



Published in final edited form as:

Psychol Res. 2009 July ; 73(4): 441–462. doi:10.1007/s00426-009-0242-2.

Tool use and the distalization of the end-effector

Michael A. Arbib¹, James B. Bonaiuto², Stéphane Jacobs³, and Scott H. Frey^{3,4}

¹ Department of Computer Science and Neuroscience Graduate Program, University of Southern California, Los Angeles, CA 90089-2520, USA, e-mail: arbib@usc.edu

² Neuroscience Graduate Program, University of Southern California, Los Angeles, CA 90089-2520, USA

³ Psychology Department, University of Oregon, Eugene, USA

⁴ Lewis Center for Neuroimaging, University of Oregon, Eugene, USA

Abstract

We review recent neurophysiological data from macaques and humans suggesting that the use of tools extends the internal representation of the actor's hand, and relate it to our modeling of the visual control of grasping. We introduce the idea that, in addition to extending the body schema to incorporate the tool, tool use involves *distalization of the end-effector from hand to tool*. Different tools extend the body schema in different ways, with a displaced visual target and a novel, task-specific processing of haptic feedback to the hand. This distalization is critical in order to exploit the unique functional capacities engendered by complex tools.

Introduction: tool use and the body schema

Nearly a century ago, Head and Holmes (1911) introduced the notion of a multimodal body schema, a dynamical multi-sensory representation of the body that is constantly updated to code the location and configuration of the body parts in space. They proposed that this representation may incorporate objects when relevant for motor control. Subsequent experimental work has elegantly demonstrated that the body schema is modified when using simple tools (rakes, sticks) that extend our reach (Berti & Frassinetti, 2000; Farnè et al., 2005; Iriki et al., 1996; Maravita & Iriki, 2004; Maravita et al., 2002). Neurophysiological data from macaques suggests that these changes involve parietal (VIP; Iriki et al., 1996) regions that are known to support multimodal representations of peripersonal space. However, expansion of the body representation alone cannot explain how the end-effector properties of a tool are represented (Johnson-Frey, 2004). A screwdriver, for instance, not only extends the reach of the user, but also expands the functional capacities of the hand.

In this paper, we focus on the use of hand-held tools. To be more precise, we focus on tools for which the degrees of freedom of the end-effector are directly controlled by the degrees of freedom of the hand, even if the correspondence is indirect. Our thesis is that the representation of such manually controlled tools cannot simply be reduced to alterations in the representation of peripersonal space. Instead, we propose that effective control of tools involves both a tool-specific expansion of the body schema and a mapping that captures how movements of the hands are transformed into actions of the tool's end-effector (representations of so-called motor-to-mechanical transformations; Frey, 2007). This may involve both a new locus of attention for visual feedback and a dramatic reinterpretation of manual hapsis. We further

suggest that these processes will involve regions of the brain that are involved in the control of distal manual actions, for instance the representation of grasp.

Nonetheless, much of what we say about hand-held tools has implications for the general case—a back hoe might be seen as akin to a spade, but with much greater force amplification and a much reduced correspondence between the body movements of the operator and the motions of the end-effector.

The body schema as an assemblage

The process whereby a tool becomes an extension of the hand to perform a specific task can be related to the flexible view of the *body schema* offered by Head and Holmes (1911): “Anything which participates in the conscious movement of our bodies is added to the model of ourselves and becomes part of those schemata: a woman’s power of localization may extend to the feather of her hat.” What is telling here is this (rather dated) reference to the feather on the woman’s hat—the body schema is not a fixed entity, but changes with one’s circumstances, whether one is wearing a feather-tipped hat, using a tool, or driving a car. The notion, then, is that the body schema may better be viewed as an assemblage of schemas of much finer granularity—such as interlinked perceptual and motor schemas for recognition of aspects of objects and deployment of actions (Arbib, 1981)—which can be integrated in different combinations as context and circumstances change. Each tool extends the body schema in a different way, depending on the functional properties of the end-effector employed. Thus, with regard to the internal representations underlying motor control, a tool extends a hand’s functionality rather than replacing the hand.

Holmes, Calvert, and Spence (2004)—a different Holmes, 93 years later!—had subjects use a tool to interact with an object by using its tip or its handle. When subjects used the tip of the tool to interact with the object, visuo-tactile interactions were stronger at the tip of the tool (visual stimuli presented at the tip of the tool interfered more strongly with tactile stimuli to the hand). However, when subjects used the handle of the tool to interact with an object, visuo-tactile interactions were strongest near the hands and decreased with distance along the shaft of the tool. In line with this, Farnè et al. (2005) showed that the extension of near space is contingent upon the effective size of the tool, not its absolute length. Indeed, use of two different rakes, one 30-cm long and the other 60 cm but with its tines attached to the middle of its shaft, 30 cm away from the hand, was associated with similar expansion of peripersonal space. These studies suggest that the changes in body schema that accompany tool use are very much dependent on the functional end of the tool. In the next section, we consider tools with more complex end-effector properties.

Distalizing the end-effector

We now chart the difference between grasping, manipulation, and tool use. *Grasping* involves recognizing certain surfaces of an object as *affordances* for grasping (the targets for placement of finger tips or other parts of the hand), moving the hand toward the object with simultaneous pre-shaping of the hand appropriately to (but somewhat further apart than) these affordances, and finally enclosing the target areas with the appropriate portions of the hand to establish the grasp as the hand reaches the object (Iberall et al., 1986; Jeannerod et al., 1995; Jeannerod & Biguer, 1982). *Manipulation* of an object requires that one first grasp it, then use the hand to change the state of the object in some way. Overall success may require intermittent release and regrasp of the object and may also involve bimanual coordination (consider opening a child-proof aspirin bottle).

We now turn to the use of tools. As it is being picked up, the tool is like any other object being manipulated. Moreover, the way many tools are grasped depends on how we intend to use them

(Johnson-Frey, 2003, 2004; Johnson & Grafton, 2003). Note that this phase may exploit more than one affordance of the tool—grasping a screwdriver combines a power grasp of the handle with a side opposition to direct the shaft (Fig. 1). But once the tool is ready to be used an important change takes place. The “end-effector” is no longer the hand, but some part of the tool that will be used to act upon some other object. For example, once the screwdriver is grasped appropriately, the blade of the shaft becomes the end-effector, to be matched against the affordance provided by the slot in the screw head. Or consider tightening a nut and bolt with either the hand or a pair of pliers. Here, the jaws of the pliers function just like the opposed fingers of the hand both for grasping the nut and for holding it while turning it. By contrast, the relation between hand and tool motion is less direct in using a screwdriver. The screwdriver does not grasp the screw—indeed, the equivalence is in terms of what the screw does (moving round and down into the substrate) rather than what the endeffector is doing. Here, a case might be made for a different correspondence—between the motion in using a screwdriver and the motion in turning a screw after inserting one’s thumb nail in the screw slot. However, in much tool use there is no direct spatiomotor correspondence between tool motion and that of the hand in some related task.

Different parts of the hand serve as end-effector for different manual actions so that the problem of how to control the arm/hand with respect to a task-dependent endeffector is not unique to tool use. However, in manipulation and tool use, there is a “change of focus” from the initial grasping of a tool or object to manipulation of the object or use of the tool in which the end-effector migrates distally from the hand. Note that a shift in attention is critical here—from the interaction of the hand with the affordances for grasping offered by the tool to the interaction of the tool with affordances of some object for the action of the tool. Yet it is still the hand that has to be controlled (and involves different parts of the hand in different tasks), but with a displaced visual target and a novel, task-specific processing of haptic feedback required for motor control.

There is also the fact of what we shall call, quite grandly, “iterative distalization”, extending the example of simple distalization where the end-effector, or locus of control, moves outward from the body to the end of the screwdriver blade. To iterate this, we may consider that the screw is the endeffector. As our skill increases, we repeatedly distalize our visual attention from the hand to the blade of the screwdriver to the progress of the screw head toward becoming flush with the surface into which the screw is advancing, while at the same time the haptic feedback to our hand becomes associated with the object of our visual attention. This is not to deny that certain tasks can be carried out without visual attention, or that we may increasingly rely on other sources of feedback rather than vision as our skill increases.

Recall our focus on tools for which the degrees of freedom of the end-effector are directly controlled by the degrees of freedom of the hand. Contrast the case where someone pokes a chain of standing dominoes and the falling of each one triggers the motion of the next domino in line. Here, the events subsequent to the initial poke require no further action or attention on the part of the pocker. By contrast, we must move the hand to turn the screwdriver, and we must turn the screwdriver to turn the screw in such a way that the screw moves into the wood. The same action undertaken with different biological effectors may display certain kinematic regularities (for handwriting, see Wing, 2000). These have also been observed during grasping actions performed with a tool or the hand (Gentilucci et al., 2004). This suggests that some abstract effector-independent representation provides criteria for success in achieving a goal that guides learning. This includes, but goes well beyond, the notion of motor equivalence or a direct correspondence between end-effectors. Consider using a knife to slice bread. The goal here is to break a loaf into pieces, but tearing a loaf apart with two hands does not have motor equivalence with using a bread knife. When, in later sections, we consider the use of pliers (Umiltà et al., 2008) or a grasping tool (Jacobs, Danielmeier, & Frey, 2009) we *will* see a

correspondence between the motions of the end-effectors—whether tool or hand—but we must note that these are special cases. In general there will be no such equivalence.¹ In some cases we may speak of goal state equivalence rather than motor equivalence, whereas in other cases the machine may carry out actions that are unavailable to the non-augmented human.

To clarify this distinction, consider the goal of writing a cursive lower case letter “a”. As is well known (but see Wing, 2000 for a scholarly study), one can use a long pole with a paint brush on the end to write on a wall in what, despite little skill, is recognizably one’s own handwriting even if one has never attempted the task before. The “secret” is that we have an abstract effector-independent representation in terms of the relative size and position of marker strokes which is highly automatized in the case of a pen or pencil but which can be invoked to control the novel end-effector of the pole-attached paintbrush, using visual feedback to match this novel performance against our visual expectations. Thus, the generic goal of writing the “a” will invoke an automatized effector-specific motor schema if the tool is a pen or pencil, but can control other tools via an “effector-independent” representation. The latter creates expectations that can be used for feedback control of a specific, albeit unfamiliar, effector—focusing on the end-effector of the writing action, specifying the trajectory and kinematics of its movement, whether it be the pencil, the piece of chalk on a blackboard, etc. With increased practice, “internal models” (more on these below) will be developed for the use of the novel effector, increasing the speed and accuracy of the movement.

In summary, we agree that an important aspect of tool use can be understood as expanding the body schema, and that the parietal and frontal areas involved in constructing representations of peripersonal space are likely candidates for these experience-dependent changes. However, tools such as pliers not only expand peripersonal space, but also transform the actions of the hand into different behaviors at the end-effector (e.g., power grip to precision grip). We demonstrate below how such changes associated with the use of complex tools (Frey, 2007; Johnson-Frey, 2004) induce experience-dependent changes in the parieto-frontal grasp network.

Basic parieto-frontal interactions for visually directed hand movements

Many neurons in premotor area F5 (Fig. 2) fire in association with specific types of manual action, such as precision grip, finger prehension, and whole hand prehension (Rizzolatti et al., 1988) as well as tearing and holding. Some neurons in F5 discharge only during the last part of grasping; others start to fire during the phase in which the hand opens and continue to discharge during the phase when the hand closes; finally a few discharge prevalently in the phase in which the hand opens. Grasping appears, therefore, to be coded by the joint activity of populations of neurons, each controlling different phases of the motor act. Fagg and Arbib (1998) (see Fig. 6 below) model the interactions between F5 and the anterior area of the intraparietal sulcus (AIP, Fig. 2) that yield these phasic relations.

Some neurons in F5 discharge when the monkey grasps an object using different effectors, for example either hand, or the mouth (Rizzolatti et al., 1988) suggesting that these neurons code the *goal* of the motor act (Rizzolatti et al., 1988). However, when turning to motor control, one

¹In the study of imitation (Dautenhahn & Nehaniv, 2002), the *correspondence problem* for the imitator is to decide how to match endeffectors of the imitator’s body to its own. We have brains and bodies that allow us to master the use of varied tools, but that rests, for most of us, on early training from caregivers that enables us to learn how to attend to a range of necessary affordances and master a set of basic uses of our hands and body, providing the effectivities which are dual to the affordances provided by perception (Turvey, 1992; Turvey et al., 1981; Zukow-Goldring & Arbib, 2007) This allows us to “get the idea” of novel tools by a mix of verbal instruction, complex imitation, mechanical reasoning (Goldenberg & Hagmann, 1998), and trial-and-error—the latter being crucial as we hone skill in using a tool through extended practice with it.

must address the issue of where and how this generic constraint gets turned into the specific commands for the movement of the selected effector.

Computational analyses (Arbib et al., 1985; Oztop et al., 2004) suggest that the motion of the hand in an object-centered reference frame must be coded at some level, and indeed that this may be an essential part of learning for the mirror neuron system. While this proposal has played an important role in computational modeling and robotics, a formal test of the associated neural encoding was never carried out. However, it does lead to the hypothesis that, as we turn from grasping an object to using a tool to grasp an object, many neurons will code the movement of the endeffector of the tool, rather than that of the hand, in a reference frame centered on the relevant affordance of the manipulandum. Just as grasping with hand versus mouth, this requires a dramatic change in motor control to achieve the same goal.

Many studies (e.g., Alexander & Crutcher, 1990; Crutcher & Alexander, 1990; Evarts, 1968; Schwartz et al., 1988) have sought to establish which movement parameters are coded by the cortical motor areas. *Reaching* neurons recorded from the premotor and primary cortical areas fall into two main categories: neurons that code movements in terms of joints or muscles, and neurons whose activity correlates with “extrinsic” variables such as direction of the object acted upon (but this could be retinotopic, or relative to the body), goal of the movement, or arm movement direction.

