Grounding Procedural and Declarative Knowledge in Sensorimotor Anticipation

GIOVANNI PEZZULO

Abstract: We propose a view of embodied representations that is alternative to both symbolic/linguistic approaches and purely sensorimotor views of cognition, and can account for procedural and declarative knowledge manipulation. In accordance with recent evidence in cognitive neuroscience and psychology, we argue that anticipatory and simulative mechanisms, which arose during evolution for action control and not for cognition, determined the first form of representational content and were exapted for increasingly sophisticated cognitive uses. In particular, procedural and declarative forms of knowledge can be explained, respectively, in terms of on-line sensorimotor anticipation and off-line simulations of potential actions, which can give access to tacit knowledge and make it explicit. That is, mechanisms that evolved for the on-line prediction of the consequences of one's own actions (i.e. forward models) determine a (procedural) form of representation, and became exapted for off-line use. They can therefore be used to produce (declarative) knowledge of the world, by running a simulation of the action that would produce the relevant information. We conclude by discussing how embodied representations afford a form of internal manipulation that can be described as internalized situated action.

1. Introduction

In recent decades the debate on representation in cognitive science and philosophy has been focused on two extremes: on the one side, arbitrary (linguistic) symbols and symbolic reasoning, and on the other side (representation-poor or representation-less) sensorimotor skills. This is unfortunate, we believe, since it could be the case that most cognition actually happens in the wide intermediate space between them, and is supported by processes that are neither symbolic or linguistic, nor purely sensorimotor. This is the realm of *embodied representations*, which are neither mere 'copies' of the external reality nor (linguistic) symbols, but can be internally manipulated (before or instead of acting in the external world) and used for most cognitive operations, including reasoning and planning. For example, a mechanic can assemble and dismantle a motor in her mind before doing it in practice; an interior designer can compare, in her mind, different possible arrangements of the furniture in a room by considering their shape, colour and

---

The author wishes to thank Professor Michael Arbib, Dr Patric Bach, Professor Rick Grush, and Dr Luca Tummolini for their invaluable advice in the revision of the manuscript, and two anonymous reviewers for insightful comments and useful criticism.

Address for correspondence: ISTC-CNR, Via S. Martino della Battaglia, 44 - 00185 Rome, Italy.
Email: giovanni.pezzulo@cnr.it

Mind & Language, Vol. 26, No. 1 February 2011, pp. 78–114.

size, and we all frequently mentally compare multiple paths to get home and select the shortest one, or the one with less traffic. The ability to manipulate internal representations is not exclusive to humans, however. Other animals, which do not use language, can internally imagine and compare multiple possibilities for action, for example several navigation paths, before attempting them in practice, and systematically predict the long term effects of their actions, in order to select from short-term and long-term goals, and to avoid potential dangers (see Pezzulo, 2008 for a review on this topic).

How do these representational processes work? Which (brain) mechanisms support them? In this paper we propose a naturalistic account of representation inspired by the *action-based view of cognition*, the idea that cognitive phenomena have a situated and embodied origin and retain the same nature even for higher-level cognitive abilities. This idea is currently gaining consensus in the fields of cognitive neuroscience and psychology thanks to a great deal of supporting empirical data, including evidence that sensorimotor areas in the human brain are activated during 'higher-level' cognitive tasks such as memory retrieval and language processing (Barsalou, 2008; Carpenter *et al*., 1999; Gallese, 2003; Glenberg and Kaschak, 2002; Pulvermüller, 1999; Scorolli *et al*., 2009). For example, it has been shown that salient properties of objects such as their appearance and use, are represented in the sensory and motor brain areas that were active when that object information was acquired (Martin, 2007). To explain these findings, which are difficult to accommodate in the traditional view of representation in cognitive science, it has been suggested that these processes are supported by *embodied representations* which, unlike traditional symbols, arise in modality-specific systems in the brain and not in amodal modules. At the same time, as with traditional symbols, embodied representations can be internally re-enacted and manipulated (Barsalou, 1999).

Another influential stream of research is focused on the fact that the same neural circuits underlie a diverse range of individual and social cognitive capabilities, such as action execution, understanding the actions of others, imitation, perspective taking, and communication. In all these cognitive phenomena a crucial role is played by common *anticipatory and simulative mechanisms*, which permit the prediction of the effect of one's own actions as well as that of external events, by re-enacting the sensorimotor brain mechanisms originally serving for situated (inter)action, and provide common representational structures for such individual and social cognitive capabilities (Arbib, 2002; Decety and Grèzes, 2006; Iacoboni, 2003; Rizzolatti *et al*., 2001; Wolpert *et al*., 2003).

Overall, based on the afore-mentioned evidence (and others), an *action-based view of cognition* is emerging which considers a substantial part of thinking and cognition to be dependent on *motor processes*, the most important mechanism of which seems to be the re-enactment of the motor apparatus 'in simulation' (Grush, 2004; Jeannerod, 2001, 2006). Mental simulations realize *covert* forms of action and perception: they exploit the same neural machinery of goal-directed action and perception, though in the absence of external stimuli and without any actual movement, and permit engagement in simulated interactions with the external environment. This simulation produces a motor understanding of objects and events which differs from a mere perceptual understanding, since it permits an immediate 'grasp' over entities of the external environment, including other people, in terms of one's own bodily states or dispositions and possibilities for action (hence the name action-based). This kind of 'direct' understanding apparently resembles the concept of *affordances* in (Gibson, 1979). One fundamental difference, however, is that it does not rely on direct perception, but is produced by internal modelling and the re-enactment of motor processes and/or of bodily states (see Pezzulo and Castelfranchi, 2009). Since these processes include both embodied and representational aspects, we call them *embodied representations*, and consider them an alternative to traditional symbolic/linguistic representations.

1.1 Three Theses

The first objective of this paper is to clarify the origin and nature of embodied representations. We trace the roots of representation in pragmatic, situated activity and its anticipation, arguing that anticipatory and simulative processes determine representational content:

> *an organism's knowledge and representation ability originates from—and is grounded in—anticipation of sensorimotor interaction.*

The core idea is that representations develop from the results of action-control mechanisms that project different courses of action into the future, and retain the same embodied format even in their more sophisticated forms. For example, knowledge of the weight of a cup of wine is coded in the procedural, interaction-oriented format that serves to grasp and lift it. Nonetheless, we argue that such grounded representations can be *detached* and reused outside their original context of acquisition. Indeed, most of our knowledge, and particularly *declarative* knowledge, cannot be tied to any particular context of action. For example, the belief 'the cup is half full' can support different (proximal or distal) actions and decisions, such as drinking, deciding to buy more wine, or using the cup as a paperweight. How is it possible to pass from the 'motor grasp' offered by procedural representations to the more complex and open-ended 'epistemic grasp' over reality offered by declarative knowledge?

We argue that one critical step is the reuse of the same mechanisms used for action control off-line, in *simulation* of action and perception (i.e. imagining lifting a cup). This permits embodied representations to be elicited in the same way they are elicited in actual action, but produces declarative knowledge (i.e. knowledge *that* the weight of said cup is such and such) and makes it available outside the original sensorimotor context. Therefore, the second thesis of the paper is that embodied representations provide a satisfactory account of both procedural and declarative knowledge representation:

> *anticipatory mechanisms, as used in situated action execution and control, produce procedural knowledge; their re-enactment (in simulation of possible situated action) can produce declarative knowledge.*

Finally, we discuss how such declarative forms of knowledge enable sophisticated forms of cognition and internal manipulations that differ substantially from symbolic manipulations of arbitrary tokens, or 'symbol crunching', implemented in traditional AI systems (Newell and Simon, 1972):

> *embodied representations afford a form of internal manipulation that can be described as internalized situated action.*

A terminological note: the terms 'anticipatory mechanisms' and 'simulative mechanisms' are sometimes used as synonyms. Here we will refer to *anticipatory mechanisms* as those (motor) brain structures that allow the on-line prediction of the effects of one's own actions or other regularities in the environment. We will instead refer to *simulative mechanisms* (or, as a synonym, *emulative mechanisms*) as those that allow anticipation on a longer time scale based on the re-enactment of anticipatory mechanisms (for example, for planning and realizing distal effects). These mechanisms are used off-line (e.g. to simulate possible future actions, or counterfactuals) and can produce long-term predictions by 'chaining' multiple short-term predictions (and at the same time inhibiting motor commands). Throughout the paper we will consider *internal forward models* (Wolpert and Ghahramani, 2004) as prototypical exemplifications of anticipatory mechanisms (if used on-line) or simulative mechanisms (if used off-line). We will argue that in both cases embodied representations are produced whose representational content consists in anticipation of the effects of interactions, including their goals.

The paper is structured as follows. In Section 2 we discuss three possible approaches to the naturalization of representation: the sensorimotor view, the recoding view, and the action-based view, and review the relevant empirical literature on how representational processes could be implemented in brain mechanisms. In Section 3, Section 4, and Section 5 we discuss the three aforementioned theses. In Section 6 we draw our conclusions and discuss the transition from embodied to linguistic representations as two steps of a process of 'internalization' of the external reality.

2. Naturalizing Representation: Three Approaches

Recently there has been a renewed interest in the situated and embodied roots of thinking and cognition not only among those researchers who tend to deny the existence of representations, but also among those who consider cognition to be essentially based on internal manipulation of representations (Clark, 1998). As a consequence, there have been a number of attempts to reformulate the traditional concept of *representation* used in philosophy and cognitive science. The amodality, arbitrariness, and discreteness of these representations have been widely challenged (see Markman and Dietrich, 2000 for a review). Here we discuss the three most popular candidate solutions: the sensorimotor, recoding, and action-based approaches.

2.1 The Sensorimotor Approach

The sensorimotor approach overturns the classical idea of cognitive science that cognition equates to reasoning, and tends to collapse cognition and sensorimotor interaction, thus minimizing the role of representation. According to this view, most (if not all) internalized knowledge exists in sensorimotor format and is grounded bottom-up in an agent's sensory states (see e.g. Harnad, 1990). Variants of this view are the *enactive* and eliminativist approaches, which tend to view cognition as not mediated by any representation (Churchland, 1981; Maturana and Varela, 1980).

According to the sensorimotor view and its variants, for situated action and most (if not all) cognition it is sufficient to 'enact' or 'master skills', with little internalization (or no internalization at all) of the external world (see e.g. Gibson, 1966; O'Regan and Noë, 2001). Epistemic processes and states, which are traditionally considered presuppositions for action, are almost totally collapsed on pragmatic actions: one might argue that there are nonetheless some forms of implicit information or representation, and this is merely a terminological dispute.

Here an example may help clarify how epistemic processes can depend on pragmatic ones. In *active perception* (Ballard, 1991) acting serves also to gather information, and behaviour drives perceptual processes toward information that is needed for acting. Organisms, with their actions, modify the environment and in doing so partially determine their next stimuli, in particular stimuli that are necessary for triggering the next action.[1] Evolution (or learning) has shaped effective action-perception loops in living organisms, and in most cases this makes it unnecessary to represent information internally. Current research in evolutionary robotics has shown that artificial creatures can successfully solve quite complex navigation tasks (Beer, 1997; Nolfi, 2005) by learning (or evolving) appropriate sensorimotor routines that allow picking up sensory information at the right time. A related view is Brooks' (1991) idea of *intelligence without representation*: the environment is the best model of itself and, when needed, information can be accessed via sensors rather than internally represented. Again, a tight action-environment coupling substitutes representation and the formation of internal models, and the environment can serve as a memory, too.

---

[1] This is essentially the objection of Dewey (1896, p. 363) to the stimulus-response paradigm: 'What we have is a circuit, not an arc or broken segment of a circle. This circuit is more truly termed organic than reflex, because the motor response determines the stimulus, just as truly as sensory stimulus determines movement.'

