

# Prospection in cognition: the case for joint episodic-procedural memory in cognitive robotics

David Vernon<sup>1\*</sup>, Michael Beetz<sup>2</sup> and Giulio Sandini<sup>3</sup>

<sup>1</sup> Interaction Lab, School of Informatics, University of Skövde, Skövde, Sweden, <sup>2</sup> Institute for Artificial Intelligence, University of Bremen, Bremen, Germany, <sup>3</sup> Department of Robotics, Brain and Cognitive Sciences, Istituto Italiano di Tecnologia, Genova, Italy

## OPEN ACCESS

### Edited by:

Guido Schillaci,  
Humboldt University of Berlin,  
Germany

### Reviewed by:

John Nassour,  
Technische Universität Chemnitz,  
Germany  
Arnaud Blanchard,  
University of Cergy-Pontoise, France  
Alessandro Di Nuovo,  
Plymouth University, UK

### \*Correspondence:

David Vernon,  
School of Informatics, University of  
Skövde, P.O. Box 408, Skövde  
SE-54128, Sweden  
david.vernon@his.se

### Specialty section:

This article was submitted to  
Humanoid Robotics, a section of the  
journal *Frontiers in Robotics and AI*

**Received:** 03 March 2015

**Accepted:** 06 July 2015

**Published:** 23 July 2015

### Citation:

Vernon D, Beetz M and Sandini G  
(2015) Prospection in cognition: the  
case for joint episodic-procedural  
memory in cognitive robotics.  
*Front. Robot. AI* 2:19.  
doi: 10.3389/frobt.2015.00019

Prospection lies at the core of cognition: it is the means by which an agent – a person or a cognitive robot – shifts its perspective from immediate sensory experience to anticipate future events, be they the actions of other agents or the outcome of its own actions. Prospection, accomplished by internal simulation, requires mechanisms for both perceptual imagery and motor imagery. While it is known that these two forms of imagery are tightly entwined in the mirror neuron system, we do not yet have an effective model of the mentalizing network which would provide a framework to integrate declarative episodic and procedural memory systems and to combine experiential knowledge with skillful know-how. Such a framework would be founded on joint perceptuo-motor representations. In this paper, we examine the case for this form of representation, contrasting sensory-motor theory with ideo-motor theory, and we discuss how such a framework could be realized by joint episodic-procedural memory. We argue that such a representation framework has several advantages for cognitive robots. Since episodic memory operates by recombining imperfectly recalled past experience, this allows it to simulate new or unexpected events. Furthermore, by virtue of its associative nature, joint episodic-procedural memory allows the internal simulation to be conditioned by current context, semantic memory, and the agent's value system. Context and semantics constrain the combinatorial explosion of potential perception-action associations and allow effective action selection in the pursuit of goals, while the value system provides the motives that underpin the agent's autonomy and cognitive development. This joint episodic-procedural memory framework is neutral regarding the final implementation of these episodic and procedural memories, which can be configured sub-symbolically as associative networks or symbolically as content-addressable image databases and databases of motor-control scripts.

**Keywords:** autonomy, cognitive system, development, episodic memory, ideo-motor theory, internal simulation, procedural memory, prospection

## Introduction

The goal of this article is to argue the case of the use of joint episodic memory to facilitate prospection and goal-directed action in cognitive robotics. The article begins with insights from the biological sciences regarding the prospective nature of action, leading to a discussion of the role of memory in prospection, and the realization of prospection through internal simulation. This sets the scene

for the introduction of ideo-motor theory, *vis-à-vis* sensory-motor theory, and an explanation of the importance of joint perceptuo-motor representations. This is then followed by two examples of how these principles have been applied in cognitive architectures and an argument in favor of explicit perceptuo-motor memory – joint episodic-procedural memory – over perceptuo-motor mappings. We finish with a description of a simple proof-of-principle example implementation of joint episodic-procedural memory for overt attention.

## The Goal-Directed and Prospective Nature of Action

Evidence from many different fields of research, including psychology and neuroscience, suggests that the movements of biological organisms are organized as actions and not reactions (von Hofsten, 2004). While reactions are elicited by earlier events, actions are initiated by a motivated subject, they are defined by goals, and they are guided by prospective information (Vernon et al., 2010). For example, when performing manipulation tasks or observing someone else performing them, subjects fixate on the goals and sub-goals of the movements not on the body parts, e.g., the hands or the objects (Johansson et al., 2001; Flanagan and Johansson, 2003). This happens only if a goal-directed action is implied. When showing the same movements without the context of an agent, subjects fixate the moving object instead of the goal.

Evidence from neuroscience also shows that the brain represents movements in terms of actions even at the level of neural processes [see Vernon et al. (2010), Chapter 4]. For example, the primate brain has two areas devoted to controlling movements: the premotor cortex and the motor cortex. The premotor cortex is the area of the brain that is active during motor planning and it influences the motor cortex which then executes the motor program comprising an action. The premotor cortex receives strong visual inputs from a region in the brain known as the inferior parietal lobule. These inputs serve a series visuomotor transformations for reaching (Area F4) and grasping (Area F5). Single neuron studies have shown that most F5 neurons code for specific goal-directed actions, rather than their constituent movements. Furthermore, several F5 neurons, in addition to their motor properties, respond also to visual stimuli. These are referred to as visuomotor neurons. The significance of this is that the premotor cortex of primates encodes actions (including implicit goals and expected states) and not just movements. The goal, therefore, is the fundamental property of the action rather than the specific motoric details of how it is achieved.