Here, a brief aside will be useful to clarify the difference between “coordinates” and “reference frame”. For example, many visual areas code the position of an object in a retinotopic frame, where the peak activity of a cell occurs only if its feature is located near a particular point on the retina. The experimentalist may use an (x,y) or (r,θ) coordinate system to specify points on the retina, but the brain makes no use of such a system. As we shall see below, some F5 neurons can be characterized within a framework of hand postures, but the firing of the cells in no way relates to a coordinate system for this space, e.g., specification of all finger angles. Instead, the brain works through population codes (Ma et al., 2006; Pouget et al., 1999) in which the necessary information is distributed across the firing of many neurons, which can then be “read out” by other sets of neurons—but, again, not in terms of a conventional set of numerical coordinates.

Since we are much concerned with area F5 of the macaque, it is important to stress that there are many different types of neurons in this area, only some of which are directly related to movement. Here, we are primarily concerned with the control of hand movements, but must note that there is a somatotopy for motions of other regions of the body, with the hand area grading into the oro-facial area. The firing of motor-related neurons correlates with the animal’s execution of a relatively limited set of actions. An important subset of these neurons are the *mirror neurons* which are active not only when the macaque executes a certain class of actions but also when the macaque observes another macaque or a human performs a more or less similar action. It has been found that visual activation of macaque mirror neurons in observing another’s manual action requires either vision of both hand and object, or recent vision of how the hand is moving toward the object, or (for a subclass, the audiovisual neurons) the sound of the action if that is sufficiently distinctive (Bonaiuto et al., 2007; Kohler et al., 2002; Umiltà et al., 2001). By contrast, *canonical neurons* are those which fire during execution but not during the observation of others. Thus, when Raos et al. (2006) say that discharge of canonical neurons precedes the beginning of movement by several hundreds of milliseconds and continues during hand movements, we stress that different neurons fire for different phases of this period. Only about half of canonical neurons have visual responses (Rizzolatti et al., 1988).

There are neurons that code force during *grasping* both in the ventral premotor cortex and in M1/F1 (Hepp-Reymond et al., 1994; Maier et al., 1993) and PMv projects heavily to the hand representation in F1 (Dum & Strick, 2005). As already noted, some neurons in F5 discharge when the monkey achieves the same goal despite the effector used, and thus might encode the goal of the motor act of grasping. Raos et al. (2006) found that “grip posture rather than object shape” determined F5 activity, and that F5 neurons selective for both grip type and wrist orientation maintained this selectivity when grasping in the dark. However, the contrast between grip posture and object shape seems mistaken once we note that visual input about object shape is used to determine the affordance for grasping, but that working memory for affordance/grip posture may substitute for visual input when reaching in the dark. Simultaneous recording from F5 and F1 showed that F5 neurons were selective for grasp type and phase, while an F1 neurons might be active for different phases of different grasps (Umiltà et al., 2007). This suggests that F5 neurons encode a high-level representation of the grasp motor schema while F1 neurons (or, at least, some of them) encode the component movements or components of a population code for muscle activity of each grasp phase.

Increasingly, research on PMv has stressed its role in transformation from a visual to a motor reference frame using experimental paradigms that dissociate the two. Kakei et al. (2001, 2003) recorded from F5 and F4 during a wrist movement task with the wrist starting in different positions, allowing them to dissociate muscle activity, direction of joint movement (intrinsic coordinates), and the direction of the hand movement in space (extrinsic). They found three types of neurons: those that coded the direction of movement in space (extrinsic-like), those that encoded the direction of movement in space but whose magnitude of response depended on the forearm posture (gain modulated extrinsic-like), and those whose activity covaried with muscle activity (muscle-like). Most directionally tuned PMv neurons were extrinsic-like (81%) or gain modulated extrinsic-like (12%), while M1 contained more equal numbers of each type. Based on these findings, they propose a model where projections from PMv to M1 perform the transformation from extrinsic to intrinsic coordinate frames.

Kurata and Hoshi (2002) recorded from PMv (F5 and F4)² and M1 during a visually guided reaching task after the animals had been trained with prisms that shifted the visual field left or right. They found that PMv mainly contained neurons whose activity depended on the target location in the visual field with or without the prisms, and M1 contained these neurons as well as those sensitive to target location in motor coordinates only and showed different activity for both visual and motor coordinates. This is consistent with the distribution of extrinsic-, gain modulated extrinsic-, and muscle-like neurons in PMv and M1 described by Kakei et al. (2001, 2003). In a similar experimental setup involving manipulation of visual information, Ochiai, Mushiake, and Tanji (2005) recorded from PMv neurons while monkeys captured a target presented on a video display using a video representation of their own hand movement. The video display either presented the hand normally or inverted horizontally in order to dissociate its movement from the monkey’s physical hand movements. PMv activity reflected the motion of the controlled image rather than the physical motion of the hand, supporting the proposed role of PMv as an early stage in extrinsic-to-intrinsic coordinate frame transformation. Furthermore, half of the direction-selective PMv neurons were also selective for which side of the video hand was used to contact the object regardless of starting position, suggesting that these visuomotor transformations are body part specific.

Grasping actions in both monkeys and humans are coded within a circuit involving the inferior parietal lobule (IPL) and ventral premotor cortex (PMv). In monkeys, this pathway includes the anterior intraparietal area (AIP) and interconnected premotor area F5 (PMv), and is implicated in transforming objects’ intrinsic spatial properties into motor programs for grasping

²Unfortunately, the data do not support ready discrimination between F4 and F5 cells.

(Fogassi et al., 2005; Jeannerod et al., 1995; Luppino et al., 1999). Functional neuroimaging studies in humans suggest that the anterior IPL, within and along the intraparietal sulcus (aIPS), and ventral precentral gyrus (putative PMv) may constitute a homolog of the monkey parieto-frontal grasp circuit (Binkofski et al., 1999; Ehrsson et al., 2000; Frey et al., 2005).

In this section, we have focused on the role of the visual system in the control of grasping. While the visual system is crucial to much of the planning of movements, the actual execution of the movement will make strong use of haptic feedback (both tactile input and proprioceptive sensing of the posture of arm and hand). Thus, the present review should be complemented by a similar analysis of the role of the somatosensory system—but this lies beyond the scope of the present paper. In any case, the data reviewed in this paper suggest that premotor cortex codes reaching and grasping movements in extrinsic coordinates rather than as detailed motor commands to the muscles, but that in some cases (and not in others) neurons may be effector-independent. This suggests that the pathway from premotor cortex to primary motor cortex is gated by signals encoding task-or effector-specific data.

Two case studies on experience-dependent changes in grasp representations with the use of complex tools

Having provided a conceptual framework for “Tool Use and the Body Schema” and surveyed a range of data on “Basic Parieto-Frontal Interactions for Visually Directed Hand Movements” we now provide a relatively interpretation-free summary of the findings in two papers which offer some insights into the neural mechanisms underlying use of a tool for grasping. The next section will present a brief view of models we have developed which make contact with the ideas and data of the first two sections. The final section will then integrate results from the case studies into our modeling framework, providing in the process a critique and ideas for further experiments.

Neurons coding for pliers as for fingers in the monkey motor system

Umiltà et al. (2008) title their paper “When pliers become fingers in the monkey motor system”, but we suggest that a better title would be “Tool use and the goal-centered representation of the end-effector in ventral premotor and primary motor cortex of macaque”, while noting that primary motor cortex also includes hand-related cells. This relates their findings to our general view that *different tools extend the body schema in different ways, and that a tool in general extends a hand’s functionality rather than becoming the hand.*

Umiltà et al. (2008) offer data in support of the claim that neurons in F5 code the end-effector’s motion, rather than that of the hand.³ To support this view, they recorded F5 and F1 neuron activity in monkeys trained to grasp objects using two tools: “normal pliers” and “reverse pliers” (Fig. 3). With normal pliers the hand has to be first opened and then closed to grasp an object, while with reverse pliers the hand has to be first closed and then opened. In either case, the pliers execute a precision pinch while the hand executes or releases a power grasp. Grasping an object with the two tools required an opposite sequence of *hand* movements to achieve the *same* movements of the end-effector relative to the object to be grasped. They could thus check whether the firing of a neuron correlated better with hand motion or end-effector motion.

Umiltà et al. (2008) selected neurons that discharged in association with hand grasping as well as grasping with the two tools for quantitative study, and assert that F5 neurons with visual

³Note that even in grasping, the hand can be reconfigured to provide different end-effectors, e.g., the pads of thumb and index finger in a precision pinch, the palm and inner finger surfaces in a power grasp. So the hypothesis that F5 neurons encode the motion of the end-effector in extrinsic coordinates is still consistent with these new data and previous research in F5 (Rizzolatti et al., 1988; Raos et al., 2006; Umiltà et al., 2007) that relate neuron activity to grasp type and phase and wrist orientation.

properties were not included in the present study. This means that they excluded at least half of F5 canonical neurons (Rizzolatti et al., 1988) and all mirror neurons. This seems unfortunate since mirror neurons eventually respond to the sight of pliers use even when monkeys have no motor experience using the tools (Ferrari et al., 2005). They state that they tested only those F5 neurons whose discharge was not influenced either by object presentation (canonical neurons) or action observation (mirror neurons) so that, in the neurons of their sample, “the coded relation between end-effector and goal concerned only the activity related to actual grasping movements”. This raises the following caveat: the neurons selected were only a subset of those whose firing was related to grasping with at least one of the hand, the normal pliers and the inverse pliers. We must thus consider (Sect. “The Case Studies in Perspective”) the possible role of other F5 and F1 cells in the observed behaviors. For the neurons studied, they concluded that F5 activity correlated with the movements of the end-effector. However, in M1/F1 some neurons discharged in relation to hand movements, while the firing of others was related to the motion of the end-effector.

When tested with normal pliers (for which pliers tips open when the hand opens and close when the hand closes) 18 (32.7%) neurons started to discharge during hand opening, reaching their maximum firing rate just before or during hand closure; 28 (50.9%) neurons discharged almost exclusively during hand closure; 7 (12.7%) started to discharge with hand closure and kept firing during the subsequent holding phase. Finally, 2 (3.6%) neurons fired only during the hand-opening phase. Crucially, neurons that discharged when the hands was opening with normal pliers discharged when the hand was closing when tested with the reverse pliers. Similarly, neurons that discharged when the hand was closing with normal pliers, discharged when the hand was opening with the reverse pliers. Thus, it is the temporal organization of end-effector movement necessary to reach the goal and not hand movement per se that was coded in the observed F5 neurons.

However, two distinct functional categories of neurons were found in M1/F1. Neurons of the first category (F1em) fired in relation to end-effector movements ($n = 26$), while neurons of the other category (F1hm) discharged in relation to hand movements ($n = 32$). Like the F5 neurons in this study, most of the end-effector related neurons showed their highest firing rate during closing of the end-effector (see Figure 5 of Umiltà et al., 2008). Perhaps closing is the most-represented phase because it must be carried out with the most precision.

Among the F5 neurons studied by Umiltà et al. (2008), most fired mainly or exclusively when the end-effector was closing on the object regardless of whether the hand was opening or closing. Other neurons discharged in relation to the grasping initial phase. Thus, grasping was achieved by (or “in correlation with”) an activation of populations of neurons each coding a specific temporal phase of the motor act which correlated with the end-effector movement rather than the hand movements necessary to achieve it. While some F1 neurons behaved in this fashion, others maintained their relation with a given hand movement regardless of whether this movement was performed during the initial or final phase of the motor act.

The F5 and F1em neurons that discharged when the monkeys used the tool also discharged when the monkey grasped objects with its hand, in spite of the fact that in this case grasping was achieved positioning the fingers in variable positions, not necessarily congruent with the axis alignment used with tools. It is plausible, therefore, that the effector alignment, a fundamental computational operation for grasping achievement, is not carried out in F5. It is plausible that this operation is performed in area F2, where neurons have been described that become active in relation to specific orientations of visual stimuli and to corresponding hand-wrist movements (Raos et al., 2004). That same paper showed that 66% of grasp neurons in F2 were highly selective for grasp type and that 72% were highly selective for wrist orientation. Raos et al. (2006) show that F5 neurons combine selectivity for grip type and wrist orientation,

and that 21 out of the 38 they tested for wrist orientation selectivity showed high selectivity for a particular orientation. The dorsal premotor cortex (including F2 and F7) is typically implicated in reach target selection (Caminiti et al., 1991; Cisek & Kalaska, 2005; Crammond & Kalaska, 1994, 2000; Weinrich & Wise, 1982) as well as wrist movements (Kurata, 1993; Riehle & Requin, 1989). The most plausible hypothesis that reconciles these findings is that the dorsal premotor cortex is involved in coding reach direction and the ventral premotor cortex is involved in coding grasps, and that interconnections between F2 and F5 (Marconi et al., 2001) allow the two regions to converge on a wrist orientation appropriate for the selected reach direction and grasp type.

Umiltà et al. (2008) reject the suggestion that F5 neuron discharge codes the prediction of the incoming sensory event. In the case of hand and normal pliers grasping, the expected sensory event would be an increase of tactile and proprioceptive input from the fingers and hand palm, while in the case of reverse pliers, a decrease in tactile hand input from the palm and an increase of proprioceptive input from the dorsal hand muscle. However, this argument only applies to hand proprioception, not to visual feedback on the relation between the end-effector and the object to be grasped. Moreover, Ferrari et al.'s (2005) finding that mirror neurons can learn to respond to observation of tool use suggests that they may be capable of providing visual feedback. This argument also does not rule out the more plausible hypothesis that F5 activity does not represent sensory prediction itself, but activates a *sensory expectation*, namely prediction of incoming sensory events in other areas, such as SII. This is actually what our FARS model predicts (Fagg & Arbib, 1998; see the complete FARS diagram below). The projections between F5 and SII would need to be modified in tool use. This seems to require F5 neurons selective for each tool type and others selective for grasping in order to activate distinct sensory predictions in SII. However, since Umiltà et al. (2008) only selected neurons active for both tool types and grasping for analysis, they provide no data about F5 neurons that may have these properties.