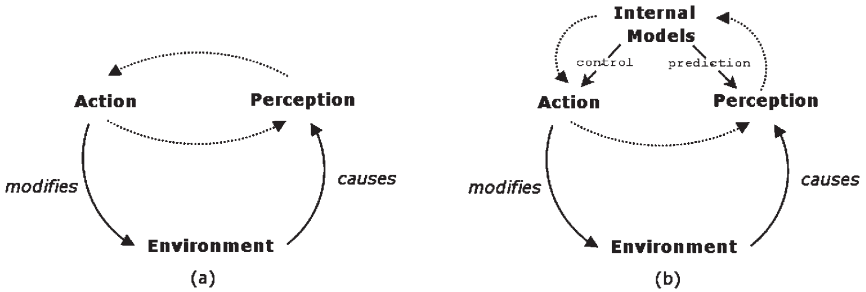

Figure 1 *(a) The action-perception loop in the sensorimotor view. (b) Action-perception loop augmented with prediction in action-based view*

Figure 1(a) shows the action-perception loop in the sensorimotor view. Solid edges stand for the double causal impact between an organism and its environment: the organism modifies its environment by acting, and in turn the environment causes the organism's perceptions. Dotted edges represent the fact that actions and perceptions influence each other indirectly; for example, an organism's actions partially determine its next stimulus. Overall, the sensorimotor view de-emphasizes representations and reduces them to internalized sensory stimuli.

2.1.1 Possible Limitations. This view has received several criticisms, in particular for its difficulty to 'scale up' and explain complex cognitive functions of organisms which go beyond the here-and-now of sensorimotor interaction. Although it appears plausible that sensorimotor interaction is required for understanding those concepts (such as 'chair', 'red' and 'heavy') that have a strong sensory component, one might argue that it is not sufficient for explaining abstract concepts (such as 'true' or 'change'), which apparently have a conceptual and not sensorimotor nature. In addition, this approach is focused on the here-and-now of interaction and this makes it difficult to account for *distal* events and goal-directed action anticipatory effects in purposeful action; that is, the fact that (distal) goals influence backward (proximal) actions. For example, when grasping a cup of wine the choice of grip varies depending on the distal goal (e.g. drinking or throwing it) although this is not dictated by the needs of the current (grip) action. This effect, which is evident in everyday life, has been studied extensively. In a study by Rosenbaum *et al.* (1990) participants were asked to grasp a cylinder (placed horizontally) and to move it to a target to their left or right, and spontaneously used overhand or underhand grip when reaching either target with the right or left end of the cylinder respectively. In other words, subjects selected awkward hand postures in the holding phase to permit a more comfortable final posture (this is why this effect is named *end-state comfort effect*). This study (and numerous others, see Rosenbaum *et al.*, 2006 for a review) indicates that during movement planning subjects predict (at least part of) the goal state, including the final limb configuration, and this (goal) information influences the plan.

2.2 Recoding Approach

An alternative approach has emerged in the literature in an attempt to explain complex cognitive capabilities: the idea that a *recoding* from a sensorimotor to a 'conceptual' format is necessary for flexible cognitive processing, and might happen in increasingly high-level cortical brain areas. For example, Mahon and Caramazza (2005) describe neurobiological findings that are difficult to explain using only motor theory of cognition. In particular, the existence of a category-based memory organization of natural objects and artefacts seems to contrast with the idea of a distributed sensorimotor, property-based organization (Martin, 2007), and supports the idea of localization of higher-level cognitive functions in specialized brain areas. The recoding approach denies or minimizes the embodied and sensorimotor aspects of representation. According to this view, representations implied in cognition have little (or nothing) in common with those used for guiding action, for two reasons: (1) recoding changes their status from modal to amodal, and from sensorimotor to arbitrary format; (2) recoding changes their status from implicit to (increasingly) explicit, accessible, and communicable. In this view behaviour can be described as a *perception → representation → action* pipeline. Perception transforms sensory stimuli into internal (recoded) representations, the cognitive system operates on these representations to produce plans, and the action system transforms plans into sequences of motor acts. Therefore, sensorimotor and representational processes are distinct; even when empirical evidence (e.g. obtained with brain imaging techniques) indicates an involvement of sensorimotor areas in higher-level cognitive tasks, this is a mere epiphenomenon without any causal role.

The process of recoding and the existence of dedicated brain areas for conceptual representations is close to the view of Karmiloff-Smith (1992) that in cognitive development, cognitive processing undergoes stages of *representational redescription* (RR) in which knowledge in the system is redescribed in increasingly abstract and accessible forms: Implicit, Explicit 1, Explicit 2 and Explicit 3 (with different levels of accessibility and communicability). Here the lowest level (Implicit) stands for procedural knowledge, and the highest levels are amodal and arbitrary symbols that are explicitly (and consciously) accessible for processing and communication, much like those postulated in the 'physical symbol system' (Newell, 1990) and 'language of thought' theories (Fodor, 1975).

2.2.1 Possible Limitations. The possible existence of dedicated brain areas for higher-level cognition, and of 'recoded' representations in the brain, are largely empirical and not philosophical problems. Here we focus on three related philosophical issues instead.

The first problem is that amodality and arbitrariness of representations have not only strengths, but weaknesses as well. Their strength is that they easily account for productivity, systematicity, and compositionality (Fodor, 1975), which are key features of the most sophisticated symbolic capabilities of the human species. Their weakness is that it is very difficult to explain them in a naturalistic perspective. The idea that (several if not all) cognitive capabilities involve manipulation of amodal representations seems to be too strong and too weak at once. Too strong, since there is evidence that modal (embodied) representations play a role in several cognitive capabilities, and thus amodality and arbitrariness may not be required. Too weak, since there is no real evidence of amodal representations in the brain (Barsalou, 1999): from an empirical point of view, proving their existence is much more problematic than assessing embodied representations.

The second problem, which further complicates the first, is that we currently lack a naturalistic account of the recoding/redescription process, or of how amodal representations could have arisen, while preserving intentionality and aboutness (Pezzulo and Castelfranchi, 2007). One possibility, which we discuss in Section 6.1, is a crucial role of language in shaping symbolic representations, but we argue that this happens on the basis of an already existing 'representational' brain substrate (see Dehaene and Cohen, 2007 on this theme).

The third problem is of a different nature. We think that in traditional cognitive science (cf. Fodor, 1975) representation has been almost collapsed into symbolic or linguistic processing. Unfortunately, part of the mystery of grounding resides in the overemphasis of these *explicit* forms of representations (Bickhard, 1993). This trend influenced AI, too, at least in its first decades, but as a consequence several forms of unboundedness arose such as the frame problem, or the difficulty for an artificial system to determine which part of its knowledge is relevant, given its current context and goals (McCarthy and Hayes, 1969). The 'novel' cognitive science of the last two decades has focused instead on sensorimotor or dynamic processes (essentially non-representational) as the only alternative to symbols and linguistic processes.

In this paper we follow instead a different strategy. We argue that there are genuine representational processes that are based on embodied representations. Such processes are sufficient for explaining a wide range of cognitive functions, including those that are supposed to be the exclusive province of amodal, arbitrary and explicit symbols, such as planning and reasoning. If this is true, the role of recoded or redescribed representations, including linguistic representations, can be de-emphasized. Moreover, we believe that our analysis can help understand complex kinds of representational processes (mediated by linguistic symbols), which should be better analyzed in continuity with the simpler sensorimotor processes we study here. To do so, we introduce the action-based approach.

2.3 Action-based Approach

The *action-based* approach shares with the sensorimotor approach the emphasis on organism-environment interaction, and gives a central role to the coupling between perception and action. However, some of its features make it better suited to explain higher order cognitive phenomena without postulating any recoding. Distinguishing the action-based from the sensorimotor approach is an acknowledgment of the essentially *goal-directed nature of action*. Converging evidence indicates that several cognitive capabilities across the individual and social domains, including action planning and execution, imagery, understanding others' actions, and imitation are essentially goal-directed (Barsalou, 2003; Decety and Grèzes, 2006; Frith, 2007; Gallese and Metzinger, 2003; Jeannerod, 2001; Wohlschlaeger *et al.*, 2003). For example, empirical findings indicate that goal representations have a crucial role in the planning and control of action, and action understanding and imitation are performed at the goal rather than the movement level. Perhaps the most striking evidence comes from the discovery of the 'mirror neuron system' (Rizzolatti *et al.*, 1996; Rizzolatti and Craighero, 2004), which determines a direct, 'resonant' understanding of goal-directed action, proximal and distal (Fogassi *et al.*, 2005) via a significant recruitment of the motor system.

2.3.1 Internal Modelling Supports Goal-directed Action. Overall, these studies have highlighted that one important part of the epistemic content of objects and events refers to action possibilities and goals. It is still unclear, however, what (brain and computational) mechanisms allow this. One idea, currently gaining consensus, is that goal-directed action cannot be supported by simple *stimulus* → *response* associations, but only by an *ideomotor* organization of action, which couples actions with their predicted effects; briefly, by *action* → *effect* ideomotor codes. Crucially, action-effect coupling can be used bidirectionally: actions can be used to predict sensory effects, and sensory effects to plan, trigger, and control goal-directed action. Support from this view comes from numerous psychological (Hommel *et al.*, 2001; Prinz, 1997) and neurophysiological (Kalaska *et al.*, 1997) studies.

To understand goal-directed action it is, however, necessary to further discuss how such action codes are organized, and how they are selected for execution. In this sense, the control-theoretic notion of *internal models* has been proposed (Kawato, 1999; Wolpert and Ghahramani, 2004), which, when applied to the study of living organisms, come in (at least) two varieties. The *inverse model* describes how the brain learns to control its actions, and the *forward model* describes how it predicts the sensory effects of such controls. Specifically, inverse models (or controllers) calculate the next motor command on the basis of action goals, actual and predicted stimuli; forward models (or predictors) calculate the next stimuli on the basis of an *efferent copy* of the motor command produced by the controller (von Holst and Mittelstaedt, 1950). Inverse and forward models are analogue to *effect* → *action* and *action* → *effect* ideomotor codes, respectively.

A complete agent architecture, which is able to produce a multitude of behaviours under different contextual conditions, can be composed of several couples of inverse and forward models, which we call here *schemas* (Arbib, 1981) (more on this below). Based on current action goals and contextual conditions, the 'best' schema(s) can be selected for execution and the commands of their controllers executed. Internal models have been studied in relation to action control (Doya, 1999; Kawato, 1999; Miall and Wolpert, 1996; Wolpert *et al.*, 1995), visuomotor control (Mehta and Schaal, 2002), imagery (Jeannerod, 2001; Kosslyn, 1994), stabilization of perception (Haarmeier *et al.*, 2001), and the realization of fast reaching movements (Desmurget and Grafton, 2000). Moreover, one key hypothesis bridging individual and social cognition is that internal models can be used to infer the goals of others in the same way in which they support the execution of one's own goals, thus opening the door to social abilities such as imitation and mindreading (Frith *et al.*, 2000; Wolpert *et al.*, 2003).

Overall, internal models produce an augmented action-perception loop: see Figure 1(b). Perceptual states do not directly influence action (as in the stimulus-response paradigm), but internal models play a mediating role. In this paper, we put forward the idea that this mediating role is *representational*, and in particular, that internal modelling produces embodied representations (Gallese and Metzinger, 2003; Grush, 2004; Jeannerod, 2001, 2006; Wolpert *et al.*, 2003). Here the idea is that by internally predicting or simulating actions and their effects we get an epistemic grasp over the external reality, which is internally represented in terms of action possibilities and action goals. In Gallese's (2000, p. 31) words: 'To observe objects is therefore equivalent to automatically evoking the most suitable motor program required to interact with them. Looking at objects means to unconsciously "simulate" a potential action. In other words, the object-representation is transiently integrated with the action-simulation (the ongoing simulation of the potential action).'

