- Csibra, G., Gergely, G., Bíró, S., Koós, O. and Brockbank, M. (1999) Goal attribution without agency cues: the perception of 'pure reason' in infancy. *Cognition*, **72**, 237–267.
- Csibra, G., Bíró, S., Koós, S. and Gergely, G. (2003) One-year-old infants use teleological representations of actions productively. *Cognitive Science*, **27**, 111–133.
- Custance, D., Whiten, A. and Fredman, T. (1999) Social learning of an artificial fruit task in capuchin monkeys (*Cepus apella*). *Journal of Comparative Psychology*, **113**, 13–23.

Please note that the information is not complete

- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. and Rizzolatti, G. (1992) Understanding motor events: a neurophysiological study. *Experimental Brain Research*, **91**, 176–180.
- Fadiga, L., Fogassi, L., Pavesi, G. and Rizzolatti, G. (1995) Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, **73**, 2608–2611.
- Falck-Ytter, T., Gredebäck, G. and von Hofsten, C. (2006) Infants predict other people's action goals. *Nature Neuroscience*, **9**, 878–879.
- Ferrari, P. F., Gallese, V., Rizzolatti, G. and Fogassi, L. (2003) Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience*, **17**, 1703–1714.
- Ferrari, P. F., Rozzi, S. and Fogassi, L. (2005) Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *Journal of Cognitive Neuroscience*, **17**, 212–226.
- Flanagan, J. R. and Johansson, R. S. (2003) Action plans used in action observation. *Nature*, **424**, 769–771.
- Fogassi, L. and Gallese, V. (2002) The neural correlates of action understanding in non-human primates. In Stamenov, M. I. and Gallese, V. (eds), *Mirror Neurons and the Evolution of Brain and Language*, pp. 13–35. John Benjamins, Amsterdam.
- Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F. and Rizzolatti, G. (2005) Parietal lobe: from action organisation to intention understanding. *Science*, **308**, 662–667.
- Gallese, V. (2006) Intentional attunement: a neurophysiological perspective on social cognition and its disruption in autism. *Brain Research*, **1079**, 15–24.
- Gallese, V. and Goldman, A. (1998) Mirror neurons and the simulation theory of mind reading. *Trends in Cognitive Sciences*, **12**, 493–501.
- Gallese, V., Fadiga, L., Fogassi, L. and Rizzolatti, G. (1996) Action recognition in the premotor cortex. *Brain*, **119**, 593–609.
- Gallese, V., Fogassi, L., Fadiga, L. and Rizzolatti, G. (2002) Action representation and the inferior parietal lobule. In Prinz, W. and Hommel, B. (eds), *Attention and Performance XIX. Common Mechanisms in Perception and Action*, pp. 334–355. Oxford University Press, New York.
- Gallese, V., Keysers, C. and Rizzolatti, G. (2004) A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, **8**, 396–403.
- Gergely, G. and Csibra, G. (2003) Teleological reasoning about actions: the naïve theory of rational action. *Trends in Cognitive Sciences*, **7**, 287–292.
- Gergely, G. and Csibra, G. (2006) Sylvia's recipe: The role of imitation and pedagogy in the transmission of human culture. In Enfield N.I. and Levinson, S.L. (eds.), *Roots of Human Sociality: Culture, Cognition, and Human Interaction* pp. 229–255. Berg, Oxford.
- Gergely, G., Nádasdy, Z., Csibra, G. and Bíró, S. (1995) Taking the intentional stance at 12 months of age. *Cognition*, **56**, 165–193.
- Gergely, G., Bekkering, H. and Király, I. (2002) Rational imitation in preverbal infants. *Nature*, **415**, 755.
- Hamilton, A., Wolpert, D. and Frith, U. (2004) Your own action influences how you perceive another person's action. *Current Biology*, **14**, 493–496.
- Hamilton, A. F. de C., Joyce, D. W., Flanagan, J. R., Frith, C. D. and Wolpert, D. M. (2007) Kinematic cues in perceptual weight judgement and their origins in box lifting. *Psychological Research*, **71**, 13–21.
- Hommel, B., Müsseler, J., Aschersleben, G. and Prinz, W. (2001) The Theory of Event Coding (TEC): a framework for perception and action planning. *Behavioral and Brain Sciences*, **24**, 849–937.
- Horner, V. and Whiten, A. (2005) Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition*, **8**, 164–181.
- Huang, C-T. and Charman, T. (2005) Gradations of emulation learning in infants' imitation of actions on objects. *Journal of Experimental Child Psychology*, **92**, 276–302.

