

Exploring the brain basis of joint action: Co-ordination of actions, goals and intentions

Roger D. Newman-Norlund , Matthijs L. Noordzij , Ruud G. J. Meulenbroek & Harold Bekkering

To cite this article: Roger D. Newman-Norlund , Matthijs L. Noordzij , Ruud G. J. Meulenbroek & Harold Bekkering (2007) Exploring the brain basis of joint action: Co-ordination of actions, goals and intentions, Social Neuroscience, 2:1, 48-65, DOI: [10.1080/17470910701224623](https://doi.org/10.1080/17470910701224623)

To link to this article: <https://doi.org/10.1080/17470910701224623>



Published online: 21 Feb 2007.



Submit your article to this journal [↗](#)



Article views: 563



View related articles [↗](#)



Citing articles: 4 View citing articles [↗](#)

Exploring the brain basis of joint action: Co-ordination of actions, goals and intentions

Roger D. Newman-Norlund

Nijmegen Institute for Cognition and Information, and F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

Matthijs L. Noordzij

F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

Ruud G. J. Meulenbroek

Nijmegen Institute for Cognition and Information, Nijmegen, The Netherlands

Harold Bekkering

Nijmegen Institute for Cognition and Information, and F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

Humans are frequently confronted with goal-directed tasks that can not be accomplished alone, or that benefit from co-operation with other agents. The relatively new field of social cognitive neuroscience seeks to characterize functional neuroanatomical systems either specifically or preferentially engaged during such joint-action tasks. Based on neuroimaging experiments conducted on critical components of joint action, the current paper outlines the functional network upon which joint action is hypothesized to be dependant. This network includes brain areas likely to be involved in interpersonal co-ordination at the action, goal, and intentional levels. Experiments focusing specifically on joint-action situations similar to those encountered in real life are required to further specify this model.

INTRODUCTION

Human beings possess a remarkable capacity to work co-operatively towards shared goals. Such a capacity is critical to the survival of social species in which individuals rely on one another to accomplish tasks that would be impossible to complete in isolation. Although possession of this faculty is not limited to humans; ants collaborate to build immense dwellings, and wolves and dolphins co-operatively hunt in packs (Shane, Wells, & Worsig, 1986; Nudds, 1978; Bonabeau,

Theraulaz, Deneubourg, Aron, & Camazine, 1997); it has been suggested that the ability to simultaneously co-ordinate co-operative behaviors at higher representational levels is so limited (Tummolini, Castelfranchi, Pacherie, & Docik, 2006). Indeed, humans seem able to effortlessly align their *actions*, *goals*, and *intentions* with other humans during social interactions. And it is this alignment that makes possible the smooth performance of innumerable joint actions, such as passing scissors to a child, helping build a house, or taking part in a drug raid, that we often take

Correspondence should be addressed to: Roger D. Newman-Norlund, Radboud University Nijmegen, NICI, NL-6500 HE, Nijmegen, The Netherlands. E-mail: rogern@nici.ru.nl

The present study was supported by the EU-Project "Joint Action Science and Technology" (IST-FP6-003747).

for granted but that significantly increase the odds of our survival in an inherently social world.

The main goal of the present report is to delineate the subset of brain areas likely to support the performance of joint action on the basis of relevant neuroimaging experiments. As a first step in this process, it is imperative to agree on what qualifies a specific interaction as a “joint action.” This very question has been the topic of serious scientific debate for some time. Because it defines joint action from the viewpoint of the individual in the same way that most neuroimaging research focuses on individual behaviors, we have chosen, in the present paper, to base our definition of joint action on Bratman’s concept of shared co-operative activities (SCAs; 1992). Additionally, Bratman’s discussion of SCAs touches upon all three levels at which mutually acting agents are presumed to integrate their behaviors, i.e., (1) intention, (2) goal, and (3) action. According to his definition, SCAs require that: (1) the interactants be mutually responsive to each other; (2) that both actors be committed to the joint activity; and (3) that the actors co-ordinate their plans of action and intentions in a way that requires understanding of their respective roles (role reversal). While the second element, which requires that interactants share a commitment to working together, is perhaps not that interesting to the present paper, the first and third elements of this definition correspond nicely to *co-ordination of actions* and *co-ordination of intentions*, respectively. Within his definition, Bratman (1992) also refers to the importance of integrating behavior through the *meshing of subplans*, which can alternatively be read as the *co-ordination of goal-subgoal* hierarchies. Although different researchers have chosen to focus on different definitions of joint action in formulating their own models/theories of joint action (Knoblich & Jordan, 2003; Tummolini et al., 2006), we believe: (1) that a complete model of joint action must minimally account for co-ordination at the levels of action, goal, and intention; and (2) that Bratman’s definition of joint action most closely approximates this requirement. A discussion of the possible roles of specific brain areas in cognitive processes implied by these three levels of co-ordination will be the major focus of the current paper.

Although joint action, per se, has been the subject of few neuroimaging experiments, many of the basic cognitive processes that are likely to support it are reasonably well understood.

A review of such experiments allows us here to specify a subset of brain areas likely to be involved preferentially during the execution of joint actions as opposed to similar solo actions (Figure 1). This subset includes brain areas that: (1) are specially sensitive to human form and biological motion; (2) represent and monitor self and other generated actions; (3) mediate the interaction of action observation and action execution; (4) represent and integrate representations at the level of physical (object), action (functional) and mental goals (Bekkering & Wohlschläger, 2002); (5) separate self and other generated actions and their consequences; (6) respond to unique attentional demands of joint-action engagements; and (7) support reasoning processes unique to joint-action situations. Of course, different specific instances of joint action will rely on each of these cognitive processes, and their corresponding functional substrates, to different extents.

PERCEIVING OTHERS

During joint action it is often beneficial to simultaneously co-represent the actions of another person along with your own. In the context of a normal human action, this co-representation focuses primarily on monitoring body parts and their motion paths. Such co-representations are especially important in joint-action situations in which close temporo-spatial co-ordination is required for success. For example, when performing a spicy salsa, it is important to monitor your own actions, as well as those of your partner. Such co-representations would allow one dancer to adjust his or her performance, specifically his or her immediate actions, goals, or intentions, in order to accommodate errors or eccentricities in the performance of their partner.

There is now considerable evidence that a specific network of brain areas differentiates human bodies moving according to biological constraints from the rest of the environment (Grossman & Blake, 2002). This area includes the extrastriate body part area (EBA), which responds to human body parts, as well as the superior temporal sulcus (STS), which responds to both biological motion and the intentions of human agents. Downing, Bray, Rogers, and Childs (2004) have suggested that “stimulus categories [such as human body parts] favored by a focal, selective cortical area will tend to be selected [for

attention] with higher priority compared to other object classes" (p. 29). In joint-action situations, situations typically involving interaction with the specific stimulus category of other humans, we will argue that these two areas may help jointly acting individuals focus attention and neuronal resources in a co-ordinated manner optimal to achieving alignment of actions and action goals. It is our claim that activation in these areas may relate to important facets of joint action that, critically, may exist to differing degrees in different tasks, and that no functional model of joint action would be complete without.

Let us first examine the EBA. This area is located adjacent to motion sensitive area V5 and preferentially responds to human body images represented as silhouettes, photographs, line drawings, or stick figures (Downing, Jiang, Shuman, & Kanwisher, 2001). In an experiment by Downing and colleagues (2001), the response of the EBA was strongest to whole bodies, weaker for hands, and weakest for control objects and faces. These data suggest that the human form is processed by dedicated neural tissue within the human cortex. Critical to our understanding of the role of the EBA in joint action is a neuroimaging experiment conducted by Astafiev, Stanley, Shulman, and Corbetta (2004). This experiment demonstrated that activity in the EBA, which, based on previous experiments, was expected to be solely influenced by visual presentation of body parts, was in fact significantly modulated by movement execution. The authors tentatively concluded that the EBA may be part of a system responsible for both action and perception of action, a finding that found further support in an experiment that suggested that EBA activation is modulated by motor control (Hamilton, Wolpert, Frith, & Grafton, 2006). The fact that the EBA specifically processes body parts, and that this processing is modulated by motor execution is critical to understanding any role it may play in joint action. Imagine, for example, a situation in which an actor is simultaneously executing actions and monitoring the actions of another actor (remember the salsa dancer). In situations requiring close temporal co-ordination, in which the movements of the actors are highly coupled, a brain mechanism capable of relating the two sets of movements would be particularly useful. Although more research must be done to test this theory, we suggest that the EBA possesses basic response characteristics fitting with this capability and may

represent the brain basis of our ability to relate the movements of others directly with our own movements in joint-action situations. Note, that this ability should not generalize to other types of interactions, as the EBA responds only to observation of objects critical to human social interactions, namely human body parts.

