Embodied Cognition and the Simulation of Action to Understand Others

Scott T. Grafton

University of California, Santa Barbara

Understanding the goals or intentions of other people requires a broad range of evaluative processes including the decoding of biological motion, knowing about object properties, and abilities for recognizing task space requirements and social contexts. It is becoming increasingly evident that some of this decoding is based in part on the simulation of other people's behavior within our own nervous system. This review focuses on aspects of action understanding that rely on embodied cognition, that is, the knowledge of the body and how it interacts with the world. This form of cognition provides an essential knowledge base from which action simulation can be used to decode at least some actions performed by others. Recent functional imaging studies or action understanding are interpreted with a goal of defining conditions when simulation operations occur and how this relates with other constructs, including top-down versus bottom-up processing and the functional distinctions between action observation and social networks. From this it is argued that action understanding emerges from the engagement of highly flexible computational hierarchies driven by simulation, object properties, social context, and kinematic constraints and where the hierarchy is driven by task structure rather than functional or strict anatomic rules.

Key words: **action understanding; simulation; embodied cognition; mirror neuron system; social-network; parietal cortex**

Introduction

Decoding action performed by other people, from the simplest eye or hand movements to the most complicated gymnastic movement, is likely to involve a multitude of cognitive processes that are supported by distinct underlying neural substrates. This review focuses on those decoding processes that are most likely to be related to embodied cognition. Emphasis is given to recent empirical studies that demonstrate the role of simulation as a mechanism for action decoding. An increasing number of studies are beginning to define the level of detail where an action can be simulated, the neural substrates that support this simulation, and most importantly, learning-dependent changes within these systems as subjects acquire physical competency with new behaviors. These findings fit within a local framework based on embodied cognition that merges into a larger framework that also incorporates nonembodied processes for action understanding.

Embodied Cognition: What Can We Do with Ourselves?

Embodied cognition carries many definitions, but the one used here is the existence of a memory system that encodes knowledge of a person's physical competencies and a person is capable of interacting with the physical world. This knowledge can be used in many ways. The notion of embodied cognition has motivated a range of psychological theories, including explanations on the origins of language (Fischer & Zwaan 2008), as a roadmap for understanding cognitive or motor development (Thelen

---

Address for correspondence: Scott T. Grafton, M.D., UCSB Brain Imaging Center, The Sage Center for the Study of Mind, UC Santa Barbara, Santa Barbara, CA 93105. Voice: +805-975-5272; fax: +805-893-4303. grafton@psych.ucsb.edu

The Year in Cognitive Neuroscience 2009: Ann. N.Y. Acad. Sci. 1156: 97–117 (2009).
doi: 10.1111/j.1749-6632.2009.04425.x

et al. 2001), and as a means for explaining the coupling between physical cues and emotional inference in people (Niedenthal 2007). In the current review embodied cognition is used to focus on the interpretation experiments where people relate knowledge of their own body to understand other people's body movements. This focus harkens back to the original formulation of embodied cognition that emerged from the artificial intelligence community two decades ago (Brooks 1991). The main point from this work was that to perform intelligent actions, an individual's knowledge—in this case a computer's—of what the body can do matters. While we can readily use disembodied computers to manipulate symbolic information, intelligence requires interaction with the world. Once an artificial system is enabled with sensory and motor attributes, the very nature of what one considers intelligent changes. Intelligence can then be understood by observing what an agent perceives and how it acts, not by what it computes. While the focus in robotics is often on overcoming the limitations that a physical plant can impose, physical systems can provide new capacities that create intelligence. In this view embodied cognition is not a description of the limitations in cognition constrained by a physical plant. Rather, it emphasizes what is enabled cognitively by having a body. This can include the gains that come from relatively low-level physical attributes, such as the spring-like quality of our musculoskeletal system to facilitate walking, or the softness of our fingertips to aid in grasping a stiff object. Or it can emerge when knowledge about the skills needed for one object or context can be generalized to new conditions. Ironically, this knowledge turns out to be exceedingly difficult for robots to acquire by unsupervised learning and relatively trivial for humans. It is from this human perspective that we can consider cases where embodied cognition leads to gains of cognitive function.

There is growing evidence from studies in athletes that physical experience can improve perceptual ability (Tenenbaum et al. 1994). For example, professional basketball players can consistently predict if a free throw shot will be successful sooner than the general population. Basketball coaches, with moderate physical experience and extensive visual familiarity, fall in the middle for this predictive ability. The players draw on subtle physical cues such as the flexion of the wrist and ankle more than the nonexpert does (Aglioti et al. 2008).

Historically, embodied cognition of action has been investigated with experiments demonstrating interference effects on self-generated actions that are influenced by the observation of another person's limb postures or actions; alternatively, this cognition has been studied when the decoding of another person's action is influenced by the simultaneous performance of a related action (Craighero et al. 2002; Craighero et al. 1999; Hamilton et al. 2004). Interference effects such as these provide indirect evidence that some knowledge about action is embodied and, furthermore, that there may be some form of simulation taking place during action observation. However, as pointed out by Jacob and Jeannerod, either mental or motor simulation could be a means for decoding action understanding and this might occur at more than one level of abstraction (2005). Defining the level of simulation becomes a central issue for interpreting both behavioral and physiological studies.

Only recently have action interference studies attempted to isolate the level of abstraction in simulation. As an extreme case, simulation could occur based on very abstract representations requiring pure mentalizing. One person could formulate a mental impression of what another person might be thinking. It has recently been shown that this is more likely to occur for novel situations (Brass et al. 2007). At the opposite end of the spectrum, simulation might occur by matching perceived actions with internal models of the same actions. The strong version of this is direct matching, where kinematically identical movements are simulated in the observer as part of the action decoding process. In the real world there is much less opportunity for direct matching and so if simulation is to occur, a weaker form is needed. The weak form requires a comparison of analogous movements that are not necessarily identical between the observer and observed motor systems. For example, observing another's actions from the first- or third-person perspective can be simulated at a similar level of abstraction (Anquetil & Jeannerod 2007). And in a weight judgment task, subjects rely on certain visual heuristics such as the lifting movement of the actor's hand to determine an object's weight and largely ignore other kinematic features such as how the object is moved around, body posture, or how an object is put down (Hamilton et al. 2007).

To make stronger inferences about the kind of simulation that takes place in action observation many groups have begun to incorporate electrophysiological methods aimed at defining the functional anatomy of embodied cognition and the circumstances where there is strong overlap between action, perception, and understanding. One electrophysiological approach for demonstrating simulation within the motor system is based on studies that identify modifications of motor corticospinal excitability. In transcranial magnetic stimulation (TMS) studies the relative sensitivity of the corticospinal tract is assessed by stimulating over motor cortex and measuring changes in the threshold needed to evoke responses in the hand (Aziz-Zadeh et al. 2002; Baldissera, Cavallari et al. 2001; Fadiga et al. 1995). This excitability increases when subjects observe actions in others. The mere observation of static photographs of hands positioned with a pincer grip causes an increase in corticospinal excitability compared to observation of hand positioned at rest. This increased excitability was specific for the muscle that would be activated during actual execution of the observed action (Urgesi et al. 2006). Action observation can also amplify learning- or experience-dependent changes of corticospinal excitability (Aglioti et al. 2008; Celnik et al. 2006). These studies are consistent with simulation at the level of the motor cortex and are helpful in demonstrating the specificity of actions that might lead to simulation. However, they do not localize simulation circuitry outside of the primary motor cortex.

The current review considers recent evidence from cognitive neuroscience that continues to build a conceptual framework for how action understanding is achieved through simulation within a framework of embodied cognition. Five issues that arise from this new research are considered:

- What are the neural underpinnings of action observation and simulation?
- Does physical expertise alter the neural substrates of action understanding?
- Is there hierarchical structure in the decoding of observed actions?
- What simulation is applicable to the social world?
- What is the relationship between embodied cognition and the mirror neuron system?