In primates, two classes of visuomotor neurons can be distinguished within area F5: canonical neurons and mirror neurons (Rizzolatti and Fadiga, 1998). The activity of both canonical and mirror neurons correlates with two distinct circumstances. In the case of canonical neurons, the same canonical neuron fires when a monkey sees a particular object and also when the monkey actually grasps an object with the same characteristic features. On the other hand, mirror neurons (Gallese et al., 1996; Rizzolatti et al., 1996; Rizzolatti and Craighero, 2004) are activated both when an action is performed and when the same or similar action is observed being performed by another agent. These neurons are

specific to the goal of the action and not the mechanics of carrying it out. So, for example, a monkey observing another monkey, or a human, reaching for a nut will cause mirror neurons in the premotor cortex to fire; these are the same neurons that fire when the monkey actually reaches for a nut. However, if the monkey observes another monkey making exactly the same movements but there is no nut present – there is no apparent goal of the reaching action – then the mirror neurons do not fire. Similarly, different motions that comprise the same goal-directed action will cause the same mirror neurons to fire. It is the action that matters: mirror neurons are not active if there is no explicit or implied goal. Since goals focus on the future, not the present, this again demonstrates the importance of prospection in action.

Finally, there is another reason why actions are guided by prospective information as opposed to instantaneous feedback data. Often, events in the agent's world may precede the feedback signals about them because the delays in the control pathways of biological systems may be substantial. If you cannot rely on feedback, the only way to overcome the problem is to anticipate what is going to happen next and to use that information to control one's behavior.

Prospection, then, is central to cognition. The question is how this prospection is achieved. The answer is, somewhat surprisingly, memory.

## Memory

Memory facilitates the persistence of knowledge and forms a reservoir of experience. Without it, it would be impossible for the system to learn, develop, adapt, recognize, plan, deliberate, and reason (Vernon, 2014). Memory functions to preserve what has been achieved through learning and development, ensuring that, when a cognitive system adapts to new circumstances, it does not lose its ability to act effectively in situations to which it had adapted previously. But memory has another role in addition to preserving past experience: to anticipate the future. It forms the basis for one of the central pillars of cognitive capacity, i.e., the ability to simulate internally the outcomes of possible actions and select the one that seems most appropriate for the current situation. Viewed in this light, memory can be seen as a mechanism that allows a cognitive agent to *prepare to act*, overcoming through anticipation the inherent “here-and-now” limitations of its perceptual capabilities.

We can distinguish memory in many ways (Squire, 2004; Wood et al., 2012). For example, it can be distinguished on the basis of the nature of what is remembered and the type of access we have to it. Specifically, memory can be categorized as either *declarative* or *procedural*, depending on whether it captures knowledge of things – facts – or actions – skills. Sometimes they are characterized as memory of knowledge and know-how: “knowing that” and “knowing how.”<sup>1</sup> This distinction applies mainly to long-term memory but short-term memory too has a declarative aspect. Declarative memory is sometimes referred to as *propositional memory* because it refers to information about the agent's world

<sup>1</sup>The distinction between *knowing that* and *knowing how* was made in 1949 by Gilbert Ryle in his book *The Concept of Mind* (Ryle, 1949)

that can be expressed in the form of propositions. This is significant because propositions are either true or false. Thus, declarative memory typically deals with factual information. This is not the case with skill-oriented procedural memory. As a consequence, declarative memories, in the form of knowledge, can be communicated from one agent to another through language, for example, whereas procedural memories can only be demonstrated.

Two different types of declarative memory can be distinguished. These are *episodic memory* and *semantic memory*. Episodic memory (Tulving, 1972, 1984) plays a key role in cognition and in the anticipatory aspect of cognition in particular. It refers to specific instances in the agent's experience while semantic memory refers to general knowledge about the agent's world which may be independent of the agent's specific experiences. In this sense, episodic memory is autobiographical. By its very nature in encapsulating some specific event in the agent's experience, episodic memory has an explicit spatial and temporal context: what happened, where it happened, and when it happened. This temporal sequencing is the only element of structure in episodic memory. Episodic memory is a fundamentally *constructive* process (Seligman et al., 2013). Each time an event is assimilated into episodic memory, past episodes are reconstructed. However, they are reconstructed a little differently each time. This constructive characteristic is related to the role that episodic memory plays in the process of internal simulation that forms the basis of prospection, the key anticipatory function of cognition.

In contrast, semantic memory "is the memory necessary for the use of language. It is a mental thesaurus, organized knowledge a person possesses about words and other verbal symbols, their meaning and referents, about relations among them, and about rules, formulas, and algorithms for the manipulation of the symbols, concepts, and relations."<sup>2</sup>

Episodic memory and semantic memory differ in many ways. In general, semantic memory is associated with how we understand (or model) the world around us, using facts, ideas, and concepts. On the other hand, episodic memory is closely associated with experience: perceptions and sensory stimulus. While episodic memory has no structure other than its temporal sequencing, semantic memory is highly structured to reflect the relationships between constituent concepts, ideas, and facts. Also, the validity (or truth, since semantic memory is a subset of propositional declarative memory) of semantic memories is based on social agreement rather than personal belief, as it is with episodic memory.<sup>3</sup> Semantic memory can be derived from episodic memory through a process of generalization and consolidation. Episodic memory can be both short-term and long-term while semantic memory and procedural memory are long-term.

## Memory and Prospection

Memory plays at least four roles in cognition: it allows us to remember past events, anticipate future ones, imagine the

viewpoint of other people, and navigate around our world. All four involve *self-projection*: the ability of an agent to shift perspective from itself in the here-and-now and to take an alternative perspective. It does this by *internal simulation*, i.e., the mental construction of an imagined alternative perspective (Schacter et al., 2008). Thus, there are four forms of internal simulation (Buckner and Carroll, 2007):

1. Episodic memory (remembering the past).
2. Navigation (orienting yourself topographically, i.e., in relation to your surroundings).
3. Theory of mind (taking someone else's perspective on matters).
4. Prospection (anticipating possible future events).

Each form of simulation has a different orientation (past, present, or future) and each refers to the perspective of either the first person, i.e., the agent itself, or another person.