There is now a large literature concerning the role of the STS in the processing of biological motion stimuli. The core finding of this research is that a variety of biological motion stimuli consistently activate a posterior region (or surrounding regions) of the STS. These stimuli include moving eyes, hands, mouths, and bodies (Bonda, Petrides, Ostry, & Evans, 1996; Grossman et al., 2000; Howard, Brammer, Wrigh, Woodruff, Bullmore, & Zeki, 1996; Puce, Allison, Bentin, Gore, & McCarthy, 1998; Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001), as well as reduced human stimuli known as point-light figures (Beauchamp, Lee, Haxby, & Martin, 2003; Giese & Poggio, 2003; Grèzes, Fonlupt, Bertenthal, Delon-Martin, Segebarth, & Decety, 2000; Grossman & Blake, 2002; Jellema & Perrett, 2002; Puce & Perrett, 2003; Thompson, Clarke, Stewart, & Puce, 2005). Consistent with its general role in the processing of biological motion, right STS activation occurs when observing both goal-directed and non-goal-directed movements. However, it is also true that the STS activation is stronger when viewing incorrect as opposed to correct movements (Pelphrey, Morris, & McCarthy, 2004). This suggests that the STS may be sensitive to the intentions of observed agents. This hypothesis finds further support in an experiment showing right STS involvement in the attribution of animacy to moving shapes (Schultz, Friston, O'Doherty, Wolpert, & Frith, 2005). Also significant to the STS's role in joint action is the fact that the response/sensitivity of STS neurons to specific types of biological motion presentations can be enhanced by learning (Grossman, Blake, & Kim, 2004). This may allow for specific types of biological movements to "earn" preferential access to brain resources through experience. Thus, the STS is sensitive to biological motion, the intentions of others, and can be trained over time.

On the basis of these findings, we suggest that the STS is a critical component of the brain basis of joint action. Specifically, we suggest that it is ideally placed to gate information into the rest of the cortex during joint-action engagements. For example, information concerning stimuli with biological characteristics may be preferentially

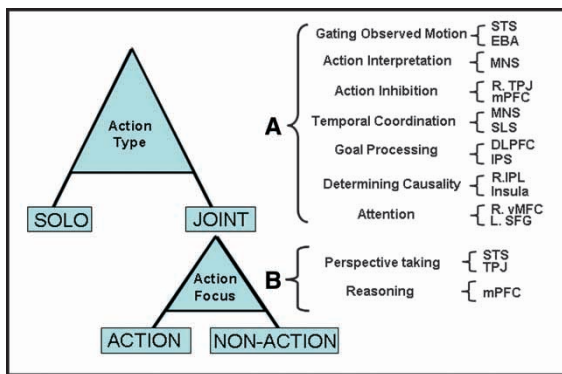


Figure 1. Graphical representation of the dimensions along which brain activation patterns may be differentiated based on the requirements of a given joint-action situation. (A) It is suggested that joint action (relative to solo action) will result in greater activation within these brain areas to the extent that each cognitive process is required. (B) Currently, the role of mentalizing in joint action is unclear. It is suggested that action-oriented and non-action-oriented situations may rely differentially on brain areas involved in perspective taking and reasoning. (STS=superior temporal sulcus, EBA=extrastriate body area, MNS=mirror neuron system—bilateral inferior parietal lobe, bilateral inferior frontal gyrus, SLS=sequence learning system—including parietal cortex, prefrontal cortex, anterior striatum, anterior cerebellum, pre-supplementary motor area and supplementary motor area, R.TPJ=right temporoparietal junction, aMFC=anterior medial frontal cortex, DLPFC=dorsolateral prefrontal cortex, IPS=intraparietal sulcus, R.IPL=right inferior parietal lobe, R.vMFC=right ventral medial frontal cortex, L.SFG=left superior frontal gyrus).

granted access to parietal and premotor “mirror areas” where it could be used in calculations related to action planning (Csibra, 2005). Alternatively, such information could be shunted to prefrontal areas that may be critical for mentalizing about the goals and intentions of others (see reasoning section). The fact that STS neurons can learn over time provides a powerful mechanism whereby actors participating in a novel joint-action task may be able to improve their performance/teamwork by selectively tuning their resources to different instances of observable behavior produced by their partner. Taken together, these data suggest that activation in the EBA and STS may be higher during performance of joint actions as opposed to solitary actions (Figure 1).

ACTION OBSERVATION AND MIRROR NEURONS

Brain representations of self-generated and observed actions, though traditionally thought of as

separate, are now believed to be largely overlapping. The basis of this claim is the rapidly growing corpus of research describing the properties of “mirror neurons” in the premotor cortex (Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Ferrari, Gallese, Rizzolatti, & Fogassi, 2003; Ferrari, Rozzi, & Fogassi, 2005; Rizzolatti, Fadiga, Fogassi, & Gallese, 1996; Rizzolatti, Fogassi, & Gallese, 2001) and the parietal cortices (Fogassi, Gallese, Fadiga, & Rizzolatti, 1998). These neurons are significant because they respond during execution of goal-directed actions as well as during perception of the same goal-directed action being executed by another agent. It has been suggested that mirror neurons provide a common substrate for action observation and execution that supports our ability to infer the goals and intentions of others (Fogassi et al., 2005; Iacoboni et al., 2005).

Over the past decade scientists have conducted a number of important experiments identifying key characteristics of mirror neurons that support this hypothesis. Apparently there exist multiple subtypes of mirror neurons in the inferior frontal and inferior parietal areas, with response properties that depend on a number of factors including the used effector, presence/absence of an object during actions (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996), whether an action is presented live or via video (Ferrari et al., 2003), and whether an action is meaningful (Gallese et al., 1996). For example, “broadly congruent” mirror neurons may respond to non-identical observed and executed actions such as (1) observation of an object being placed on a table and (2) execution of bringing that object towards the mouth (Di Pellegrino et al., 1992). One argument regarding the response properties of this class of mirror neurons is that they respond to all action components functionally related to a similar goal state (i.e., eating; Fogassi & Gallese, 2002; Rizzolatti et al., 2001). Together, these data suggest a role of human mirror areas in goal processing that roughly corresponds to the second facet of joint action in the definition adopted in this paper. The idea that mirror neurons function to recognize action goals, as opposed to specific actions, is also supported by the fact that there is generally a weak congruence between the motor and perceptual properties of these neurons. The percentage of mirror neurons with a strict one–one motor–visual congruence has been estimated at only one third (Csibra, 2005; Rizzolatti & Craighero, 2004). This means that the other two thirds of mirror

neurons do something other than directly match observed and executed actions. In addition to this circumstantial evidence, a number of experiments have directly demonstrated that some mirror neurons are sensitive to goals. For example, select mirror neurons located in the monkey inferior parietal lobule are responsive to the goal of an action as opposed to the action itself (Fogassi et al., 2005). And fMRI research on human subjects has demonstrated that a specific site in the left anterior inferior parietal sulcus is sensitive to the functional goals of grasping actions but not to the movements themselves (Hamilton et al., 2006). Lastly, Umiltà et al. (2001) showed that some mirror neurons fire at the presumed goal of an action even if the goal is occluded. This suggests that mirror neurons may support the ability to simulate the outcomes of ongoing events. Importantly, this sensory input into these mirror neurons need not be visual, as a new class of mirror neurons, responsive to both actions and sounds associated with them, has been described by other researchers (Gazzola, Aziz-Zadeh, & Keysers, 2006; Keysers et al., 2003; Kohler et al., 2002). That mirror neurons appear to code all manner of sensory stimuli associated with a specific goal has led to the hypothesis that they are critical multi-modal mediators of the goal-representations governing our behavioral space (Metzinger & Gallese, 2003). This function of mirror neurons is precisely what makes them critical to joint-action situations and constitutes our theoretical justification for inclusion of human mirror areas in the proposed functional model of joint action. Human brain areas exhibiting “mirror like” properties currently include the left inferior frontal cortex (Brodmann areas 44 and 45), the proposed human homologue of monkey area F5, and the left inferior parietal cortex (Buccino et al., 2004a; Decety et al., 1997; Grèzes, Costes, & Decety, 1998; Iacoboni et al., 1999; Koski et al., 2002; Manthey, Schubotz, & von Cramon, 2003; Rizzolatti & Craighero, 2004; Rizzolatti et al., 1996).