2.3.2 Internal Modelling Produces Embodied Representations. Figure 2 highlights the differences between a stimulus-response system (a) and one endowed with internal models (b). In the latter, the internal models realize an *inner loop*, which parallels actual sensorimotor interaction and mimics its input-output properties. Such loops can function on-line with action and off-line. Internal modelling provides the foundation of the *emulation theory of representation* proposed by Grush (2004, p. 1):

> . . . in addition to simply engaging with the body and environment, the brain constructs neural circuits that act as models of the body and environment. During overt sensorimotor engagement, these models are driven by efference copies in parallel with the body and environment, in order to provide expectations of the sensory feedback, and to enhance and process sensory information. These models can also be run off-line in order to produce imagery, estimate outcomes of different actions, and evaluate and develop motor plans.

Inner loops are extremely relevant for our analysis since they provide an *internalization* of the external reality; or, rather, of certain relevant characteristics

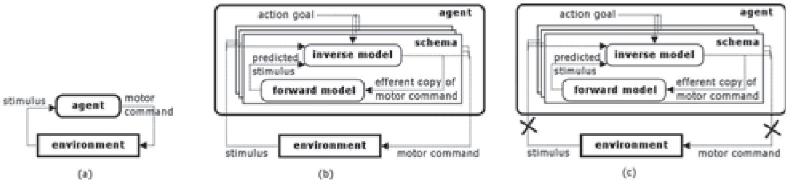

Figure 2 *Comparison between purely stimulus-response systems (b) and those endowed with anticipatory capabilities, which run an 'internal loop' on-line with action (b), or off-line (c)*

of the organism-environment engagement (see Section 3). Such internalization is genuinely representational, since it includes indications for future actions and at the same time affords internal manipulation prior to and independent of any actual action. In fact, not only is the inner loop active while the agent takes action, but it can also be *decoupled* from the original sensorimotor loop and run off-line, in simulation of potential action(s) to support motor planning, preparation, and imagery (Grush, 2004; Jeannerod, 2001, 2006; Pezzulo and Castelfranchi, 2007, 2009).

As shown in Figure 2(c), in our control-theoretic model, decoupling is realized by suppressing incoming sensory inputs and inhibiting outgoing motor commands, while at the same time allowing the inverse and forward models to continue generating motor commands and predictions. This *mental simulation* process, which reuses off-line the mechanisms of on-line prediction, was firstly studied by Jeannerod (1997) in the context of motor planning, when an action is first simulated off-line, then executed on-line. Other, related studies (Decety, 1996; Decety *et al.*, 1989; Decety and Grèzes, 2006) have shown that the timing of observed, imagined and executed actions is the same, leading to the conclusion that they share the same neural substrate. Similar results have been reported in classical studies on mental rotation of objects (Shepard and Metzler, 1971). The most likely conclusion is that the same action system permits action execution, observation, and off-line simulation, and so the same representational processes are in play. Therefore, by broadening the scope of his analysis, Jeannerod (2001, 2006) has argued that actions have the same content, a *motor representation*, in all their manifestations: when they are executed, planned, imagined, or observed, and that in all these cases it is the *mental simulation* process that produces the (same) motor representation. This means that motor representations are detachable and reusable in a number of cognitive processes.

Further evidence that action representations can be *decoupled* from the current sensorimotor context comes from studies of frontal and parietal lobe patients. Schwoebel *et al.* (2002) reports a patient with bilateral parietal lesions who was unable to refrain from executing imagined (hand) movements, and was unaware of these movements. Lhermitte (1983) first described *utilisation behaviour* in frontal lobe patients, or the failure to inhibit prepotent action processes (e.g. grasping) elicited by seen objects. These studies support the idea that mental simulation consists in an off-line reuse of the same representation involved in on-line motor execution; in addition, mental simulation may require an inhibitory mechanism, which is impaired in these patients.

2.3.3 Other Kinds of Forward Models.

Up to this point we have focused on forward models that embed *action* → *effect* relationships, but this is just one specific case. Schubotz (2007) has studied how the prediction of external events that we cannot reproduce ourselves (e.g. the movements of animals or objects, the listening of a melody) can be realized by using internal models that predict the exteroceptive or interoceptive changes these events produce on our own bodies instead of the effects of our own actions. By using such internal models, we can re-enact models of our body and its transformations to *simulate external events*, not limited to actions we can actually execute; for example, they permit perceptual predictions and simulations by modelling changes in the patterns of perceptual stimuli. Coupled with the capability to discriminate self-produced effects from the rest of sensory experience (which is based on prediction too; see Blakemore *et al.*, 1998; Jeannerod, 2003), internal models of external events extend our representational power from the realm of action-related concepts to an 'objective' understanding of the external reality.

2.3.4 Possible Limitations. The studies mentioned above indicate that mental simulations can enable several cognitive capabilities in the individual and social domains that are (or at least seem to be) representational in nature. However, the nature and limits of embodied representations and simulative processes remain unclear. In particular, consider the popular distinction between *procedural representation (or knowledge)*, 'how to do something', and *declarative representation* (or *knowledge*),[2] 'knowledge of what is true' (Anderson, 1993). One might argue that embodied representations are sufficient for explaining *procedural* forms of knowledge, but cannot support higher-level cognitive functions such as planning and reasoning, in which it is necessary to explicitly manipulate knowledge stored in a declarative format.

Indeed, contrary to the standard view of declarative knowledge as organized in symbolic structures like semantic networks (Norman, 1970; Quillian, 1969), embodied representations are realized in motor brain areas and are modal (or multimodal) and not amodal, and have a dynamic, 'transient' (non-enduring) nature. It remains to be explained whether this view accounts for conceptual forms of knowledge, and their flexible manipulation. A recent review by Barsalou (2008) reports empirical support for this hypothesis with a variety of techniques, from behavioural studies to brain imaging, but at the moment there is ample debate on this topic. This paper offers some philosophical arguments on how embodied representations can determine both procedural and declarative forms of knowledge.

3. Embodied Representations and their Representational Content

In this Section we further elaborate the idea introduced in Section 2 that inner loops are full-fledged (embodied) representational processes. This idea implies an account of 'intentionality' (Brentano, 1985) that differs from the sensorimotor perspective. According to the latter perspective, internal representations should be *grounded* bottom-up: they have to refer to the external reality and be constrained by physical

---

[2] The terms 'knowledge' and 'belief' are usually used in two senses: the former implicit (e.g. we act on the basis of implicit presuppositions) and the latter explicit (e.g. knowledge we use for reasoning and that we can report explicitly). Here we use 'belief' and 'knowledge' as synonyms.

interaction with it. More precisely, to do so they must refer to sensory states, which *cause* and permit discriminate among them[3]. For example, the representation *abc* refers to the sensorimotor state *xyz* if *xyz* causes *abc.* According to Harnad (2001, sec. 26): 'Symbols must be grounded in the capacity to discriminate and identify the objects, events and states of affairs that they stand for, from their sensory projections.'

This bottom-up form of intentionality is sufficient to support the *stimulus → response* structure of reactive action, but not the *ideomotor* structure of goal-directed action that is instead assumed in the action-based perspective. Internal representations, like sensory states, merely provide information about the perceptual situation but do not give any cue about action outcomes, which are necessary for steering goal-directed action. Therefore, another element is necessary: prediction, and in particular the kind of prediction of action effects that is realized through internal modelling. Figure 3 illustrates the different views of grounding of the sensorimotor and action-based approaches respectively. (a) The sensorimotor view individuates only a bottom-up pathway of environment-organism causation. (b) The action-based view is based on the *causal-predictive loop* realized by internal modelling. The causal part represents the fact that the representation *abc* is (totally or partially) caused by the sensorimotor state *xyz*. The predictive part represents the fact that one produces expectations about the effects of one's own actions (e.g. my sensorimotor state is *xyz* and I expect that, as a consequence of my action α, it will become *wxz*).

For goal-directed action, representations should encode patterns of (potential) actions and their outcomes; this is a form of *motor understanding*. In such a view, representations could be better conceived as *indications for action* required for an effective interaction instead of 'internal copies' of the external reality (Bickhard, 1993). Consequently, we argue, the primary representational content of knowledge about objects and events consists of a set of dispositions that entail numerous expectations. That is, representations are about (external) realities in virtue of the fact that they afford successful interaction with them and produce expectations of such interaction effects. In turn, the machinery behind such dispositions consists of action

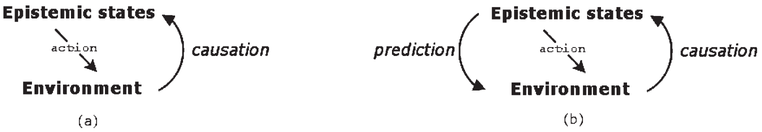

Figure 3 *(a) Environment-organism causation in sensorimotor and causal accounts; (b) Prediction of action's effects permits to 'close the circle' in action-based view and to realize a causal-predictive loop*

---

[3] The role of *world → representation* causation plays a key role in the informational semantic view of Dretske (1981) and the teleosemantic view of Millikan (2004), too, who however elaborate this idea in different directions. We do not enter in the details here (but see Grush, 1997 for a discussion on this topic).

schemas whose success or failure confirms their corresponding presuppositions and expectations and guarantees aboutness of representation. For example, the belief 'the cat is on the mat' consists in a set of dispositions, such as 'if I look at the mat I will see the cat', and 'if I move the mat I move the cat', whose expectations are confirmed by action success. However, there is more to the situation. Since action schemas include numerous control parameters, such as the colour and weight of a cat or a mat, those elements can be considered part of the content of the representation, too. This makes it possible to formulate the belief 'the black cat is on the mat'. As we will see in Section 4, it is often (if not always) the case that only part of such content is elicited or used in a given context of interaction. For example, when lifting an object size knowledge, but not colour knowledge, is typically elicited.

Overall, this is a pragmatist's account of knowledge and its mental content, based on the idea that any access to the world's verifiable structure is mediated by action, actual or potential, and knowledge verification criteria are pragmatic. At the same time, the coupling with action schemas explains the causal power of beliefs and knowledge without postulating dualism. One advantage of this view is that it offers a solution to the problem of the *normativity* of representations, since their correctness or incorrectness depend on pragmatic success or failure, and action success is also a reality check for their aboutness. Another problem easily accommodated within this view is *misrepresentation*: the problem of how one can think anything at all about something if one thinks something false about it (Grush, 1997). From a pragmatist's perspective, representational failure is not a failure of referring to the external reality, but simply a failure of picking up the best indications for actions. Such an error can eventually be corrected, as long as the interaction proceeds. Finally, a relevant implication of this view is that maintaining true representations is a pragmatic necessity of schemas and not a necessity dictated by cognitive needs. All schemas that are used in the same environmental conditions tend to have the same representational content, and no extra cognitive bias (e.g. maintaining coherence between schemas) is required. Indeed, living organisms are selected by evolution to produce good indications for action, since accuracy of their internal modelling processes is necessary for the success of their schemas, and ultimately for their survival (Bickhard, 1993; Pezzulo and Calvi, 2007).

3.1 Detached Representations

This is not the whole story however. To pass from a simple motor understanding to a fully representational, *epistemic* grasp over the external reality, representations must be *detachable* from the here-and-now so as to afford internal, endogenous (not stimuli-driven) regeneration and manipulation independent of the current sensorimotor flow. One key characteristic of representations is that they are *internally manipulable*, and can be used for attempting actions internally, before or instead of acting in the external reality, and in diverse goal and sensory contexts, i.e. even outside the context in which they were learned. This necessitates that representations be (temporarily) *detached* (or *decoupled*) from the current sensorimotor loop. At the same time, detached representations need to maintain their intentionality and grounding. How can our account of representation based on internal modelling satisfy the two apparently conflicting constraints of detachment and grounding?