Prospection – the mental simulation of future possibilities – plays a central role in organizing perception, cognition, affect, memory, motivation, and action (Seligman et al., 2013). Prospection is referred to in various ways, e.g., *episodic future thinking*, *memory of the future*, *pre-experiencing*, *proscopic chronesthesia*, *mental time travel*, and just plain *imagination* and it can involve conceptual content and affective – emotional – states (Buckner and Carroll, 2007).

Recent evidence suggests that all four kinds of internal simulation involve a single core brain network and this network overlaps what is known as the *default-mode network*, a set of interconnected regions in the brain that is active when the agent is not occupied with some attentional tasks (Østby et al., 2012).

It is significant that all four forms of simulation are constructive, i.e., they involve a form of imagination. There is a difference between knowing about the future and projecting ourselves into the future. The latter is experiential and the former is not. Thus, episodic memory (memory of experiences) and semantic memory (memory of facts) facilitate different types of prospection. Episodic memory allows you to re-experience your past and pre-experience your future. There is evidence that projecting yourself forward in time is important when you form a goal, creating a mental image of yourself acting out the event and then episodically pre-experiencing the unfolding of a plan to achieve that goal. This use of episodic memory in prospection is referred to as *episodic future thinking*, a term coined by Cristina Atance and Daniela O'Neill to refer to the ability to project oneself forward in time to pre-experience an event (Atance and O'Neill, 2001; Szpunar, 2010).

The constructive aspect of episodic memory, whereby old episodic memories are reconstructed slightly differently every time a new episodic memory is assimilated or remembered, is particularly important in the context of internal simulation of events that have not been previously experienced. While episodic memory certainly needs some constructive capacity to assemble individual details into a coherent memory of a given episode, the *constructive episodic simulation hypothesis* (Schacter and Addis, 2007a,b; Schacter et al., 2008; Szpunar, 2010) suggests that its role in prospection involving the simulation of multiple possible futures imposes an even greater need for a constructive capacity because of the need to extrapolate beyond past experiences. In

<sup>2</sup>This quotation explaining the characteristics of semantic memory appears in Endel Tulving's 1972 article (Tulving, 1972), p. 386 and is quoted in his *Précis* (Tulving, 1984). While this definition of semantic memory dates from 1972, it is still valid today. It also explains the linguistic origins of the term.

<sup>3</sup>Semantic memory and episodic memory can be contrasted in many other ways: twenty-seven differences are listed in Tulving (1983), p. 35.

other words, simulating multiple yet-to-be-experienced futures requires flexibility in episodic memory. This flexibility is possible because episodic memory is not an exact and perfect record of experience but one that conveys the essence of an event and is open to re-combination.

It is also significant that when humans imagine the future, they not only anticipate an event, but they also anticipate how they feel about that event. These are referred to as *hedonic* consequences of the event: whether we feel good about it or bad about it, whether it is associated with pleasure or pain, and lack of concern or fear. Thus, the pre-experience of prospection also involves “pre-feeling.” The brain accomplishes prospection by simulating the event and the associated hedonic experience (Gilbert and Wilson, 2007). While pre-feeling is not always reliable because contextual factors also play a part in the hedonic experience, this hedonic aspect of episodic memory is important because it reflects the affective nature of cognition and opens up a plausible way to factor emotional drives and value systems into the operation of memory, prospection, and action selection.

## Internal Simulation and Action

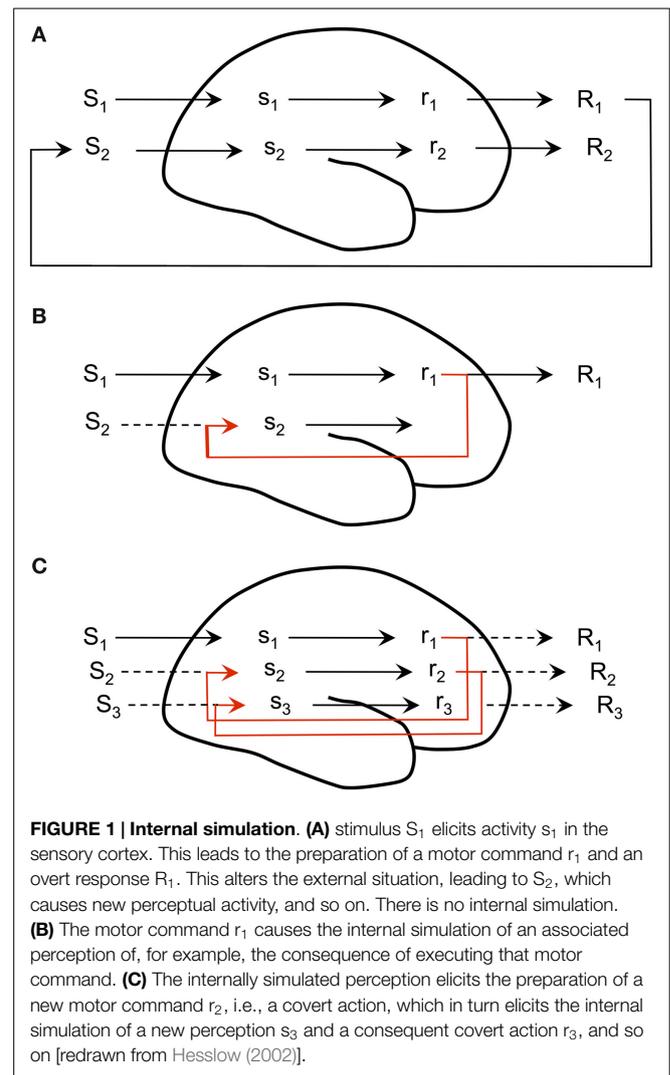
In the preceding section, we considered internal simulation entirely in terms of memory-based self-projection, using re-assembled combinations of episodic memory to pre-experience possible futures, re-experience (and possibly adjust) past experiences, and project ourselves into the experiences of others. However, we know that action plays a significant role in our perceptions so the question then is: does action play a role in internal simulation? The answer is a clear “yes” (Hesslow, 2002, 2012; Svensson et al., 2007). Internal simulation extends beyond embodied interaction and includes simulated interaction, particularly embodied interaction. Although the terms simulation, internal simulation, and mental simulation are widely used, you will also see references being made to *emulation*, very often when the approach endeavors to model the exact mechanism by which the simulation is produced (Grush, 2004).