A recent computational model of joint action proposed by Cuijpers, van Schie, Koppen, Erlhagen, and Bekkering (2006) makes the argument that joint action can only be understood if goals and goal processing are given a central role. Indeed, it has been argued that humans can not help but interpret observed actions in terms of their functional significance (Csibra & Gergely, 2007). But precisely how and why might goals and goal processing be differentially relied upon in

joint-action situations? To date, much has been made of the possibility that mirror neurons support our ability to simulate the future states of observed actions, and that this ability permits us to deduce the goals of other actors. Of course this could be critical for joint-action tasks in which goals are constantly in flux, and alignment of the goal hierarchies of two or more actors is critical to success. For example, when running down the soccer field with a partner, it is helpful to be able to deduce the goal of your teammate based on their actions and your previous experience with them. This will allow you to be in the right position at the right time (i.e., your goal) to receive the ball and continue the drive towards the net. As another example, imagine that you are working with a partner to build a model airplane. In order to assist in this task, it is important to know your partner’s current goal (i.e., building a wing). By monitoring your partner’s actions in the context of this presumed goal you set your system up to detect unexpected events (i.e., grasping a part of the fuselage) that can then lead to a reconsideration of your partner’s goal (i.e., they are attaching the broadest parts of the wing to the fuselage first). Csibra and Gergely (2007) have characterized “action-to-goal” and “goal-to-action” inferences as examples of the inverse problem. Making these inferences is only possible because we make assumptions about the world and which actions should lead to certain goals. Although this type of processing may apply to all sorts of dynamic event perception (Wilson & Knoblich, 2005), it clearly plays a vital role in joint action. We suggest that the human mirror system’s capacity to simulate the consequences of observed actions (Umiltà et al., 2001) plays a critical role in both “action-to-goal” and “goal-to-action” inferences. Additionally, we highlight the importance of “goal-to-action” inferences in particular, as this seems especially relevant for joint actions.

A question critical to our understanding of the role of mirror neurons in joint action concerns the malleability of their response properties. Preliminary data suggest that the response properties of mirror neurons can, indeed, be changed. For example, after extensively training a monkey on the use of a novel tool, a new subclass of “tool-responding mirror neurons” arises in its premotor cortex (Ferrari et al., 2005). These results suggest that, through experience-based learning, the mirror-neuron system may function to extend our action-understanding capacities (Ferrari et al.,

2005). Such a function could be critical when learning specific actions associated with specific items encountered in novel joint-action situations. It could, for instance, support the development of efficient co-ordination strategies in joint-action tasks involving man-made implements such as drills, wrenches, and hammers. Motor-facilitation effects normally produced by observation of actions performed by other people, could, via this learning mechanism, be extended into the technological realm in which an increasing number of examples of joint action can easily be found.

A final characteristic of mirror neurons that seems particularly relevant when considering joint-action tasks is the powerful influence of the degree of perceived similarity between actor and observer on their activity. An fMRI experiment conducted by Grèzes, Frith, and Passingham (2004) revealed that mirror neuron areas including premotor and inferior parietal cortices are more quickly activated when watching movies of oneself performing an action than when watching movies of other actors performing a similar action. Additionally, Buccino et al. (2004a) found that these mirror areas were activated when observing biting or silent speech reading, but not when observing a dog barking (Buccino, Binkofski, & Riggio, 2004b). Motor areas are not activated by biologically impossible movements such as an arm rotating past its range (Stevens, Fonlupt, Schiffrar, & Decety, 2000), nor do grasping movements performed by robots activate one's motor system (Gallese et al., 1996; Kilner, Paulignan, & Blakemore, 2003; Tai, Scherfler, Brooks, Sawamoto, & Castiello, 2004). Based on the assumption that mirror-neuron activity is important for joint action, the perceived similarity of co-actors should influence the efficiency of their working relationship and be measurable via specific performance parameters. In other words, these findings lead directly to specific, testable predictions regarding the ability of particular types of dyads (i.e., robot-human, and male-female) to efficiently perform joint-action tasks.

Based on these known properties of mirror neurons, we suggest that brain areas exhibiting mirror like properties should be more active during joint action than during similar solitary actions (see Figure 1). These neurons represent relationships between observed and executed actions, and their activity is modulated by the degree to which actor and observer are perceived

as similar. Additionally, their response properties are highly malleable. This allows for the possibility that they support learning in joint-action situations. Experiments contrasting activation within the human mirror neuron system (MNS) during joint and solo action as well as experiments designed to measure learning related changes in brain activity associated with performance gains in these two situations may address these questions.

TEMPORAL CO-ORDINATION

Clearly, certain joint-action situations are successful only to the extent that co-actors are able to achieve close temporal co-ordination of their actions. For example, carrying a piano up a narrow staircase requires both actors to quickly and accurately compensate for tilting errors induced by their partner. Similarly, dance partners performing an American Smooth must be able to co-ordinate their actions with those of their partner in order to minimize dysfluencies that could negatively impact the judges' opinion of their technical or artistic abilities. In some circumstances, it is sufficient to make such adaptive movements in reaction to observed errors. These reactive adaptive movements may be mediated by mental simulations performed by the MNS. Based on their observations of errors in the current state of the system/partner, such simulations might allow partners to predict the consequent behavior of their partner and adjust appropriately.

However, the real power of our ability to perform rapid adaptive movements may arise out of our ability to predict future behavior on the basis of past information. Such adaptation goes beyond the reactive sort discussed above in that it requires the detection of errors prior to their actual occurrence. It is currently unclear whether or not the MNS supports this ability. If one adopts the view that mirror neurons are not specialized to enable imitation, but rather govern specific categories of associative learning involving interactions between observed and executed actions (Brass & Heyes, 2006), then one may indeed hypothesize that mirror neurons are involved in this function. Specifically, they may be capable of linking observation of particular movement errors with execution of actions designed to remedy disturbances that these errors introduce into the system. For example, when a

male dancer observes that his partner has mis-spotted on one of her spins he may automatically engage a specific motor plan (i.e., guiding her back into position with his hand) in order to get his partner back on track.

Although mirror neurons possess all the response properties necessary to guide temporal co-ordination—they can link observed and executed actions (Rizzolatti & Craighero, 2004), and their response properties can be changed through experience (Ferrari et al., 2005)—it is hard to ignore the possibility that a second possible functional conglomerate may support the establishment and maintenance of close temporal co-ordination during joint action. Specifically, we refer to the sequence learning system (SLS). This can be further subdivided into brain systems devoted to the implicit and explicit acquisition of sequences. Both may be relevant to the study of joint action, as jointly acting individuals may improve their temporal co-ordination on the basis of explicit knowledge regarding their partner (i.e., “my partner always moves left when I move right”) or implicit knowledge, a type of learning that allows individuals to gain knowledge through experience rather than relying on recall of explicit rules. The results of numerous lesion and neuroimaging experiments agree that a specialized cortical/subcortical network supports our capacity to explicitly and implicitly learn complex spatial and motor sequences (Grafton, Hazeltine, & Ivry, 1995, 1998; Jackson, Jackson, Harrison, Henderson, & Kennard, 1995; Pascual-Leone et al., 1993; Willingham & Koroshetz, 1993). According to the model proposed by Hikosaka, Nakamura, Sakai, and Nakahara (2002), knowledge governing observed spatial sequences are usually explicitly learned and place high demands on the attentional system. Acquisition of such visual sequences involves the associative portions of the basal ganglia and cerebellum, as well as prefrontal and parietal cortices. Regarding our present focus on joint action, this spatial learning system seems a good candidate for the acquisition of patterns (based on action observation) that might support temporal co-ordination in jointly acting dyads. Discernable action patterns, linked to specific motor responses via mechanisms housed in the supplementary and pre-supplementary motor areas, might provide the functional basis of temporal co-ordination in jointly acting dyads (see Hikosaka et al., 2002). Unfortunately, the involvement of these brain areas in action-

observation-based learning of the sort encountered in joint action remains entirely unexplored.