The Neural Underpinnings of Action Understanding

The Action Observation Network

There is now overwhelming evidence from functional magnetic resonance imaging (fMRI) of blood oxygen level-dependent (BOLD) signals that observation by one person of another who is in motion will engage in the observer a widespread, bilateral network of cortical brain regions in a highly reproducible manner (Buccino et al. 2001; Frey & Gerry 2006; Grafton et al. 1996; Grezes & Decety 2001; Hari et al. 1998; Johnson-Frey et al. 2003; Rizzolatti et al. 1996; Saygin et al. 2004; Ulloa & Pineda 2007; Zentgraf et al. 2005). We refer to this as the action observation network (AON), shown in Fig. 1. It includes the bilateral posterior superior temporal sulcus (STS), inferior parietal lobule (IPL), inferior frontal gyrus (IFG), dorsal premotor cortex, and ventral premotor cortex. What is striking about this network is that a

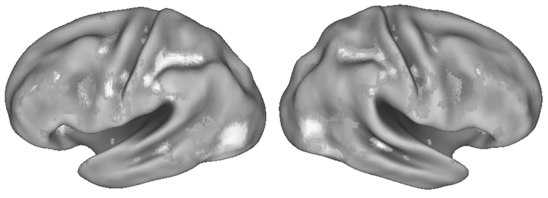

Figure 1. An example of the action observation network. Colored areas indicate cortical regions that are activated when subjects make an inference about why a person being observed is grasping an object in a typical context where that object might be found. Comparing blood oxygen level-dependent (BOLD) activity during these images relative to a blank screen captures visual evoked responses beginning with early visual processing as well as the decoding of the possible meaning of the action.

perceptual stimulus can lead to such a large-scale recruitment of the brain with a complete disregard of textbook divisions between sensory and motor portions of the cortex. Understanding or predicting the temporal and spatial structure of a stimulus is just as likely to recruit a premotor area (Schubotz & von Cramon 2001, 2004). This distributed network likely supports many subtasks, including the transformation of perceptions to action, the simulation of observed movements in relationship to known movements, and the storage of physical knowledge (both of self and objects) that can be used for simulation. What is the evidence that all or part of the AON might support these processes and, in particular, simulation?

An early argument in support of simulation was motivated by the fact that some of the areas of the AON overlap strongly with areas associated with movement execution, particularly within the IFG and IPL (see Grèzes et al. 2001 for review). This overlapping cortex, with shared substrates for execution and observation, is referred to as the mirror neuron system (MNS) (Rizzolatti & Craighero 2004). Because some of these core areas are also used for action execution it has been proposed many times that there exist underlying mirror neurons that could serve as the critical functional substrate for simulation (Gallese et al. 2004). It should be noted however that invoking a mirror neuron system is an anatomic rather than a functional explanation for most imaging experiments and it does no more than constrain the number of locations where simulation might occur, and it does so for only a subset of tasks where simulation might occur. Given this limitation a number of new and innovative experimental approaches can be interpreted as evidence of simulation without the need to invoke the MNS heuristic to explain the data.

Simulation and Somatotopy

One fMRI approach to study simulation is based on the idea that observation of actions performed by different body parts should recruit different parietal and premotor areas, in a roughly somatotopic distribution. This general finding has now been observed in multiple studies (Buccino et al. 2001; Buccino et al. 2004; Gazzola et al. 2006; Goldenberg & Karnath 2006). The logic is motivated by the fact that motor execution areas have a roughly somatotopic representation, so simulation within these circuits should also be somatotopic. It should be noted, however, that differences in activation as a function of somatotopy might be confounded by other task variables such as familiarity or complexity. It is also important to note that not all components of the AON demonstrate a somatotopic arrangement. While there are distinct localizations of fMRI-BOLD responses when a subject is observing face versus hand motions within the IPL, the same is not true for the STS. This is possible evidence for modularity within the AON. Areas such as STS may provide lower level input to IPL and other parts of the AON involved with simulating human actions at a higher level of analysis (Thompson et al. 2007).

Simulation and Conspecifics

Another neuroimaging paradigm in support of simulation is based on mapping fMRI-BOLD responses of an observer within the AON during the viewing of conspecifics versus other species as they perform an action

(Buccino, Lui et al. 2004). The key finding is that when observing a nonconspecific such as a dog performing actions that the observer can perform, there is more activation of the AON compared to what is seen in the observer for actions not commonly performed (barking). Similarly, observation of a humanoid like action by a robot can recruit the AON network more than nonhumanoid actions can. This also suggests that simulation might be driven more by the familiarity and feasibility of an action rather than direct matching to an image of a human body (Demiris & Simmons 2006).

Simulation and Objects

It has long been known that the mere observation of a graspable object will activate components of the action observation network including dorsal premotor cortex and the anterior intraparietal sulcus (Grafton et al. 1997), and the influence of the object on grasp planning is within the AON (Grezes et al. 2003). Having a graspable object in view also modulates visual attention, even when the object is irrelevant, and these effects are mediated through dorsal parietal and premotor circuits (Handy et al. 2003). Experiments in which subjects vary in their knowledge of tools or objects could potentially be used as paradigms to test for simulation (Frey & Gerry 2006). However, tool knowledge is not necessarily matched to simulation capacity because the familiarity for tools might be in terms of object properties rather than hand actions for manipulating the object. In addition, the semantic knowledge and pragmatic knowledge of a tool introduces additional complexities about what subjects are using to understand actions are topics that will not be broached in this review (Beauchamp et al. 2003; Chao & Martin 2000; Weisberg et al. 2007).

Simulation and Action Familiarity

Another inroad to understand simulation is based on action familiarity. If an action itself (rather than the dog or robot performing the action) is more familiar, then there should be greater activity in a simulation circuit when this action is observed. The challenge in designing experiments that tap into familiarity of human action is to find behaviors with more or less novelty. One approach for manipulating familiarity that avoids object-centered action has been to measure brain activity during observation of dancers. When dancers watch someone else perform a familiar dance there is greater activation within the AON compared to watching a dancer perform an unfamiliar form (Calvo-Merino et al. 2005). This suggests that the AON, and in particular the premotor and parietal cortex, demonstrate stronger simulation for movements that are physically familiar. There are two concerns that undermine the specificity of this interpretation. The first concern is that in this experiment the familiar dances were both physically and visually familiar. The second is that the familiar dance movements had names and the unfamiliar did not. Subverbal naming might be occurring in this task and thus interact with visual or physical familiarity. These concerns were directly addressed in an elegant follow-up study (Calvo-Merino et al. 2006). Men and women of a professional ballet company observed movies of ballet dance movements that were all strongly familiar in terms of visual recognition naming. The critical manipulation was watching movements performed by men or women. In ballet many of the movements are gender-specific. There was stronger activation within the AON when watching dance movements by same sex dancers compared to the opposite sex, presumably because these were the movements with which the observer had more physical experience. These familiarity effects were most robust in inferior parietal and premotor areas.

Simulation and Action Competence

If activity in the AON is stronger during simulation of movements that are physically familiar then there should be learning-dependent changes in the network as a function of physical training. This idea was tested in a group of modern dancers, who learned to perform a new dance piece. They were scanned on multiple occasions over several months as they gained physical competency. During scanning they watched short video clips of the piece they were learning and another piece they did not practice. In addition, for each clip they were instructed to imagine performing the movements. They also rated how well they thought they could perform each segment themselves (Cross et al. 2006). There was stronger activity within the AON when watching dance movements that had been learned after training. In addition, activity in the AON corresponded to self-rated competency on a movement-by-movement basis.

The previous studies suggest there is a form of motor resonance that develops in the AON when people observe others performing physically familiar actions. In an interesting extension of this idea, this finding was inverted to test if motor resonance within the AON could serve as a substrate for learning new movements based on observation alone. To do this, subjects learned dance steps in a computer game similar to "Dance Dance Revolution." Both symbolic cues and a physical model were used to learn specific dance steps that were accompanied by unique music (Cross et al. 2008). Some dance pieces were learned with actual physical practice. Others were learned only by observation of the model and symbols, without any physical movements. On follow-up testing, subjects could perform those dances for which they had physical practice or observational learning better than novel dances. As in the previous studies, there was an experience-dependent change of activity within the AON for dance steps that had been physically practiced and, to a lesser degree, for observational learning. This finding suggests that the simulation operations within the AON may be used both for decoding observed actions and also to model new movements in the observer. These learning-related changes were limited to parietal and premotor cortex. The extrastriate body area and STS areas demonstrated no change with learning and were only sensitive to the presence or absence of a physical model. This finding provides additional evidence for modularity within the AON.

In a follow-up analysis the same data set was tested to determine if the presence of the physical model was necessary for learning-dependent changes to take place in the AON. The key finding was that symbolic cues alone were sufficient to induce learning-related change in the AON. Furthermore, once a dance was learned, the symbols and music alone served as potent stimuli for activating the AON. In other words, an external physical model is not required for motor simulation to occur once a particular body movement of the observer has been associated with symbols and music. (Emily Cross, personal communication) It has been argued this rhythmic coupling of external cues and induced body motion is a core substrate for the pleasures of dancing to music (Janata & Grafton 2003).