Although they initially serve for on-line action control, inner loops can be *detached* from the current sensorimotor loop, making it possible to run them off-line; see Figure 2(c). By chaining multiple short-term predictions together, this mechanism allows *simulations* of possible goal-directed actions to be generated, or multiple scenarios to be imagined and compared without executing any actual action. This elicitation-without-actual-action gives an 'epistemic grasp' over the external environment that goes beyond mere observation since it embeds interactive elements: it gives reference to possible actions and their outcomes. A complementary aspect is that, unlike sensory states, predictions and simulations are not only automatically elicited by external events but can be *endogenously generated* when needed. This gives autonomy from the external environment, which is a necessary precondition for true cognition. For example, it makes it possible to set up and pursue goals which are not dictated by current affordances.

Importantly, mental simulations of actions are realized with a (partial) re-enactment of the motor processes involved in the same overt actions—a feature that distinguishes them from related ideas widespread in cognitive science such as *small scale models* (Craik, 1943) or *mental models* (Johnson-Laird, 1983). As discussed in (Pezzulo and Castelfranchi, 2007, 2009), internal modelling loops could have originated for the purpose of adaptive interaction of the organism with its environment and not for cognition. The process was then exapted for realizing increasingly sophisticated representation-based cognitive tasks. For example, predictive capability could have originated in the need to compensate for feedback delays (Desmurget and Grafton, 2000) and then became used for anticipating future relevant events. Similarly, the ability to run off-line simulations during action preparation (Jeannerod, 2001) or for forecasting future dangers (Damasio, 1994) could have been exapted to allow endogenous internal manipulations, such as reasoning and selecting distal goals, which are the hallmark of representation-based cognition. Although off-line inner loops have gained increased autonomy from the here-and-now, due to their origin as control mechanisms they maintain reference to (possible) situated action, and their basic functioning remains the same (e.g. they allow comparing predictions and perceptions). For this reason they remain *grounded* and constrained by the external environment.

3.1.1 What Counts as a Representation, and What Does Not? We have argued that representations are dynamical processes produced by the re-enactment of sensorimotor brain structures. However, this does not mean that all dynamical processes are representational. For example, central pattern generators (Grillner, 1981) and attractors in dynamical systems (Port and van Gelder, 1995) are not representations. These steering mechanisms cannot be *decoupled* from the current sensorimotor context and cannot be *used as* representations to stand in for something else, or internally manipulated instead of acting directly on the external environment.

Similarly, internal states such as retinal images, internal (hidden) units of neural networks, and endogenous regenerations of perceptual stimuli, although they (often) co-occur with external events, should be called 'information states' and not 'representations', unless they also afford detachment and internal processing independent of the most immediate action affordances.

4. An Action-Based View of Procedural and Declarative Knowledge

We have argued that inner loops, having the two characteristics of detachment and grounding, are prototypical *embodied representations*, grounded in the motor system but which nevertheless afford internal manipulation. The aim of this section is to clarify how such view accounts for procedural and declarative forms of knowledge.

In order to frame the discussion, however, it is necessary to draw a more complete picture of the functioning of inner loops for goal-directed action, and to extend the schema model introduced in Section 2 to account for more complex motor programs such as drinking a cup of wine. It is in fact implausible that such a motor program is realized by one single schema. It has been proposed that the brain maintains a *vocabulary* of actions that provide basic functionalities such as finding, reaching, grasping and lifting a cup, and that such actions can be recombined flexibly (Fadiga and Craighero, 2004; Rizzolatti *et al.*, 1988). In a schema-based view, we can treat each basic action as a set of schemas that can realize the action's goal under different contexts. Figure 4 illustrates a sample schema-based computational model for drinking a cup of wine which is composed of two sets of perceptual schemas (top) and five sets of motor schemas (bottom); (see Fagg and Arbib, 1998; Pezzulo, 2009; Pezzulo and Calvi, 2006; Roy, 2005 for related schema-based models). Sets of schemas are required since their parameters can be specialized for different sensory contexts (e.g. lifting full or empty cups, grasping big or small cups, etc.). Both perceptual and motor schemas can be conceptualized as coupled inverse and forward models, as in Figure 2(b); the difference is in the inputs they receive (as a rough approximation, visual for perceptual schemas, and visual, tactile and kinaesthetic for motor schemas) and the effectors they govern (as a rough approximation, head and eye movements for perceptual schemas, and arm and hand movements for motor schemas). As shown in the figure, each schema can be dissected into simpler schemas; for example, grasping involves grip preshaping, hand rotation and hand closure. For the sake of simplicity here we focus on coarse-grained schemas.[4]

One important aspect of this computational architecture is that schemas compete for selection, the reason being that access to sensors and effectors as well as to the

---

[4] It is widely accepted that action control structures are arranged in a hierarchical fashion, with more abstract and effector-independent schemas at the top, and motor programs for muscle movements at the bottom of the hierarchy (Hamilton and Grafton, 2007). For the sake of simplicity here we do not consider this aspect (but see Haruno *et al.*, 2003).

processing capabilities of the brain is limited. For example, it is impossible to look in two different directions at the same time, or to drink and throw the same cup of wine at the same time. For this reason, only a few schemas get the chance to be selected. The higher their *activity level*, the higher the chance of selection. In such computational models, activity level accounts for both the 'achievability' (i.e. how much the schema is expected to be successful given the current context) and the 'desirability' (i.e. how much the schema's goal is desirable given the agent's current internal motivations) dimensions. To account for the 'achievability' dimension, schemas are activated in two ways: (1) when their forward models generate accurate sensory predictions (compared to actual stimuli), and (2) when other schemas create the appropriate contextual conditions and activate them (dotted lines in Figure 4). This mechanism should ensure that active schemas are closely tied to the current context. Biological systems have other sources of activation, too, that account for the 'desirability' dimension, most notably their motivations. For example, thirst and the desire to hit somebody bias the selection of drinking versus throwing schemas. For the sake of simplicity we can assume here that one action goal (drinking) is active, and the problem is how to select the right sequence of schemas to realize it.

Another key aspect is that active schemas can input other schemas, and so transfer information to them (solid lines in Figure 4). For example, perceptual schemas can provide information about cup location, size and orientation to appropriate motor schemas. This aspect is important in regulating schema sequences, but it also has epistemic effects, as we will see.

4.1 Procedural Knowledge is Implicitly Produced when Acting

The goal-directed structure of schemas makes it possible to reverse the classical *perception* → *representation* → *action* pipeline scheme. Since actions produce expectations, knowledge is not only a presupposition of action, but can be discovered *by* acting, too. For example, lifting a cup gives weight information and affords its categorization it as 'full' or 'empty'. The process through which this occurs involves the formulation of expectations based on prior knowledge and their verification through action. In Bayesian terms, this is a *generative* process which goes from *priors* to *posteriors* (Friston, 2003; Wolpert and Ghahramani, 2004).

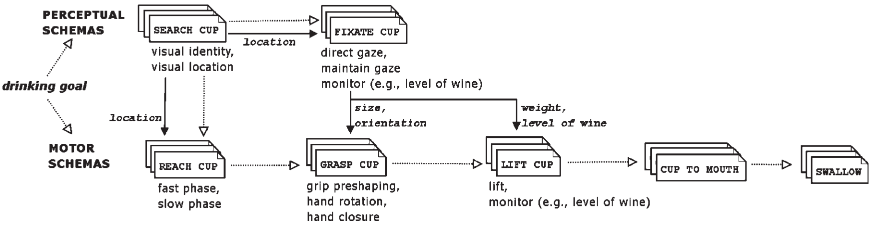

Figure 4 *Drinking a cup of wine: a schema-level view*

Suppose that someone provided with the above computational/neural architecture wants to drink a cup of wine (we focus on the lifting part). Suppose that the cup looks empty (or one is told it is). Such visual (or linguistic) information may lead to the selection of the appropriate internal models, which govern the lifting of full versus empty cups, and which, in turn, will predict different sensory weight effects on the acting body parts.[5] This *prior* (visual) information can be continuously updated and integrated based on new evidence. One source of evidence is action success or failure: a successful lifting implicitly verifies the internal model's presuppositions and supports prior weight information. A failed lifting, instead, disconfirms it. If the action involves multiple steps, prediction accuracy can be used as a further source of evidence in order to gradually update information. This mechanism works as follows: all schemas (including those which are not associated to the currently running motor programs) are allowed to predict sensory feedback and are assigned more or less access to the effectors, depending on their prediction accuracy. As a result, not only does the action schema that predicts better allow the appropriate lift, but also the selected forward models disambiguate the context and provide an embodied representation of the object's weight (the *posterior*), by predicting the sensory effects of lifting the object. Overall, the success of a schema indicates that the specific (actual or expected) states of affairs, encoded in its assumptions and expectations, are true.

Moreover, the success of a grasping schema implicitly indicates that a cup is in touch and ready to be lifted, and provides extra information about its weight, size, etc. Not only is such information essential for the functioning of single schemas, but also the whole schema structure supporting goal-directed action requires it. In some cases this information can be passed between schemas. For example, the *fixate cup* schema can pass information about the level of wine to motor schemas, which can therefore avoid spilling. This process supports division of labour and permits costly operations, such as context estimation, or extraction of visual information, to be executed only once (Vetter and Wolpert, 2000). However, an explicit encoding or passage of information is not always necessary, as some information can simply be used in the schema structure as an *indication for action*. For example, the success of the *grasp cup* motor schema creates the appropriate environmental conditions for the next schemas to be executed, and implicitly signals constraining conditions (e.g. the fact that the cup is in touch). Interactivism (Bickhard, 1993) suggests a similar perspective: if an active interaction succeeds, it implicitly determines knowledge that can be used to steer further interaction; if it fails, then the indication for action (and the content of the representation) is false. Suppose, for instance, that the failure of a grasping action derives from the object lying out of reach. Although this

---

[5] Experiments performed by Tucker and Ellis (2004) have shown that appropriate motor programs for interacting with objects are indeed automatically selected on the basis of visual cues, such as an object's handle. A related finding is the fact that 'canonical neurons' in monkey F5 area neurons appear to code action plans in relation to seen objects (Rizzolatti *et al*., 1988).

information is not technically 'inside' the failing schema, nevertheless it is revealed by the failure, and we can *act on its basis* (for example, decide to move toward the object to be grasped). Overall, the main characteristics of procedural knowledge are its dynamic and transient nature, since it is produced only when necessary, and is grounded and situated in the current sensorimotor context. However, under certain conditions, part of said knowledge can be made explicit, in order to be used outside the original sensorimotor context.

4.2 Accessing Procedural Knowledge via Simulation and Making it Explicit

Like procedural knowledge, declarative knowledge is not necessarily mediated by (amodal) symbols, but can be produced *on the fly* by eliciting the 'accessible' subset of procedural information stored in internal models (we call it *tacit information*). In other terms, the internal models used in motor control, motor planning, or the prediction of external events, are also available to some extent across other sensorimotor or cognitive systems, and can be intentionally re-enacted so that their representational content can be retrieved, giving access to the same kind of knowledge used for taking action.