Currently, there is no way to determine whether the MNS or the sequence learning system primarily supports tight temporal co-ordination that characterizes some joint-action situations (see Figure 1). However, this issue could be easily addressed in future experiments. For example, longitudinal neuroimaging experiments in which jointly acting individuals gradually acquire the ability to accurately predict their partner's actions farther and farther into the future might provide insight into this question. Additionally, comparison of functional activations during participation in joint action that either does or does not require close temporal co-ordination might be used to address this issue.

BEYOND IMITATION

The idea that action observation leads to a default tendency to imitate is supported by some experimental data. Children have been shown to spontaneously imitate all sorts of actions including facial expressions and lip movements (Field, Woodson, Greenberg, & Cohen, 1982). In adults, Maeda, Kleiner-Fisman, and Pascual-Leone (2002) showed that observation of hand movements results in a decrease in the amount of brain stimulation required to induce a hand movement in the observer. This modulation of motor excitation unfolds over time in synchrony with perceived movements (Gangitano, Mottaghy, & Pascual-Leone, 2001), occurs without effort, and is independent of planning (Vogt, Taylor, & Hopkins, 2003).

A recent theory regarding joint action posits that interaction with other agents often involves the selection of a non-imitative (“complementary”) as opposed to an imitative response tendency (Van Schie & Bekkering, 2007). For example, when someone hands you a cup using the handle (using a precision grip), the logical reaction is to receive the cup via its body (using a power grip). In such cases, it is not beneficial to imitate the observed action. Rather, a complementary action, specifically designed to make joint action possible, is appropriate. Although it is currently unclear what mechanisms support our ability to perform complementary action, one possibility is that the natural tendency towards action imitation must be inhibited in such cases. So, then, what is the brain basis of this natural

imitative tendency, and what brain areas are likely to be involved in its suppression during joint action?

Functional magnetic resonance imaging (fMRI) experiments designed to study general inhibitory mechanisms (Garavan, Ross, Murphy, Roche, & Stein, 2002; Hester et al., 2004) suggest that right prefrontal regions play a critical role in response inhibition. Further evidence on brain mechanisms subserving imitative inhibition comes from studies by Brass and colleagues (Brass, Zysset, & von Cramon, 2001; Brass, Derrfuss, & von Cramon, 2005). In a series of experiments using a paradigm devised to examine the inhibition of imitative actions (finger movements), Brass and colleagues found evidence that brain tissue in the anterior frontal-medial cortex and the right temporo-parietal junction is involved in inhibition of imitative actions (Brass et al., 2001, 2005). These data suggest that activation in inferior frontal and right parietal areas plays an important role in joint-action situations requiring inhibition of imitative responses (see Figure 1).

Although the production of non-imitative complementary actions common to joint-action encounters might involve inhibition of brain areas normally involved in imitation, another possibility is that these actions are supported by the mirror neuron system. How might this be accomplished? One possibility is that “broadly congruent” mirror neurons, which constitute some 60% of mirror neurons and respond to non-identical observed and executed actions (Fogassi & Gallese, 2002), effectively link observed actions to appropriate complementary action responses. Future experiments might test this hypothesis by examining the functional correlates of imitative and complementary actions performed during joint-action engagements.

GOAL REPRESENTATIONS

A critical aspect of some theories of joint action is that goals are critical mediators of perception–action linkages. Behavioral experiments have now demonstrated that goals may be more important than means (Van Schie & Bekkering, 2007) in determining the dynamics of joint-action engagements. The typical finding of these experiments is that movements cued by action goals result in faster reaction times than movements cued by action means.

Currently, it is unclear precisely how goals (be they object based, functional, or intentional) are represented in the human brain. However, one experiment has examined this issue in a rather clever manner. Using a task that parametrically varied the intensity of goal-processing demands using a Tower of Hanoi task, Fincham, Carter, van Veen, Stenger, and Anderson (2002) found that goal-processing intensive trials elicited greater activation in a number of brain sites. First, they found that activation in the right dorsolateral prefrontal cortex (DLPFC) was greater during goal-intensive (requiring a higher number of subgoals to achieve the final goal) trials. Although there was a possible confound with working-memory load in this task, it is consistent with an idea shared by many researchers: that the prefrontal cortex is a crucial component of the goal-processing network (Hasselman, 2005; Metzinger & Gallese, 2003). In addition to the prefrontal cortex, Fincham and colleagues also found that goal intensity was positively correlated with brain activity in left inferior parietal and bilateral premotor cortices. Interestingly, a recent experiment conducted by Hamilton and Grafton (2006) demonstrated that physical, object related goals may be processed by a specific site in the left intraparietal sulcus. Although, in this experiment, the goal of the action was directly related to the object involved, making it difficult to rule out the possibility that activations were related to action-object affordances, the location of one of the signal peaks, specifically the left middle anterior intraparietal activation, overlaps with the goal-intensity modulated site reported by Fincham and colleagues (2002). Taken together, these imaging data support the conclusion that goals, or at least some aspects of goals (such as the complexity of the subgoal tree), are represented in the prefrontal, parietal, and premotor cortices (see Figure 1).

The experiment involving the brain correlates of goal intensity mentioned above is particularly relevant to our discussion of how brain activation might differ between solo and joint action. In situations where two people are working together to accomplish a task, the case often arises that the immediate action goals of the two participants differ. To the extent that one partner is required to/benefits from forming a simultaneous brain representation of both their own goal and the goal of their partner, we might expect greater activation at sites shown to process goal intensity, such as the right DLPFC and left intraparietal sulcus. Cur-

rently, the interaction between these brain areas, as well as their precise contributions to goal representations, are not entirely clear. Additional imaging experiments, for instance event-related fMRI experiments that image human brains during joint-action tasks involving the simultaneous monitoring of multiple-action goals, may allow us to more accurately describe the role of these brain areas in both goal processing and joint action.

A SENSE OF AGENCY

Another line of research relevant to the brain basis of joint action deals with our sense of agency, or how we perceive the effect of our actions on the environment. In any goal-directed situation, it is imperative that an actor has a clear understanding concerning this relationship. This may be especially true in goal-directed situations involving multiple participants. In such cases the actors must simultaneously monitor their own effect on task relevant variables as well as the effects exerted by their partner. Only by successfully disambiguating the relative contributions of each other and by systematically attributing the proper result to the proper agent, can actors accurately and efficiently decide what to do next.

Data obtained from patient and brain-imaging experiments strongly implicate the right inferior parietal cortex in processes responsible for the human sense of agency (see Figure 1). For example, schizophrenic patients' delusions of "alien control" ("The book opened itself, I didn't have anything to do with it.") are to be mediated by activation in the right parietal lobe (Spence et al., 1997). This area appears to be sensitive to the attribution of movements to external agents.

Ruby and Decety (2001) found that activation in this area was greater when participants imagined someone else making a movement as compared to imagining themselves making the same movement. McGuire, Silbersweig, and Frith (1996) found that activation in this area was greater when listening to an unfamiliar voice versus one's own voice; Farrer and Frith (2002) found increased right inferior parietal activation when participants drove a computer graphic of a car but were told that someone else was actually controlling it. With myriad confirmatory results from other experimenters (Blakemore, Oakley, & Frith, 2003; Blanke & Arzy, 2005; Farrer et al., 2003), there is growing consensus in the field that the right parietal cortex is critical to the sense of agency

and self-identification (Jeannerod & Pacherie, 2004). Jeannerod (1999) has suggested that this function of the parietal cortex may arise as it remaps actions performed by others onto one's own internal models of these movements.

A relatively new finding is that assigning oneself as the cause of an observed change in the environment preferentially activates the anterior and posterior insular cortex (Farrer et al., 2003; Farrer & Frith, 2002). Farrer et al. (2003) scanned subjects while parametrically varying the degree of similarity between an actual hand movement and movement of a virtual hand seen via a monitor but rotated by 0, 25, or 50 degrees. This manipulation led to a parametric modulation of insular activation such that activation increased with decreasing discrepancy between observed and executed actions. According to Farrer and Frith (2002), information critical to the determination of causality of one's actions converges in the insula. This information includes somatosensory feedback signals, auditory and visual information about the state of the environment, and motor feedback (efference signals). These three types of information are thought to be critical for attributing causality to one's actions.