Hierarchy and Action Understanding

The experiments described so far are a representative rather than exhaustive review of attempts to use functional imaging to capture simulation processes during action observation. Taken in aggregate they provide relatively strong evidence that some form of simulation is occurring within the AON during action observation. They also begin to demonstrate granularity in the functional specialization within the AON proper, with some areas specifically linked to physical models (STS) and others more sensitive to learning-dependent experience (inferior parietal lobule). Two interrelated issues emerge from this granularity. Is specialization within the AON related to the kind of simulation that is taking place (e.g., direct matching, goal-oriented, or abstract) or to other processes such as decoding of biological motion or contextual information? If there is specialization related to the level of simulation, is there evidence for any functional hierarchy?

Lesions of the AON

One way to identify specialization within the AON is to test for behavioral deficits in the face of real or virtual focal lesions to the brain. In patients with ischemic stroke of the cerebral cortex a deficit in biologic motion processing can be present with lesions of the STS. This is consistent with classic models of visual processing where the human homologue of area MT and the extrastriate body area are essential for the perceptual decoding of body-specific movement. What is striking is that a lesion in premotor cortex of the AON will also cause a deficit in the processing of biological motion (Saygin 2007). While this could be due to a low level disruption in the decoding of motion information, a more likely interpretation is that the premotor segment of the AON is capable of forming more abstract relations between motions and the meaning of the motion. In support of this latter interpretation, deficits in the nonlinguistic matching of pictures of actions (such as licking something) with appropriate objects (an ice cream cone) are also associated with inferior frontal cortex lesions (Saygin et al. 2004). Studies of patients with limb apraxia have identified gesture comprehension deficits, suggesting that there is an anatomic overlap for the neural substrates of both production and comprehension (Ferro et al. 1983; Heilman et al. 1982; Rothi et al. 1985). Precise lesion localization has only recently emerged to show that the lesions are consistently left-lateralized and independent of language comprehension (Buxbaum et al. 2005). Precise localization based on lesion overlap analysis shows that gesture comprehension deficits correlate with damage to the opercular and triangularis portions of the inferior frontal gyrus, two regions of the AON. Remarkably, no such relationship was observed with lesions centered on the inferior parietal cortex, an area robustly activated by action observation and associated strongly with apraxia. The findings suggest that left inferior frontal cortex, involving regions that are also involved in planning and performing actions, is causatively associated with deficits in the recognition of the correct execution of meaningful gestures (Pazzaglia et al. 2008). A continuing concern with gesture comprehension experiments is whether the behavioral tasks are tapping into semantic comprehension or are due to deficits of embodied cognition irrespective of semantics.

Innovative studies in normal subjects help to build a case for an action specific network involving inferior frontal cortex. Transient disruption of inferior premotor cortex by TMS impairs judgments in the weight of an object another person is lifting and also leads to increased response times in a delayed match to sample task when the matching is for body actions (Pobrinc & Hamilton 2006; Urgesi et al. 2007). In contrast STS disruption slowed matching of body form. These real and virtual lesion data suggest that the anterior lesions of the AON may be particularly important for the storage or retrieval of action vocabularies, that is, the relations between known body movements and their probable meaning in terms of action consequences.

Simulation and Context

How a body moves in the world or interacts with an object is constrained by the features of the environment or object. This interaction between the body and world gives some insight into what a person is doing, but not necessarily why they are performing an action. A fundamental question is how contextual information is integrated with simulation to generate an improved estimate of another person's actions. One approach to study this is to add contextual information that alters the meaning of an observed action (Iacoboni et al. 2005). For example, an object such as a coffee cup can be grasped in different contexts that imply different intentions of the actor. If it is the start of a meal, the intention is to drink, if the dishes are a mess, the intention is to clean up. Note that in this experiment the intentionality is centered on the hand—object interaction, that is, Why is a person grasping a cup in a particular way? Adding this contextual information amplifies the decoding of why an object is manipulated. This further recruits parietal and premotor areas within the AON. Whether this activation represents a simulation or supports a role for embodied cognition is unclear.

Hierarchical Processing of Action Observation: Evidence from Repetition Suppression

What is the evidence for hierarchical processing within the AON? The answer depends in part on what is meant by "hierarchy." The origin of the word *hierarchy* can be found in sermons from the 1380s that codified the different angels into three levels, the Seraphin, Cherubyn, and Trones. The notion of successive orders or grades, one above the other, was eventually generalized to rulers, species, and ultimately any persons, things, or concepts. When we watch someone perform an action, we group the movements into coherent subunits or parts with some of them subordinate to others. In the making of coffee the individual task requirements (scooping, boiling water, filtering) are subordinate to the goal of brewing (Cooper & Shallice 2006). This whole—part structure can be readily identified in imitation studies that show a bias towards goal rather than parts, and the same studies also show emergence of this bias with development (Bekkering et al. 2000). Existing computational models that use hierarchy to understand structured behavior generally assume that the part—whole organization of an action or its recognition is mirrored by the neural representations that underlie the hierarchical organization (Botvinick 2008; Cooper & Shallice 2006). A major challenge is to determine if putative hierarchies as defined by experimental manipulation and fMRI are consistent with this assumption. This would require successive ordering or levels of functional anatomy that directly map to a part—whole structure. Alternatively, the functional anatomy could simply reflect differences in other aspects of task structure, such as inherent complexity, task maintenance, scheduling, or other nonhierarchical processes such as increasing temporal delay. In other words, demonstrating that different parts of the brain are activated as a function of task complexity or abstraction by itself is insufficient toward establishing that the underlying functional anatomy is hierarchical (Badre & D'Esposito 2007). It is also necessary to demonstrate that there is a nesting of part—whole processes, with higher levels dependent on lower levels. This nesting can be very challenging to establish with conventional cognitive subtraction paradigms. By using alternative fMRI techniques such as repetition suppression the whole—part structure and nesting of interrelated processes can be addressed more directly.

We used repetition suppression (RS) methods to distinguish different levels of action understanding. RS has been extensively used in studies of visual processing (Grill-Spector & Malach 2001; Kourtzi & Kanwisher 2000), where it is also referred to as "fMRI-adaptation." It is based on the trial-by-trial reduction of a physiologic response to repeated stimuli. RS is not unique to fMRI and can be observed with electroencephalography and even at the level of single neurons. It can be associated with changes of behavior, such as reaction time priming (Maccotta & Buckner 2004; Wig et al. 2005); however, this behavioral effect is not required. Most importantly, RS can identify changes within a class of stimuli or a level within a potential representational hierarchy rather than between classes of stimuli. For example, areas that recognize the identity of a particular face will show RS effects that are view-, context-, or hairstyle independent.

Suppression occurs when two successive stimuli are represented in the same neural population, and release from suppression occurs when two successive stimuli are represented in different populations. Thus, the method assumes the existence of population coding within brain regions, for which there is extensive evidence in many parts of the cortex (Britten et al. 1993; Georgopoulos et al. 1986). RS also requires the population response to change when the same stimulus feature is repeated. The details of this change are not known. It could be due to a reduction of neuronal firing, a decrease in firing duration, or a sharpening of neuronal tuning curves (Grill-Spector et al. 2006; Krekelberg et al. 2006). Irrespective of these different mechanisms for RS, at the population level their effects on the BOLD-fMRI signal are all the same.

To identify possible topologies in the AON corresponding to a possible hierarchy, a library of stimulus sets designed to induce RS based on different features of simple hand–object actions were developed. In the initial experiment, subjects observed movies of a hand–arm reaching and grasping either of two objects from a first person perspective (de C. Hamilton & Grafton 2006). The actor grasped one of the objects, then lifted it and transported it to the midline and the trial ended. The position of the two objects and which object was grasped were independently manipulated. In this way repetition of trajectory and repetition of hand–object grasp could be separated. The two objects in each trial were closely paired in terms of size and shape, such as a cookie or a computer diskette, so that grip configuration was similar. Subjects performed an incidental monitoring task, and new objects were introduced frequently to avoid the problem of long-term adaptation to a single pair of objects.

The main finding was a strong RS effect in the left anterior intraparietal sulcus (aIPS) when a given object was grasped repeatedly. aIPS was not sensitive to repetition of trajectory. RS effects for trajectory were observed in left lateral occipital sulcus and right superior precentral sulcus. If a single network decoded all of these trajectories and hand–object features similarly then RS would look the same for the two features. The fact that differences emerged provides evidence for dissociable systems for decoding limb trajectory information and hand–object interactions.