Consider again the case of lifting a cup. Although a seemingly less direct application, an *off-line simulation* of lifting the same cup can be used, too, for accessing action-related information; the schema's tacit knowledge is made available in the form of sensory effects predicted by the internal model, if the object were actually to be lifted. This information can be available to other processes, for example, for answering the question: 'Is your red cup heavier of lighter than your blue cup?', or 'Can you reach the red cup?'. In the first example, the visual images of the two cups activate internal models for grasping, which will predict different sensory weight effects, thus making this information accessible (see Hamilton *et al.*, 2007 for a discussion of motor simulation in perceptual weight judgments). In the second example, a simulation involving the 'reaching' schema(s) can give access to the 'reachability' dimension. A recent illustration of this situation comes from a study by Coello *et al.* (2008): participants were asked to judge (perceptually) whether a visual target was reachable or not with their right hand. Significantly, the subjects' performance was facilitated (especially for stimuli located at the boundary of peripersonal space) when their motor cortex was stimulated using TMS. This study shows that mental simulations elicited by visual stimuli not only prepare the body to act, but also reveal relevant information, for example about the feasibility of potential actions. In this way, motor representations participate in the visual determination of what is reachable. More generally, consider that motor simulation can involve multiple schemas: for reaching, grasping, locating the cup in the visual scene, and so on. When combined, they can give access to multiple categories, such as 'weight', 'colour', and 'distance', by eliciting the appropriate motor processes and parameters. Critically, declarative knowledge produced in this way (*that* the cup is red, or big) has the same *content* as its corresponding procedural knowledge.

Anyway, unlike actual action much of the extra information that is obtained via action failures is not available when an action is simulated, and so knowledge obtained by internal simulation can be less reliable.

This mechanism is surprisingly versatile and gives access to knowledge that is much harder, if not impossible, to obtain with different means. Consider the huge amount of 'knowledge in the fingers' that is required to play the piano or to use a keyboard. In most cases, the easiest way to access information, such as what finger to use to press the 'Ctrl' button on your keyboard, is running an internal simulation of that very action, which counts as a truly *epistemic action*—that is, an action that is not conducted for the sake of producing a pragmatic effect but for the sake of obtaining knowledge. Part of the power of this mechanism consists in the coupling of inverse and forward models, which have preconditions and effects that are implicitly verified by action success; this process can be 'inverted' and exploited for inferring what action to perform (either actually or 'in simulation') to obtain knowledge about preconditions or expectations; for example, what motor command to execute to produce a given key pressure, or what action to do to decide if a cup is heavy or light.

Up to this point we have provided examples of on-line and off-line motor (hand) control, but the same mechanism can be used for steering visual simulations, as well, since 'looking' is merely another action that can be executed and simulated. When one is asked, 'What is the colour of your kitchen furniture?', 'How many chairs are there in your apartment?', or, 'Where is the book I gave you yesterday?', one can scan the visual scene in one's imagination by re-enacting the motor processes usually involved in overt perception, including those governing eye movements. Since the demands of visual processing are the same as any other form of motor control (e.g. overcoming delays in feedback), it is plausible that similar solutions based on internal modelling could be implemented in the brain. Indeed, recent research reveals the presence of internal models for moving visual targets in the cerebellum (Cerminara *et al.*, 2009), which might support visual simulations. Numerous experiments with eye tracking methods indicate that visual scans are invoked during cognitive tasks in which visual scenes are elicited. In one experiment reported in (Spivey *et al.*, 2004), when participants are asked to imagine a perceptual situation (e.g. people in different floors inside a building), they do several saccades in directions in which the described objects are more likely to be in reality (e.g. upward or downward). Similarly, Estes *et al.* (2008) report that reading object words (i.e. 'plane', 'soccer ball') activates typical gaze directions (up/down). In our view, visual simulations, which determine eye movements even when there is nothing to look at in the external environment (Spivey and Geng, 2001), could give access to task information encoded in internal models (in addition, they could implement embodied problem solving strategies; see Section 5).

Note that the neural correlates of visual simulations of the spatial environment and of visuomotor control (e.g. lifting cups) could be different. Based on empirical evidence, a distinction has been proposed between two visual systems: a dorsal, action-related pathway, which uses an egocentric frame of reference and a ventral, perception-related pathway, which usually uses an allocentric frame of reference and whose content is available to consciousness[6] (Milner and Goodale, 1995, 2008). The existence of multiple neural pathways for vision is compatible with our claims as long as they are both supported by internal models which can be re-enacted in simulation, as in the examples we discussed above. Indeed, it is possible that, when processing different types of sensory information, the visual system exploits different kinds of schemas, which determine different representations in different frames of references, and support different forms of simulation (of hand-cup dynamics, of visual scans of objects in the scene, and so on).[7]

Most of our examples of simulations involve single actions. However, complex sequences of actions can be simulated, too. For example, one can simulate the whole path home to recall whether or not a given shop is on the route. In terms of our computational model, such simulations re-enact whole portions of a schema structure (see the example in Figure 4).

Finally, it is worth noting that our account is not restricted to declarative knowledge having obvious links to action; conceptual and linguistic knowledge can be addressed as well. Recent research in the embodied aspects of language has shown that language processing activates motor areas in the brain (Pulvermüller, 1999). Although general frameworks have been proposed that attempt to ground almost all language processing in embodied action (Glenberg and Gallese, submitted; Simmons *et al*., 2008), most empirical research has focused on sentences that have explicit action components, such as 'open the drawer' (Glenberg and Kaschak, 2002); in these domains it has been found a significant interference of motor and linguistic representations. However, we note that communication is another form of action (composed of *speech acts*, see Austin, 1962), which could be supported by appropriate internal models allowing communicative intentions to be achieved. Such models predict the effect of one's own utterances, or the understanding of another's

---

[6] Note that there is an open discussion on frames of reference. Recent evidence from multiple research groups challenges the bold egocentric/allocentric distinction (Pisella *et al*., 2006), and numerous studies indicate goal-directed frames of reference in the superior temporal sulcus (Jellema and Perrett, 2006) and the mirror neuron system (di Pellegrino *et al*., 1992). See also Briscoe, 2009 for the alternative hypothesis that both pathways use an egocentric frame of reference.

[7] Why should different sets of internal models for vision be needed? Here we agree with Milner and Goodale (1995, 2008) that, in evolutionary terms, this could reflect different computational demands: on-line action control for the dorsal stream, and distal goal-directed action for the ventral stream. This could have determined different encodings in the two visual pathways, and possibly different frames of reference, so that ventral vision specialized to provide a visual experience of the world that can be used for planning actions, or stored for future reference, whereas dorsal vision specialized to provide the direct foundation for on-line action control. Our disagreement with Milner and Goodale (1995, 2008) is about the nature of semantic representation. These authors endorse the classical view that visual experience and visual (semantic) representation consist in *internalized percepts* obtained via the ventral stream, while we argue that they consist in *internalized perception* with the reenactment of internal models; see Section 5.

communicative intentions, so we can apply to language and communication the same theoretical framework that we have proposed for action control. For example, when one is asked: 'What is the capital city of Jamaica?' or 'What is the past tense of "to read"?' the answers *Kingston* and *read* could be obtained through an internal simulation process which recruits internal models used in communication or sentence processing. A theoretical architecture for speech perception based on internal modelling has recently been proposed by Pickering and Garrod (2007) that could support the kind of mental simulations we envision here. In addition, it could be the case that internal models supporting goal-directed action control and goal-directed communication are interrelated, the former providing the foundation for the latter. All these ideas are however highly speculative and still require empirical confirmation.

4.2.1 Why is Re-Enactment Necessary for Obtaining Declarative Knowledge from Internal Models? A fundamental aspect of our view of declarative knowledge is that it derives from procedural knowledge encoded in internal models. For example, knowledge of action-outcome contingencies is encoded in internal forward models. Internal models are not 'transparent' data structures (such as 'lookup tables' commonly used in computer science) where information, for example the effects of a given action, can be searched directly. They do not support the retrieval of information by simply using search keys as with databases with key-value data structures. There is no way to know, for example, the outcome (e.g. sensory effects) of actions by simply searching a lookup table with 'action α' as the key. The internal model has to be re-enacted in simulation to provide the action outcome.[8]

4.2.2 What Procedural Information is Accessible via Simulation, and What is Not? One critical issue to assess is the level of accessibility of procedural knowledge, and what mental simulations we can actually perform.[9] Recent empirical evidence indicates that the mechanism we have described of *access via simulation* is not limited to providing knowledge about the effects of possible actions (Davidson and Wolpert, 2005), but also informs one's idiosyncratic performance (Knoblich and Flach, 2003) and one's own mental states (Goldman, 2006; Metzinger, 2003); others' mental states (Gallese *et al.*, 1996; Gallese and Goldman, 1998; Wilson

---

[8] A caveat: it is certainly possible that simpler, associative mechanisms permit storage and retrieval of declarative knowledge (e.g. of action effects) 'by rote' from semantic or episodic memory, in a process which resembles a search in a lookup table. For example, most habitants of Jamaica could use such associations for answering questions about their capital city. However, we argue that this is not the standard method for dealing with declarative knowledge, especially outside the linguistic domain. The first reason is that such associative structures are inflexible, while most knowledge needs to be situated to be used effectively. The second reason is parsimony: most of this knowledge is already present in internal models.

[9] Although related, this is not the same problem of conscious awareness, since most simulations are automatic and not conscious; see Pacherie, 2007 on this topic.

and Knoblich, 2005), expectations (Bosbach *et al*., 2005), emotions and aesthetic experiences (Freedberg and Gallese, 2007). This mechanism is not limited to the realm of intentional agents, but the motor apparatus is also involved in understanding external events and the dynamics of the surrounding environment (Schubotz, 2007), such as the movement of objects. At the same time, recent research has focused on how to scale up the same approach to abstract concepts such as 'grasping an idea' (Gallese and Lakoff, 2005), 'truth', 'freedom' and 'invention' (Barsalou and Wiemer-Hastings, 2005).

All these examples show that the mechanism of *access via simulation* could be a widespread method for accessing and producing knowledge, and represents a valid alternative to the traditional idea of 'storage and retrieval'. Indeed, at least part of the knowledge required to act, or to make decisions, does not have to be retrieved from memory, but can be produced on-line by reenacting appropriate schemas. As Barsalou (1999) clearly argues, the advantage of this mechanism is that information obtained by simulation is always *situated* and contextually appropriate (see also Rosenbaum and Dawson, 2004).

Note that like representation, accessibility of procedural information could have emerged for practical reasons rather than for cognition. It has been proposed that one of the main roles of accessibility is recognizing the effects of one's own actions as self-generated (Blakemore *et al*., 1998; Jeannerod, 2003), and deficits of accessibility may have dramatic results such as delusions of control (Frith *et al*., 2000). Once equipped with mechanisms for accessing its procedural knowledge, any organism can use them to generate declarative knowledge 'on the fly' when needed.

4.2.3 How Specific and 'Deep' are Simulations? Another aspect to clarify is how specific simulations are. Does simulated grasping recruit the same motor processes involved in actual grasping, or does it just activate the motor system in a more general way? In Section 2.3 we mentioned studies that indicate that (in some cases) the timing of executed and mental actions is the same. These studies highlight the presence of fine-grained simulations. This is not always the case, however. For example, when planning our route home, the duration of the action is not replicated in our simulation, nor do we take into account the fine-grained details such as the number of steps, which can be filled in on-line during action performance (see Hommel *et al*., 2001, on the division of labour between off-line planning and on-line action). The same happens when the objective is inferring or imitating someone else's goal, and (in the majority of cases) it is unnecessary to replicate exactly the observed movements. This leads to the conclusion that mental simulations could be a versatile mechanism whose level of specificity is adaptable. Since action is represented in the brain as a motor hierarchy (Hamilton and Grafton, 2007), it is plausible that re-enactment can be more or less 'deep', and involve schemas at different levels of the motor hierarchy: from quite abstract, effector-independent action schemas (Schmidt, 1975) down to muscle activation (Fadiga *et al*., 2005) depending on the demands of the task.

4.3 Procedural versus Declarative Revisited

At this point, we can reformulate the difference between procedural and declarative knowledge as related to embodied simulations. Contrary to traditional accounts, in our view both depend on non-symbolic representational processes based on anticipation. The differences between them are their *levels of access* and their *use*. Procedural knowledge is only accessible and used on-line by the process that generates it (e.g. the inner loop for the control of an arm). By contrast, declarative knowledge is elicited by off-line simulation of action, is used outside the original sensorimotor context and can support a wider range of actions and decisions (more on this below). This distinction could reflect the different demands and constraints of situated action and reasoning: the former could be too fast to permit explicit knowledge to form and be processed outside the action context (see also Jeannerod, 2006 on this topic).