## The Simulation Hypothesis

There are a number of simulation theories, but perhaps the most influential is what is known as the *simulation hypothesis* (Hesslow, 2002, 2012). It makes three core assumptions:

1. The regions in the brain which are responsible for motor control can be activated *without* causing bodily movement.
2. Perceptions can be caused by internal brain activity and not just by external stimuli.
3. The brain has associative mechanisms that allow motor behavior or perceptual activity to evoke other perceptual activity.

The first assumption allows for simulation of actions and is often referred to as *covert action* or *covert behavior*. The second allows for simulation of perceptions. The third assumption allows simulated actions to elicit perceptions that are like those that would have arisen if the actions had actually been performed. There is an increasing amount of neurophysiological evidence in support of all three assumptions (Svensson et al., 2013). If we link these assumptions together, we see that the simulation hypothesis



shows how the brain can simulate extended perception-action-perception sequences by having the simulated perceptions elicit simulated action which in turn elicits simulated perceptions, and so on. **Figure 1** summarizes the simulation hypothesis, showing three situations, one where there is no internal simulation, one where a motor response to an input stimulus causes the internal simulation of an associated perception, and one where this internally simulated perception then elicits a covert action which in turn elicits a simulated perception and a consequent covert action, and so on.

## Motor, Visual, and Mental Imagery

Action-directed internal simulation involves three different types of anticipation: implicit, internal, and external (Svensson et al., 2009). *Implicit anticipation* concerns the prediction of motor commands from perceptions (which may have been simulated in a previous phase of internal simulation). *Internal anticipation* concerns the prediction of the proprioceptive consequences of carrying out an action, i.e., the effect of an action on the agent's own body. *External anticipation* concerns the prediction of the consequences for external objects and other agents of carrying

out an action.<sup>4</sup> Implicit anticipation selects some motor activity (possibly covert, i.e., simulated) to be carried out based on an association between stimulus and actions; internal anticipation and external anticipation then predict the consequences of that action. Collectively, they simulate actions and the effects of actions.

Covert action involves what is referred to as *motor imagery* and simulation of perception is often referred to as *visual imagery*. Perceptual imagery would perhaps be a better term since there is evidence that humans use imagery from all the senses. In a way, motor imagery is also a form of perceptual imagery, in the sense that it involves the proprioceptive and kinesthetic sensations associated with bodily movement. However, reflecting the interdependence of perception and action, covert action often has elements of both motor imagery and visual imagery and, *vice versa*, the simulation of perception often has elements of motor imagery. Visual imagery and motor imagery are sometimes referred to collectively as *mental imagery* (Wintermute, 2012). Moulton and Kosslyn (2009) identify several different types of perceptual imagery and distinguish between two different types of simulation: *instrumental simulation* and *emulative simulation*. The former concerns itself only with the content of the simulation while the latter also replicates the process by which that content is created in the simulated event itself. They refer to this as *second-order simulation*.

## Joint Perceptuo-Motor Representations

In the foregoing, we remarked on the fact that mental imagery, viewed as another way of expressing the process of internal simulation, comprises both visual imagery (or perceptual imagery) and motor imagery. More importantly, though, we noted that these two forms of imagery are tightly entwined: they complement each other and the simulation of perception and covert action both involve elements of visual and motor imagery.

Classical treatments of memory usually maintain a clear distinction between declarative memory and procedural memory, in general, and between episodic memory and procedural memory, in particular. However, contemporary research takes a slightly different perspective, binding the two more closely, e.g., the mirror neuron system, in particular. While it is still a major challenge to understand how these two memory systems are combined, this coupling is the basic idea underpinning joint perceptuo-motor representations: representations that bring together the motoric and sensory aspects of experience in one framework, such as that anticipated in the simulation hypothesis.

In this section, we look at four approaches that have been developed to address joint perceptuo-motor representations. First, we look at two approaches to implementing ideo-motor theory in cognitive robotics: Shanahan's Global Workspace Theory architecture and Demiris's HAMMER architecture. We follow this by highlighting two additional approaches that endeavor to integrate perceptuo-motor representations more tightly: the Theory of Event Coding (TEC) and Object-Action Complexes. Since none of these explicitly incorporate episodic or procedural memory, we then suggest a way of drawing the principal ideas of each together

in a form of explicit joint episodic-procedural memory. We then argue that this joint episodic-procedural memory allows several of the challenges of cognitive robotics to be addressed.

Before discussing these, to provide the necessary context for prospective perceptuo-motor representations, we first address the difference between sensory-motor theory and ideo-motor theory.

## Sensory-Motor Theory and Ideo-Motor Theory

Broadly speaking, there are the two distinct approaches for planning actions: *sensory-motor* action planning and *ideo-motor* action planning (Stock and Stock, 2004). Sensory-motor action planning treats actions as reactive responses to sensory stimuli and assumes that perception and action use distinct and separate representational frameworks. The sensory-motor view builds on the classic unidirectional data-driven information-processing approach to perception, proceeding stage by stage from stimulus to percept and then to response. It is unidirectional in that it does not allow the results of later processing to influence earlier processing. In particular, it does not allow the resultant (or intended) action to impact on the related sensory perception.