What is clear from these data is that activation in the right inferior parietal lobule appears to be negatively correlated with a sense of agency, while activation in the insula seems to be positively correlated with feelings of control. The right inferior parietal lobe may represent the remapping of observed actions onto our own internal model based on pure visuo-spatial mapping, while the insular activation may determine agency based on multimodal information. A major limitation of these data is that the interaction between activation in the right inferior parietal cortex and the insula remains unclear. One possibility is that activation in the right inferior parietal lobe is somehow linked to activation in the insula. Alternatively, activation in these areas may be mediated by inputs from a separate brain region. Future experiments might clarify this relationship. Regardless of the specifics, current data predict that both right inferior parietal and insular cortices should be recruited during joint action but not necessarily during similar solo action.

ATTENTIONAL CONSIDERATIONS

It is clear that the dynamics of joint-action situations place extra and perhaps unique

demands on our attentional system. For example, simultaneously monitoring your own actions and the actions of another agent requires greater effort than simply monitoring your own actions. Such additional effort should lead to increased demands on relevant attention networks. Joint action also introduces unique types of attentive processes. For example, during joint action it is often beneficial to selectively attend to stimuli that move in accordance with the biological constraints of your own body (such as eyes, lips, heads, and hands). Additionally, joint action may require co-actors to simultaneously attend to multiple sets of movements generated both by oneself and other relevant participants. And if the processing of self- and other-generated movements relies on partially similar networks (i.e., the human MNS), it is possible that the simultaneous directing of attention towards multiple sets of human movements (i.e., divided attention) requires different brain networks than attending to a single set of human movements. Indeed, regarding the attentional dynamics of joint action, there currently seem to be more questions than answers.

When considering which brain areas will be involved in the attentional aspects of a joint action (or any given task), it may be useful, as a first step, to refer to a meta-analysis of brain areas involved in attention (Cabeza & Nyberg, 2000). This meta-analysis recognized patterns of brain activation associated with five distinct types of attention: sustained, selective, stimulus-response compatibility, orientation, and divided. This meta-analysis may be useful as an initial point of reference, particularly with respect to brain mechanisms subserving the selective and divided subtypes of attention (which seem most relevant to understanding joint attention, see above). However, such an approach can hardly be considered comprehensive as we presently do not know exactly what types of attention are engaged by humans in joint-action situations. For example, how might divided attention be implemented during joint-action engagements? Do actors in a joint-action situation attend simultaneously to their own movements and the movements of others, or do they serially switch their attention from one set of movements to another? Or is there a functional correlate of selective attention to human movements? Additionally, we know that activity in specific brain areas is correlated with increasing attentional demands when tracking objects (Culham, Cavanagh, & Kanwisher,

2001). But is the same network responsible for tracking multiple human agents that might be present in a joint-action engagement? Such questions are critical to our understanding of the brain basis of joint action.

Despite the importance of attentional mechanisms in joint-action situations, only one experiment has attempted to examine the fMRI correlates of joint attention. Williams, Waiter, Perra, Perrett, and Whiten (2005) designed an experiment in which a sense of “joint attention” was induced in participants via their interaction with a video recording of a person making eye movements towards a cued location. In one condition, the person in the video and the participant oriented to (looked at) the same visual cue, while in the other condition they looked at a different visual cue. Areas exhibiting relatively greater activity in the joint-attention condition included the right ventral MFC and the left superior frontal gyrus (BA 10). The MFC area is associated with theory of mind (ToM; see following section), while BA 10 has been implicated in tasks involving the integration of multiple cognitive operations at an abstract level (Ramnani & Owen, 2004); see Figure 1. Data from this inaugural experiment suggest that the phenomenon of joint attention may rely on the integration of knowledge of others with other, perhaps environmental, information.

The neural correlates of joint attention are just beginning to be understood. Evidence from meta-analyses of the fMRI correlates of attention and the first experiment to directly examine joint attention using fMRI provide a very preliminary outline of brain systems areas likely to be involved in joint attention. Clearly, more research is necessary to clarify the roles of these brain areas during joint-action tasks.

REASONING PROCESSES

Although many of the interactions between two people in perceptual-motor tasks can be described on a neural level as brain mechanisms involved with action, perception, and attention, there are aspects of joint action in which brain mechanisms involved with reasoning processes might be very important. In any given situation where two people are jointly performing a task, it is likely that they have different (possibly complementary or opposite) beliefs and resulting action intentions. It therefore seems that

successful joint action would depend on the ability of people to infer what type of beliefs and action intentions other people have (Tomasello, Carpenter, Call, Behne, & Moll, 2005). Surprisingly, research has shown that this supposed “mind-reading” ability (or alternatively “mentalizing”/“theory of mind”/“intentional stance”) is probably already in operation from about 18 months of age (Frith & Frith, 2003). The litmus test for the presence of mentalizing in both behavioral (mostly with children) and neuroimaging (only with adults) tasks has been the ability to predict another person’s behavior on the basis of this person’s false belief (originally proposed by Dennet, 1978; for the classic experimental paradigm see Wimmer & Perner, 1983). In recent years a network of brain sites has been proposed that underlie mentalizing (for reviews see Amodio & Frith, 2006; Frith & Frith, 1999, 2003, 2006; Lieberman, 2007; Saxe, 2006). The primary brain areas of the proposed mentalizing network are the medial prefrontal cortex (mPFC), the superior temporal sulcus (STS), and the temporo-parietal junction (TPJ); see Figure 1. Here we will discuss whether having a ToM is really crucial for successful joint action, and how and to what extent the mentalizing brain network plays a specific role in joint action.

Sebanz, Bekkering, and Knoblich (2006) argued that in many cases joint intentionality (as proposed by Bratman, 1992) need not be a part of successful joint action. Instead they proposed that joint action can emerge when two people have the same goal at the same time. This opens up the possibility that the mentalizing ability and the mentalizing brain network do not play a critical role in joint action. To examine this possibility the following example is instructive: Imagine either two people building a large wardrobe together (joint action situation) or one person building this wardrobe while another person watches television in the same room (joint inaction situation). It is clear that in both situations a given person might reason about the mental states of the other person. Namely in the joint-action situation both partners might reason about the technical skills of the other and how this relates to their task division, while in the joint-inaction situation the person building the wardrobe might contemplate the lazy nature of the other and keep wondering why the other person doesn’t help out. This example makes clear that it is probably not the case that the mentalizing system does not play a role in joint action at all, it might just not be a

very specific role. Or in other words, in any situation where two or more people are together, the mentalizing network in the brain might be active, regardless of whether they are involved in joint action or not. However, there are experiments that try to refine the function of the component brain areas that make up the mentalizing network, and the findings from these experiments could prove to be specifically relevant for joint-action situations.

An experiment by Mitchell, Banaji, and Macrae (2005) correlated brain activity recorded during a mentalizing task (judging how happy a person was to have their picture taken) with self/other similarity ratings obtained from a post-scanning questionnaire. They found that activity in ventral mPFC was significantly correlated with similarity ratings during the mentalizing task, but not during a face-symmetry judgment task (Macrae, Moran, Heatherton, Banfield, & Kelley, 2004; Schmitz, Kawahara-Baccus, & Johnson, 2004; Vogeley et al., 2004). These findings suggest that self-reflection-based mentalizing occurs only when mentalizing about persons considered to be similar to oneself. Interestingly, Mitchell et al. (2005) found that the dorsal mPFC did not correlate with self/other similarity ratings, suggesting that this area instead instantiates more universally applicable social-cognitive processes that can aid mentalizing even when simulation is inappropriate (e.g., for dissimilar others). For a joint-action situation it can be relevant how similar the two partners judge themselves to be to the other on, for instance, their technical skills. It might also be that two people engaged in joint action automatically judge their similarity on dimensions that are relevant for the joint-action situation (e.g., technical skills), while the same individuals might judge their similarity on completely different dimensions when they are in a joint-inaction situation (e.g., facial similarity). Based on this difference the mPFC might contribute differently to solo or joint-action situations (see Figure 1).