Grasping an object, irrespective of how or why it is being grasped, is an action goal in and of itself. The RS results suggest that object-centered action goals are localized to aIPS. This is a generalization of the finding that this area matches visual or tactile features of an object to an appropriate grip. fMRI studies comparing reach and grasp identify greater activity in aIPS for grasp (Binkofski et al. 1998; Culham et al. 2003; Frey et al. 2005). In addition, isolated lesions to this region disrupt grasp control but not reach kinematics (Binkofski et al. 1998; Frak et al. 2006). This evidence leads to the conclusion that the area is a homologue to area AIP identified in nonhuman primates including macaque monkeys (Borra et al. 2008). A traditional view of AIP function is that it is a repository of grip apertures generated from object features. Object shape, size, and orientation have been found to be encoded not only in early visual areas, but also by neurons in monkey area AIP (Murata et al. 2000). Neurons representing three-dimensional shape have been found in the caudal intraparietal sulcus (area CIP) (Sakata et al. 2005; Tsutsui et al. 2005) as well as in an anterior section of the lateral bank of the intraparietal sulcus, area AIP (Sakata et al. 2005) of monkeys. Object-specific firing occurs with or without vision of the grasping hand (Murata et al. 1996). However, in human studies, repetition suppression of objects does not modulate activity in aIPS, whereas the grasp on an object does (Shmuelof & Zohary 2005), and these grasp RS effects occur irrespective of the location of the object in space (Shmuelof & Zohary 2006).

In the RS studies described so far it is assumed that the trajectory information carries lower level information that is subordinate to the more interesting and behaviorally relevant information about what is being grasped. Thus, it can be inferred that these distinctions support a relative functional hierarchy during action observation that is based on distinctions between reach kinematics and the goal of the action defined by what object is grasped. In this case, area aIPS is supraordinate in function to the areas decoding an object identify, limb trajectory, or where an object is located in space. Note that unlike the strict hierarchy of angels described above, the functional anatomic hierarchy implied in this action understanding experiment is relative. Under different circumstances the trajectory information might be more important for action understanding than what is being grasped.

To further test for potential hierarchies in the AON, a second RS experiment was performed (Hamilton & Grafton 2007). Subjects observed movies of an actor reaching and grasping either a wine bottle or a dumbbell placed on end. The trial ended after the object was lifted and placed in a new location. In this study the viewpoint was from the side rather than first person perspective. As in the first experiment there could be repeated observation of grasping for one object or the other, thus allowing an independent replication of the goal—object RS effect determined in the first experiment. Independent of this, there could be repetition for whether the grasp was performed with a power grip along the thick part of the object or a pincer grip along the thin part of the object. In this case the type of grasp did not imply any specific cues as to the actor's intentions because the same action, a transport of the object, was always performed. If aIPS is driven by the specific details of how a hand interacts with an object irrespective of the intention of the grasp then there should be RS when the grip type was repeated. On the other hand, if there was no RS in aIPS for grip type, then the effect observed in the first experiment should be related to a higher order representation of goal—object interactions.

The RS effect for what object was grasped was again localized to left aIPS (extending into the adjacent IPL) and to a lesser degree within right aIPS. These findings provide a strong replication of the first study. Using a statistical threshold appropriate for an exploratory analysis, RS for goal—object was also found in another part of the AON, the left IFG. The RS effect of how the object was grasped identified three clusters in the supplementary motor areas, middle frontal gyrus, mid intraparietal sulcus, and inferior and middle occipital regions. Critically, there was no RS for grasp-type in the anterior IPS or in IPL in either the left or right hemisphere.

As in the first study, it is assumed that the neural substrates for decoding what was being grasped in this task are supraordinate to those decoding how it is being grasped. In this experiment the nature of the grasp had no predictive value on a subsequent action. This implies a functional/anatomic hierarchy where the decoding of grasp, irrespective of object identify, occurs across lateral occipital regions that contribute to a visual analysis of hand—object kinematics for both how the hand approaches an object, the specific grip on the object and the subsequent movement of the object. In contrast, both experiments support a role of aIPS for representing an object as a goal for the hand rather than as a node dedicated to the sensorimotor transformation of visual features to a particular hand shape.

Although grasping an object can be considered an action goal, for most activities this is but an interim step toward object manipulation to achieve more complex goals, as in tool use (Frey 2008). Thus, the two RS experiments described above are not adequate for assessing intentionality at a level of why an actor is performing a particular action with an object. This was studied in a third RS experiment that considered physical outcomes of actions (Hamilton & Grafton 2008). To localize action outcome effects with RS, participants observed actors manipulating objects or tools in a way that led to a specific outcome. For example, in one trial they might see the actor reach and grasp the sliding top of a wooden box and either push or pull the lid. Depending on the starting position of the lid, the outcome of the movement was to either open or close the box. Using RS, the outcome (open or close the box) was independently manipulated relative to the means to accomplish the outcome (push or pull the lid). Over the course of imaging, the subjects watched a battery of movies capturing many familiar behaviors including turning a stove on or off, switching a light on or off, tying or untying a string, drawing or erasing with a pencil, and hammering a nail or a nut by two different means.

An RS effect for outcomes was found in the bilateral IPL and the IFG. That is, the response to an outcome was suppressed when the same outcome was repeated on a second trial, regardless of the means used to generate the outcome. Analysis of the responses to each of the individual sets of movies indicated that the RS effects for outcome in parietal and frontal areas were not driven by a single action or outcome, but generalized across a wide variety of actions.

In the first two experiments the grasping of a particular object could be considered a type of low level outcome in the sense that the final goal—object interaction was the purposeful movement. If so, then there might also be evidence for an RS effect in aIPS for the more complex action outcomes in the third experiment. To test this prediction a region-of-interest analysis was performed in the aIPS. Within this region-of-interest, a weak but significant suppression for repeated outcomes was also detected, supporting a model in which aIPS is involved in action outcomes for tasks spanning a range of complexity. It has recently been proposed that left aIPS carries specific information about the interaction between a grasp and the specific requirements of how to use a tool, which would also be consistent with these findings (Valyear et al. 2007).

An analysis of RS effects for means, comparing repeated movements and novel movements, identified weak effects in left middle intraparietal sulcus, left lateral occipital cortex, and left STS. These results show that visual areas support the general analysis of movement features. Again, if it is assumed that understanding the means to an action are functionally subordinate to understanding the outcome, then the results can be interpreted as supporting a relative hierarchy within the AON.

Hierarchy for Imagined Actions

Relative hierarchy between kinematics and goals has recently been defined using a completely different yet complementary approach to the passive observation RS methods described in the previous section. Instead of passive watching, subjects imagined imitating intransitive gestures made of point light videos of a human actor (Lestou et al. 2008). Subjects observed the human forms performing actions such as kicking, waving, knocking, and throwing. Direct task contrasts of the videos with motion relative to static frames identified the extrastriate body area and areas of STS associated with body movement. Consistent with many prior studies of imagined movement, the contrast of imagined movement relative to passive observation activated bilateral PMv and IPL (Filimon et al. 2007; Grafton et al. 1996; Lotze et al. 1999; Stephan et al. 1995). Using these two contrasts as localizers, the authors tested for RS effects in all of these areas as a function of repetition for different goals or movements used to generate these goals. In this experiment the authors used intransitive actions. Thus, goal of the action was also the “meaning” of the action, such as waving hello, rather than a hand—object interaction coupled to a particular object or use as studied in the other RS studies described above. To manipulate the kind of movement made, the authors played movies backwards or changed the temporal properties of a movie without impairing the observer’s ability to detect the meaning of the action. All areas of the AON generated by the localizer tasks demonstrated robust RS effects when either a goal or movement to achieve a goal was repeated. By examining the RS effects in closer detail it was found that parietal areas and the STS were particularly sensitive to differences between movements with different meaning, whereas premotor areas were more sensitive to differences between movements even when they shared similar meanings. The results are very interesting in light of the previous experiments showing goal and outcome associated RS effects in parietal and premotor areas. However, the imagined imitation adds complexity because it is unclear what cognitive strategy was used to represent the goal. The use of imagined imitation of intransitive gestures introduces some ambiguity because in this case the action goal, that is, the meaning of the action, can be represented either by a semantic route or via direct kinematic matching (Rumiati et al. 2005). In addition, the use of imagined action, while typically constrained to activations associated within the AON, introduces top-down strategies to represent the particular action and maintain it in working memory. Despite these potential problems the results are remarkably consistent with the simpler, object-centered studies based on passive observation. Both found a closer interaction between task goal and parietal activity and task kinematics and premotor activity. What is new in this experiment is the sensitivity within the STS to repetition of action meaning. It is possible that for imitation of intransitive movements the body movement features are the key and possibly exclusive source of insight into what the actor is doing. In terms of a relative hierarchy, the STS then serves a supraordinate or equivalent role to parietal cortex in action understanding.