Ideo-motor action planning, on the other hand, treats action as the result of internally generated goals. It is the idea of achieving some action outcome, rather than some external stimulus, that is at the core of how cognitive agents behave. This reflects the view of action described above, with action being initiated by a motivated subject, defined by goals, and guided by prospection. The key point of the ideo-motor principle is that the selection and control of a particular goal-directed movement depends on the anticipation of the sensory consequence of accomplishing the intended action: the agent images (e.g., through internal simulation) the desired outcome and selects the appropriate actions in order to achieve it.

There is an important difference, though, between the concrete movements comprising an action and the higher-order goals of an action. Typically, actors do not voluntarily pre-select the exact movements required to achieve a desired goal. Instead, they select prospectively guided intention-directed goal-focused action, with the specific movements being adaptively controlled as the action is executed. Thus, ideo-motor theory should be viewed both as an anticipatory idea-centered way of selecting actions and as a way of bridging the higher-order conceptual representations of intentions and goals<sup>5</sup> with the concrete adaptive control of movements when executing that action (Ondobaka and Bekkering, 2012).

In contrast to sensory-motor models, ideo-motor theory assumes that perception and action share a common representational framework. Because ideo-motor models focus on goals, and because they use a common joint representation that embraces both perception and action, they provide an intuitive explanation of why cognitive agents, humans in particular, are so adept at and predisposed to imitation (Iacoboni, 2009). The essential idea is that when I see somebody else's (goal-directed) actions and

<sup>4</sup>The terms *internal anticipation* and *external anticipation* are also referred to as *bodily anticipation* and *environmental anticipation* (Svensson et al., 2013).

<sup>5</sup>Michael Tomasello and colleagues note that the distinction between intentions and goals is not always clearly made. Taking their lead from Michael Bratman (1998), they define an intention as a plan of action an agent chooses and commits itself to in pursuit of a goal. An intention therefore includes both a means (i.e. an action plan) as well as a goal (Tomasello et al., 2005).

the consequences of these actions, the representations of my own actions that would produce the same consequences are activated.

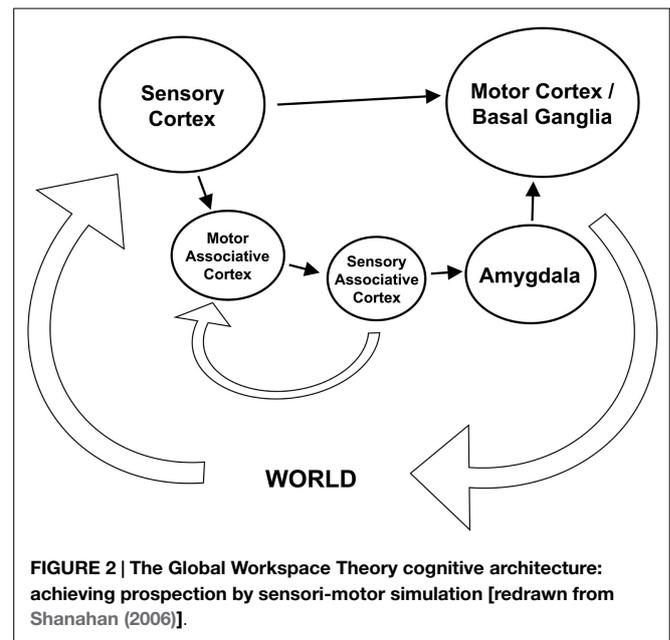
At first glance, *ideo-motor theory* seems to present a puzzle: how can the goal, achieved through action, cause the action in the first place? In other words, how can the later outcome affect the earlier action? This seems to be a case of *backward causation*. The solution to the puzzle is *prospection*. It is the anticipated goal state, not the achieved goal state, that impacts on the associated planned action. Goal-directed action, then, is a center-piece of *ideo-motor theory*, which is also referred to as the *goal trigger hypothesis* (Hommel et al., 2001).

Before proceeding to consider two cognitive architectures that build on *ideo-motor theory*, we mention *cognitive maps* to highlight the importance of joint perceptuo-motor representations in animal and robot cognition. The idea of a cognitive map was introduced by Tolman as a geometric representation to support navigation in biological agents (Tolman, 1948). While there is a certain lack of consensus on what exactly constitutes a cognitive map (Bennett, 1996; Eichenbaum et al., 1999), most agree that it involves metric information rather than purely topological information to encode spatial relationships in an allocentric framework and that it exploits path integration, at least partially, to effect navigation (Gallistel, 1989, 1990; Stachenfeld et al., 2014); for an alternative perspective, see Gaussier et al. (2002). In any case, a cognitive map combines memories of environmental cues (or perceptual landmarks) with geometrical properties of space that are specified by the remembered landmarks (Metta et al., 2010). Based on the existence of the hippocampus place cells (O'Keefe, 1976), O'Keefe and Nadel suggested that the hippocampal formation provides the neural basis for the cognitive map (O'Keefe and Nadel, 1978).

However, the hippocampus does not just create and store cognitive maps but it also plays a part in episodic memory, e.g., helping to minimize the similarities between new representations and representations that already exist in memory (McNaughton et al., 2006). As with episodic memory, it is also responsible for associating information in ways that allow flexible use of past experiences to guide future actions (*flexible memory expression*) (Eichenbaum et al., 1999; McNamara and Shelton, 2003). Furthermore, it has a role as a prediction mechanism for novelty detection and especially as a way to merge planning and sensory-motor function in a single coherent system (Gaussier et al., 2002). As McNaughton et al. note, “our current understanding of [the hippocampal formation] underscores the growing paradigm shift in the neurosciences away from thinking about neural coding as being driven primarily by bottom-up, sensory inputs, but rather as a reflection of rich and complex internal dynamics” (McNaughton et al., 2006). Taken together, the characteristics of cognitive maps and the operation of the hippocampal formation echo the arguments being put forward in this paper about the importance of joint perceptuo-motor representations in cognition.