A fronto-parietal circuit representing grasping actions involving the hand or a novel tool

As detailed above, numerous studies have shown that tool use is accompanied by its incorporation into the neural representation of the body, and we have just seen (Umiltà et al., 2008) that, after extensive training, neurons in monkey area F5 that used to code manual grasping, may also come to code grasps performed with a tool—independently of the hand movements involved. Together, these findings suggest that the capacity to use such simple tools might emerge from experience-dependent changes within cerebral areas involved in the control of manual actions. In line with this, Jacobs et al. (2009) sought to test the hypothesis that in humans, following training, grasping with a formerly novel tool (Fig. 4, top left) would be coded in areas previously shown to be involved in the control of manual prehension. These areas constitute a putative functional homolog of the macaque AIP-F5 grasp circuit, namely the anterior intraparietal sulcus (aIPS) (Frey et al., 2005) and inferior portion of ventral premotor cortex (Binkofski et al., 1999).

They used fMRI to compare planning the grasping of an object directly with the hand and with use of the tool as shown in Fig. 4. The opposition axis between the jaws of the tool was perpendicular to that of the hand. Stimuli consisted of a half pink and half tan sphere with indentations on each side for “finger placement” (Fig. 4, top right). Subjects learned to use the tool outside the scanner, but while in the scanner-made judgments about how they would use the tool. Note the double change from hand to tool here: not only in orientation, but in the use of a power grasp of the hand to achieve a precision pinch with the tool. Thus (as with Umiltà et al., 2008), recognizing the same affordance initiates two very different paths to motor control.

Learning to use the tool—Participants first practiced grasping the stimulus object with either hand, or the tool held in either hand. They performed 96 trials in each of those four conditions, across which the orientation of the object varied randomly (between 0 and 345° in 15° increments). On each trial, participants were instructed to reach for and grasp the object using the most comfortable posture (i.e., over- or under-hand), using a precision grip *of the end-effector* (whether tool or hand) by placing their thumb and forefinger or the jaws of the tool on the object's indentations. Note that the alignment between the handle and the jaws of the tool is fixed. Thus, for both hand alone and hand with tool, there is a choice of comfort angle in terms of which groove will be the target for the thumb or upper pincer, respectively. The initial practice session was sufficient for subjects to learn how to accurately plan the angle in the prospective grasp selection (PGS) task in the scanner (described below). Participants' grip preferences reflected the 90° offset between the opposition axes of the hand and tool, but also reflected a further change when using the tool which Jacobs et al. attribute to the dynamical constraints imposed by the weight of the tool (Fig. 5).

As distinct from overt grasp selection (OGS)—actually executing the grasp outside the scanner—subjects in the fMRI scanner performed PGS, making a decision about grasping but not executing the action. Subjects had to decide where (i.e., on which colored indentation) their thumb or the target jaw (i.e., upper or lower) of the tool would be on the object if they were to grasp it, and then press the appropriate foot pedal (left or right) to respond. Importantly, the design of the experiment was such that participants had to wait for a response cue indicating which pedal corresponded to which color (pink and tan squares, position varied randomly) to appear after the stimulus object to give the desired answer. This way, the authors were able to separate neural activations accompanying grasp selection from those associated with foot press selection and execution.

The behavioral data show that participants in the scanner expressed grip preferences based on the use of either the hands or the tool that accurately corresponded to those learned earlier during actual prehension (OGS). This indicates that subjects based their planning of grasping actions with the tool on an internal representation that accurately captured properties of the novel implement and the constraints its use imposed. While the authors state that the PGS task was identical to the OGS task “except that participants remained still while planning grasping actions to decide whether an over- or under-hand grip would be preferred when grasping stimulus objects using the hands or tool”, this overlooks an important difference. While action selection is common to PGS and OGS, different motor controllers are involved: OGS directly activates the hand control (whether for a direct or tool-mediated grasp, which can be very different in detailed muscle control) whereas PGS requires (perhaps) a side path which symbolically recognizes the subgoal choice then executes a learned symbol–motor association with the foot, rather than using verbal expression. However, the published results compensate for this by only presenting imaging results for PGS in the period of time corresponding to the planning of grasps, before foot response selection, and execution.

fMRI activity for planning grasping actions with the hands or tool—Performance of the PGS task was associated with a remarkably similar pattern of increased fMRI activity as compared to resting baseline across all four conditions: planning grasping the target object with the hand versus tool, using left or right hand (). aIPS and putative PMv showed consistently increased activity regardless of whether grip preferences were based on the hand or the tool. This suggests that after training, planning grasping actions with either the hands or the tool engages the same parieto-frontal grasp circuit (Johnson, 2000). In all conditions, the effects in aIPS and PMv were stronger in the left hemisphere irrespective of the side involved (left, right), but we will discuss some differences below.

In addition to the grasp circuit, increased activity was also detected in more dorsal regions, including areas—namely, caudal superior parietal lobule (SPL) and dorsal premotor cortex (PMd)—that may constitute the human homolog of the *reach* circuit described in monkeys. In the present case, these activations might thus reflect the internal simulation of reaching actions (Johnson et al., 2002), or may be associated with the selection of grasping action (i.e., over- or under-hand) requested by the PGS task (Schluter et al., 1998, 2001). Given the earlier discussion on F2 and F5 selecting for reach direction and grasp type, respectively, and coordinating the wrist angle between them, it may be that the reach and grasp must both be simulated in order to select an appropriate wrist orientation. However, just what has to be simulated is a matter of debate, and we shall offer a new hypothesis in the Sect. “Executability—a new type of internal model” below. Increases in activity were also found in the pre-supplementary motor area (pre-SMA), which is involved in representing conditional visuo-motor associations (Picard & Strick, 2001), and bilaterally in the lateral cerebellum. Other subcortical structures related to motor control (basal ganglia, pulvinar nucleus of the thalamus) also showed increased activity in some of the conditions.

Most crucially for the interpretation of their results, Jacobs et al. (2009) found no significant difference related to the use of the tool as opposed to the hand, or to the side (left, right) involved in the task, even when using more lenient statistical thresholds. Further, a more sensitive analysis focused on regions of interest functionally defined in bilateral aIPS and in left PMv (right PMv was never significantly activated in the whole brain analysis) showed that while all areas were more active when the contralateral upper limb was involved in the task, none of them showed differential activity depending on whether the hand or the tool was used. This analysis further confirmed the lateralization of the network to the left hemisphere: left but not right PMv showed significant activation relative to baseline while aIPS was consistently activated in the left hemisphere, while increase in activity reached significance in right aIPS only when the contralateral left limb was used. However, we should note that this involvement of the same regions in each case does not preclude subtle differences in the activity of individual neurons, which can be detected in the macaque with single-cell recording but not in the human with fMRI.

A theoretical framework

In this section, we review the FARS (Fagg-Arbib-Rizzolatti-Sakata) model of the role of canonical neurons in grasping (Fagg & Arbib, 1998) as well as the critical notion of internal models. We will then show how the notion of executability from our augmented competitive queuing model provides a key element for one of the extensions of the FARS model needed to address the data reviewed above.

The FARS model of the role of canonical neurons in grasping

The FARS model, as diagramed in Fig. 6, is organized around the dorsal and ventral streams as described in the monkey, but is considered to match essential processes in the human brain as well. The areas shown include AIP (anterior intraparietal cortex), area F5 (of the ventral pre-motor cortex), inferotemporal cortex (IT), and regions providing supporting input to F5, namely F6 (pre-SMA), area 46 (dorsolateral prefrontal cortex), and F2 (dorsal premotor cortex). The dorsal stream (from visual cortex to frontal cortex via parietal cortex) is said to extract affordances for objects and use these to set parameters for possible interactions with the object; the ventral stream (via IT) can categorize objects and pass these to prefrontal cortex (PFC) to analyze which actions are appropriate to meet current goals. The overall diagram represents FARS as a conceptual model, but a crucial subset of the regions and connections have been subjected to rigorous computer implementation and simulation (Fagg & Arbib, 1998). In Fig. 6, cIPs extracts information about oriented surface patches of objects. According to the FARS model, AIP uses visual input to extract affordances from these data which highlight

various surface assemblages of the object that are relevant to grasping it in various ways. The parallel pathway via VIP and F4 for controlling reaching is part of the conceptual model but not part of the detailed implementation. Various constraints are applied on selection of a grasp for execution. In the original FARS model (as in Fig. 6) the constraints act on F5, while in the modified version they have been posited to act on AIP. F6 is posited to provide task constraints; F2 instructional stimuli for conditional (e.g., visuomotor) association, and 46 provides working memory of parameters for recent grasps. Reciprocal connections between AIP and F5 inform AIP of the status of execution via F5, thus updating AIP's active memory.

The infant learning to grasp model (ILGM)

The infant learning to grasp model (ILGM) (Oztop et al., 2004) provides a learning mechanism to complement the FARS model, though it only addresses the earliest stages of such learning—the initial open-loop reaching strategies, subsequent feedback adjustments, and emergent feedfor-ward strategies employed by infants in the first postnatal year (Figure 2; Oztop et al., 2004). Infants quickly progress from a crude reaching ability at birth to finer reaching and grasping abilities around 4 months of age, learning to overcome problems associated with reaching and grasping by interactive searching (Berthier et al., 1996; von Hofsten, 2004; von Hofsten & Ronnqvist, 1993). At first, the poorly controlled arm, trunk, and postural movements make it very difficult for the young infant; nonetheless, experience eventually yields a well-established set of grasps, including the precision grip, with preshaping to visual affordances by 12–18 months of age (Berthier et al., 1999; Berthier & Keen, 2006).

ILGM is a systems level model based on the broad organization of primate visuomotor control with visual features extracted by the parietal cortex and used by the premotor cortex to generate high-level motor signals that drive lower motor centers thereby generating movement. It demonstrates how infants may learn to generate grasps to match an object's affordances. It assumes initial reaches are inherently variable, providing a basis for reinforcement learning, where the reinforcement signal, related to the extent to which the current action yields a stable grasp of an object, is called *Joy of Grasping*. The requirements of the model are that infants will (1) execute a reflex grasp when triggered by contact of the inside of the hand with an object encountered while reaching, (2) can sense the effects of their actions, and (3) will use reinforcement learning to adjust their movement-planning parameters.

The reinforcement from object contact is used to modify grasp generation mechanisms within the premotor cortex. The model consists of four modules: the input module (Parietal Cortex), the grasp learning module (Premotor Cortex), the movement generation module (Motor Cortex and Spinal Cord), and the grasp evaluation module (Somatosensory Cortex). The computational layers of the grasp learning module are based on the *Preshape*, *Approach Vector* and *Orient* grasping schemas proposed by Iberall and Arbib (1990). These layers encode a minimal set of kinematic parameters specifying basic grasp actions. Thus, the grasp-learning module formulates a grasp plan and instructs the movement generation module, located in both the spinal cord and motor cortex. The movement generation module completes task execution (Jeannerod et al., 1995). The sensory stimuli generated by the execution of the plan are then integrated by the movement evaluation module, located in the primary somatosensory cortex. Output of the somatosensory cortex, the reinforcement signal, is used to adapt the connection strength between parietal-premotor connections (Fig. 7). Implementation of the reinforcement signal is defined simply in terms of grasp stability: a grasp attempt that misses the target or yields an inappropriate object contact produces a negative reinforcement signal. For the sake of simplicity, the brain areas that relay the output of the somatosensory cortex to the parietal and premotor areas have not been modeled. However, it is known that primate orbitofrontal cortex receives multimodal sensory signals (conveying both pleasant and non-pleasant sensations), including touch from the somatosensory cortex; moreover, orbitofrontal cortex

may be involved in stimulus-reinforcement learning, affecting behavior through the basal ganglia (Rolls, 2004).

During early visually-elicited reaching while infants are reaching toward visual or auditory targets, they explore the space around the object and occasionally touch the object (Clifton et al., 1993). ILGM models the process of grasp learning starting from this stage. We represent infants' early reaches using an object-centered reference frame and include random disturbances during initial reach trials by ILGM so that the space around the object is explored. We posit that the *Hand Position* layer specifies the hand approach direction relative to the object, restricting the angles from which the object can be grasped and touched on the basis, e.g., of the location of the object relative to the infant's shoulder; the *Wrist Rotation* layer learns the possible wrist orientations given the approach direction specified by the Hand Position layer; while the *Virtual Finger* layer indicates which fingers will move together as a unit given an input (Arbib et al., 1985).

Simulations with the ILGM model explained the development of units with properties similar to F5 canonical neurons. The important point is that ILGM discovered the appropriate way to orient the hand toward the object as well as the disposition of the fingers to grasp it. Such results provide computational support for the following three hypotheses:

- (H1) In the early postnatal period, infants acquire the skill to orient their hand toward a target, rather than innately possessing it.
- (H2) Infants are able to acquire grasping skills before they develop an elaborate adult-like object visual analysis capability.
- (H3) Task constraints due to environmental context (or the action opportunities afforded by the environment) are factors shaping infant grasp development. To this we add a fourth hypothesis:
- (H4) Grasping performance in the absence of object affordance knowledge, 'affordance-off stage', mediates the development of a visual affordance extraction circuit.

which we now address. Oztop et al. (2007) developed a grasp affordance learning model (GAEM) which uses the grasp plans discovered by ILGM to adapt neural circuits that compute affordances of objects relevant for grasping. The main idea is that, by experiencing more grasps, infants move from an initially unspecific vision of objects to an "informed" visual system which can extract affordances, the features of objects that will guide their reach and grasp. The model emphasizes the role of the anterior intra-parietal sulcus (AIP) in providing information on the affordances of an object for use by the canonical neurons of the grasping system of premotor area F5 in the macaque. Thus, the model does not address object recognition but instead models the emergence of an affordance extraction circuit for manipulable objects.