Potential Pitfalls of the RS Method

There are three important issues that might affect interpretation of RS experiments. The first is the incorrect assumption that the different levels of behavior that are examined for RS effects must be independent of each other for this approach to be informative. This assumption is clearly not true at a behavioral level, as there is much evidence that the kinematics of performed hand actions are altered by the goal of the action (Ansuini et al. 2006; Gentilucci et al. 1997). Thus, in RS imaging experiments there is no requirement that the different levels must be independent parameters of interest. The RS experiments simply identify brain areas most sensitive to one or more levels of control, and this experimental method is biased toward the detection of modularity across functional levels. The results of the RS experiments do not exclude the existence of shared processing across action levels. Obviously, these relative levels of control all interact with each other.

Given the tight coupling between kinematics and goals, a second issue is whether RS effects for outcome are actually due to more subtle differences of kinematics rather than the outcome itself. However, if this were the case, we would expect to see similar RS for both the kinematic effects for different outcomes and also for kinematic differences irrespective of outcomes. This was not the case.

The third issue that arises from these experiments is the possibility of an attentional confound. Parietal cortex, in particular on the right, has been associated with spatial attention (Corbetta et al. 2002), so it becomes important to consider whether the RS effects we observed could be due to manipulation of attention. To test this consideration, participant's attention was manipulated to different task features during action observation (Hamilton & Grafton 2007). The authors did not find any differences in the RS as a function of guided attention. Furthermore, if attention alone were responsible for the RS we identified, we would expect to see RS in the same "attentional" brain region for every contrast. The fact that a distinct set of brain regions for grasps, goals, and outcomes was found is evidence that the RS effects are truly reflecting neuronal population coding in different brain areas, rather than a result of attention.

The Social Milieu

The previous section makes a case for specialization and hierarchy within the AON for a number of simple limb movements based on gestures or hand–object manipulation. In the real world, actions are far more complex, social, and contextualized. A simple change in another person's eye gaze, shift of their posture, or hand gesture, all within rich contextual cues, can convey meaning, agency, and intention. The need to understand in this more complex environment is paramount. A false interpretation of a potential enemy could lead to a faux pas or even death. A fundamental challenge in social neuroscience is to elucidate the mechanisms that allow this process to unfold automatically across a broad range of contexts and contingencies. Recent evidence suggests that action understanding in these more complex scenarios is achieved by more than one solution, and these solutions involve a complex circuitry that extends beyond the AON described so far.

The Social-Network

Evidence for a distinct social-network (S-N) in the neocortex is motivated by studies in which subjects perform any one of a variety of social tasks demonstrating selective activation of the posterior STS, insula, medial prefrontal cortex (mPFC), posterior cingulate, and amygdala (Adolphs 2001; Frith & Frith 1999). These areas are sensitive to biological motion, animacy, the experience and recognition of affect and theory of mind attribution (Allison et al. 2000; Anderson & Phelps 2001; Mitchell et al. 2002; Pelphrey et al. 2004; Saxe et al. 2004; Wheatley et al. 2007). Observation of limb gestures with emotive content (feeling angry, not caring, etc.) is more likely to activate components of the S-N, whereas observing instrumental gestures (come here, look over there, etc.) is more likely to activate left hemisphere parietal and premotor areas of the AON (Gallagher & Frith 2004). The same network is thought to carry more semantic content related to social than mechanical knowledge (Martin & Weisberg 2003). The S-N construct integrates these different components within this circuitry to achieve a unified representation of an animate being.

Explicit Intention Understanding in the Social-Network

In some cases the decoding of intentionality of an animate being can take place within the S-N based on inferential processes rather than direct mapping or simulation. In an important review, Saxe provided a detailed analysis of behavioral errors in intentionality decoding experiments to build a case that inferential reasoning can in fact occur without simulation (Saxe 2005). This is a strong argument against the conclusion that simulation, and in particular simulation as supported exclusively within the mirror neuron system, is sufficient to cover the range of decoding operations performed in the real world.

There is strong evidence that theory of mind tasks, in which subjects are required to make explicit inferential judgments about the intentions of other people's behavior, will activate components of the S-N (Frith & Frith 2006), particularly in areas not typically associated with the AON. For example, Pelphrey showed greater activation in right STS within the S-N for reach-to-grasp actions when the observed action was unsuccessful compared to actions when the hand reached a target (Pelphrey et al. 2004). While the authors argued that this was due to incidental processing of the action, it can be argued that observing unsuccessful actions will evoke explicit intention inference on the part of the observer as he tries to understand why the subject is making an error.

This has been reiterated to a certain degree in a recent fMRI study where subjects were instructed to observe an actor performing actions with an object such as a coffee cup (de Lange et al. 2008). The actor made ordinary movements, unusual movements to accomplish ordinary intention such as sipping, or ordinary movements that resulted in atypical intentions such as placing the coffee cup against the forehead. Subjects made a judgment about whether the means were ordinary or not on some blocks, and whether the intentions were ordinary or not on other blocks of trials. Thus, the judgment task used explicit "top-down" manipulation of attention to identify differences of means and intentions. The main finding was greater activity in components of the S-N including right STS, posterior cingulate, and mPFC during the

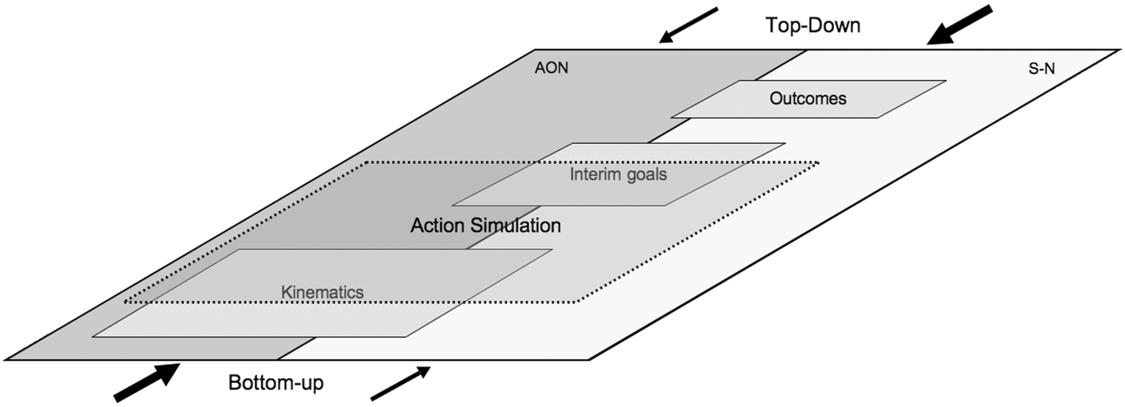

Figure 2. Relationship of action simulation (green) with underlying functional anatomy. The AON (pink) is particularly sensitive to the decoding of kinematics and interim goals, whereas the S-N (yellow) is sensitive to higher level inference about intentions, outcomes, or mental attribution. Action simulation is more closely associated with the AON, but also requires the S-N. The AON is more strongly associated with bottom-up processes and the S-N with top-down processes, but these are not exclusive inputs.

judgment of intentions. Note that the cingulate and prefrontal areas are rarely observed in studies of action observation without explicit instructions to judge a task. When judging the means of an action there was greater activity in the extrastriate body area for unusual actions and greater activity in the bilateral IFG for ordinary actions. What this shows is (1) functional distinctions between the AON and S-N, and (2) the sensitivity of the S-N to explicit top-down control.

With this experimental insight, a hierarchical model of action understanding, motivated in part by Csibra, can be associated with the underlying functional anatomy (Csibra & Gergely 2007). As shown schematically in Fig. 2, action understanding can be achieved through bottom-up processes, what Csibra has termed “resonance,” or through top-down “emulation.” Action simulation emerges most commonly at the interim goal level of representation, although it can be found also at the direct kinematic level of analysis. Not all outcomes, goals, or kinematics are simulated. At first glance the studies reviewed so far in this chapter would support a functional anatomic model where the AON, on the left side of Fig. 2, is driven entirely by bottom-up processes, and the S-N, on the right, is driven only by top-down emulation.

Bottom-Up Recruitment of the Social-Network?