Note that this account of declarative knowledge is incompatible with the traditional view of a global organization of (long-term) declarative memory composed of more or less static memory traces completely decoupled from action and perception, e.g. semantic networks (Quillian, 1969; Norman, 1970). The functioning of our schemas and internal models gives rise to multiple local procedural memories, which can become 'active' when necessary. In addition, our models suggest that there is no clear distinction between the two traditional kinds (or 'stores') of declarative memory, semantic and episodic (Tulving, 1972); note that throughout the discussion we have provided examples of both semantic and episodic content accessed by action simulation in a similar way. Although a complete review is beyond the scope of this paper, it is worth mentioning that recent theories of episodic and working memory (Glenberg, 1997), embodied working memory (Spivey *et al*., 2004), (episodic) mental time travel (Suddendorf and Corballis, 2007), embodied simulations (Barsalou, 1999), and a 'fractionated' central executive (Baddeley, 2002), all are investigating the issues of how sensorimotor and memory process are related.

5. Internal Manipulation of Embodied Representations

In most cases, retrieval of declarative knowledge via mental simulations is just the first step. A complete theory should also explain how such knowledge can be flexibly manipulated. Throughout the paper we have provided examples of mental simulations that not only elicit knowledge, but also realize internal manipulations and support planning and reasoning; for example, in Section 1 we have provided examples of a mechanic and an interior designer who use mental simulations in their work, and of mental simulations for evaluating and selecting among multiple possibilities for action. We admit that we do not have (yet) evidence that such manipulations have the same degree of productivity and systematicity as symbolic systems (Fodor, 1975); nonetheless, here we offer some tentative thoughts on this topic.

We have argued that mental simulations are covert forms of action and perception that re-enact action schemas off-line. To be successful, this process has to respect the dynamics of agent–environment engagement, at least at a given level of abstraction (see above on the discussion of how 'deep' are simulations). This means that, unlike traditional AI reasoning systems (Newell and Simon, 1972), mental simulations are not arbitrary mental operations, but are constrained by situated action possibilities and have the same 'grammatical' structure of action; that is, they are a form of *internalized situated action* (Campbell, 1999; Ito, 2008). Within these limits, however, they afford sophisticated cognitive operations in the individual and social domains. Some examples mentioned earlier include planning, understanding others' intentions, and imitation, but in principle any operation that can be performed in reality and requires internal modelling can be internally simulated. One example is constituted by visual simulations, which we consider to be a form of *internalized perception* with the same characteristics as overt perception (except its overt execution), and supporting similar operations. A recent illustration of this idea comes from a study by Wohlschlaeger (2001), in which mental object rotation affected actual rotational hand movements, leading to the conclusion that mental object rotation is a simulated action involving motor representations. In more ecological contexts, one can imagine how one would look (and feel) with a new hairstyle and, if the result is pleasant, go to the hairdresser.

Motor processes have been shown to play a role in problem solving, too. For instance, we can rotate puzzle pieces to facilitate our understanding of where to put them (Kirsh and Maglio, 1994), and use eye movements to 'compute' the relationships (e.g. 'close to' or 'far from') between objects (Pylyshyn, 2001). We argue that one can use mental (e.g. visual) simulations in a similar way, for example for mentally solving a puzzle, composing a melody, or planning holidays with a travel map (or even by imagining it). If this is true, mental simulations should have the same timing and recruit the same mechanisms as overt actions in the same environmental conditions. Indeed, in addition to evidence on mental rotation, several studies of spatial reasoning tasks reveal that the patterns of eye movements are the same when the tasks are executed overtly and in one's own mind (Grant and Spivey, 2003). This suggests that mental simulations, in addition to giving access to procedural information, can produce novel knowledge and support an embodied form of problem solving. Another intriguing, speculative possibility is that mental simulations, which are less tied to the external environment than overt actions, and can in principle revolve around one's own mental processes, could have paved the way to a conceptualization of abstract situations in terms of concrete ones by recruiting motor strategies that work well in the latter to operate in the former, including, for example, reasoning about abstract relations by reusing visual strategies that capture spatial relations in the external environment. In a similar vein, Barsalou (1999, p. 601) argues that some abstract concepts could be grounded in introspective processes, and offers an analysis of the concept of 'true' as the successful 'matching' of internal simulations with perceived events. If this hypothesis is accurate, then the neural realization of 'changing perspective on something' or 'grasping an idea' could be more literal than hitherto believed.

5.1 Detachment of Representations and Compositionality of Simulations

Another crucial issue to assess is whether or not mental simulations afford compositionality and the use of elicited knowledge outside its original context. As already discussed, one key feature of representations is that they can be *detached* from any actual action to regulate our behaviour in an open-ended way. Take as an example the piece of knowledge that the red cup in front of me is big. This is implicitly used by action schemas for hand preshaping when grasping—its prototypical context of use. It can also be elicited with off-line simulations, for example when planning to grasp the red cup before actually grasping it—again, in the same context. However, there are daily examples of mental simulations in which the same knowledge is used for taking action in diverse goal and sensory contexts, too: for example, for deciding to buy small or large gift box, or for using the cup to store pencils. Here the situation is different, since the simulation that gives access to such knowledge and the action to be taken belong to two different contexts of use. In terms of our computational architecture, the set of (grasping and lifting) schemas that elicit size information is not the same as that used in the successive choice (what to buy) or action (putting pencils in it). Another example is using size information to respond to questions such as, 'is the cup light or heavy?'; here size information is elicited in a sensorimotor context and then used in a communicative context. To afford an open-ended use of representations in diverse contexts, and compositionality of simulations (i.e. simulations that use knowledge elicited by another simulation to produce novel knowledge), knowledge elicited by one simulation has to be broadly available across other simulations. Up to the current moment, however, it has been unclear what information is 'penetrable' in this way. An experiment performed by Chang *et al.* (2008) has shown that in forming grasping forces a hand can pass the 'veridical' weight information of an object to another hand, and numerous studies have revealed that conscious motor images are so detailed that they can be used, for example, by athletes for their training (see Jeannerod, 2006). In other cases motor competence can be cognitively impenetrable. For instance, while our motor interactions with falling objects are very accurate, we are inaccurate in perceptually judging and explicitly reporting Newtonian dynamics (Zago and Lacquaniti, 2005). In addition, it is unclear what the mechanism behind information transfer might be. One recent hypothesis is that this could be done via a *broadcast* mechanism, making the content of a (limited) 'global workspace' available to several processes and schemas at once (Baars, 2002). This mechanism could make the results of a mental simulation available to other processes, and afford compositionality and flexible forms of thinking, not unlike symbolic manipulations, still maintaining a grounding of thinking in situated action.

6. Conclusions

Recent evidence in cognitive psychology and neuroscience indicates that anticipatory brain mechanisms (based on enactment or re-enactment of motor brain areas) are involved in (and sufficient for) several higher-level cognitive abilities in the individual and social domains. In this paper we have discussed one epistemological implication of the action-based view: representation derives from—and is grounded in—situated action and its anticipation. Unlike traditional symbolic/linguistic theories, we have argued that several cognitive functions can be realized without any 'recoding' of modal into amodal ones such as physical symbols (Newell and Simon, 1972). Our claim is that anticipatory and simulative processes that originated from the control of goal-directed actions are the first form of embodied representations. In other words, representing consists of anticipatory and simulative processes: emulations of an action's effects, the environment, the others, etc. This happens both on-line (during action), thanks to anticipatory mechanisms, and off-line, thanks to their re-enactment and chaining, which produces long-term simulations. Procedural and declarative knowledge is produced in the first and second cases respectively. More precisely, procedural knowledge is entrained by action, which implicitly checks presuppositions and expectations; indeed, action success is the ultimate proof that the representational content of our presuppositions and expectations was appropriate. Declarative knowledge can be created on the fly with action simulations, which determine a re-enactment of the same sensorimotor structures that govern actual action execution and elicit the same tacit information. Such declarative knowledge can be flexibly manipulated, for example for planning and reasoning in diverse contexts of use, without any recoding. This process, which can be described as *simulated situated action*, is very versatile, although (probably) not as much as the use of linguistic symbols.

To conclude this section, it is worth noting that throughout the paper we have emphasized the individual dimension of action and knowledge. However, as discussed earlier, action-based theories posit a strong link between individual and social cognition, which could be supported by a common neural substrate; for example, motor activation can be induced socially, by simply observing others acting. Since we argued that motor representations are a product of the mechanisms that govern goal-directed action, the fact that these mechanisms are also active in social cognition implies that part of our knowledge, procedural and declarative, has social origins and is *interaction-oriented* rather than merely *action-oriented*.

6.1 A Coda. Steps of Internalization: From Anticipation to Language

In both evolutionary and developmental terms, the representational content of motor acts and perceptual processes has a sensorimotor grounding which is intimately related to (and used for) action itself. However, this is not the whole story about representation. A complete theory must also understand how 'action-oriented representation' grounds language or symbolic capabilities of increasing abstraction and how these in turn serve to enrich the more basic perceptual and motor schemas.

Although we have emphasized embodied representations we do not deny that there are other, more complex forms of representation, or other sources of knowledge that differ from (actual or simulated) direct experience. For example, Sloman and Chappell (2005) like us refuse a purely bottom-up, causal notion of grounding and propose a form of 'structural grounding' called *symbol tethering* wherein knowledge may be derived from internal, coherence-based constraints instead of by (real or simulated) action. Moreover, most of our knowledge has a social origin. Humans can know that a door is open because they are told, and they form beliefs about non-directly accessible parts of the environment (abstract concepts, theoretical entities, past history) thanks to culture-mediated practices such as linguistic communication and writing. It is possible that part of this knowledge is still grounded in the sensorimotor experience of others, thanks to a kind of division of labour. When we accept the fact that the door is open, we rely on the fact that the verifiable content of such a proposition is accessible and verifiable (at least in principle) by some of our or of others' actions. As we have suggested above, however, sociality could have favoured the emergence of other forms of knowledge, including certain abstract concepts, that are grounded in social (communicative) actions rather than, or in addition to, situated action (see e.g. Borghi and Cimatti, 2009). This idea is based on the hypothesis, sketched in Section 4.2 that communication is made of speech acts, and exploits the same mechanisms of situated actions, and namely internal modelling.

This thesis contrasts with another, more radical view: that representation and knowledge are *essentially* cultural and linguistic. For instance, Vygotsky (1978) proposes that epistemic structures are first produced by agents in their environments as a support to their situated action, and then *internalized* in the agent's minds. We certainly agree that tools and other cultural artefacts have greatly enhanced cognition (Clark and Chalmers, 1998; Tomasello *et al*., 1993), and that language—the ultimate human artefact—is an aid to thinking in several possible ways: it helps in shaping (individually or culturally) and memorizing knowledge and categories, as well as in acquiring and transmitting novel concepts, and allows a powerful form of symbolic reasoning, i.e. *reasoning with words*. Here, however, the pertinent question is *when* does representation begin: before or after language? Are language (and sociality) responsible for representational capabilities *tout court* (and not only for the emergence of advanced symbolic capabilities, a view that we accept)? We argue that *representation does not begin with language* and that linguistic/symbolic representations are not the first 'interesting' form of representation. It is certainly true that language has enhanced symbolic capabilities and is necessary for some cognitive tasks in humans; in addition, language is a powerful stimulus for eliciting mental simulations that in turn make a lot of tacit knowledge declarative. Nonetheless, we, along with many others (e.g. Dehaene and Cohen, 2007), have provided reasons for believing that language develops on top of an extant representational (brain) substrate. It follows, then, from this discussion that our view of 'internalization' is wider than Vygotsky's (1978). Indeed, biological organisms internalize their external environment in several ways; for instance, reverberant neural architectures tend to self-generate stimuli that could resemble external stimuli. However, we consider internal modelling to be the first form of *representational internalization*, which produces then two relevant precursors of linguistic/symbolic representations: (1) implicit knowledge realized by praxis, and (2) explicit knowledge accessed via simulation. Internalization by language or other forms of cultural transmission is the next step, and language is an especially powerful tool for formulating complex assertions (e.g. negations, counterfactuals), exploiting and recombining them productively, and for producing persistent signs.[10] However, the basic epistemic content of assertions is still there in embodied representations.