During infant motor development, it is likely that ILGM and GAEM learning coexists—AIP units that are being shaped by GAEM learning provide 'better' affordance input for ILGM. In turn ILGM expands the class of grasp actions (learns new ways to grasp) providing more data points for GAEM learning. When this dual learning system stabilizes, GAEM and ILGM will be endowed with a set of affordance extraction (Visual Cortex→AIP weights) and robust grasp planning (ILGM weights) mechanisms. However, Oztop et al. (2007) adopt a *staged learning* approach solely for ease in implementing GAEM learning. In other words, in the simulations presented GAEM adaptation is based on the successful set of grasps generated by ILGM—once ILGM learning is complete and weights are fixed).

To model the unprocessed/unspecific raw visual input, the model represents the visual sensation of object presentation as a matrix encoding visual depth which can be considered a

simple pattern of retinal stimulation. The ILGM model showed how grasp plans could be represented by F5 canonical neurons without competition among the learned grasps. There is indirect competition, however. ILGM uses a rank-order coding scheme where the neuron that fires first inhibits all the others and forms a population code around itself. Neurons receiving large inputs are more likely to fire first, so their competitive strength is governed by learned connection weights. This is how the model can learn two different grasps for the same affordance. Which grasp is selected at any point in time depends on which neurons fire first. The GAEM model proposes that the successful grasp plans (and associated grasp configurations) represented by F5 canonical neurons could form a self-organizing map (SOM) where the grasp plans are clustered into set of grasps with similar kinematics (e.g., common grip types and fixed aperture range.) Note that the clusters are formed via self-organization. Parallel to the self-organization, a vertical supervised learning takes place that aims at learning the mapping from the visual input to the cluster (see Fig. 8), modeling learning along the dorsal pathway, namely Visual Cortex→AIP→F5. Note that a single object may afford more than one grasp plan, and thus each cluster forms an independent subnetwork. In this sense, the self-organization could be seen as the competition of sub-networks to learn the data point available to GAEM at a given instant.

After GAEM learning, objects with varying dimensions were presented and resulting AIP unit responses were recorded. For cylinders with various diameters and heights, learning yielded a varied set of units—some coded the height of the cylinder, some encoded diameter; others encode a range of width or height values (such units are rarer than the width and height encoding units), and finally, there are a few units that prefer certain width and height combinations.

The ACQ model of opportunistic scheduling of the next action

We have developed a model called *augmented competitive queuing* (ACQ) to analyze some intriguing results about sequential behavior and its rapid reorganization under certain conditions.⁴ However, in the present paper our interest in ACQ focuses not on sequencing as such, but rather how the choice of the next action may depend on current circumstances. This model is implemented as a set of interacting functional units called *schemas* (Arbib, 1981). This allows some components to be represented as (possibly state-dependent) mappings from input to output variables, while those that are the focus of this study are implemented as neural networks. A stripped-down version of the ACQ model is shown in Fig. 9.

The heart of the model is given by the three boxes at lower left. The external world is modeled as a set of environmental variables. The *executability signal* provides a graded signal to the cells in the Actor, each of which represents an action. This signal is highest when the action is easy to execute, decreasing as the effort required increases; and reaching zero when the action cannot be executed in the current circumstances. In the study of Jacobs et al., the key “high-level variable” is the orientation of the end-effector. The choice of action (in the hand case, “thumb up” or “thumb down”) rests on the relative executability of the task for the current orientation.

The *desirability signal* specifies for each motor schema the reinforcement that is expected to follow its execution, based on the current internal state (subgoals and motivational state) of the organism. We then calculate *priority* as a combination that increases with both *executability* and *desirability*. The Actor box then simply uses a noisy winner-take-all (WTA) mechanism to select for execution a motor schema with the highest (or near highest) priority.

⁴An informal description of ACQ was published in Arbib and Bonaiuto (2008), an extended description of the model with further simulation results will appear in Bonaiuto and Arbib (2009).

Using the general approach of temporal difference (TD) learning (Sutton & Barto, 1998), the Adaptive Critic learns to transform primary reinforcement (which may vary from task to task but will in general be intermittent, e.g., the payoff for lengthy foraging may come only when the food is consumed) into *expected reinforcement* (so that actions are desirable not in terms of immediate payoff, but in terms of the statistics on eventual reinforcement when this action is executed in similar circumstances, with later payoffs discounted more than imminent ones).

What we have discussed so far defines a system for *planning* behavior—choosing an action and setting key high-level parameters thereof. The output of the Actor projects to a lower-level motor control structure which controls the execution of the selected, parameterized motor schema.

An important aspect of the model is that it offers a new interpretation of the Mirror System as fulfilling two roles not normally associated with it: recognizing when an intended action fails and recognizing when an executed action may appear similar to an unintended action. Where previous discussions of the mirror system have it coding a single action—either that intended by the agent, or the action of another as recognized by the agent—our new model (Bonaiuto & Arbib, 2009) suggests that during self-action the mirror system is activated by analysis of visual input and can thus recognition may encompass apparent as well as intended actions. Moreover, we add that comparison of the mirror system signal with an efference copy of the Actor output can indicate if the intended action was performed successfully. As a result, the Adaptive Critic, receiving input from the Mirror System, can apply TD learning (Sutton & Barto, 1998) to update estimates of expected reinforcement (desirability) for (1) the intended action (i.e., that specified by the output from the Actor) unless it was unsuccessful and not recognized by the Mirror System, as well as (2) any actions that appear to have been performed even if they differ from the intended action. Similarly, such estimates of whether or not an action was executed successfully can be used to update estimates of executability (action success) for the intended action when performed for the currently perceived environmental state.

The case studies in perspective

Our aim in this final section is twofold: to offer a novel perspective on the case studies by analyzing them in the light of the models reviewed in the previous section, and to use this analysis to set challenges for further modeling.

The FARS and MNS models of the previous section focus on the execution of actions in which the hand, appropriately configured, serves as the end-effector. We must now address the challenge of suggesting how these models might be expanded to address the data on neuron firing in tool use provided by Umiltà et al. (2008). In doing this, we will postulate that homologous mechanisms operate in the experiments of Jacobs et al. (2009)—but with one highly non-trivial difference, that in the human case verbal instruction and symbolic representations can greatly speed the learning process. However, the analysis of these more or less specifically human abilities lies outside the scope of this article. But before proceeding to our analysis of brain mechanisms underlying tool use, we first offer two general observations:

1. Learning to use a tool must involve both the dorsal and ventral stream of FARS: the dorsal path must learn to generalize affordances or learn new ones; the ventral stream must learn to recognize new objects and relate them to goals, invoking the appropriate affordances for use of the tool rather than the hand alone. This requires recognizing the tool and using the hand to grasp it appropriately and then using the hand “implicitly” while controlling the end-effector “explicitly”.
2. We need much more data on what happens during training. For example, Umiltà et al. (2008) offer no detailed data from the training phase, but such data would be

extremely helpful in testing models that extend ILGM/GAEM to tool use. Their supplementary material says training took 6–8 months but tells us nothing about the training method and the changes in both behavior and cell response that occurred during this protracted period. For example, are the macaques trained with each type of pliers in block or interleaved trials? Did improvement with one type interfere with learning the other type until they learn to switch contexts? In any case, we may contrast the 6–8 months of training with the 192 trials (96 trials with the tool in one hand, and 96 in the other) that sufficed for the human subjects of Jacobs et al. (2009) to master a similar task.

Linking codes for end-effector to codes for the hand

Umiltà et al. (2008) found that activity in their selected F5 neurons correlated with the movements of the end-effector, the jaws of the pliers. Similarly, some neurons recorded from M1/F1 discharged in relation to the motion of the end-effector (F1em), but for others the firing was related to hand movements (F1hm). From a motor control point of view, then, F5 plus F1em neurons provide end-effector specifications that must be routed in a tool-dependent way to the F1hm movements that help control the actual musculature. Depending on the context (hand, normal pliers, reverse pliers) the activity encoding the desired motion of the end-effector would be differentially transformed into appropriate motion of the hand for achieving the goal in that context. As a consequence, the neurons' properties change and their discharge may *in certain contexts* become related to a movement opposite to that normally effective. This hypothesis could also explain the stronger discharge displayed by F1em neurons during grasping with the normal tool than with the reverse one. This stronger activation could be due to the maintenance when using the reverse pliers of a feedback input from motor neurons related to normal grasping, a feedback that would have to be inhibited in this case. This maintenance in F1, but not in F5, could depend on the greater functional proximity of F1em neurons to the primary motor cortex neurons that control hand movements.

In this regard, recall from Asanuma and Arissian (1984) that the combination of dorsal column section and sensory cortex removal produced severe motor deficits. These consisted of loss of orientation within extrapersonal space and loss of dexterity of individual fingers (deficits which did not recover within the 4–5 weeks of observation). Although direct sensory input to motor cortex from the thalamus plays an important role in the control of voluntary movements, its loss can be compensated for by input from somatosensory cortex. Sakamoto, Arissian, and Asanuma (1989) studied the role of corticocortical input projecting to the motor cortex in cats learning motor skills. After unilateral removal of somatosensory areas and a part of area 5, the cat was placed in a box and trained to pick up a small piece of food from a beaker in front of the box. Since the beaker and the edge of the box had a space in between, the cat had to develop a new motor skill to bring the food back to the box across the space. This skill consisted of combined supination and flexion of the paw to hold the food over the gap. In all three cats, the training period necessary for acquisition of the motor skill for the forelimb contralateral to the lesioned brain was significantly longer than the period necessary for the forelimb ipsilateral to the lesioned cortex. Ablation of the remaining projection area after completion of the training did not impair the learned motor skill. The results suggest that, prior to the lesion, the input from the lesioned area to the motor cortex participated in learning motor skills. Perhaps learning the skill requires a *task-dependent coupling* of cells responsive to visual feedback of the end-effector (“goal-related” cells) and those controlling the appropriate hand movements with somatosensory cortex providing the training input.

Executability—a new type of internal model

A major principle of nervous system design is that of the “internal model” (a tradition stretching from at least Craik, 1943, in a way that enriches the notion of body schema which we have

traced back to Head & Holmes, 1911). The notion is that we consider part of the body (configured in a certain way, and possibly extended by a tool or other accoutrements, in interaction with part of the external world) as a control system: control signals issued by the nervous system will yield different patterns of behavior depending on its current state. To develop this point, consider two types of model which have figured prominently in the recent literature on sensorimotor learning. A *forward model* in the CNS is a neural representation of the relation between the control signals and resulting behavior, and thus serves to create neural code for expectations as to the outcome of an action. Conversely, an *inverse model* takes as input neural code for a desired behavior and returns a possible set of control signals for achieving that goal (Jordan & Rumelhart, 1992; Wolpert & Kawato, 1998). However, the brain can also make use of models which represent far less of the causal structure of part of the body schema. In particular, we see *executability*—a key element of the ACQ model—in this light: through long practice, we have in neural terms (and thus not necessarily accessible to consciousness) an estimate (a model, if you will) of how much effort, and with how much chance of success, we can execute an action in given circumstances. In ACQ, this combines with desirability to affect the decision as to whether or not to execute the action. Once the decision to execute is made, the brain must then invoke the appropriate inverse model to determine what commands will lead to successful execution of the action.

After practice, participants in the imaging study of Jacobs et al. (2009) expressed grip preferences based on the use of a hand or the tool that accurately corresponded to those demonstrated earlier during actual prehension. Jacobs et al. thus suggest that grip selection decisions relied on internal models that accurately captured the specific properties of the newly mastered device and the physical constraints experienced during its use. But what sort of internal model? Note that their fMRI study requires a judgment to be made that is relevant to *planning* a grasp, but does not exercise any internal model acquired with increasing skill in using the tool for *execution* of the grasp. We now suggest that it is neither a forward nor an inverse model of the actual control system for grasping, but is rather the executability for the four grasps, each refined to take the orientation parameter into account. The fact that subjects could successfully switch between the hand and tool during training suggests that the appropriate controllers for using the tool were learned at that time. However, the invocation of these latter controllers was not necessary for successful completion of the task in the scanner. An estimate of executability of “thumb up” and “thumb down” will suffice.

How could such an executability model be acquired? Let us return to the example of “writing” with a paintbrush at the end of a long pole. We suggested that initially, execution of the task relies on an effector-*independent* representation based on the visual appearance of the strokes which comprise a letter, which can serve to provide visual feedback for execution of the novel task with different effectors. However, with practice, a feedforward model can be acquired for the performance, with an accompanying transition from a slow effortful process in which errors are frequent and feedback plays the crucial role, to an increasingly fast and accurate process in which feedforward dominates over feedback (see Hoff & Arbib, 1993 for an example of a control system for coordination of reach and grasp in which feedback and feedforward are fully integrated). The resultant controller is effector-*dependent*. Once we are committed to the end-effector, there may be several courses of action open for its use. Learning, we claim, not only yields a controller for each action, with its accompanying forward and inverse models, but also yields an executability model, i.e., the repeated execution of an action under varying circumstances yields a neural representation of its executability (whether or not it can be executed; the amount of effort or likelihood of success if it is executed) under a range of circumstances. Indeed, we saw that ACQ includes the provision that the mirror system can signal whether or not it observes that an intended action was performed successfully. These signals can then be used to update estimates of executability (action success) for the intended action when performed for the currently perceived environmental state. What we suggest, then,

is that two learning processes go hand in hand—the learning whereby the controller (inverse model) for an action improves its performance and the learning whereby the system tracks the success with which the controller executes the action in various circumstances. We then see the imaging results of Jacobs et al., as showing the role of the executability model and other parts of the planning process, dissociated from the activity of the controller were the action to be physically performed.