Is inference supported by the S-N always driven by top-down processes? One way to test this is to demonstrate automatic bottom-up activation within the S-N without manipulating a subject’s attention or their mental state during explicit reasoning. Two recent studies provide strong, complementary evidence that intentions can be decoded in the S-N by bottom-up processing. Both studies relied on context effects. In the first, the authors manipulated the plausibility of an action that on its own was quite unusual, such as turning a light switch on with the knee rather than the hand (Brass et al. 2007). This makes no sense under normal circumstances and perfect sense if the hands are carrying a large object. Subjects observed these two sorts of actions as well as a third, implausible condition such as performing the same action with the hands loaded with only a small object that could otherwise be held in one hand. Parametric modulation of BOLD signal activity across these three conditions was present in the STS and posterior STS as well as the mPFC components of the S-N. The subjects were instructed to monitor for interruptions in the movies so that task related modulations of BOLD responses were likely to be incidental.

In the second study, a moving object rather than human form was used as a stimulus (Wheatley et al. 2007). Context in this case influenced whether the object was perceived as an inanimate object, such as a spinning top moving amongst toys, or an ice skater moving around on an ice pond. When context induced a perception of animacy there was robust activation within components of the S-N, including the right STS, mPFC and posterior cingulate. The perception of inanimate motion induced deactivation within these same areas.

There are many other examples of highly prosocial actions driven by bottom-up processes that are increasingly associated with activity in the S-N. Examples include the automatic, imitative behavior people adopt with others (Chartrand & Bargh 1999), the emergence of joint attention by multiple observers (Pelphrey et al. 2003), and the emergence of shared effort by multiple actors working on common goals (Sebanz et al. 2006). The relevance of these examples to implicit social processing has been recently reviewed in detail elsewhere (Frith & Frith 2008).

Top-Down Modulation in the AON

The flip side of the previous argument is whether more abstract interpretive processes can influence simulation by direct matching. To test this, Liepelt and colleagues used a behavioral task and showed that motor priming and influenced the attribution of intention of another persons while they were performing a simple action (Liepelt et al. 2008). Their data support a model in which direct matching can be modulated top-down by the observer's interpretation of whether the observed movement is intended or not. In other words, it is likely that no functional anatomic restriction of the AON to bottom-up processes, or the S-N to top-down processes, exists. Furthermore, while simulation is typically induced by bottom-up resonance, it can also come through emulation.

Action Observation and the Mirror Neuron System

What additional explanatory power is gained by invoking a mirror neuron system description for action observation and simulation? The case is clear that some but not all of the areas of the AON overlap with brain areas associated with motor execution. Knowing this anatomic specificity, can we make a stronger model of action understanding? Before answering this it is important to address the two main assumptions inherent in the MNS explanation: (1) there are mirror neurons in humans, and (2) the functional overlaps observed on imaging between execution and observation are driven by populations of mirror neurons. Knowing the truth of these assumptions allows us to claim that the mirror neuron is a putative substrate for simulation processes. To date the first assumption has not been confirmed. The second has mostly been evaluated through experiments where there might be mirror neurons and researchers employ reverse inference. For example, many fMRI studies have used imitation tasks to define the MNS using either finger or facial movements (Heiser et al. 2003; Iacoboni et al. 1999; Koski et al. 2003; Koski et al. 2002). Both parietal and inferior frontal areas are active during imitative execution or observation. However, imitation requires additional task demands including executive control working memory and semantic as well as motoric routes to action (Rumiati et al. 2005; Tessari & Rumiati 2004). Furthermore, the presence of activity in the MNS during action observation is insufficient evidence toward a conclusion that the same area would be active during execution in the same subject. This is a form of reverse inference that, from a statistical perspective, is insufficient if one is to conclude that there are specific neural substrates for action and perception based on observation alone (Poldrack 2006, 2008). Ideally, it would be desirable to compare actions decoded by observation and generated by execution in the same subjects and then to experiment to further this argument (Dinstein et al. 2007); however, the confines of fMRI constrain the types of movements that can exist. A strong test of the second assumption would be to demonstrate repetition suppression within mirror neuron areas for observing and then to execute the same action, but not a different action. However, no study has found evidence for this cross agent RS in the putative MNS (Dinstein et al. 2007; Dinstein et al. 2008).

If we put these criticisms aside and allowing for the existence of a human MNS, the existence of a human mirror neuron becomes a useful but not necessarily essential neural mechanism that could be used to explain how simulation might occur. These cells carry information about an action that is agent-independent (Gallese et al. 1996). In this sense they are one form of canonical neuron that represents a class of information at a higher level of abstraction. The key point is that there are other types of neurons capable of higher level representations spanning goals and kinematics (Shen & Alexander 1997), intentions (Cisek & Kalaska 2004), or categories and concepts (Miller et al. 2002), and these other types of neurons could also support simulation. Furthermore, the growing number of human imaging studies showing activation during action observation and simulation in the cortex outside of the putative human MNS suggests that even if there is an MNS system, its existence is an insufficient explanation for all facets of action understanding in humans.

Conclusion

We navigate through a complex world and choose our behavior based on insights derived from another's movements, goals, intentions, and emotions. The review does not make the claim that embodied cognition is the only way to understand another's behavior. It is tempting to conclude that simulation for the decoding of action is restricted to the AON and the theory of mind attribution within a S-N. These latter theories are sometimes portrayed as oppositional in that intention is established when either one network or the other is engaged. While there is strong evidence that both of these functional systems are engaged in action understanding, the posing of the functional anatomy as an either/or problem undermines the complexity of the problem at hand. This oppositional construct is likely to be a result of experimental design rather than a physiological or anatomic mechanism. I have tried to make the case that these are compatible, possibly synergistic systems. On both computational and conceptual grounds, the decoding of others' intentions requires the resolution of a cascade of operations including the evaluation of physical actions, the clarification of who the actor is, establishing the context in which an action is performed, and generating possible aims of the actor. The studies on the S-N and AON reviewed above show that they are not mutually exclusive, not limited to top-down or bottom-up processes, and not strictly hierarchical with respect to each other. Each can be supraordinate in terms of a decoding hierarchy, depending on the nature of the action and context. These relationships are shown in Fig. 2.

More broadly, growing evidence from cognitive neuroscience supports the case for the existence of embodied cognition as a distinct knowledge domain localized within the AON and S-N. Much as we think of episodic memory as a special knowledge system with dedicated encoding and retrieval processes, the same argument can be made for embodied cognition. What is particularly interesting about embodied cognition is that it involves a direct interaction between what we do and what we see. It is based on our skills in the organization of action and the complementary ability to perceive action. As in other knowledge domains, there is capacity for generalization, abstraction, and knowledge-sharing irrespective of language.

Acknowledgments

The author thanks Antonia Hamilton, Emily Cross, Stephanie Ortigue, and the staffs at the Dartmouth and UCSB Brain Imaging Centers. Supported in part by PHS grants NS33504 and NS44393, the Dana Foundation, and a grant DAAD19-03-D-0004 to the Institute for Collaborative Biotechnologies from the U.S. Army Research Office.

Conflicts of Interest

The author declares no conflicts of interest.

References

Adolphs, R. (2001). The neurobiology of social cognition. *Curr. Opin. Neurobiol.*, *11*, 231–239.

Aglioti, S. M., Cesari, P., Romani, M., et al. (2008). Action anticipation and motor resonance in elite basketball players. *Nat. Neurosci.*, *11*, 1109–1116.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends Cogn. Sci.*, *4*, 267–278.

Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, *411*, 5.

Anquetil, T., & Jeannerod, M. (2007). Simulated actions in the first and in the third person perspectives share common representations. *Brain Res.*, *1130*, 125–129.

Ansuini, C., Santello, M., Massaccesi, S., et al. (2006). Effects of end-goal on hand shaping. *J. Neurophysiol.*, *95*, 2456–2465.

Aziz-Zadeh, L., Maeda, F., Zaidel, E., et al. (2002). Lateralization in motor facilitation during action observation: A TMS study. *Experimental Brain Research*, *144*, 127–131.

Badre, D., & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *J. Cogn. Neurosci.*, *19*, 2082–2099.

Baldissera, F., Cavallari, P., Craighero, L., et al. (2001). Modulation of spinal excitability during observation of hand actions in humans. *Eur. J. Neurosci.*, *13*, 190–194.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., et al. (2003). fMRI responses to video and point-light displays of moving humans and manipulable objects. *J. Cogn. Neurosci.*, *15*, 991–1001.