A related issue concerns how linguistically-mediated and symbolic representation and intelligence could have developed in evolutionary continuity. A recent research program (Arbib, 2005; Rizzolatti and Arbib, 1998) focuses on how language can be described in continuity with simpler forms of interaction, suggesting that area F5 of the monkey's brain (where mirror neurons are found) is a precursor of Broca's area (devoted to language and speech processing), and that language could have inherited the 'grammatical' structure of action. These studies have just scratched the surface of this topic however. We believe that the whole process of how representation first derived from simple sensorimotor interaction and then evolved to sophisticated, linguistically and culturally mediated forms, is a long and complex story, that can be analyzed in terms of a process of *internalization* that proceeds from the simplest forms—first of all anticipation and internal simulations of potential actions, with their simple 'grammatical' structure—toward more complex ones, possibly linguistically and culturally mediated. This paper has focused on the first part of this process, with the aim of shedding light on how representation-mediated cognition could have worked *before* language.

We believe that this perspective is especially important in current cognitive science work, in order to reconfigure the traditional debate between defenders of symbolic representations and anti-representationalists. It is equally important for the realization of future artificial cognitive systems, which—in our opinion—should focus on *embodied representations* as their central construct instead of mere sensorimotor processes or symbolic manipulations.

Istituto di Linguistica Computazionale 'Antonio Zampolli' - CNR, Via Giuseppe Moruzzi, 1 - 56124 Pisa, Italy; and
Istituto di Scienze e Tecnologie della Cognizione - CNR, Via S. Martino della Battaglia, 44 - 00185 Rome, Italy.

References

Anderson, J. 1993: *Rules of the Mind.* Lawrence Erlbaum Associates, Hillsdale NJ.

Arbib, M. 1981: Perceptual structures and distributed motor control. In V. Brooks (ed.), *Handbook of Physiology—The Nervous System II. Motor Control.* Bethesda, MD: American Physiological Society, 1449–80.

---

[10] This is not to say that all cognition is about internalization. External scaffolding and the off-load of cognitive work onto the environment are important as well (see e.g. Clark, 1998; Wilson, 2002).

Arbib, M. 2002: The mirror system, imitation, and the evolution of language. In K. Dautenhahn and C. Nehaniv (eds), *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press.

Arbib, M. 2005: From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28, 105–21.

Austin, J. L. 1962: *How to do Things with Words*. Oxford: Oxford University Press.

Baars, B. J. 2002: The conscious access hypothesis: origins and recent evidence. *Trends in Cognitive Sciences*, 8, 47–52.

Baddeley, A. D. 2002: Fractionating the central executive. In D. Stuss and R. T. Knight (eds), *Principles of Frontal Lobe Function*. Oxford: Oxford University Press, 246–60.

Ballard, D. 1991: Animate vision. *Artificial Intelligence*, 48, 1–27.

Barsalou, L. 1999: Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–600.

Barsalou, L. 2003: Situated simulation in the human conceptual system. *Language and Cognitive Processes*, 18, 513–62.

Barsalou, L. 2008: Grounded cognition. *Annual Review of Psychology*, 59, 617–645.

Barsalou, L. and Wiemer-Hastings, K. 2005: Situating abstract concepts. In D. Pecher and R. Zwaan (eds), *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thought*. Cambridge: Cambridge University Press, 129–63.

Beer, R. 1997: The dynamics of adaptive behavior: a research program. *Robotics and Autonomous Systems*, 20, 257–89.

Bickhard, M. H. 1993: Representational content in humans and machines. *Journal of Experimental and Theoretical Artificial Intelligence*, 5, 285–333.

Blakemore, S-J., Wolpert, D. M. and Frith, C. D. 1998: Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1 (7), 635–40.

Borghi, A. and Cimatti, F. 2009: *Words as tools and the problem of abstract words meanings. Proceedings of the 31st Annual Conference of Cognitive Science Society*. Amsterdam: Cognitive Science Society, 2304–09.

Bosbach, S., Cole, J., Prinz, W. and Knoblich, G. 2005: Inferring another's expectation from action: the role of peripheral sensation. *Nature Neuroscience*, 8, 1295–97.

Brentano, F. 1985: *Psychology from an Empirical Standpoint*, trans. A. C. Rancurello, D. B. Terrell and L. L. McAlister. London: Routledge.

Briscoe, R. 2009: Egocentric spatial representation in action and perception. *Philosophy and Phenomenological Research*, 79 (2), 423–60.

Brooks, R. A. 1991: Intelligence without representation. *Artificial Intelligence*, 47, 139–59.

Campbell, J. 1999: Schizophrenia, the space of reasons, and thinking as a motor process. *The Monist*, 84 (4), 609–25.

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W. F. and Thulborn, K. R. 1999: Time course of fMRI-activation in language and spatial networks during sentence comprehension. *NeuroImage*, 10, 216–24.

Cerminara, N. L., Apps, R. and Marple-Horvat, D. E. 2009: An internal model of a moving visual target in the lateral cerebellum. *Journal of Physiology*, 587 (Pt 2), 429–42.

Chang, E. C., Flanagan, J. R. and Goodale, M. A. 2008: The intermanual transfer of anticipatory force control in precision grip lifting is not influenced by the perception of weight. *Experimental Brain Research*, 185, 319–29.

Churchland, P. 1981: Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, 78 (2), 67–90.

Clark, A. 1998: *Being There. Putting Brain, Body, and World Together.* Cambridge MA: MIT Press.

Clark, A. and Chalmers, D. J. 1998: The extended mind. *Analysis*, 58, 10–23.

Coello, Y., Bartolo, A., Amiri, B., Devanne, H., Houdayer, E. and Derambure, P. 2008: Perceiving what is reachable depends on motor representations: Evidence from a transcranial magnetic stimulation study. *PLoS ONE*, 3 (8), e2862.

Craik, K. 1943: *The Nature of Explanation.* Cambridge: Cambridge University Press.

Damasio, A. R. 1994: *Descartes' Error: Emotion, Reason and the Human Brain.* New York: Grosset/Putnam.

Davidson, P. and Wolpert, D. 2005: Widespread access to predictive models in the motor system: A short review. *Journal of Neural Engineering*, 2, S313–S319.

Decety, J. 1996: Do imagined and executed actions share the same neural substrate? *Cognitive Brain Research*, 3, 87–93.

Decety, J. and Grèzes, J. 2006: The power of simulation: imagining one's own and other's behavior. *Cognitive Brain Research*, 1079, 4–14.

Decety, J., Jeannerod, M. and Prablanc, C. 1989: The timing of mentally represented actions. *Behavioral Brain Research*, 34, 35–42.

Dehaene, S. and Cohen, L. 2007: Cultural recycling of cortical maps. *Neuron*, 56, 384–98.

Desmurget, M. and Grafton, S. 2000: Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Science*, 4, 423–31.

Dewey, J. 1896: The reflex arc concept in psychology. *Psychological Review*, 3, 357–70.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. and Rizzolatti, G. 1992: Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91, 176–80.

Doya, K. 1999: What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks*, 12, 961–74.

Dretske, F. 1981: *Knowledge and the flow of information.* Cambridge MA: MIT Press.

Estes, Z., Verges, M. and Barsalou, L. W. 2008: Head up, foot down: object words orient attention to the objects typical location. *Psychological Science*, 19, 93–7.

Fadiga, L. and Craighero, L. 2004: Electrophysiology of action representation. *Journal of Clinical Neurophysiology*, 21, 157–69.

Fadiga, L., Craighero, L. and Olivier, E. 2005: Human motor cortex excitability during the perception of others' action. *Current Opinion in Neurobiology*, 15, 213–218.

Fagg, A. and Arbib, M. 1998: Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks*, 11, 1277–1303.

Fodor, J. 1975: *The Language of Thought.* Cambridge, MA: Harvard University Press.

Fogassi, L., Ferrari, P., Chersi, F., Gesierich, B., Rozzi, S. and Rizzolatti, G. 2005: Parietal lobe: From action organization to intention understanding. *Science*, 308, 662–67.

Freedberg, D. and Gallese, V. 2007: Motion, emotion and empathy in esthetic experience. *Trends in Cognitive Science*, 11 (5), 197–203.

Friston, K. 2003: Learning and inference in the brain. *Neural Networks*, 16(9), 1325–52.

Frith, C. D. 2007: *Making up the Mind. How the Brain Creates our Mental World.* Oxford: Blackwell.

Frith, C. D., Blakemore, S. J. and Wolpert, D. M. 2000: Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society B Biological Sciences*, 355 (1404), 1771–88.

Gallese, V. 2000: The inner sense of action: agency and motor representations. *Journal of Consciousness Studies*, 7, 23–40.

Gallese, V. 2003: A neuroscientific grasp of concepts: from control to representation. *Philosophical Transactions of the Royal Society B Biological Sciences*, 358, 1231–40.

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

Gallese, V. and Goldman, A. 1998: Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2 (12), 493–501.

Gallese, V. and Lakoff, G. 2005: The brains concepts: the role of the sensory-motor system in reason and language. *Cognitive Neuropsychology*, 22, 455–79.

Gallese, V. and Metzinger, T. 2003: Motor ontology: the representational reality of goals, actions, and selves. *Philosophical Psychology*, 13, 365–88.

Gibson, J. 1966: *The Senses Considered as Perceptual Systems.* Boston, MA: Houghton Mifflin.

Gibson, J. 1979: *The Ecological Approach to Visual Perception.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Glenberg, A. M. 1997: What memory is for. *Behavioral and Brain Sciences*, 20, 1–55.

Glenberg, A. M. and Gallese, V. submitted: Action-based language: a theory of language acquisition, comprehension, and production.

Glenberg, A. M. and Kaschak, M. P. 2002: Grounding language in action. *Psychonomic Bulletin & Review*, 9, 558–65.

Goldman, A. 2006: *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading.* Oxford: Oxford University Press.

Grant, E. R. and Spivey, M. J. 2003: Eye movements and problem solving: guiding attention guides thought. *Psychological Science*, 14, 462–6.

Grillner, S. 1981: Control of locomotion in bipeds, tetrapods, and fish. In V. Brooks, (ed.), *Handbook of Physiology. Section* 1. *The nervous system II Motor control*, 3rd edn. Bethesda, MD: American Physiological Society, 1179–1236.

Grush, R. 1997: The architecture of representation. *Philosophical Psychology*, 10, 5–25.

Grush, R. 2004: The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27, 377–96.

Haarmeier, T., Bunjes, F., Lindner, A., Berret, E. and Thier, P. 2001: Optimizing visual motion perception during eye movements. *Neuron*, 32, 527535.

Hamilton, A. F. de C. and Grafton, S. T. 2007: The motor hierarchy: from kinematics to goals and intentions. In P. Haggard, Y. Rossetti and M. Kawato (eds), *Sensorimotor Foundations of Higher Cognition.* Oxford: Oxford University Press.

Hamilton, A. F. de C., Joyce, D. W., Flanagan, J. R., Frith, C. D. and Wolpert, D. M. 2007: Kinematic cues in perceptual weight judgement and their origins in box lifting. *Psychological Research*, 71, 13–21.

Harnad, S. 1990: The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42, 335–46.