In fact, we have already published a model of stroke rehabilitation (Han et al., 2008) which employs an executability model in just such a way, though it was not described in these terms. The model contains a left and a right motor cortex, each controlling the opposite arm, and a single action choice module (the executability model). The action choice module learns, via reinforcement learning, the executability value for using each arm for reaching in specific directions. Each motor cortex uses a neural population code to specify the initial direction along which the contralateral hand moves toward a target. The motor cortex learns to maximize executability by minimizing directional errors and maximizing neuronal activity for each movement. After a simulated stroke, the action choice module rapidly learns that there are directions for which reaching with the affected arm is no longer as readily executable, and so the unaffected arm will be used instead (the use of either hand is equally desirable in terms of the expected reinforcement). However, when we simulated constraint-induced therapy (Taub et al. 2002), which forces the use of the affected limb by restraining the use of the less affected limb, learning within the motor cortex contralateral to the affected limb led to increasing executability of reaching to the various directions with that hand. As the action choice module (the executability model) came to estimate these improved values, a threshold of rehabilitation might be reached (depending on the severity of the lesion) at which the affected hand could now be chosen over the unaffected hand for certain directions after the constraint was removed.

Returning to the Jacobs et al. (2009) study, we note as a challenge for further experimentation that this runs counter to a more standard view (aspects of which are reviewed by Oztop et al., 2006) which would argue that prior to action the brain runs multiple simulations of the forward and inverse models for grasping with the different orientations of the hand or tool (as appropriate) in order to determine the most comfortable end position. We are arguing that the notion of executability provides a shortcut that is vastly more economical of neural processing than such extensive simulation, since it encapsulates the results of many prior simulations and actions rather than requiring extensive, multiple runs of possibly relevant forward and inverse models each time. It is worth noting, however, that executability—just like the forward and inverse models relating muscle (or higher-level) commands to desired motion of the end-effectors—is indeed effector-specific, since “comfort angles” for the hand translate into different angles of the end-effector relative to the target, depending on whether hand or tool is used to effect the grasp.

The neural estimate of an action’s executability in a situation increases with success in performing the action (an “immediate payoff”), rather than success in achieving an external reward. If the affordance for an action cannot be perceived in the current situation, that action will not be performed and its executability in that situation is zero. If an affordance for an action can be perceived, its executability depends on the organism’s effectivities and their efficiency in acting on this affordance (in terms of probability of success and effort expended). As mentioned above, the dorsal stream is typically implicated in extracting affordances for action from visual stimuli. Specifically, area AIP is involved in extracting affordances for grasping and V6A/VIP for reaching. Given the role of AIP→F5 projections in selecting (under PFC modulation) which affordance to act upon and which action to select to act upon it, we speculate that learning of executability may be subserved by dopamine-dependent plasticity in parietal-premotor projections based on dopaminergic responses to an intrinsic reward (like successfully completing an action; cf. “joy of grasping” in the ILGM model).

Lateralization

The lateralization of planning (as distinct from motor control, which is not assessed here) in the data of Jacobs et al. (2009) is an interesting replication of prior wisdom, but the partial lateralization demands further analysis from a modeling perspective. The increases found in the pre-supplementary motor area (pre-SMA), which is involved in representing conditional visuo-motor associations, e.g., color and type of movement (Picard & Strick, 2001) may reflect the association between the stimulus object and a particular grasping action.

Jacobs et al. (2009) find that while the regional activation (as distinct from the neuron-by-neuron processes observable in macaque neurophysiology) underlying grip planning within bilateral aIPS and putative left PMv did not differentiate between the use of the hand or tool, they showed some degree of specificity for the contralateral limb. This may indicate that these areas are involved in coding grasps at two different levels, effector-specific and effector-independent. This is an important point that needs development. When planning bimanual actions, how much is “overall” and how much is contra-lateral for the individual hands? The interaction between the side and the use of the tool or hand was non-significant in aIPS on both sides but was significant in putative left PMv, with responses in this region greater when the contralateral hand was used, but not differing between sides when the tool was involved. The reasons for this last point are unclear at present. Since the subjects are much more trained in use of their hand for grasping than the tool, one speculation would be that perhaps the left PMv has become specialized for grasping with the hand, but since using the tool is a relatively new skill, its representation is still distributed bilaterally in the PMv. However, Jacobs et al. have no evidence of the involvement of the right PMv in the PGS task, regardless of the effector used. Speculatively, it could be just under threshold in the whole brain analysis but still show increased activity with respect to resting baseline in more sensitive ROI analysis (as for right aIPS in the Tool Left condition), but there is still probably a strong asymmetry in favor of the left hemisphere. Much research has consistently shown left hemisphere specialization for planning familiar praxic behaviors, whether for tool use pantomime or intransitive gesture (Króliczak & Frey, 2009). However, we know that a young child who receives a left hemispherectomy can nonetheless acquire language. This suggests that the two hemispheres initially have somewhat similar circuitry, and it is through learning that each hemisphere becomes more or less specialized for its own range of functions.

The importance of the cerebellum

The models reviewed earlier have placed the major emphasis on interactions between cortical regions (see Fig. 2), though various models (not spelled out here) have emphasized the role of the basal ganglia in the selection of the next action to perform. However, we must now stress that the cerebellum may play an important role in the studies of interest here. Jacobs et al. (2009) detected bilateral increases of activity in the lateral cerebellum and relate this to work implicating the cerebellar hemispheres in manual grasping (Grafton et al., 1996) and the acquisition of novel tool use skills (Imamizu et al., 2007).

To see what this suggests by way of augmenting the models of the previous section, we recall the experiments of Martin et al. (1996a, b) and our related modeling (Arbib et al., 1995; Spoelstra & Arbib, 2001). Martin et al. studied adaptation to prisms when throwing a dart at a target. The crucial observations were this:

- a. Adaptation to the prisms was followed by a process of counter-adaptation once the prisms were removed.
- b. The cerebellum is necessary for this learning.
- c. The learning does not involve global adaptation to the shift of the visual field induced by donning or doffing the prisms, but rather involves adaptation of a visuo-motor

transformation specific to a class of actions. For example, a subject may adapt to the prisms for throwing over arm yet show no adaptation for throwing underarm, whereas another subject may show some transfer of adaptation between the two tasks.

- d. Whereas each episode of donning or doffing the prisms initially requires a number of trials of adaptation, a subject who repeatedly experiences these two conditions will come to automatically change the adaptation state each time the prisms are taken on or off.

Our model explained all these effects by having specific microcomplexes in the cerebellum modulate activity in appropriate circuitry of the posterior parietal lobe. Afferents to the cerebellum encoding parameters directly related to throwing are richly represented in the mossy fibers afferent to the cerebellum whereas those representing prism on versus prism off are sparsely represented among a plethora of cortical state signals. Thus (to oversimplify), the adaptation each time the prisms are donned or doffed proceeds fairly rapidly, but adjusting the sparse synapses encoding whether the prisms are on or off requires protracted learning to extract the signal from the noise. Similarly, we associate underarm and over arm throws with different cerebellar microcomplexes, and thus the degree of overlap between these microcomplexes can explain the degree of transfer between prism adaptation for the different throws.

The “prism on” versus “prism off” switching the sensorimotor transformation based on this high-level cue is reminiscent of the “hand”/“normal pliers”/“reverse pliers” switch in Umiltà et al. (2008), and thus offers further motivation to consider the role of cerebellum. Note that we are looking at skills where cerebellum projects back to cerebral cortex to adjust its control, rather than projecting “downstream” to brainstem or spinal cord. Kurata and Hoshi (2002) focus on cellular activity in PMv and M1 during prism adaptation, posing the challenge of understanding what role cerebellum plays in affecting the changes in cell firing they observe.

The fMRI results of Jacobs et al. (2009) suggest that planning grasping actions based on the hands or tool involved the very same regions of parietal and premotor cortex, most notably putative homologs of the macaque grasp circuit (aIPS and PMv). However, humans have facility in tool use and monkeys do not, so the fact remains that the human homologs have capacities that the regions of monkey brain do not. So how can we devise experiments that chart the difference? One potential line of work regarding this point is the mechanical problem solving capacities of humans, which allows us to infer function from the structure of the tool and is probably determinant for our ability to easily learn the use of novel tools, and which is disrupted in apraxic patients with impairment of familiar tool use (Goldenberg & Hagmann, 1998).

From FARS and ILGM to effector-dependent and independent representations

Umiltà et al. (2008) demonstrated that single F5 neurons that represent manual prehension may come to also code grasping actions performed with tools after extensive training (They did not study cells that did not adapt to tool use.). Movements appropriate for the effector involved (i.e., hand, normal or reversed pliers) may emerge from modulation of the interactions between these effector-independent neurons and movement-specific cells found in the primary motor cortex (M1). For Jacobs et al. (2009), while behaviors specific to the hands or tool were observed and associated with activations within the grasp circuit, overt execution was not required and there was no significant change in neural activity within M1. This is consistent with our executability hypothesis, which suggests that effector-specificity arises earlier in the grasp circuit, in aIPS or PMv. This converges with our discussion (at the end of “Executability—a new type of internal model”) suggesting that effector-specific representations must be involved to explain the behavioral data on planning, maybe coexisting with effector-

independent representations within the same network of areas for visually monitoring the approach of the end-effector to the target.

One can hypothesize that PMv plays a similar role here for Jacobs et al. as in Umiltà's experiment: coding an endeffector specific representation of the pliers grasp, whatever pliers are used. Notice the way in which the neurons they recorded are "end-effector specific". If we accept that their F5 neurons encode the motion of the end-effector they may appear effector *independent* but in fact their activity must be gated by task knowledge (hand, normal pliers, reverse pliers) to respond to effector-specific input and be routed to the appropriate F1/Mi neurons, with the latter varying with the choice of end-effector. However, Jacobs et al.'s fMRI data do not allow us to disentangle whether these areas represent grasping actions in an exclusively effector-independent or effector-specific manner, or whether both types of representations might coexist there within distinct but overlapping neural networks. The fact that neurons controlling the movement of the end-effector can be found within monkey M1 together with neurons specific for hand movements (Umiltà et al., 2008) suggests that coexistence of both types of representation in the same brain area is possible. Furthermore, Jacobs et al. (2009) also underlined the fact that overall activation of aIPS and PMv was not sensitive to the use of the tool as opposed to the hand although specificity for the contralateral limb suggests that co-localization of both levels of representation is possible.

The finding that some tool use behaviors arise from changes within areas involved in manual actions is in line with the observation that tool use is accompanied by an extension of the body-schema, as shown by the expansion of the multi-sensory representations of the forelimb occurring after the use of a rake to encompass this novel tool (Iriki et al., 1996). While this perceptual integration might be fundamental for the development of certain tool use behaviors, it is certainly not sufficient, as we argued earlier, to allow the acquisition of tool use skills where the tool modifies the functionality of the hand, even in such "simple" cases as those illustrated by the studies of Umiltà et al. (2008) and Jacobs et al. (2009). Future work will be necessary to investigate where the transformations relating the movements of the hand to the movements of the end-effector (e.g., tool) occur, and also what differentiates tool use behaviors where the tool provides capacities more or less analog to hand function, like those exposed here, from those where the tool allows actions such as sawing, drilling, and cutting that are well beyond what the non-augmented human can do.

In addition to their role in sensory-motor transformations, premotor and parietal areas within the grasp network also appear to code actions at a level of abstraction that is independent of the specific movements involved (Fogassi et al., 2005; Tunik et al., 2005). This raises the issue of how many of the parameters are coded here, and how the remaining parameters are filled in "downstream" on the basis of the "abstract" parameters and the choice of endeffector. In terms of ACQ, we ask where is the coding of the subgoal/reinforcement criterion against which the desirabilities of actions are to be evaluated, to what extent do current sensory-motor data affect the estimate of executability, and how then are these data channeled to the control system that actually executes the movement.

We close by looking at the implications of such data for future modeling which addresses the issues raised in our opening discussion of distalization and the body schema. We have used the example of writing to suggest how an abstract representation can provide the mechanisms for feedback control prior to the "automatization" that yields internal models that support feedforward control. The essential idea was that, in this case, a visual model of the relative size and placement of the strokes in writing a letter could be used to provide feedback for any end-effector for which control mechanisms were already available for making a visible stroke which matched parametric descriptions. The resultant performance would be slow and would have poor co-articulation between the strokes, but would exploit the abstract representation to drive

not only the initial performance but also with repeated practice, the necessary learning signals to drive the formation of the model for a skill. Note the similarity with and differences from our earlier work on ILGM/GAEM. There we used reflex grasping to generate grasps that could be evaluated by “joy of grasping”, which evaluated each grasp by the stability of the resultant hold on the object. Here, we use explicit visual feedback not only for correcting the current performance but also as training signal for better attuning motor control parameters to match desired to performed stroke trajectory. A second-order issue is the co-articulation which speeds and smoothes the transition from one stroke to the next. The cerebellum is heavily involved in both cases.

The other key point is to look afresh at the FARS model and note the key role of attention that is implicit in the model. The visual input is provided by the affordance that is to be acted on. The motor output is provided by the motion of the hand—but since the hand can be configured for a precision grip, a side opposition, a power grasp and more, the actual control involves multiple models, one appropriate to each hand configuration. Moreover, the appropriate model involves not only the ability to make on-line feedback corrections based on visual assessment of the hand as a specifically configured end-effector toward the chosen affordances, but also the haptic expectation (linked to SII in FARS) whereby the contact of the end-effector with the object can be evaluated and adjusted based on haptic feedback. Now note that neural networks that correspond to the FARS model are in “general position” linking the visual attention to the relation of the end-effector provided by any tool to the appropriate affordances to the specific model of haptic feedback linked to the use of that particular tool.