- Iacoboni, M. (2005) Understanding of others: imitation, language, empathy. In Hurley, S. and Chater, N. (eds), *Perspectives on Imitation: From Mirror Neurons to Memes*, Vol. 1, pp. 77–100. MIT Press, Cambridge, MA.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C. and Rizzolatti, G. (1999) Cortical mechanisms of human imitation. *Science*, **286**, 2526–2528.
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C. and Rizzolatti, G. (2005) Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, **3**, 529–535.
- Jacob, P. and Jeannerod, M. (2005) The motor theory of social cognition: a critique. *Trends in Cognitive Sciences*, **9**, 21–25.
- James, W. (1890) *The Principles of Psychology*. Henry Holt, New York.
- Jeannerod, M. (1994) The representing brain. Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, **17**, 187–245.
- Jeannerod, M., Arbib, M. A., Rizzolatti, G. and Sakata, H. (1995) Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neurosciences*, **18**, 314–320.
- Johnson-Frey, S. H., Maloof, F., Newman-Norlund, R., Farrer, C., Inati, S. and Grafton, S. T. (2003) Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron*, **39**, 1053–1058.
- Kahneman, D. and Treisman, A. (1984) Changing views of attention and automaticity. In Parasuraman, R. and Davies, D. R. (eds), *Varieties of Attention*, pp. 29–61. Academic Press. New York.
- Kamewari, K., Kato, M., Kanda, T., Ishiguro, H. and Hiraki, K. (2005) Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. *Cognitive Development*, **20**, 303–320.
- Keysers, C. and Perrett, D. I. (2004) Demystifying social cognition: a Hebbian perspective. *Trends in Cognitive Sciences*, **8**, 501–507.
- Kilner, J. M., Paulignan, Y. and Blakemore, S. J. (2003) An interference effect of observed biological movement on action. *Current Biology*, **13**, 522–525.
- Kilner, G., Vargas, C., Duval, S., Blakemore, S-J. and Sirigu, A. (2004) Motor activation prior to observation of predicted movement. *Nature Neuroscience*, **7**, 1299–1301.
- Knoblich, G. and Jordan, J. S. (2002) The mirror system and joint action. In M. I. Stamenov and V. Gallese (eds), *Mirror Neurons and the Evolution of Brain and Language*, pp. 115–124. John Benjamins, Amsterdam.
- Koski, L., Wohlschlager, A., Bekkering, H., Woods, R. P., Dubeau, M-C., Mazziotta, J. C. and Iacoboni, M. (2002) Modulation of motor and premotor activity during imitation of target-directed actions. *Cerebral Cortex*, **12**, 847–855.
- Luo, Y. and Baillargeon, R. (2005) Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science*, **16**, 601–608.
- Meltzoff, A. N. (1988) Infant imitation after a one week delay: long term memory for novel acts and multiple stimuli. *Developmental Psychology*, **24**, 470–476.
- Murata, A., Fadiga, L., Figassi, L., Gallese, V., Raos, V. and Rizzolatti, G. (1997) Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology*, **78**, 2226–2230.
- Nishitani, N. and Hari, R. (2000) Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the USA*, **97**, 913–918.
- Onishi, K. H., Baillargeon, R. and Leslie, A. M. (2007) 15-month-old infants detect violations in pretend scenarios. *Acta Psychologica*, **124**, 106–128.
- Oztop, E., Wolpert, D. and Kawato, M. (2005) Mental state inference using visual control parameters. *Cognitive Brain Research*, **22**, 129–151.