Bekkering, H., Wohlschlager, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Q. J. Exp. Psychol. A*, *53*, 153–164.

Binkofski, F., Dohle, C., Posse, S., et al. (1998). Human anterior intraparietal area subserves prehension: A combined lesion and functional MRI activation study. *Neurology*, *50*, 1253–1259.

Borra, E., Belmalih, A., Calzavara, R., et al. (2008). Cortical connections of the macaque anterior intraparietal (AIP) area. *Cereb Cortex*, *18*, 1094–1111.

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci. (Regul. Ed.)*, *12*, 201–208.

Brass, M., Schmitt, R. M., Spengler, S., et al. (2007). Investigating action understanding: Inferential processes versus action simulation. *Curr. Biol.*, *17*, 2117–2121.

Britten, K. H., Shadlen, M. N., Newsome, W. T., et al. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.*, *10*, 1157–1169.

Brooks, R. (April 1991). Intelligence without reason. Paper presented at the *Proceedings of the International Joint Conference on Artificial Intelligence*, San Mateo, CA.

Buccino, G., Binkofski, F., Fink, G. R., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *Eur. J. Neurosci.*, *13*, 400–404.

Buccino, G., Binkofski, F., & Riggio, L. (2004). The mirror neuron system and action recognition. *Brain Lang.*, *89*, 370–376.

Buccino, G., Lui, F., Canessa, N., et al. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: An fMRI study. *J. Cogn. Neurosci.*, *16*, 114–126.

Buxbaum, L. J., Kyle, K. M., & Menon, R. (2005). On beyond mirror neurons: Internal representations subserving imitation and recognition of skilled object-related actions in humans. *Brain Res. Cogn. Brain Res.*, *25*, 226–239.

Calvo-Merino, B., Glaser, D. E., Grezes, J., et al. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex*, *15*, 1243–1249.

Calvo-Merino, B., Grèzes, J., Glaser, D. E., et al. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Curr. Biol.*, *16*, 1905–1910.

Celnik, P., Stefan, K., Hummel, F., et al. (2006). Encoding a motor memory in the older adult by action observation. *NeuroImage*, *29*, 677–684.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *NeuroImage*, *12*, 478–484.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *J. Pers. Soc. Psychol.*, *76*, 893–910.

Cisek, P., & Kalaska, J. F. (2004). Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature*, *431*, 993–996.

Cooper, R., & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychol. Rev.*, *113*, 887–916; discussion 917–831.

Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. *J. Cogn. Neurosci.*, *14*, 508–523.

Craighero, L., Bello, A., Fadiga, L., et al. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia*, *40*, 492–502.

Craighero, L., Fadiga, L., Rizzolatti, G., et al. (1999). Action for perception: A motor-visual attentional effect. *J. Exp. Psychol. Hum. Percept. Perform.*, *25*, 1673–1692.

Cross, E., Kraemer, D., Hamilton, A., et al. (2008). Sensitivity of the action observation network to physical and observational learning. *Cerebral Cortex*, *12*.

Cross, E. S., Hamilton, A. F., & Grafton, S. T. (2006). Building a motor simulation *de novo*: Observation of dance by dancers. *NeuroImage*, *31*, 1257–1267.

Csibra, G., & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta. Psychologica*, *124*, 60–78.

Culham, J. C., Danckert, S. L., DeSouza, J. F., et al. (2003). Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Exp. Brain Res.*, *153*, 180–189.

de C. Hamilton, A. F., & Grafton, S. T. (2007). The motor hierarchy: From kinematics to goals and intentions. In P. Haggard, Y. Rossetti & M. Kawato (Eds.), *Sensorimotor Foundations of Higher Cognition* (Vol. Attention and Performance XXII, pp. 381–408). Oxford: Oxford University Press.

de Lange, F. P., Spronk, M., Willems, R. M., et al. (2008). Complementary systems for understanding action intentions. *Curr. Biol.*, *18*, 454–457.

Demiris, Y., & Simmons, G. (2006). Perceiving the unusual: Temporal properties of hierarchical motor representations for action perception. *Neural Networks: Official J. Inter. Neural Network Soc.*, *19*, 272–284.

Dinstein, I., Hasson, U., Rubin, N., et al. (2007). Brain areas selective for both observed and executed movements. *J. Neurophysiol.*, *98*, 1415–1427.

Dinstein, I., Thomas, C., Behrmann, M., et al. (2008). A mirror up to nature. *Curr. Biol.*, *18*, R13–18.

Fadiga, L., Fogassi, L., Pavesi, G., et al. (1995). Motor facilitation during action observation: A magnetic stimulation study. *J. Neurophysiol.*, *73*, 2608–2611.

Ferro, J. M., Martins, I. P., Mariano, G., et al. (1983). CT scan correlates of gesture recognition. *J. Neurol. Neurosurg. Psych.*, *46*, 943–952.

Filimon, F., Nelson, J. D., Hagler, D. J., et al. (2007). Human cortical representations for reaching: Mirror neurons for execution, observation, and imagery. *NeuroImage*, *37*, 1315–1328.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *Quart. J. Exp. Psych. (2006)*, *61*, 825–850.

Frak, V., Paulignan, Y., Jeannerod, M., et al. (2006). Prehension movements in a patient (AC) with posterior parietal cortex damage and posterior callosal section. *Brain Cogn.*, *60*, 43–48.

Frey, S. H. (2008). Tool use, communicative gesture and cerebral asymmetries in the modern human brain. *Philos. Trans. R. Soc. Lond., B., Biol. Sci.*, *363*, 1951–1957.

Frey, S. H., & Gerry, V. E. (2006). Modulation of neural activity during observational learning of actions and their sequential orders. *J. Neurosci.*, *26*, 13194–13201.

Frey, S. H., Vinton, D., Newman-Norlund, R., et al. (2005). Cortical topography of human anterior intraparietal cortex active during visually-guided grasping. *Cogn. Brain Res.*, *23*, 397–405.

Frith, C. D., & Frith, U. (1999). Interacting minds—A biological basis [see comments]. *Science*, *286*, 1692–1695.

Frith, C. D., & Frith, U. (2006). How we predict what other people are going to do. *Brain Res.*, *1079*, 36–46.

Frith, C. D., & Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, *60*, 503–510.

Gallagher, H. L., & Frith, C. D. (2004). Dissociable neural pathways for the perception and recognition of expressive and instrumental gestures. *Neuropsychologia*, *42*, 1725–1736.

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

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends Cogn. Sci.*, *8*, 396–403.

Gazzola, V., Aziz-Zadeh, L., & Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Curr. Biol.*, *16*, 1824–1829.

Gentilucci, M., Negrotti, A., & Gangitano, M. (1997). Planning an action. *Exp. Brain Res.*, *115*, 116–128.

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, *233*, 1416–1419.

Goldenberg, G., & Karnath, H. O. (2006). The neural basis of imitation is body part specific. *J. Neurosci.*, *26*, 6282–6287.

Grafton, S. T., Arbib, M. A., Fadiga, L., et al. (1996). Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Exp. Brain Res.*, *112*, 103–111.

Grafton, S. T., Fadiga, L., Arbib, M. A., et al. (1997). Premotor cortex activation during observation and naming of familiar tools. *NeuroImage*, *6*, 231–236.

Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Hum. Brain Mapp.*, *12*, 1–19.

Grezes, J., Tucker, M., Armony, J., et al. (2003). Objects automatically potentiate action: An fMRI study of implicit processing. *Eur. J. Neurosci.*, *17*, 2735–2740.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends Cogn. Sci. (Regul. Ed.)*, *10*, 14–23.

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta. Psychol. (Amst.)*, *107*, 293–321.

Hamilton, A., Wolpert, D., & Frith, U. (2004). Your own action influences how you perceive another person's action. *Curr. Biol.*, *14*, 493–498.

Hamilton, A. F., & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *J. Neurosci.*, *26*, 1133–1137.

Hamilton, A. F., Joyce, D. W., Flanagan, J. R., et al. (2007). Kinematic cues in perceptual weight judgment and their origins in box lifting. *Psychol. Res.*, *71*, 13–21.

Hamilton, A. F. d. C., & Grafton, S. T. (2008). Action outcomes are represented in human inferior frontoparietal cortex. *Cereb. Cortex*, *18*, 1160–1168.

Handy, T. C., Grafton, S. T., Shroff, N., et al. (2003). Graspable objects grab attention when the potential for action is recognized. *Nat. Neurosci.*, *6*, 421–427.

Hari, R., Forss, N., Avikainen, S., et al. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proc. Natl. Acad. Sci. USA*, *95*, 15061–15065.