References

- Alexander GE, Crutcher MD. Neural representations of the target (goal) of visually guided arm movements in three motor areas of the monkey. *Journal of Neurophysiology* 1990;64:164–178. [PubMed: 2388063]
- Arbib, MA. Perceptual structures and distributed motor control. In: Brooks, VB., editor. *Handbook of physiology—the nervous system II. motor control*. American Physiological Society; 1981. p. 1449-1480.
- Arbib MA, Bonaiuto JB. From grasping to complex imitation: modeling mirror systems on the evolutionary path to language. *Mind & Society* 2008;7:43–64.
- Arbib, MA.; Iberall, T.; Lyons, D. Coordinated control programs for control of the hands. In: Goodwin, AW.; Darian-Smith, I., editors. *Hand function and the neocortex*. Berlin: Springer; 1985. p. 111-129.
- Arbib, MA.; Schweighofer, N.; Thach, WT. Modeling the cerebellum: from adaptation to coordination. In: Glencross, DJ.; Piek, JP., editors. *Motor Control and Sensory-Motor Integration: Issues and Directions*. North Holland: Elsevier; 1995. p. 11-36.
- Asanuma H, Arissian K. Experiments on functional role of peripheral input to motor cortex during voluntary movements in the monkey. *Journal of Neurophysiology* 1984;52(2):212–227. [PubMed: 6481430]
- Berthier NE, Keen R. Development of reaching in infancy. *Experimental Brain Research* 2006;169(4): 507–518.
- Berthier NE, Clifton RK, Gullapalli V, McCall DD, Robin DJ. Visual information and object size in the control of reaching. *Journal of Motor Behavior* 1996;28(3):187–197. [PubMed: 12529202]
- Berthier NE, Clifton RK, McCall DD, Robin DJ. Proximodistal structure of early reaching in human infants. *Experimental Brain Research* 1999;127(3):259–269.
- Berti A, Frassinetti F. When far becomes near: remapping of space by tool use. *Journal of Cognitive Neurosciences* 2000;12(3):415–420.
- Binkofski F, Buccino G, Dohle C, Seitz RJ, Freund HJ. Mirror agnosia and mirror ataxia constitute different parietal lobe disorders. *Annals of Neurology* 1999;46(1):51–61. [PubMed: 10401780]
- Bonaiuto JB, Arbib MA. What did i just do? A new role for mirror neurons. 2009submitted

- Bonaiuto JB, Rosta E, Arbib MA. Extending the mirror neuron system model. I: Audible actions and invisible grasps. *Biological Cybernetics* 2007;96:9–38. [PubMed: 17028884]
- Caminiti R, Johnson PB, Galli C, Ferraina S, Burnod Y. Making arm movements within different parts of space: the premotor and motor cortical representation of a coordinate system for reaching to visual targets. *Journal of Neuroscience* 1991;11(5):1182–1197. [PubMed: 2027042]
- Cisek P, Kalaska JF. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron* 2005;45(5):801–814. [PubMed: 15748854]
- Clifton RK, Muir DW, Ashmead DH, Clarkson MG. Is visually guided reaching in early infancy a myth? *Child Development* 1993;64(4):1099–1110. [PubMed: 8404258]
- Craik, KJW. *The nature of explanation*. Cambridge: Cambridge University Press; 1943.
- Crammond DJ, Kalaska JF. Modulation of preparatory neuronal activity in dorsal premotor cortex due to stimulus-response compatibility. *Journal of Neurophysiology* 1994;71(3):1281–1284. [PubMed: 8201421]
- Crammond DJ, Kalaska JF. Prior information in motor and premotor cortex: activity during the delay period and effect on pre-movement activity. *Journal of Neurophysiology* 2000;84(2):986–1005. [PubMed: 10938322]
- Crutcher MD, Alexander GE. Movement-related neuronal activity selectively coding either direction or muscle pattern in three motor areas of the monkey. *Journal of Neurophysiology* 1990;64:151–163. [PubMed: 2388062]
- Dautenhahn, K.; Nehaniv, C., editors. *Imitation in animals and artifacts*. Cambridge, MA: The MIT Press; 2002.
- Dum RP, Strick PL. Frontal lobe inputs to the digit representations of the motor areas on the lateral surface of the hemisphere. *Journal of Neuroscience* 2005;25(6):1375–1386. [PubMed: 15703391]
- Ehrsson HH, Fagergren A, Jonsson T, Westling G, Johansson RS, Forssberg H. Cortical activity in precision- versus power-grip tasks: an fMRI study. *Journal of Neurophysiology* 2000;83(1):528–536. [PubMed: 10634893]
- Evarts EV. Relation of pyramidal tract activity to force exerted during voluntary movement. *Journal of Neurophysiology* 1968;31:14–27. [PubMed: 4966614]
- Fagg AH, Arbib MA. Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks* 1998;11(7–8):1277–1303. [PubMed: 12662750]
- Farnè A, Iriki A, Ladavas E. Shaping multisensory action-space with tools: evidence from patients with cross-modal extinction. *Neuropsychologia* 2005;43(2):238–248. [PubMed: 15707908]
- Ferrari PF, Rozzi S, Fogassi L. Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *Journal of Cognitive Neurosciences* 2005;17(2):212–226.
- Fogassi L, Ferrari PF, Gesierich B, Rozzi S, Chersi F, Rizzolatti G. Parietal lobe: from action organization to intention understanding. *Science* 2005;308(5722):662–667. [PubMed: 15860620]
- Frey SH. What puts the how in where? Tool use and the divided visual streams hypothesis. *Cortex* 2007;43(3):368–375. [PubMed: 17533760]
- Frey SH, Vinton D, Norlund R, Grafton ST. Cortical topography of human anterior intraparietal cortex active during visually guided grasping. *Brain Research: Cognitive Brain Research* 2005;23(2–3):397–405. [PubMed: 15820646]
- Gentilucci M, Roy AC, Stefanini S. Grasping an object naturally or with a tool: are these tasks guided by a common motor representation? *Experimental Brain Research* 2004;157(4):496–506.
- Goldenberg G, Hagmann S. Tool use and mechanical problem solving in apraxia. *Neuropsychologia* 1998;36(7):581–589. [PubMed: 9723930]
- Grafton ST, Arbib MA, Fadiga L, Rizzolatti G. Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Experimental Brain Research* 1996;112:103–111.
- Han CE, Arbib &, Schweighofer N. Stroke rehabilitation reaches a threshold. *PLoS Computational Biology* 2008;4(8):e1000133. [PubMed: 18769588]
- Head H, Holmes G. Sensory disturbances from cerebral lesions. *Brain* 1911;34:102–254.

- Hepp-Reymond MC, Husler EJ, Maier MA, Qi HX. Force-related neuronal activity in two regions of the primate ventral premotor cortex. *Canadian Journal of Physiology and Pharmacology* 1994;72(5): 571–579. [PubMed: 7954087]
- Hoff B, Arbib MA. Models of trajectory formation and temporal interaction of reach and grasp. *Journal of Motor Behavior* 1993;25(3):175–192. [PubMed: 12581988]
- Holmes NP, Calvert GA, Spence C. Extending or projecting peripersonal space with tools? Multisensory interactions highlight only the distal and proximal ends of tools. *Neuroscience Letters* 2004;372(1–2):62–67. [PubMed: 15531089]
- Iberall, T.; Arbib, MA. Schemas for the control of hand movements: an essay on cortical localization. In: Goodale, MA., editor. *Vision and action: the control of grasping*. Ablex Publishing Corporation; 1990. p. 204-242.
- Iberall, T.; Bingham, G.; Arbib, MA. Opposition space as a structuring concept for the analysis of skilled hand movements. In: Heuer, H.; Fromm, C., editors. *Generation and modulation of action patterns*. Berlin: Springer; 1986. p. 158-173.
- Imamizu H, Higuchi S, Toda A, Kawato M. Reorganization of brain activity for multiple internal models after short but intensive training. *Cortex* 2007;43(3):338–349. [PubMed: 17533757]
- Iriki A, Tanaka M, Iwamura Y. Coding of modified body schema during tool use by macaque postcentral neurones. *Neuro-Report* 1996;7:2325–2330.
- Jacobs S, Danielmeier C, Frey SH. Human anterior intra-parietal and ventral premotor cortices support effector-specific representations of grasping with the hand or a novel tool. 2009submitted
- Jeannerod, M.; Biguer, B. Visuomotor mechanisms in reaching within extra-personal space. In: Ingle, DJ.; Mansfield, RJW.; Goodale, MA., editors. *Advances in the analysis of visual behavior*. Cambridge, MA: The MIT Press; 1982. p. 387-409.
- Jeannerod M, Arbib MA, Rizzolatti G, Sakata H. Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neurosciences* 1995;18(7):314–320. [PubMed: 7571012]
- Johnson SH. Thinking ahead: the case for motor imagery in prospective judgements of prehension. *Cognition* 2000;74(1):33–70. [PubMed: 10594309]
- Johnson SH, Grafton ST. From ‘acting on’ to ‘acting with’: the functional anatomy of object-oriented action schemata. *Progress in Brain Research* 2003;142:127–139. [PubMed: 12693258]
- Johnson SH, Rotte M, Grafton ST, Hinrichs H, Gazzaniga MS, Heinze HJ. Selective activation of a parietofrontal circuit during implicitly imagined prehension. *Neuroimage* 2002;17(4):1693–1704. [PubMed: 12498743]
- Johnson-Frey, SH. Cortical mechanisms of human tool use. In: Johnson-Frey, SH., editor. *Taking action: cognitive neuroscience perspectives on the problem of intentional acts*. Cambridge, MA: The MIT Press; 2003. p. 185-217.
- Johnson-Frey SH. The neural bases of complex tool use in humans. *Trends in Cognitive Sciences* 2004;8(2):71–78. [PubMed: 15588811]
- Jordan MI, Rumelhart D. Forward models: supervised learning with a distal teacher. *Cognitive Science* 1992;16:307–354.
- Kakei S, Hoffman DS, Strick PL. Direction of action is represented in the ventral premotor cortex. *Nature Neuroscience* 2001;4(10):1020–1025.
- Kakei S, Hoffman DS, Strick PL. Sensorimotor transformations in cortical motor areas. *Neuroscience Research* 2003;46(1):1–10. [PubMed: 12725907]
- Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G. Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 2002;297:846–848. [PubMed: 12161656]
- Króliczak G, Frey SH. A common network in the left cerebral hemisphere represents planning of tool use pantomimes and familiar intransitive gestures at the hand-independent level. *Cerebral Cortex*. 2009in press
- Kurata K. Premotor cortex of monkeys: set- and movement-related activity reflecting amplitude and direction of wrist movements. *Journal of Neurophysiology* 1993;69:187–200. [PubMed: 8433130]
- Kurata K, Hoshi E. Movement-related neuronal activity reflecting the transformation of coordinates in the ventral premotor cortex of monkeys. *Journal of Neurophysiology* 2002;88(6):3118–3132. [PubMed: 12466435]

- Luppino G, Murata A, Govoni P, Matelli M. Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Experimental Brain Research* 1999;128(1–2):181–187.
- Ma WJ, Beck JM, Latham PE, Pouget A. Bayesian inference with probabilistic population codes. *Nature Neuroscience* 2006;9(11):1432–1438.
- Maier MA, Bennett KM, Hepp-Reymond MC, Lemon RN. Contribution of the monkey corticomotoneuronal system to the control of force in precision grip. *Journal of Neurophysiology* 1993;69(3):772–785. [PubMed: 8463818]
- Maravita A, Iriki A. Tools for the body (schema). *Trends in Cognitive Sciences* 2004;8(2):79–86. [PubMed: 15588812]
- Maravita A, Spence C, Kennett S, Driver J. Tool-use changes multimodal spatial interactions between vision and touch in normal humans. *Cognition* 2002;83(2):B25–B34. [PubMed: 11869727]
- Marconi B, Genovesio A, Battaglia-Mayer A, Ferraina S, Squatrito S, Molinari M, et al. Eye-hand coordination during reaching. I. Anatomical relationships between parietal and frontal cortex. *Cerebral Cortex* 2001;11(6):513–527. [PubMed: 11375913]
- Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT. Throwing while looking through prisms. I. Focal olivo-cerebellar lesions impair adaptation. *Brain* 1996a;119(Pt 4):1183–1198. [PubMed: 8813282]
- Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT. Throwing while looking through prisms. II. Specificity and storage of multiple gaze-throw calibrations. *Brain* 1996b;119(Pt 4):1199–1211. [PubMed: 8813283]
- Matelli M, Luppino G. Parietofrontal circuits for action and space perception in the macaque monkey. *Neuroimage* 2001;14:S27–S32. [PubMed: 11373129]
- Ochiai T, Mushiaki H, Tanji J. Involvement of the ventral premotor cortex in controlling image motion of the hand during performance of a target-capturing task. *Cerebral Cortex* 2005;15(7):929–937. [PubMed: 15483048]
- Oztop, E.; Arbib, MA.; Bradley, N. The development of grasping and the Mirror System. In: Arbib, MA., editor. *Action to language via the mirror neuron system*. Cambridge: Cambridge University Press; 2006. p. 397–423.
- Oztop E, Bradley NS, Arbib MA. Infant grasp learning: a computational model. *Experimental Brain Research* 2004;158(4):480–503.
- Oztop E, Imamizu H, Cheng G, Kawato M. A computational model of anterior intraparietal (AIP) neurons. *Neurocomputing* 2007;69:1354–1361.
- Picard N, Strick PL. Imaging the premotor areas. *Current Opinion in Neurobiology* 2001;11:663–672. [PubMed: 11741015]
- Pouget A, Deneve S, Ducom JC, Latham PE. Narrow versus wide tuning curves: what's best for a population code? *Neural Computation* 1999;11(1):85–90. [PubMed: 9950723]
- Raos V, Umiltà MA, Gallese V, Fogassi L. Functional properties of grasping-related neurons in the dorsal premotor area F2 of the macaque monkey. *Journal of Neurophysiology* 2004;92(4):1990–2002. [PubMed: 15163668]
- Raos V, Umiltà MA, Murata A, Fogassi L, Gallese V. Functional properties of grasping-related neurons in the ventral premotor area F5 of the macaque monkey. *Journal of Neurophysiology* 2006;95(2):709–729. [PubMed: 16251265]
- Riehle A, Requin J. Monkey primary motor and premotor cortex: single-cell activity related to prior information about direction and extent of an intended movement. *Journal of Neurophysiology* 1989;61(3):534–549. [PubMed: 2709098]
- Rizzolatti G, Camarda R, Fogassi L, Gentilucci M, Luppino G, Matelli M. Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Experimental Brain Research* 1988;71:491–507.
- Rolls ET. Convergence of sensory systems in the orbitofrontal cortex in primates and brain design for emotion. *The Anatomical Record: Part A, Discoveries in Molecular, Cellular, and Evolutionary Biology* 2004;281(1):1212–1225.
- Sakamoto T, Arissian K, Asanuma H. Functional role of the sensory cortex in learning motor skills in cats. *Brain Research* 1989;503(2):258–264. [PubMed: 2605518]