Harnad, S. 2001: Grounding symbols in the analog world with neural nets—a hybrid model. *Psycoloquy*, 12 (034).

Haruno, M., Wolpert, D. and Kawato, M. 2003: Hierarchical mosaic for movement generation. In T. Ono, G. Matsumoto, R. Llinas, A. Berthoz, H. Norgren and R. Tamura (eds), *Excepta Medica International Coungress Series.* Amsterdam: Elsevier Science, 575–90.

von Holst, E. and Mittelstaedt, H. 1950: The reafference principle: interaction between the central nervous system and the periphery. *Naturwissenschaften*, 37, 464–76.

Hommel, B., Musseler, J., Aschersleben, G. and Prinz, W. 2001: The theory of event coding (tec): a framework for perception and action planning. *Behavioral and Brain Sciences*, 24, 849–78.

Iacoboni, M. 2003: Understanding others: imitation, language, empathy. In S. Hurley, and N. Chater (eds), *Perspectives on Imitation: From Cognitive Neuroscience to Social Science.* Cambridge, MA: MIT Press.

Ito, M. 2008: Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9, 304–13.

Jeannerod, M. 1997: *The Cognitive Neuroscience of Action.* Oxford: Blackwell.

Jeannerod, M. 2001: Neural simulation of action: a unifying mechanism for motor cognition. *NeuroImage*, 14, S103–S109.

Jeannerod, M. 2003: The mechanism of self-recognition in humans. *Behavioral Brain Research*, 142, 1–15.

Jeannerod, M. 2006: *Motor Cognition.* Oxford: Oxford University Press.

Jellema, T. and Perrett, D. I. 2006: Neural representations of perceived bodily actions using a categorical frame of reference. *Neuropsychologia*, 44 (9), 1535–46.

Johnson-Laird, P. 1983: *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness.* Cambridge: Cambridge University Press and Cambridge, MA: Harvard University Press.

Kalaska, J. F., Scott, S. H., Cisek, P., and Sergio, L. E. 1997: Cortical control of reaching movements. *Current Opinion in Neurobiology*, 7, 849–59.

Karmiloff-Smith, A. 1992: *Beyond Modularity: A Developmental Perspective on Cognitive Science.* Cambridge, MA: MIT Press/Bradford.

Kawato, M. 1999: Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9, 718–27.

Kirsh, D. and Maglio, P. 1994: On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18, 513–49.

Knoblich, G. and Flach, R. 2003: Action identity: evidence from self-recognition, prediction, and coordination. *Consciousness and Cognition*, 12, 620–32.

Kosslyn, S. 1994: *Image and Brain: The Resolution of the Imagery Debate.* Cambridge, MA: MIT Press.

Lhermitte, F. 1983: 'Utilization behaviour' and its relation to lesions of the frontal lobes. *Brain: A Journal of Neurology*, 106 (Pt 2), 237–55.

McCarthy, J. and Hayes, P. J. 1969: Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer and D. Michie (eds), *Machine Intelligence*, vol. 4. Edinburgh: Edinburgh University Press, 463–502.

Mahon, B. and Caramazza, A. 2005: The orchestration of the sensory-motor systems: clues from neuropsychology. *Cognitive Neuropsychology*, 22, 480–94.

Markman, A. B. and Dietrich, E. 2000: Extending the classical view of representation. *Trends in Cognitive Sciences*, 4 (12), 470–75.

Martin, A. 2007: The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25–45.

Maturana, H. R. and Varela, F. J. 1980: *Autopoiesis and Cognition: The Realization of Living*. Dordrecht: D. Reidel.

Mehta, B. and Schaal, S. 2002: Forward models in visuomotor control. *Journal of Neurophysiology*, 88, 942–53.

Metzinger, T. 2003: *Being No One. The Self-Model Theory of Subjectivity.* Cambridge, MA: MIT Press.

Miall, R. C. and Wolpert, D. M. 1996: Forward models for physiological motor control. *Neural Networks*, 9 (8), 1265–79.

Millikan, R. G. 2004: *Varieties Of Meaning.* Cambridge, MA: MIT Press.

Milner, A. D. and Goodale, M. A. 1995: *The Visual Brain in Action.* Oxford: Oxford University Press.

Milner, A. D. and Goodale, M. A. 2008: Two visual systems re-viewed. *Neuropsychologia*, 46, 774–85.

Newell, A. 1990: *Unified Theories of Cognition.* Cambridge, MA: Harvard University Press.

Newell, A. and Simon, H. A. 1972: *Human Problem Solving.* Englewood Cliffs, NJ: Prentice-Hall.

Nolfi, S. 2005: Behaviour as a complex adaptive system: on the role of self-organization in the development of individual and collective behaviour. *ComPlexUs*, 2, 195–203.

Norman, D. (ed.) 1970: *Models of Human Memory.* New York: Academic Press.

O'Regan, J. and Noë, A. 2001: A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24, 883–917.

Pacherie, E. 2007: The sense of control and the sense of agency. *Psyche*, 13, 1–30.

Pezzulo, G. 2008: Coordinating with the future: the anticipatory nature of representation. *Minds and Machines*, 18, 179–225.

Pezzulo, G. 2009: DiPRA: a layered agent architecture which integrates practical reasoning and sensorimotor schemas. *Connection Science*, 21 (4), 297–326.

Pezzulo, G. and Calvi, G. 2006: A schema based model of the praying mantis. In S. Nolfi, G. Baldassarre, R. Calabretta, J. Hallam, D. Marocco, O. Miglino, J-A Meyer, D. and Parisi, (eds), From animals to animats 9*, Proceedings of the Ninth International Conference on Simulation of Adaptive Behaviour*, vol. LNAI 4095, Berlin: Springer Verlag, 211–23.

Pezzulo, G. and Calvi, G. 2007: Schema-based design and the akira schema language: an overview. In M. Butz, O. Sigaud, G. Pezzulo and G. Baldassarre, (eds), *Anticipatory Behavior in Adaptive Learning Systems: Advances in Anticipatory Processing*, volume LNAI 4520. Berlin: Springer Verlag.

Pezzulo, G. and Castelfranchi, C. 2007: The symbol detachment problem. *Cognitive Processing*, 8 (2), 115–31.

Pezzulo, G. and Castelfranchi, C. 2009: Thinking as the control of imagination: a conceptual framework for goal-directed systems. *Psychological Research*, 73, 559–77.

Pickering, M. J. and Garrod, S. 2007: Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11 (3), 105–10.

Pisella, L., Binkofski, F., Lasek, K., Toni, I. and Rossetti, Y. 2006: No double-dissociation between optic ataxia and visual agnosia: Multiple sub-streams for multiple visuo-manual integrations. *Neuropsychologia*, 44, 2734–48.

Port, R. and van Gelder, T. 1995: *Mind as Motion: Explorations in the Dynamics of Cognition.* Cambridge MA: MIT Press.

Prinz, W. 1997: Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129–54.

Pulvermüller, F. 1999: Words in the brains language. *Behavioral & Brain Sciences*, 22, 253–336.

Pylyshyn, Z. W. 2001: Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80, 127–58.

Quillian, R. 1969: Semantic memory. In M. Minsky (ed.), *Semantic Information Processing*. Cambridge MA: MIT Press.

Rizzolatti, G. and Arbib, M. A. 1998: Language within our grasp. *Trends in Neurosciences*, 21 (5), 188–94.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G. and Matelli, M. 1988: Functional organization of inferior area 6 in the macaque monkey. ii. area f5 and the control of distal movements. *Experimental Brain Research*, 71, 491–507.

Rizzolatti, G. and Craighero, L. 2004: The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–92.

Rizzolatti, G., Fadiga, L., Gallese, V. and Fogassi, L. 1996: Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 131–41.

Rizzolatti, G., Fogassi, L. and Gallese, V. 2001: Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2, 661–70.

Rosenbaum, D. and Dawson, A. 2004: The motor system computes well but remembers poorly. *Journal of Motor Behavior*, 36, 390–92.

Rosenbaum, D., Halloran, E. S. and Cohen, R. G. 2006: Grasping movement plans. *Psychonomic Bulletin and Review*, 13 (5), 918–22.

Rosenbaum, D., Marchak, F., Barnes, H., Vaughan, J., Slotta, J. and Jorgensen, M. 1990: Constraints for action selection: overhand versus underhand grips. In M. Jeannerod (ed.), *Attention and performance XIII. Motor representation and control.* Hillsdale, NJ: Lawrence Erlbaum, 211–65.

Roy, D. 2005: Semiotic schemas: a framework for grounding language in action and perception. *Artificial Intelligence*, 167, 170–205.

Schmidt, R. A. 1975: A schema theory of discrete motor skill learning. *Psychological Review*, 82, 225–60.

Schubotz, R. I. 2007: Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, 11 (5), 211–18.

Schwoebel, J., Boronat, C. B. and Coslett, H. B. 2002: The man who executed 'imagined' movements: evidence for dissociable components of the body schema. *Brain Cognition*, 50, 1–16.

Scorolli, C., Borghi, A. M. and Glenberg, A. 2009: Language-induced motor activity in bi-manual object lifting. *Experimental Brain Research*, 193, 43–53.

Shepard, R. N. and Metzler, J. 1971: Mental rotation of three-dimensional objects. *Science*, 171, 701–03.

Simmons, W. K., Hamann, S. B., Harenski, C. L., Hu, X. P. and Barsalou, L. W. 2008: fMRI evidence for word association and situated simulation in conceptual processing. *Journal of Physiology* (Paris), 102, 106–19.

Sloman, A. and Chappell, J. 2005: The altricial-precocial spectrum for robots. *Proceedings of IJCAI*. Edinburgh: Professional Book Center, 1187–92.

Spivey, M. J. and Geng, J. J. 2001: Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research*, 65 (4), 235–41.

Spivey, M., Richardson, D. and Fitneva, S. 2004: Thinking outside the brain: spatial indices to linguistic and visual information. In J. Henderson and F. Ferreira (eds), *The Interface of Vision Language and Action*. New York: Psychology Press.

Suddendorf, T. and Corballis, M. C. 2007: The evolution of foresight: what is mental time travel and is it unique to humans? *Behavioral and Brain Sciences*, 30, 299–313.

Tomasello, M., Kruger, A. and Ratner, H. 1993: Cultural learning. *Behavioral and Brain Sciences*, 16, 495–552.

Tucker, M. and Ellis, R. 2004: Action priming by briefly presented objects. *Acta Psychologica*, 116, 185–203.

Tulving, E. 1972: Episodic and semantic memory. In E. Tulving and W. Donaldson (eds), *Organization of Memory*. New York: Academic Press.

Vetter, P. and Wolpert, D. M. 2000: Context estimation for sensorimotor control. *Journal of Neurophysiology*, 84, 1026–34.

Vygotsky, L. S. 1978: Mind in society*: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Wilson, M. 2002: Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9, 625–36.

Wilson, M. and Knoblich, G. 2005: The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131, 460–73.

Wohlschlaeger, A. 2001: Mental object rotation and the planning of hand movements. *Perception and Psychophysics*, 63, 709–18.

Wohlschlaeger, A., Gattis, M. and Bekkering, H. 2003: Action generation and action perception in imitation: an instance of the ideomotor principle. *Philosophical Transactions of the Royal Society B Biological Sciences*, 358, 501–15.

Wolpert, D. M., Doya, K. and Kawato, M. 2003: A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society B Biological Sciences*, 358, 593–602.

Wolpert, D. M. and Ghahramani, Z. 2004: Computational motor control. *Science*, 269, 1880–82.

Wolpert, D. M., Ghahramani, Z. and Jordan, M. 1995: An internal model for sensorimotor integration. *Science*, 269, 1179–82.

Zago, M. and Lacquaniti, F. 2005: Cognitive, perceptual and action-oriented representations of falling objects. *Neuropsychologia*, 43, 178–88.