### The Global Workspace Cognitive Architecture

Shanahan (Shanahan, 2005a,b, 2006; Shanahan and Baars, 2005) proposes a biologically plausible brain-inspired neural-level cognitive architecture in which cognitive functions such as anticipation and planning are realized through internal simulation of interaction with the environment. Action selection, both actual



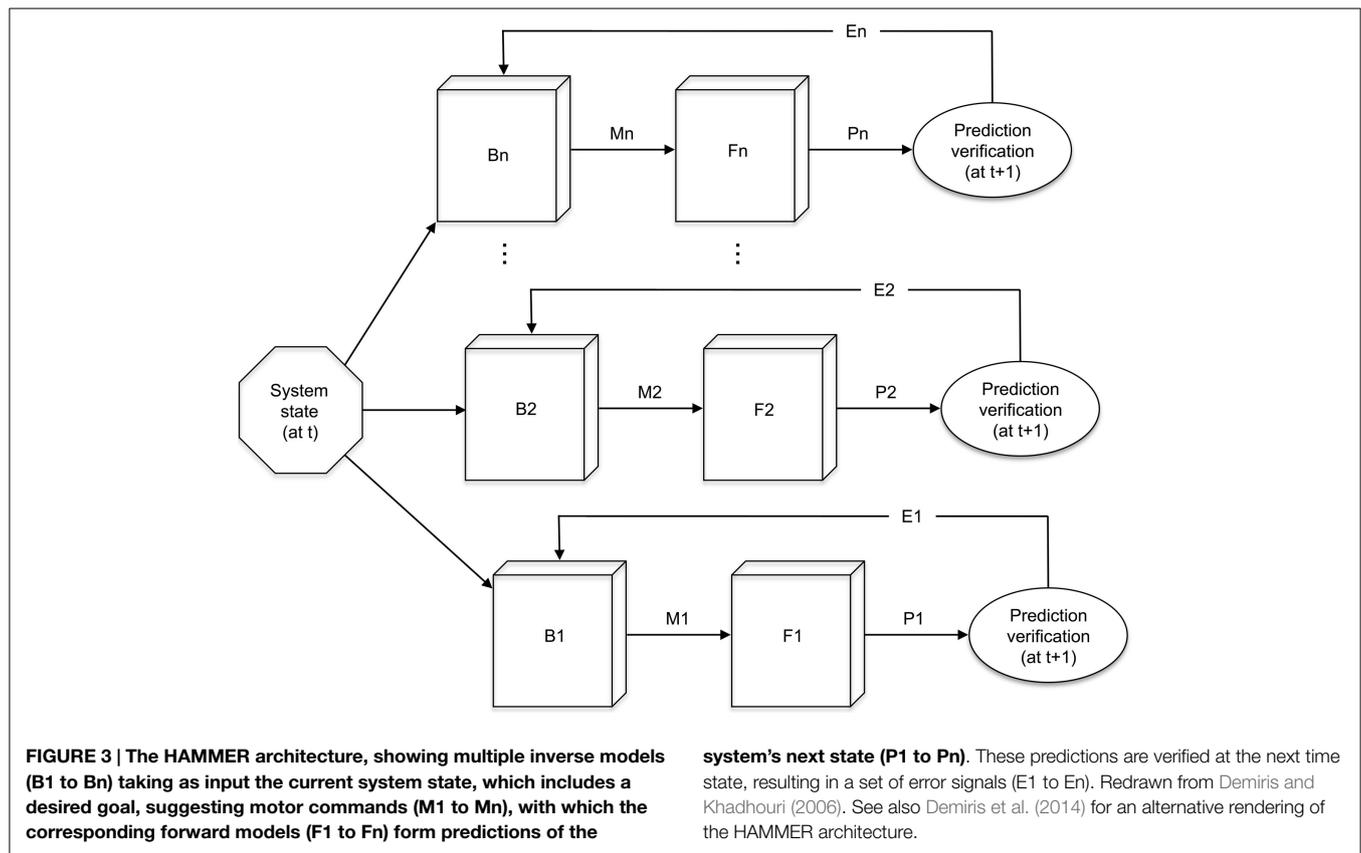
and internally simulated, is mediated by affect. The architecture is based on an external sensori-motor loop and an internal sensori-motor loop in which information passes through multiple competing cortical areas and a global workspace (Baars, 1998, 2002).

Shanahan's cognitive architecture is comprised of the following components: a first-order sensori-motor loop, closed externally through the world, and a higher-order sensori-motor loop, closed internally through associative memories (see **Figure 2**). The first-order loop comprises the sensory cortex and the basal ganglia (controlling the motor cortex), together providing a reactive action-selection sub-system. The second-order loop comprises two associative cortex elements which carry out off-line simulations of the system's sensory and motor behavior, respectively. The first associative cortex simulates a motor output while the second simulates the sensory stimulus expected to follow from a given motor output. The higher-order loop effectively modulates basal ganglia action selection in the first-order loop via an affect-driven amygdala component. Thus, this cognitive architecture is able to anticipate and plan for potential behavior through the exercise of its “imagination” (*i.e.*, its associative internal sensori-motor simulation).

### The HAMMER Architecture

While internal simulation is an essential aspect of human cognition, it is also an increasingly important part of artificial cognitive systems. For example, The Hierarchical Attentive Multiple Models for Execution and Recognition (HAMMER) architecture (Demiris and Khadhouri, 2006; Demiris et al., 2014) builds on the simulation hypothesis, accomplishing internal simulation using forward and inverse models which encode internal sensori-motor models that the agent would utilize if it were to execute that action (see **Figure 3**).

HAMMER deploys several inverse-forward pairs to simulate multiple possible futures using a winner-take-all attention process to select the most appropriate action to execute. HAMMER



includes recurrent connections, thereby allowing multi-stage extended internal simulation and mental rehearsal. This provides the architecture with a way of encapsulating the internal simulation hypothesis proposed by Hesslow (2002, 2012).