- Pacherie, E. and Dovic, J. (2006) From mirror neurons to joint actions. *Cognitive Systems Research*, 7, 101–112.
- Pelphrey, K. A. and Morris, J. P. (2006) Brain mechanisms for interpreting the actions of others from biological motion cues. *Current Directions in Psychological Science*, 15, 136–140.
- Perrett, D.I., Harries, M.H., Bevan, R., Thomas, S., Benson, P.J., Mistlin, A.J., Chitty, A.J., Hietanen, J.K. and Ortega, J.E. (1989) Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology*, 146, 87–113.
- Pobric, G. and Hamilton, A. F. de C. (2006) Action understanding requires the left inferior frontal cortex. *Current Biology*, 16, 424–429.
- Press, C., Bird, G., Flach, R. and Heyes, C. (2005) Robotic movement elicits automatic imitation. *Cognitive Brain Research*, 25, 632–640.
- Prinz, W. (2006) What re-enactment earns us. *Cortex*, 42, 515–517.
- Ramnani, N. and Miall, R. C. (2004) A system in the human brain for predicting the actions of others. *Nature Neuroscience*, 7, 85–90.
- Rizzolatti, G. and Craighero, L. (2004) The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192.
- Rizzolatti, G., Fogassi, L. and Gallese, V. (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Review Neuroscience*, 2, 661–670.
- Saxe, R. (2005) Against simulation: the argument from error. *Trends in Cognitive Sciences*, 9, 174–179.
- Scholl, B. J. (2005) Innateness and (Bayesian) visual perception: reconciling nativism and development. In Carruthers, P., Laurence, S. and Stich, S. (eds), *The Innate Mind: Structure and Contents*, pp. 34–52. Oxford University Press, New York.
- Sebanz, N., Bekkering, H. and Knoblich, G. (2006) Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, 10, 70–76.
- Slaughter, V., Heron, M. and Sim, S. (2002) Development of preferences for the human body shape in infancy. *Cognition*, 85, B71–B81.
- Sodian, B., Schoeppner, B. and Metz, U. (2004) Do infants apply the principle of rational action to human agents? *Infant Behavior and Development*, 27, 31–41.
- Southgate, V., Johnson, M. H. and Csibra, G. (unpublished) Infants attribute goals to biologically impossible actions.
- Subiaul, F., Cantlon, J. F., Holloway, R. L. and Terrace, H. S. (2004) Cognitive imitation in rhesus macaques. *Science*, 305, 407–410.
- Tai, Y. F., Scheffler, C., Brooks, D. J., Sawamoto, N. and Castiello, U. (2004) The human premotor cortex is 'mirror' only for biological actions. *Current Biology*, 14, 117–120.
- Tennie, C., Call, J. and Tomasello, M. (2006) Push or pull: imitation versus emulation in human children and great apes. *Ethology*, 112, 1159–1169.
- Tomasello, M. (1996) Do apes ape? In Heyes, C. M. and Galef, B. G. (eds), *Social Learning in Animals: The Roots of Culture*, pp. 319–346. Academic Press, New York.
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C. and Rizzolatti, G. (2001) I know what you are doing: a neurophysiological study. *Neuron*, 32, 91–101.
- Wagner, L. and Carey, S. (2005) 12-month-old infants represent probable ending of motion events. *Infancy*, 7, 73–83.
- Williamson, R. A. and Markman, E. M. (2006) Precision of imitation as a function of preschoolers' understanding of the goal of the demonstration. *Developmental Psychology*, 42, 723–731.
- Wilson, M. and Knoblich, G. (2005) The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131, 460–473.
- Whiten, A. and Custance, D. (1996) Studies of imitation in chimpanzees and children. In Heyes, C. M. and Galef, B. G. (eds), *Social Learning in Animals: The Roots of Culture*, pp. 347–370. Academic Press, New York.

- Wohlschläger, A., Gattis, M. and Bekkering, H. (2003) Action generation and action perception in imitation: an instance of the ideomotor principle. *Philosophical Transactions of The Royal Society of London, B*, 358, 501–515.
- Wolpert, D. M. and Ghahramani, Z. (2000) Computational principles of movement neuroscience. *Nature Neuroscience*, 3 (Suppl.), 1212–1217.
- Wolpert, D. M., Doya, K. and Kawato, M. (2003) A unifying computational framework for motor control and social interaction. *Philosophical Transactions of The Royal Society of London, B*, 358, 593–602.
- Zentall, T. R. (1996) An analysis of imitative learning in animals. In Heyes, C. M. and Galef, B. G. (eds), *Social Learning in Animals: The Roots of Culture*, pp. 221–243. Academic Press, New York.