- Schluter ND, Krams M, Rushworth MF, Passingham RE. Cerebral dominance for action in the human brain: the selection of actions. *Neuropsychologia* 2001;39(2):105–113. [PubMed: 11163368]
- Schluter ND, Rushworth MF, Passingham RE, Mills KR. Temporary interference in human lateral premotor cortex suggests dominance for the selection of movements: a study using transcranial magnetic stimulation. *Brain* 1998;121(Pt 5):785–799. [PubMed: 9619185]
- Schwartz AB, Kettner RE, Georgopoulos AP. Primate motor cortex and free arm movements to visual targets in three-dimensional space. 1. Relations between single cell discharge and direction of movement. *Journal of Neuroscience* 1988;8:2913–2927. [PubMed: 3411361]
- Spelstra J, Arbib MA. Cerebellar microcomplexes and the modulation of motor pattern generators. *Autonomous Robots* 2001;11:273–278.
- Sutton, RS.; Barto, AG. Reinforcement learning: an introduction. Cambridge, MA: The MIT Press; 1998.
- Taub E, Uswatte G, Elbert T. New treatments in neurore-habilitation founded on basic research. *Nature Reviews: Neuroscience* 2002;3:228–236.
- Tunik E, Frey SH, Grafton ST. Virtual lesions of the anterior intraparietal area disrupt goal-dependent on-line adjustments of grasp. *Nature Neuroscience* 2005;8(4):505–511.
- Turvey M. Affordances and prospective control: an outline of the ontology. *Ecological Psychology* 1992;4:173–187.
- Turvey M, Shaw R, Reed E, Mace W. Ecological laws of perceiving and acting. *Cognition* 1981;9:237–304. [PubMed: 7197604]
- Umiltà MA, Brochier T, Spinks RL, Lemon RN. Simultaneous recording of macaque premotor and primary motor cortex neuronal populations reveals different functional contributions to visuomotor grasp. *Journal of Neurophysiology* 2007;98(1):488–501. [PubMed: 17329624]
- Umiltà MA, Escola L, Intskirveli I, Grammont F, Rochat M, Caruana F, et al. When pliers become fingers in the monkey motor system. *Proceedings of the National Academy of Sciences of the United States of America* 2008;105(6):2209–2213. [PubMed: 18238904]
- Umiltà MA, Kohler E, Gallese V, Fogassi L, Fadiga L, Keysers C, et al. I know what you are doing: a neuro-physiological study. *Neuron* 2001;31:155–165. [PubMed: 11498058]
- von Hofsten C. An action perspective on motor development. *Trends in Cognitive Sciences* 2004;8(6):266–272. [PubMed: 15165552]
- von Hofsten C, Ronnqvist L. The Structuring of neonatal arm movements. *Child Development* 1993;64:1046–1057. [PubMed: 8404256]
- Weinrich M, Wise SP. The premotor cortex of the monkey. *Journal of Neuroscience* 1982;2(9):1329–1345. [PubMed: 7119878]
- Wing AM. Motor control: mechanisms of motor equivalence in handwriting. *Current Biology* 2000;10(6):R245–248. [PubMed: 10744963]
- Wolpert DM, Kawato M. Multiple paired forward and inverse models for motor control. *Neural Networks* 1998;11(7–8):1317–1329. [PubMed: 12662752]
- Zukow-Goldring P, Arbib MA. Affordances, effectivities, and assisted imitation: caregivers and the directing of attention. *Neurocomputing* 2007;70:2181–2193.

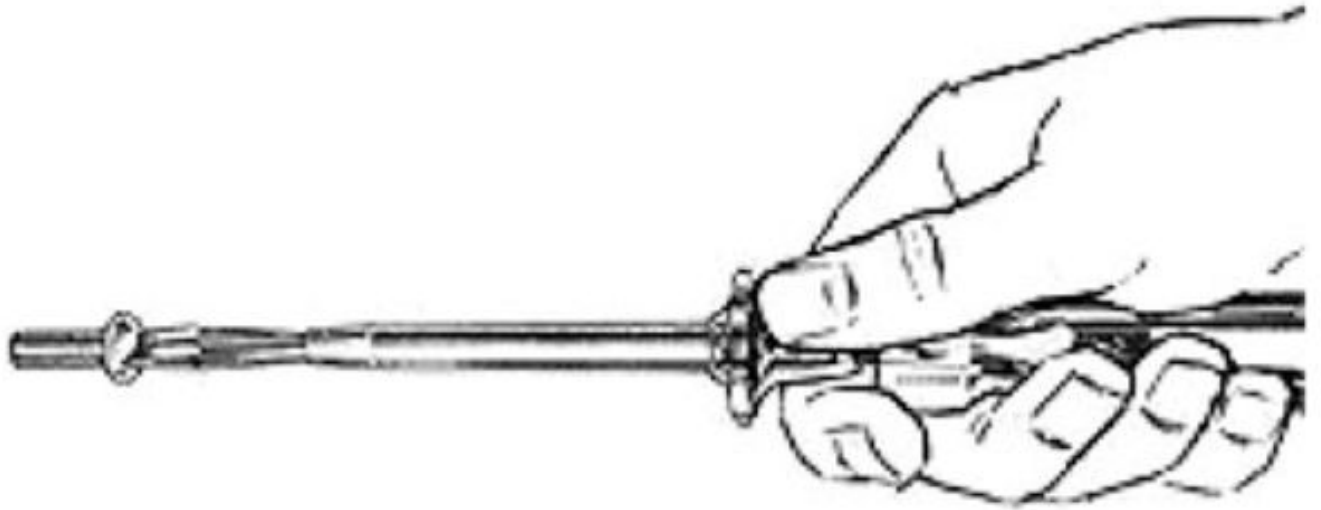


Fig. 1.
A compound of a side grasp and a power grasp for holding a screwdriver

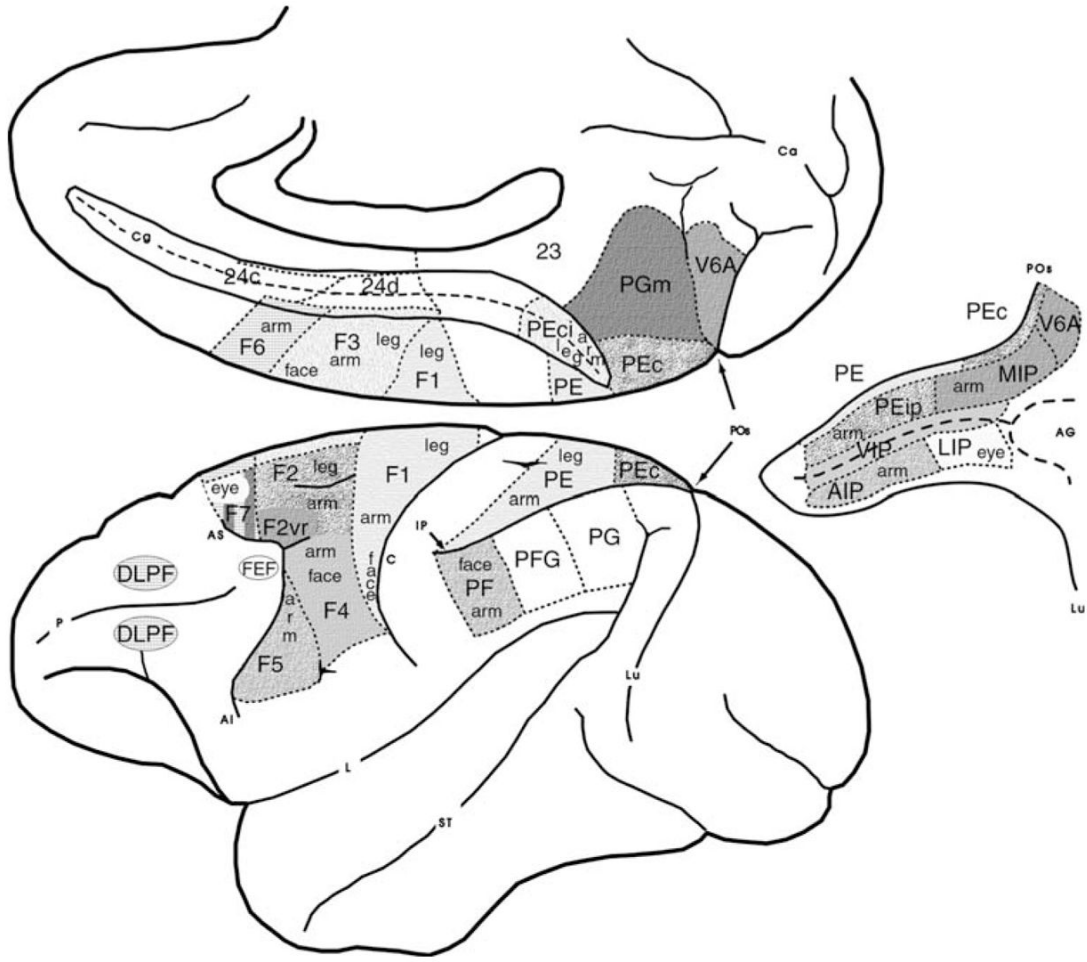


Fig. 2. A view of macaque brain areas involved in basic parieto-frontal interactions for visually directed hand movements (from Matelli & Luppino, 2001). We will see that inferotemporal cortex and prefrontal cortex play a major role in modulating these interactions

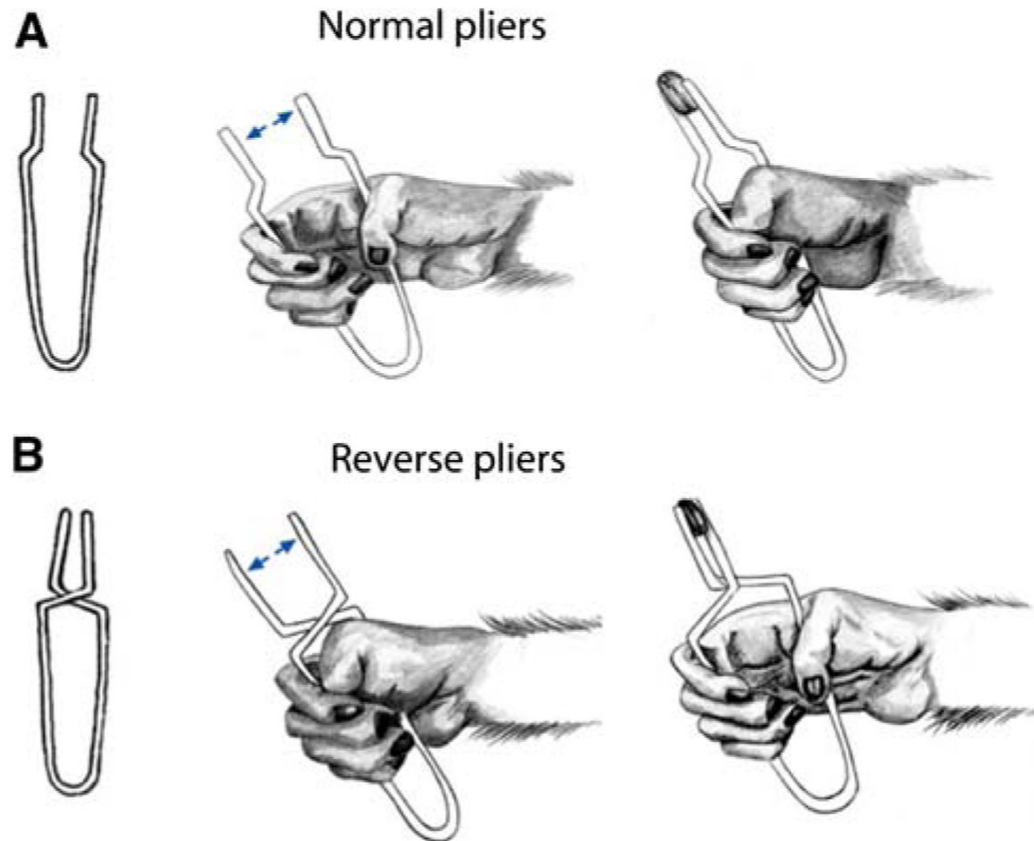


Fig. 3. Schematic illustration of the paradigm used by Umiltà et al. (2008) to dissociate motion of the end-effector (jaws of the pliers) from the hand movements required to achieving it. To grasp the object in a precision pinch using normal pliers (**a**) the monkey has to close its hand in a power grasp, while with the reverse pliers (**b**) the monkey has to release a power grasp