The inverse model takes as input information about the current state of the system and the desired goal, and it outputs the motor commands necessary to achieve that goal. The forward model acts as a predictor. It takes as input the motor commands and simulates the perception that would arise if this motor command were to be executed, just as the simulation hypothesis envisages. HAMMER then provides the output of the inverse model as the input to the forward model. This allows a goal state (demonstrated, for example, by another agent or possibly recalled from episodic memory) to elicit the simulated action required to achieve it. This simulated action is then used with the forward model to generate a simulated outcome, i.e., the outcome that would arise if the motor commands were to be executed. The simulated perceived outcome is then compared to the desired goal perception and the results are then fed back to the inverse model to allow it to adjust any parameters of the action.

A distinguishing feature of the HAMMER architecture is that it operates multiple pairs of inverse and forward models in parallel, each one representing a simulation – a hypothesis – of how the goal action can be achieved. The choice of inverse/forward model pair is made by an internal attention process based on how close the predicted outcome is to the desired one. Furthermore, it provides for the hierarchical composition of primitive actions into more complex sequences.

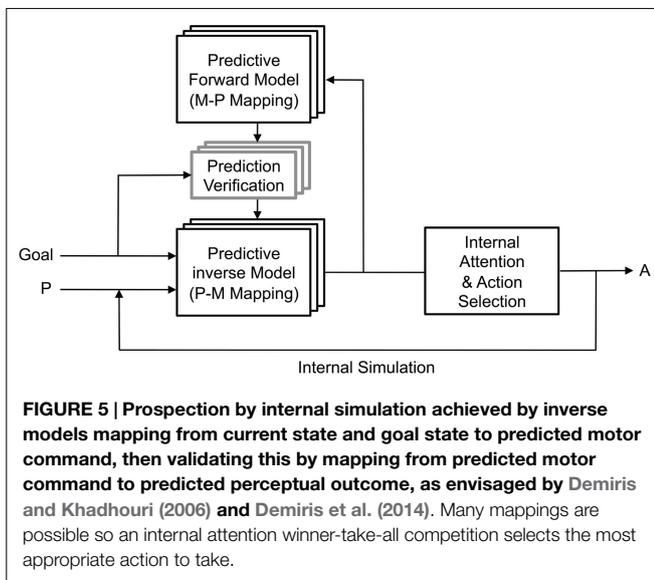
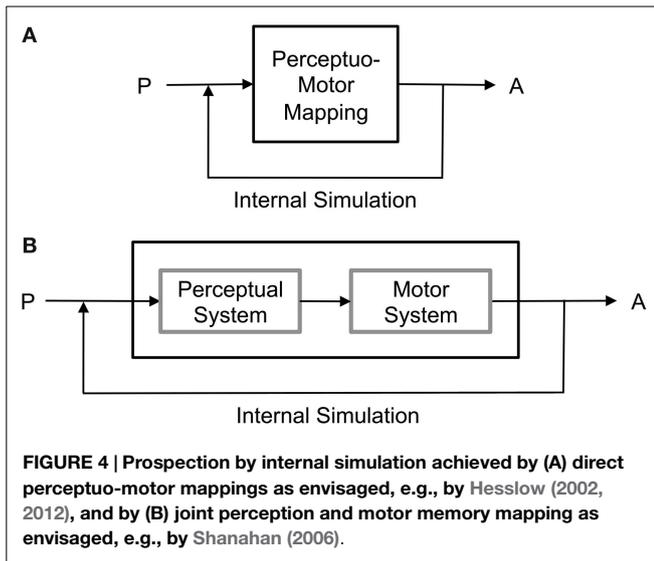
## From Perceptuo-Motor Mappings to Perceptuo-Motor Memory

Both Global Workspace Theory and HAMMER are good models of the simulation hypothesis for internal simulation as a vehicle for prospection in cognition. However, they focus on the mapping between perception and motor command, with memory being left implicit (see Figures 4 and 5).

Other models, such as the *Theory of Event Coding* (TEC) (Hommel et al., 2001) and *Object Action Complexes* (OACs) (Krüger et al., 2011) attempt to provide a tighter coupling of the perceptual and motor aspect in a joint perceptuo-motor representation.

The Theory of Event Coding (TEC) is a representational framework for combining perception and action planning. It focuses mainly on the later stages of perception and the earlier phases of action. As such, it concerns itself with perceptual features but not with how those features are extracted or computed. Similarly, it concerns itself with preparing actions – action planning – but not with the final execution of those actions and the adaptive control of various parts of the agent's body. The main idea is that perception, attention, intention, and action all work with a common representation and, furthermore, that action depends on both external and internal causes.

TEC provides a basis for combining both sensory-motor and ideo-motor action planning (Stock and Stock, 2004) and to be a joint representation that serves both sensory-stimulated action and prospective goal-directed action. The core concept in TEC is the *event code*. This is effectively a structured aggregation of distal features of an event in the agent's world. These *feature codes* can



be relatively simple (e.g., color, shape, moving to the left, falling) or more complex, such as an affordance. Also, TEC feature codes can emerge through the agent’s experience; they do not have to be pre-specified. A given TEC feature code is associated with both the sensory system and the motor system. Typically, a feature code is derived from several proximal sensory sources (sensory codes) and it contributes to several proximal motor actuators (motor codes). Each *event code* comprises several feature codes representing some event, be it a perceived event or a planned event. Feature codes associated with an event are activated both when the event is perceived and when it is planned. Because features can be elements of many event codes, the activation of a given feature effectively primes, i.e., predisposes, all the other events of which this feature is a component.