While STS has been strongly implicated in simulation theory (as a preprocessing station that then sends information to parietal and frontal cortex mirror areas), proponents of theory theory claim that this brain area is also an important part of the functional network supporting mentalizing (Frith & Frith, 2003; Samson, Apperly, Chiavarino, & Humphreys, 2004; Saxe & Kanwisher, 2003). Most relevant for the present review is that the posterior STS and also the TPJ have been

implicated in representing the world from different visual perspectives (Aichhorn, Perner, Kronbichler, Staffen, & Ladurner, 2006). The most well-known example is the above mentioned false-belief task in which participants have to realize that although the (fictional) other sees the same visual scene, this person has a different perspective on it. At this point it is not clear how activity in STS/TPJ might differ when people have to compute the perspective of another person in joint action or more general social situations. It could be that this process of computing the (visual) perspective of another person might play a specific role in joint action because at any given time during a joint action people might have to assist or even take over the task of their partner. Consequently, activity in pSTS/TPJ might be more sustained or constant than during “normal” social situations in which role reversals are less likely.

The experimental settings that have revealed the mentalizing network are somewhat far removed from a real joint-action task (i.e., two people are performing actions together). Some studies have looked at more interactive, but not action-related situations such as two people playing games that required co-operation or competition. In addition, the control condition in these studies was that one person was told that the other player was actually a computer (while in fact the behavior of the computer was the same as that of the human player). McCabe, Houser, Ryan, Smith, and Trouard (2001) employed an economic game in which mutual co-operation between players increased the amount of money that could be won. In contrast, Gallagher, Jack, Roepstorff, and Frith (2002) let participants play a competitive game where success depends on predicting what the other player would do next. Both these studies found mPFC activation when the players thought they were interacting with another person instead of with a computer. No differences were found in more posterior regions. This suggests that for interactive situations only the mPFC is involved in reasoning about another person, while the rest of the mentalizing network is more involved in processing the available sensory information.

BACK TO BRATMAN

Based on a review of neuroimaging data relevant to joint action, the present paper proposes a tentative network of brain areas that are

hypothesized to be differentially recruited during solo and joint action (see Figure 1). Within this system, specific brain areas can be linked to specific cognitive processes implied by the components of Bratman's (1992) definition of joint action.

This first stipulation of Bratman's definition of joint action refers to the fact that, at the lowest level of motor co-ordination, jointly acting individuals must be able to adapt their actions to the actions of their partner. This refers to a sort of yin–yang, or push–pull relationship that is common to joint-action engagements involving co-operative motor tasks.

Clearly, such a relationship is contingent upon neural mechanisms governing our ability to perceive relevant actions of our partners. We suggest that the flow of such information is controlled by specific brain areas during joint action, and that our brains are particularly aptly organized to deal with such interactions. In many joint-action situations it is important to pay particular attention to the body parts of your partner. Based on the experiments reviewed above revealing the basic response properties of EBA and STS, we suggest that the human brain is specifically tuned (either through nature or nurture) to process others, and that this ability arose from the critical human need to engage in social interactions. Of course, these actions are not limited to joint action as defined in the current paper, as competitive actions would be likely to rely just as heavily, if not more, on our ability to rapidly and selectively attend to and process the actions of other agents within the environment. That said, joint action certainly qualifies as a subcategory of the more general social interactions that these areas support. As such, we would expect activation within these brain areas to evince differences as a function of whether or not an actor is engaged in solitary or joint actions. Based on the reviewed experiments, their activation should be greater during joint action.

The model of action co-ordination proposed by Knoblich and Jordan (2002) suggests that the MNS is a critical component of action co-ordination. Given the known properties of mirror neurons, these authors suggest that all that is needed to enable joint action is a simple co-ordination mechanism and a way to represent joint-action effects. We agree that the human MNS is an important component of mutual responsivity. However, we suggest that the role of the MNS may be more complicated than

currently presumed. Specifically, we suggest that “broadly congruent” mirror neurons may directly support the co-ordination of actions. Previous data has shown that this class of mirror neurons can fire during observation of one action and execution of a non-identical, but logically related action. This allows for the possibility that mirror neurons represent the *relationship* between the execution of Action A by one agent and the execution of Action B by another agent. This differs from the model put forth by Knoblich and Jordan (2002), which assumes that some sort of overarching mechanism takes the input from mirror neurons and calculates co-ordination patterns and representations of joint-action effects. A clear example of this is the case of complementary action where the actions of one agent must be reciprocated with the execution of non-identical actions by the other agent (see van Schie & Bekkering, 2007). The functional correlates of this important subclass of actions remain a critical question for future experiments.

Another question that remains to be comprehensively addressed is the extent to which joint action engages the MNS. To this end, experiments must be conducted that compare solo and joint action directly. In an attempt to address this question, an experiment by Newman-Norlund and colleagues (2007) examined brain activation in a virtual bar lifting/balancing task in conditions where one person lifted and balanced the bar alone or two participants lifted and balanced the bar co-operatively. Although these conditions did not differ in success rate or error measures, fMRI data revealed significant differences between the two conditions at sites in the MNS. Specifically, the joint-action condition engaged right-hemisphere homologues of the MNS (both inferior parietal and inferior frontal gyri) to a greater extent than solo action. These preliminary data raise the intriguing possibility that actions performed alone or with a partner are differentially represented at the level of the MNS.

A critical component of Bratman’s discussion of shared co-operative activities involves the ability of co-actors to mesh their goal hierarchies in a manner that is conducive to task completion. Indeed, we suggest that jointly acting individuals must be aware of shared goals *on multiple levels*. As a bare minimum, they should be simultaneously aware of the immediate, intermediate, and ultimate action goals of the current joint operation. For example, while two participants might share the ultimate goal of building a model

plane, they might have different intermediate goals (build a wing versus build the landing gear) or immediate goals (use the glue versus use the x-acto knife). The convergence and divergence of these goals must be carefully managed in time in order to avoid conflicting goal configurations (i.e., we both want to use the glue at the same time). We have discussed how this may involve both “action to goal” and “goal to action” inferences. Such processes, presumably performed by brain areas dedicated to processing and integrating goal-related information, are crucial to the fluency of joint-action encounters.

Currently, there is a paucity of brain-imaging data relevant to the processing of goals in joint-action situations. However, existing experiments suggest that the right DLPFC and the left inferior parietal lobe are critically involved in goal processing. The DLPFC has been shown to be important in the skilful organization of movement and the governance of temporal aspects of information flow in primates (Constantinidis, Williams, & Goldman-Rakic, 2002; Goldman-Rakic, 1987; Hoshi, 2006). Also, the human homologue of this area appears to be involved in similar processes as well as working memory (Cohen et al., 1997; D’Esposito, Postle, Ballard, & Lease, 1999; Leung, Gore, & Goldman-Rakic, 2002; MacDonald, Cohen, Stenger, & Carter, 2000). Based on evidence from primate studies suggesting that the DLPFC receives projections from the inferior parietal lobule (IPL; Cavada & Goldman-Rakic, 1989), information is proposed to flow from the IPL to the DLPFC during action observation. We suggest that goal-related information originating in the IPL (Hamilton & Grafton, 2006) is relayed to the DLPFC, which functions to store and co-ordinate information regarding the current status of multiple-goal states. This prediction could be tested by examining the functional correlates of a joint-action task that could be manipulated so as to require the maintenance of one’s own goals or the maintenance of both one’s own goals and the goals of another. The current proposal would predict that activity in both areas should increase as the number of goals being simultaneously represented increases (as should be the case when knowledge of another’s goal is relevant to one’s own performance).

The third component of Bratman’s definition of joint action deals with the understanding of the intentions of other agents and the willingness to provide mutual support to one another. Given

that people have a co-operative attitude, the question remains as to how they understand the intentions of their partner. Recent neuroimaging studies provide a good starting point in terms of the brain areas that might play a role in intention-recognition processes in joint action. These studies have indicated three main areas of interest: the mPFC, the STS, and the TPJ. The mPFC seems especially important for reasoning about the mental states of oneself and others. A review by Amodio and Frith (2006) proposed that for both oneself and others the ventral mPFC is involved with emotion processing, and the dorsal mPFC with monitoring of actions, while the middle region is activated by thinking about people similar to ourselves. In contrast, STS and TPJ seem important for perspective taking and being able to understand the role of the partner in a joint-action task (Aichhorn et al., 2006). However, as mentioned above, no fMRI experiments regarding reasoning processes about intentions and mental state of other people in actual joint-action tasks have been conducted until now. The challenge for future experiments will be to answer the question of whether activity in the mentalizing brain network, specifically regarding the functions of perspective taking and reasoning, represents something qualitatively different during joint action as compared to the more general class of social interactions (not action oriented) in which humans engage.