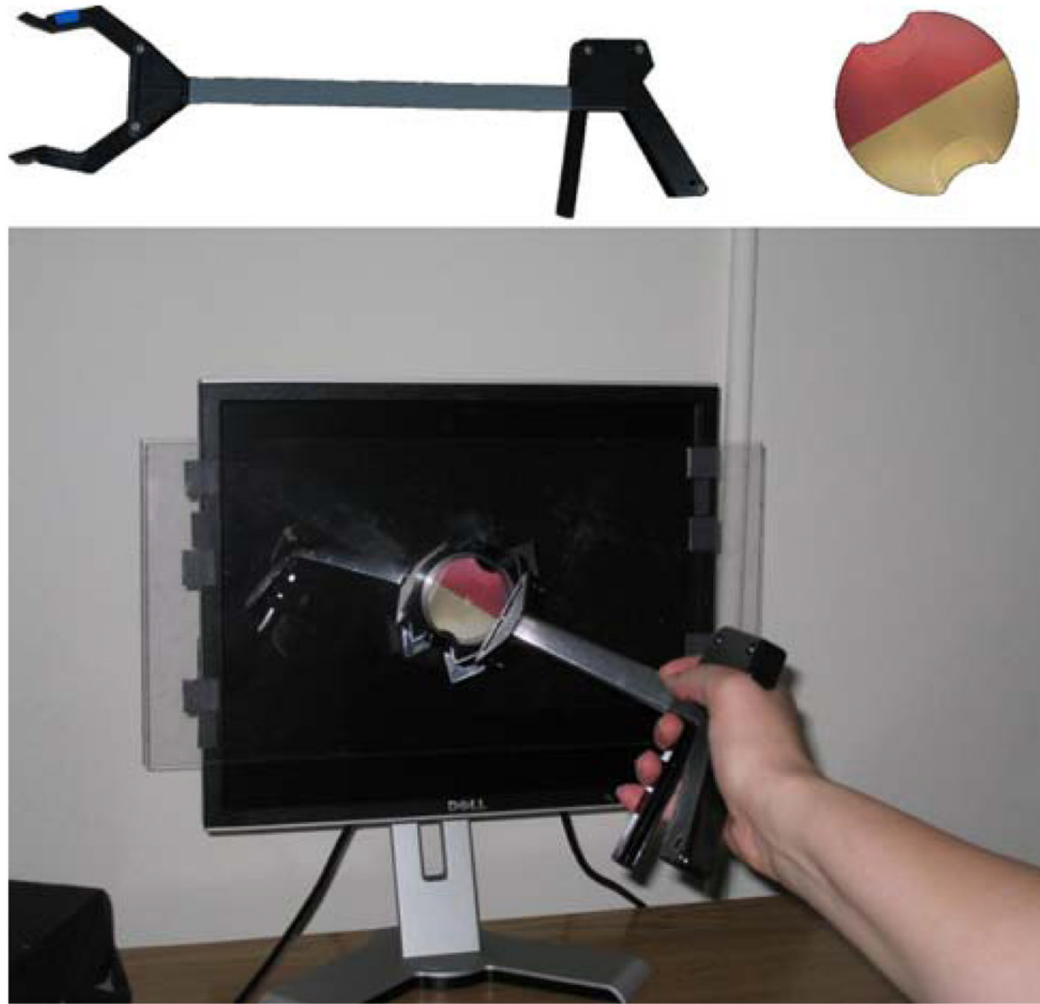


Fig. 4. Novel tool and a sample of the stimulus object. *Bottom* a participant is shown grasping the object with the tool in the right hand during the overt grasp selection (OGS) task

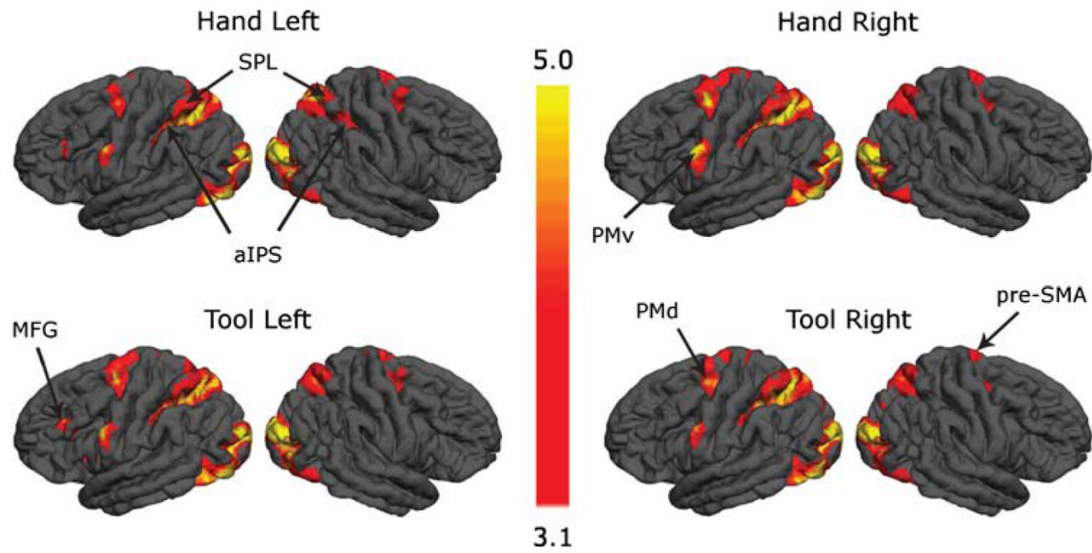


Fig. 5.

Increased neural activity associated with planning grasping actions with the hands or the tool relative to resting baseline. Selecting grips (PGS task) based on either hand or on the tool was consistently associated with activations within *left* aIPS and PMv, the putative homologs of the macaque AIP-F5 grasp circuit. In addition, activity was increased within and along the intra-parietal sulcus, as well as in bilateral superior parietal lobule (SPL) and dorsal premotor cortex (PMd), and in pre-SMA. Other structures were also activated only in some conditions, such as the *right* aIPS and the *left middle* frontal gyrus (MFG). (Jacobs et al., 2009)

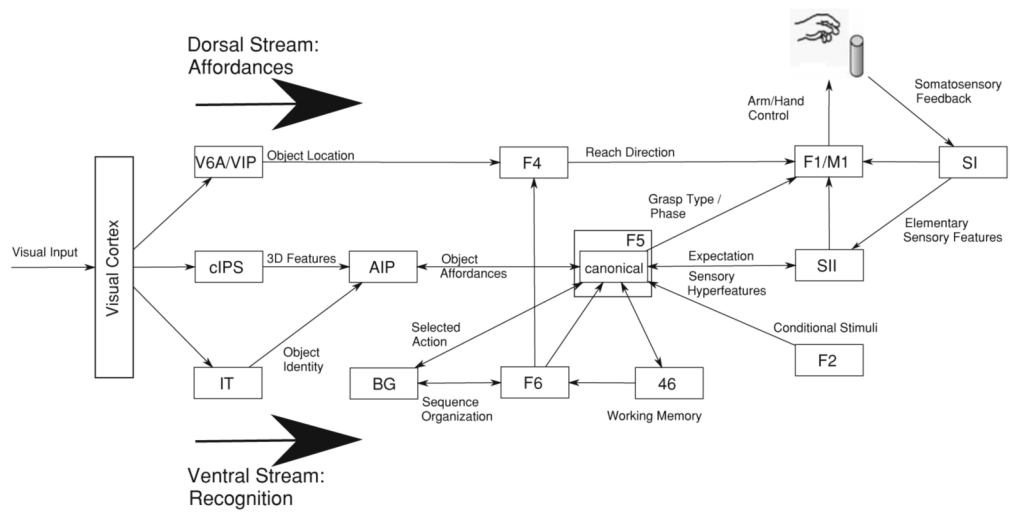


Fig. 6. The complete FARS model. (adapted from the original FARS figure of Fagg & Arbib, 1998)

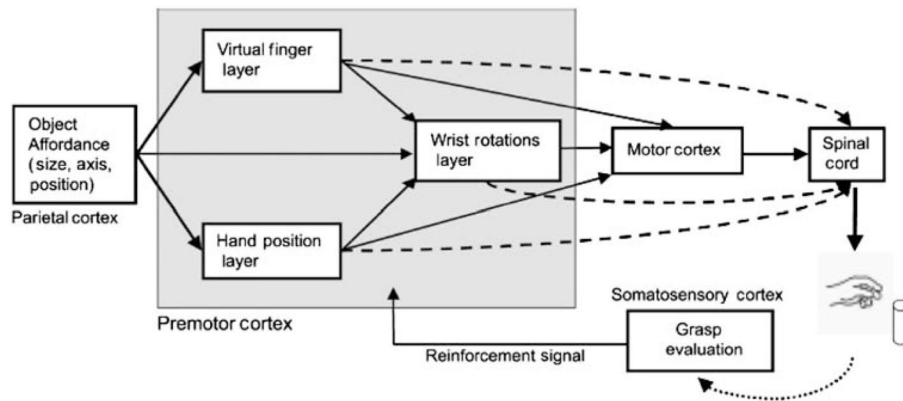


Fig. 7. The structure of the infant learning to grasp model (ILGM) (Oztop et al., 2004). The individual layers inside the *gray box* are trained based on a *Joy of Grasping* reinforcement signal arising from somatosensory feedback which increases with the stability of the current grasp

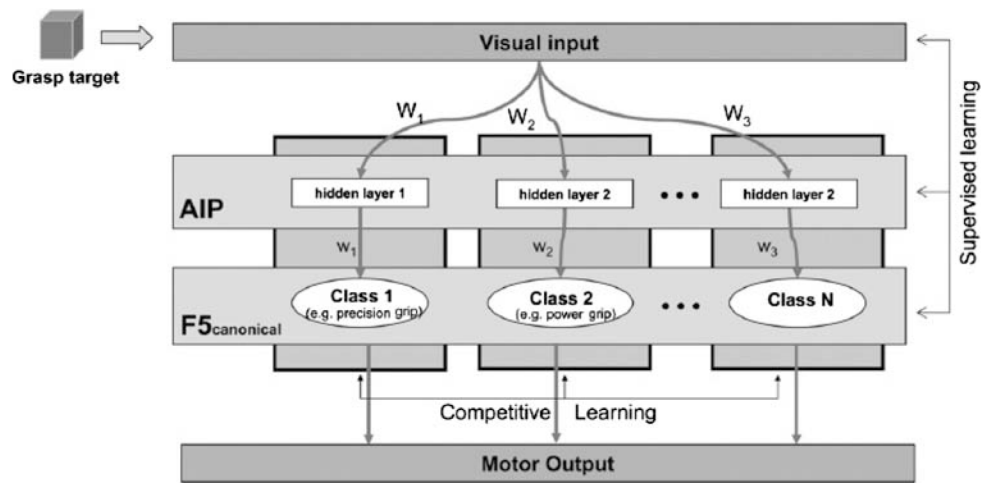


Fig. 8. The grasp affordance learning model (GAEM) derives its training inputs from ILGM. The successful grasping movements provide the basis for learning by GAEM: *vertical* learning, supervised; *horizontal* learning, competitive (Oztop et al., 2007)

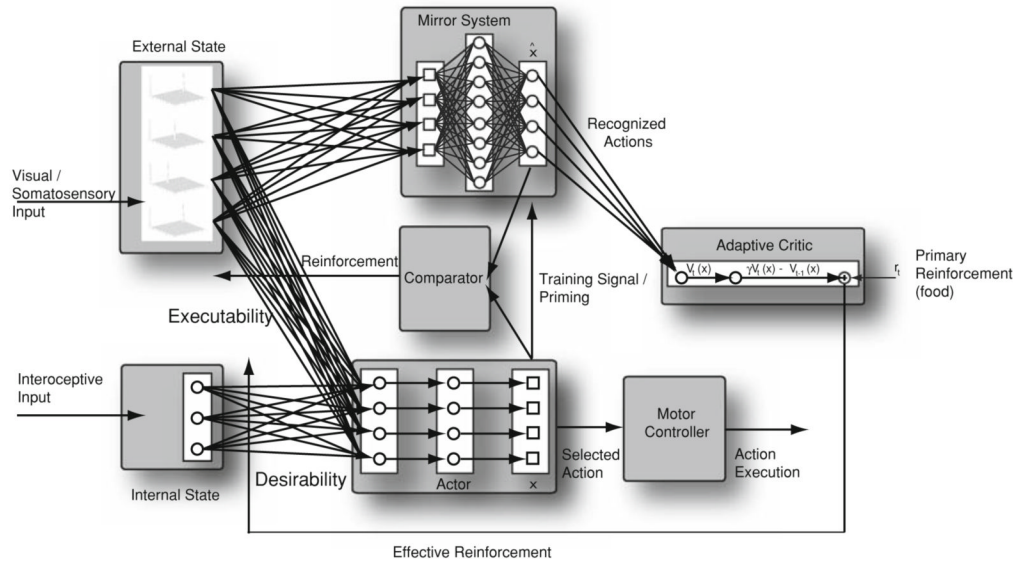


Fig. 9. A simplified version of the augmented competitive queuing (ACQ) system as structured for the sequencing of actions in general. A crucial change here is that the mirror system can support the representation of multiple actions. Thus, during a self movement, the mirror system can represent both the intended action as well as any other action whose external appearance matches the actual performance. In ACQ, the Actor selects for execution the action that has the highest current priority, defined as a combination of executability and desirability. Desirability is the expected reinforcement for executing an action in the current internal state (which may combine homeostatic variables and cognitive subgoals). Estimates of desirability are updated by the Critic, which employs temporal difference (TD) learning. A crucial innovation here is that the Critic assesses not only the current action but also those apparent actions reported by the Mirror System in making its assessments