Heilman, K. M., Rothi, L. J., & Valenstein, E. (1982). Two forms of ideomotor apraxia. *Neurology*, *32*, 342–346.

Heiser, M., Iacoboni, M., Maeda, F., et al. (2003). The essential role of Broca's area in imitation. *Eur. J. Neurosci.*, *17*, 1123–1128.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., et al. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.*, *3*, e79.

Iacoboni, M., Woods, R. P., Brass, M., et al. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: A critique. *Trends Cogn. Sci.*, *9*, 21–25.

Janata, P., & Grafton, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nat. Neurosci.*, *6*, 682–687.

Johnson-Frey, S. H., Maloof, F. R., Newman-Norlund, R., et al. (2003). Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron*, *39*, 1053–1058.

Koski, L., Iacoboni, M., Dubeau, M. C., et al. (2003). Modulation of cortical activity during different imitative behaviors. *J. Neurophysiol.*, *89*, 460–471.

Koski, L., Wohlschlager, A., Bekkering, H., et al. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cereb. Cortex*, *12*, 847–855.

Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *J. Neurosci.*, *20*, 3310–3318.

Krekelberg, B., Boynton, G. M., & van Wezel, R. J. (2006). Adaptation: From single cells to BOLD signals. *Trends Neurosci.*, *29*, 250–256.

Lestou, V., Pollick, F. E., & Kourtzi, Z. (2008). Neural substrates for action understanding at different description levels in the human brain. *J. Cogn. Neurosci.*, *20*, 324–341.

Liepelt, R., Cramon, D. Y., & Brass, M. (2008). What is matched in direct matching? Intention attribution modulates motor priming. *J. Exp. Psych. Hum. Perc. Perf.*, *34*, 578–591.

Lotze, M., Montoya, P., Erb, M., et al. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: An fMRI study. *J. Cogn. Neurosci.*, *11*, 491–501.

Maccotta, L., & Buckner, R. L. (2004). Evidence for neural effects of repetition that directly correlate with behavioral priming. *J. Cogn. Neurosci.*, *16*, 1625–1632.

Martin, A., & Weisberg, J. (2003). Neural foundations for understanding social and mechanical concepts. *Cogn. Neuropsych.*, *20*, 575–587.

Miller, E. K., Freedman, D. J., & Wallis, J. D. (2002). The prefrontal cortex: Categories, concepts and cognition. *Philos. Trans. R. Soc. Lond., B., Biol. Sci.*, *357*, 1123–1136.

Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proc. Natl. Acad. Sci. USA*, *99*, 15238–15243.

Murata, A., Gallese, V., Kaseda, M., et al. (1996). Parietal neurons related to memory-guided hand manipulation. *J. Neurophysiol.*, *75*, 2180–2186.

Murata, A., Gallese, V., Luppino, G., et al. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J. Neurophysiol.*, *83*, 2580–2601.

Niedenthal, P. M. (2007). Embodying emotion. *Science*, *316*, 1002–1005.

Pazzaglia, M., Smania, N., Corato, E., et al. (2008). Neural underpinnings of gesture discrimination in patients with limb apraxia. *J. Neurosci.*, *28*, 3030–3041.

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. *J. Cogn. Neurosci.*, *16*, 1706–1716.

Pelphrey, K. A., Singerman, J. D., Allison, T., et al. (2003). Brain activation evoked by perception of gaze shifts: The influence of context. *Neuropsychologia*, *41*, 156–170.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci. (Regul. Ed.)*, *10*, 59–63.

Poldrack, R. A. (2008). The role of fMRI in cognitive neuroscience: Where do we stand? *Curr. Opin. Neurobiol.*, *18*, 223–227.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169–192.

Rizzolatti, G., Fadiga, L., Matelli, M., et al. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Exp. Brain Res.*, *111*, 246–252.

Rothi, L. J., Heilman, K. M., & Watson, R. T. (1985). Pantomime comprehension and ideomotor apraxia. *J. Neurol. Neurosurg. Psych.*, *48*, 207–210.

Rumiati, R. I., Weiss, P. H., Tessari, A., et al. (2005). Common and differential neural mechanisms supporting imitation of meaningful and meaningless actions. *J. Cogn. Neurosci.*, *17*, 1420–1431.

Sakata, H., Tsutsui, K., & Taira, M. (2005). Toward an understanding of the neural processing for 3D shape perception. *Neuropsychologia*, *43*, 151–161.

Saxe, R. (2005). Against simulation: the argument from error. *Trends Cogn. Sci. (Regul. Ed.)*, *9*, 174–179.

Saxe, R., Xiao, D. K., Kovacs, G., et al. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, *42*, 1435–1446.

Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain*, *130*, 2452–2461.

Saygin, A. P., Wilson, S. M., Dronkers, N. F., et al. (2004). Action comprehension in aphasia: Linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia*, *42*, 1788–1804.

Saygin, A. P., Wilson, S. M., Hagler, D. J., et al. (2004). Point-light biological motion perception activates human premotor cortex. *J. Neurosci.*, *24*, 6181–6188.

Schubotz, R. I., & von Cramon, D. Y. (2001). Functional organization of the lateral premotor cortex: fMRI reveals different regions activated by anticipation of object properties, location and speed. *Brain Res. Cogn. Brain Res.*, *11*, 97–112.

Schubotz, R. I., & von Cramon, D. Y. (2004). Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery. *J. Neurosci.*, *24*, 5467–5474.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends Cogn. Sci.*, *10*, 70–76.

Shen, L., & Alexander, G. E. (1997). Preferential representation of instructed target location versus limb trajectory in dorsal premotor area. *J. Neurophysiol.*, *77*, 1195–1212.

Shmuelof, L., & Zohary, E. (2005). Dissociation between ventral and dorsal fMRI activation during object and action recognition. *Neuron*, *47*, 457–470.

Shmuelof, L., & Zohary, E. (2006). A mirror representation of others' actions in the human anterior parietal cortex. *J. Neurosci.*, *26*, 9736–9742.

Stephan, K. M., Fink, G. R., Passingham, R. E., et al. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *J. Neurophysiol.*, *73*, 373–386.

Tenenbaum, G., Levy-Kolker, N., Bar-Eli, M., et al. (1994). Information recall of younger and older skilled athletes: The role of display complexity, attentional resources and visual exposure duration. *J. Sports Sci.*, *12*, 529–534.

Tessari, A., & Rumiati, R. I. (2004). The strategic control of multiple routes in imitation of actions. *J. Exp. Psychol. Hum. Percept Perform*, *30*, 1107–1116.

Thelen, E., Schoner, G., Scheier, C., et al. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behav. Brain Sci.*, *24*, 1–34; discussion 34–86.

Thompson, J. C., Hardee, J. E., Panayiotou, A., et al. (2007). Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *NeuroImage*, *37*, 966–973.

Tsutsui, K., Taira, M., & Sakata, H. (2005). Neural mechanisms of three-dimensional vision. *Neurosci. Res.*, *51*, 221–229.

Ulloa, E. R., & Pineda, J. A. (2007). Recognition of point-light biological motion: Mu rhythms and mirror neuron activity. *Behav. Brain Res.*, *183*, 188–194.

Urgesi, C., Candidi, M., Ionta, S., et al. (2007). Representation of body identity and body actions in extrastriate body area and ventral premotor cortex. *Nat. Neurosci.*, *10*, 30–31.

Urgesi, C., Moro, V., Candidi, M., et al. (2006). Mapping implied body actions in the human motor system. *J. Neurosci.*, *26*, 7942–7949.

Valyear, K. F., Cavina-Pratesi, C., Stiglick, A. J., et al. (2007). Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp? *NeuroImage*, *36*(Suppl 2), T94–T108.

Weisberg, J., van Turennout, M., & Martin, A. (2007). A neural system for learning about object function. *Cereb. Cortex*, *17*, 513–521.

Wheatley, T., Milleville, S. C., & Martin, A. (2007). Understanding animate agents: distinct roles for the social network and mirror system. *Psych. Sci.: A J. Amer. Psych. Soc./APS*, *18*, 469–474.

Wig, G. S., Grafton, S. T., Demos, K. E., et al. (2005). Reductions in neural activity underlie behavioral components of repetition priming. *Nat. Neurosci.*, *8*, 1228–1233.

Zentgraf, K., Stark, R., Reiser, M., et al. (2005). Differential activation of pre-SMA and SMA proper during action observation: Effects of instructions. *NeuroImage*, *26*, 662–672.