GENERAL CONCLUSION

In conclusion, the present paper outlines the subset of brain areas predicted to evince differential activation patterns during solitary and joint action (see Figure 1). The extrastriate body area and superior temporal sulcus are preferentially engaged by human body parts and objects moving according to biological constraints. The formulation of theories of intention, animacy, and agency may be critically modulated by the carefully gated information obtained from efferent connections projecting from these areas. For example, information projecting to the insula and right IPL could differentiate the inputs of multiple actors on the environment, a distinction that is critical both to effective co-ordination and to the recognition of patterns within others' actions. The mirror neuron system, including Broca's area and the IPL, functions to map observed actions onto our own internal action repertoire. Such

mapping can be used to predict the actions of another agent as well as to align one's own action plans with theirs. More generally, it can lead to an alignment of goals via an alignment of activation in areas projecting directly to goal-related areas such as the DLPFC. However, the fact that the majority of mirror neurons are broadly congruent enables the system to activate complementary actions of the sort demanded by co-operative tasks. Data shows that the medial frontal and superior frontal gyri appear to be especially important in the creation and maintenance of joint-attentional frameworks critical to co-ordinated performance. Lastly, brain areas including the medial frontal/prefrontal cortex, the temporoparietal junction, and the superior temporal sulcus are consistently recruited when participants ponder the "mental state" of other people. Taken together, these areas comprise a brain network that should be particularly responsive to experimental manipulations involving various aspects of joint action. Here, we would stress that some parts of this network might be relevant for action understanding in general, but not critical for performing joint action (e.g., brain areas involved in mentalization). Future experiments may clarify the exact role of each of these areas, as well as possible interactions between them, during the joint actions that play a critical role in the creation and maintenance of humanity's social fabric.

Manuscript received 15 February 2006

Manuscript accepted 16 January 2007

REFERENCES

- Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *NeuroImage*, 30, 1059–1068.
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Neuroscience Reviews*, 7, 268–277.
- Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature Neuroscience*, 7, 542–548.
- Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). fMRI responses to video and point-light displays of moving humans and manipulable objects. *Journal of Cognitive Neuroscience*, 15, 991–1001.
- Bekkering, H., & Wohlschlaeger (2002). Action perception and imitation: A tutorial. In W. Prinz & B. Hommel (Eds.), *Common mechanics in perception and action* (pp. 294–315). New York: Oxford University Press.

- Blakemore, S. J., Oakley, D. A., & Frith, C. D. (2003). Delusions of alien control in the normal brain. *Neuropsychologia*, *41*, 1058–1076.
- Blanke, O., & Arzy, S. (2005). The out-of-body experience. Distributed self-processing at the temporo-parietal junction. *The Neuroscientist*, *11*, 16–24.
- Bonabeau, E., Theraulaz, G., Deneubourg, J.-L., Aron, S., & Camazine, S. (1997). Self-organization in social insects. *Trends in Ecological Evolution*, *12*, 188–193.
- Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal cortex and the amygdala in the perception of biological motion. *Journal of Neuroscience*, *16*, 3737–3744.
- Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: A functional double dissociation. *Neuropsychologia*, *43*, 89–98.
- Brass, M., & Heyes, C. (2006). Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*, 489–495.
- Brass, M., Zysset, S., & von Cramon, D. Y. (2001). The inhibition of imitative response tendencies. *NeuroImage*, *14*, 1416–1423.
- Bratman, M. (1992). Shared cooperative activity. *The Philosophical Review*, *101.2*, 327–341.
- Buccino, G., Binkofski, G., & Riggio, L. (2004b). The mirror system and action recognition. *Brain and Language*, *89*, 370–376.
- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., et al. (2004a). Neural circuits involved in the recognition of actions performed by nonconspecifics: An fMRI study. *Journal of Cognitive Neuroscience*, *16*, 114–126.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*, 1–47.
- Cavada, C., & Goldman-Rakic, P. (1989). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *The Journal of Comparative Neurology*, *287*, 393–445.
- Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, *386*, 604–608.
- Constantinidis, C., Williams, G. V., & Goldman-Rakic, P. S. (2002). A role for inhibition in shaping the temporal flow of information in prefrontal cortex. *Nature Neuroscience*, *5*, 175–180.
- Csibra, G. (2005). *Mirror neurons and action understanding. Is simulation involved?* (Retrieved February 1, 2007 from <http://www.interdisciplines.org/mirror>)
- Csibra, G., & Gergely, G. (2007). Obsessed with goals: Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, *124*, 60–78.
- Cuijpers, R., van Schie, H., Koppen, M., Erilagen, W., & Bekkering, H. (2006). Goals and means in action observation: A computational approach. *Neural Networks*, *19*, 311–322.
- Culham, J. C., Cavanagh, P., & Kanwisher, N. G. (2001). Attentional response functions: Characterizing brain areas using fMRI activation during parametric variations of attentional load. *Neuron*, *32*, 737–745.
- D'Esposito, M., Postle, B. R., Ballard, D., & Lease, J. (1999). Maintenance versus manipulation of information held in working memory: An event related fMRI study. *Brain and Cognition*, *41*, 66–86.
- Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., et al. (1997). Brain activity during observation of actions: Influence of action content and subject's strategy. *Brain*, *120*, 1763–1777.
- Dennet, D. C. (1978). Beliefs about beliefs. *Behavioural & Brain Sciences*, *1*, 568–570.
- Di Pellegrino, Fadiga, G., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176–180.
- Downing, P. E., Bray, D., Rogers, J., & Childs, C. (2004). Bodies capture attention when nothing is expected. *Cognition*, *93*, B27–B38.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*, 2470–2473.
- Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., & Jeannerod, M. (2003). Modulating the experience of agency: A positron emission tomography study. *NeuroImage*, *18*, 324–333.
- Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs. another person as being the cause of an action: The neural correlates of the experience of agency. *NeuroImage*, *15*, 596–603.
- Ferrari, P. F., Gallese, V., Rizzolatti, G., & 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., & 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.
- Field, T. M., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expressions by neonates. *Science*, *218*, 179–181.
- Fincham, J. M., Carter, C. S., van Veen, V., Stenger, V. A., & Anderson, J. R. (2002). Neural mechanisms of planning: A computational analysis using event related fMRI. *Proceedings of the National Academy of Sciences*, *99*, 3346–3351.
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science*, *29*, 662–667.
- 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 movements in the parietal area PF (7b) of the macaque monkey. *Society for Neuroscience Abstract*, *24*, 154.
- Frith, C. D., & Frith, U. (1999). Interacting minds—A biological basis. *Science*, *286*, 1692–1695.

- Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50, 531–534.
- Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*, 358, 459–473.
- Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage*, 16, 814–821.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593–609.
- Gangitano, M., Mottaghy, F. M., & Pascual-Leone, A. (2001). Phase-specific modulation of cortical motor output during movement observation. *Neuroreport*, 12, 1489–1492.
- Garavan, H., Ross, T. J., Murphy, K., Roche, R. A., & Stein, E. A. (2002). Dissociable executive functions in the dynamic control of behavior: Inhibition, error detection, and correction. *NeuroImage*, 17, 1820–1829.
- Gazzola, V., Aziz-Zadeh, L., & Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current Biology*, 16, 1824–1829.
- Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements and action. *Nature Neuroscience Reviews*, 4, 179–192.
- Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In V. B. Mountcastle & F. Plum (Eds.), *Handbook of psychology. Vol. 5: The nervous system, higher functions of the brain* (pp. 373–417). Bethesda, MD: American Psychological Society.
- Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of Cognitive Neuroscience*, 7, 497–510.
- Grafton, S. T., Hazeltine, E., & Ivry, R. B. (1998). Abstract and effector-specific representations of motor sequences identified with PET. *Journal of Neuroscience*, 18, 9420–9428.
- 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.
- Grèzes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C., & Decety, J. (2000). Does perception of biological motion rely on specific brain regions. *NeuroImage*, 13, 775–785.
- Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *NeuroImage*, 21, 744–750.
- Grossman, E. D., & Blake, R. (2002). Brain areas active during visual perception of biological stimuli. *Neuron*, 35, 1167–1175.
- Grossman, E. D., Blake, R., & Kim, C. (2004). Learning to see biological motion: Brain activity parallels behavior. *Journal of Cognitive Neuroscience*, 16, 1669–1679.
- Grossman, E. D., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., et al. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, 12, 711–720.
- Hamilton, F., & Grafton, S. (2006). Goal representation in human anterior intraparietal sulcus. *The Journal of Neuroscience*, 26, 1133–1137.
- Hamilton, F. de C., Wolpert, D. M., Frith, U., & Grafton, S. (2006). Where does your own action influence your perception of another person's action in the brain? *NeuroImage*, 29, 524–535.
- Hasselmo, M. E. (2005). A model of prefrontal cortical mechanisms for goal-directed behavior. *Journal of Cognitive Neuroscience*, 17, 1115–1129.
- Hester, R., Murphy, K., Foxe, J., Foxe, D., Javitt, D., & Garavan, H. (2004). Predicting success: Patterns of cortical activation and deactivation prior to response inhibition. *Journal of Cognitive Neuroscience*, 16, 776–785.
- Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current Opinions in Neurobiology*, 12, 217–222.
- Hoshi, E. (2006). Functional specialization within the dorsolateral prefrontal cortex: A review of anatomical and physiological studies of non-human primates. *Neuroscience Research*, 54, 73–84.
- Howard, R. J., Brammer, W., Wright, I., Woodruff, P. W., Bullmore, E. T., & Zeki, S. (1996). A direct demonstration of functional specialization within the motion related visual and auditory cortex of the human brain. *Current Biology*, 6, 1015–1019.
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLOS Biology*, 3, 529–535.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286, 2526–2528.
- Jackson, G. M., Jackson, S. R., Harrison, J., Henderson, L., & Kennard, C. (1995). Serial reaction time learning and Parkinson's disease: Evidence for a procedural learning deficit. *Neuropsychologia*, 33, 577–593.
- Jeannerod, M. (1999). To act or not to act: Perspectives on the representation of actions. *Quarterly Journal of Experimental Psychology*, 52, 1–29.
- Jeannerod, M., & Pacherie, E. (2004). Agency, simulation and self-identification. *Mind and Language*, 19, 113–146.
- Jellema, T., & Perrett, D. I. (2002). Coding of visible and hidden actions. In W. Prinz & B. Hommel (Eds.), *Common mechanics in perception and action* (pp. 356–381). New York: Oxford University Press.
- Keysers, C., Kohler, E., Umiltà, M. A., Fogassi, L., Nanetti, L., & Gallese, V. (2003). Audio-visual mirror neurons and action recognition. *Experimental Brain Research*, 153, 628–636.
- Kilner, J. M., Paulignan, Y., & Blakemore, S.-J. (2003). An interference effect of observed biological movement on action. *Current Biology*, 13, 522–525.
- Knoblich, G., & Jordan, 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.

- Knoblich, G., & Jordan, S. (2003). Action coordination in individuals and groups: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 29, 1006–1016.
- 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.
- Koski, L., Wohlschläger, A., Bekkering, H., Woods, R. P., Dubeau, M. C., Mazziotto, J. C., et al. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cerebral Cortex*, 12, 847–855.
- Leung, H. C., Gore, J. C., & Godlman-Rakic, P. (2002). Sustained mnemonic response in the middle human frontal gyrus during on-line storage of spatial memoranda. *Journal of Cognitive Neuroscience*, 14, 659–671.
- Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, 58, 18.11–18.31.
- MacDonald, A. W., Cohen, J., Stenger, A., & Carter, S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 9, 1835–1838.
- Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F., & Kelley, W. M. (2004). Medial prefrontal activity predicts memory for self. *Cerebral Cortex*, 14, 647–654.
- Maeda, F., Kleiner-Fisman, G., & Pascual-Leone, A. (2002). Motor facilitation while observing hand actions: Specificity of the effect and role of observers' orientation. *Journal of Neurophysiology*, 87, 1329–1335.
- Manthey, S., Schubotz, R. I., & von Cramon, D. Y. (2003). Premotor cortex in observing erroneous actions: An fMRI study. *Cognitive Brain Research*, 15, 296–307.
- McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences USA*, 98, 11832–11835.
- McGuire, P. K., Silbersweig, D. A., & Frith, D. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain*, 119, 907–917.
- Metzinger, T., & Gallese, V. (2003). The emergence of a shared action ontology: Building blocks for a theory. *Consciousness and Cognition*, 12, 549–571.
- Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, 17, 1306–1315.
- Newman-Norlund, R. D., Meulderbroek, R., & Bekkering, H. (2007). *Anatomical substrates of cooperative joint action: A continuous balancing task*. Manuscript in preparation.
- Nudds, T. D. (1978). Convergence of group size strategies by mammalian social carnivores. *American Naturalist*, 112, 957–960.
- Pascual-Leone, A., Grafman, J., Clark, K., Stewart, M., Massaquoi, S., Lou, J.-S., et al. (1993). Procedural learning in Parkinson's disease and cerebellar degeneration. *Annals of Neurology*, 34, 594–602.
- Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience*, 16, 1706–1716.
- Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, 118, 2188–2199.
- Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London*, 83, 475–487.
- Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Neuroscience Reviews*, 3, 184–194.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192.
- Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 131–141.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Neuroscience Reviews*, 2, 661–670.
- Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, 4, 546–550.
- Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience*, 7, 499–500.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, 16, 235–239.
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporoparietal junction in “theory of mind”. *NeuroImage*, 19, 1835–1842.
- Schmitz, T. W., Kawahara-Baccus, T. N., & Johnson, S. C. (2004). Metacognitive evaluation, self relevance, and the right prefrontal cortex. *NeuroImage*, 22, 941–947.
- Schultz, J., Friston, K. J., O'Doherty, J., Wolpert, D. M., & Frith, C. D. (2005). Activation in posterior superior temporal sulcus parallels parameter inducing percept of animacy. *Neuron*, 45, 625–635.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, 10, 70–76.
- Shane, S. H., Randall, S., Wells, R. S., & Worsig, B. (1986). Ecology, behaviour and social organization of the bottlenose dolphin: A review. *Marine Mammal Science*, 2, 34–63.
- Spence, S. A., Brooks, D. J., Hirsch, S. R., Liddle, P. F., Meehan, J., & Grasby, P. M. (1997). A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain*, 120, 1997–2011.
- Stevens, J. A., Fonlupt, P., Schiffrar, M., & Decety, J. (2000). New aspects of motion perception: Selective

- neural encoding of apparent human movements. *Neuroreport*, 11, 109–115.
- Tai, Y. F., Scherfler, C., Brooks, D. J., Sawamoto, N., & Castiello, U. (2004). The human premotor cortex is “mirror” only for biological actions. *Current Biology*, 14, 117–120.
- Thompson, J. C., Clarke, M., Stewart, T., & Puce, A. (2005). Configural processing of biological motion in human superior temporal sulcus. *Journal of Neuroscience*, 25, 9059–9066.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675–735.
- Tummolini, L., Castelfranchi, C., Pacherie, E., & Docik, J. (2006). From mirror neurons to joint actions. *Cognitive Systems Research*, 7, 101–112.
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al. (2001). I know what you are doing: A neurophysiological study. *Neuron*, 32, 91–101.
- Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences*, 98, 11656–11661.
- Van Schie, H. T., & Bekkering, H. (2007). *Biological and symbolic cueing of action: An advantage for cueing of action goals over action means*. Manuscript in preparation.
- Van Schie, H. T., van Waterschoot, B. M., & Bekkering, H. (2007). Understanding action beyond imitation: Reversed compatibility effects of action observation in imitation and joint action. *Brain Research*. Manuscript submitted for publication.
- Vogely, K., May, M., Ritzl, A., Falkai, P., Zilles, K., & Fink, G. R. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *Journal of Cognitive Neuroscience*, 16, 817–827.
- Vogt, S., Taylor, P., & Hopkins, B. (2003). Visuomotor priming by pictures of hand postures: Perspective matters. *Neuropsychologia*, 41, 941–951.
- Williams, J. H. G., Waiter, G. D., Perra, O., Perrett, D. I., & Whiten, A. (2005). An fMRI study of joint attention experience. *NeuroImage*, 25, 133–140.
- Willingham, D. B., & Koroshetz, W. J. (1993). Evidence for dissociable motor skills in Huntington’s disease patients. *Psychobiology*, 21, 173–182.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131, 460–473.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs-representation and constraining function of wrong beliefs in young children’s understanding of deception. *Cognition*, 13, 1031–28.