Inspired by the Theory of Event coding, an Object-Action Complex (OAC) (Krüger et al., 2011) is a triple, i.e., a unit with three components: ( $E, T, M$ ).  $E$  is an “execution specification”

(effectively an action).  $T$  is a function that predicts how the attributes that characterize the current state of the agent’s world will change if the execution specification is executed. Effectively, of  $T$  as a prediction of how the agent’s perceptions will change as a result of carrying out the actions given by  $E$ .  $M$  is a statistical measure of the success of the OAC’s past predictions. In this way, an OAC combines the essential elements of a joint representation – perception and action – with a predictor that links current perceived states and future predicted perceived states that would result from carrying out that action. To a large extent, an OAC models an agent’s interaction with the world as it executes some motor program (this is referred to as a low-level control program  $CP$  in the OAC literature). For example, an OAC might encode how to grasp an object or push an object into a given position and orientation (usually referred to as the object pose). OACs can be learned and executed, and they can be combined into more complex representations of actions and their perceptual consequences.

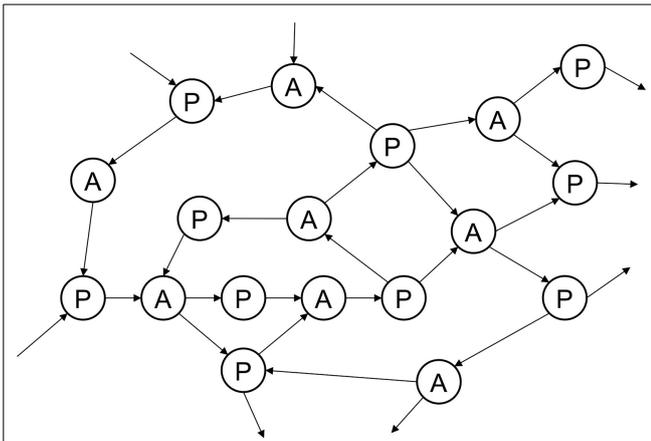
To date, neither TEC nor OAC has been embedded in the more general internal simulation framework described above. So, it is proposed here that there is a strong case for making memory – episodic and procedural – more explicit and embedding them in an internal simulation framework (such as that envisaged in the simulation hypothesis, the GWT Architecture, and the HAMMER architecture) in a way that makes their links more explicit (such as that envisaged in TEC and OAC). We address such a possible framework on the next section.

## A Network-Based Joint Episodic-Procedural Memory for Internal Simulation

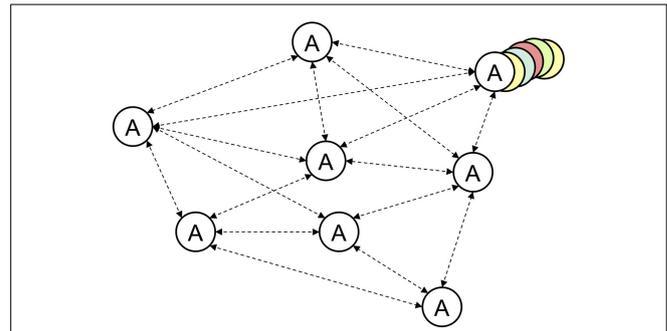
The core idea being proposed is to unwind the temporal and causal relationships between specific perceptions and actions that are implicit in the mappings of, e.g., GWT and HAMMER, and make them explicit in a weighted network of associations between perceptions and actions, in the manner of TEC and OAC (see **Figure 6**). In doing so, it makes the input to the joint perceptuo-motor mapping explicit as perceptual episodic memories and motoric procedural memories (see **Figures 7 and 8**). In the case of episodic memory, this provides a way to include other modalities including affective or hedonic memories. Procedural memory operates associatively in their own right: such procedural memories are not static but are dynamic and adapt as the action is executed.

Furthermore, such a framework allows one to expose the mapping dynamics explicitly. This may have several advantages in, for example, cognitive development which focuses on extending the timescale of the agent’s prospective capacity and expanding the agent’s repertoire of actions. Specifically, development might be facilitated by adjusting and adapting the network structure – its topology and strength of connectivity – as a function of experiential learning, intrinsic value systems (Merrick, 2010), including those derived from autonomic self-maintenance (Bickhard, 2000), and affective homeostasis and allostasis (Sterling, 2004, 2012; Morse et al., 2008; Ziemke and Lowe, 2009).

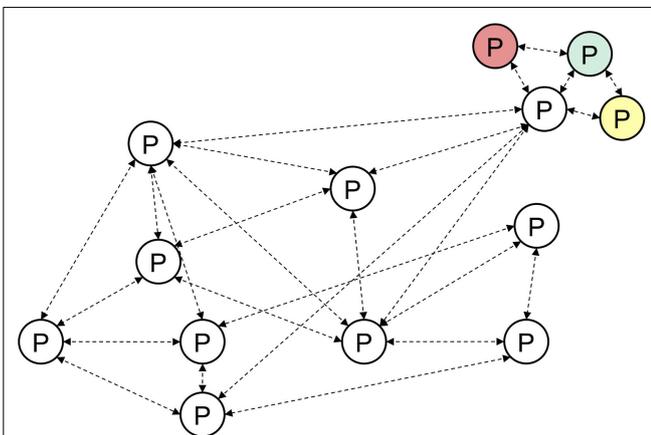
The network model of joint episodic-procedural memory facilitates prospection in three senses: prospection by predicting the



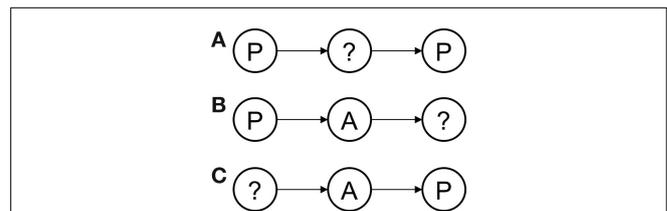
**FIGURE 6 |** Joint episodic-procedural memory as an explicit network of associations between perceptions and actions, drawn from episodic and procedural memories, unwinding the temporal and causal relationships between specific perceptions and actions that are implicit in the mappings of other perceptuo-motor representations.



**FIGURE 8 |** The procedural elements of the joint episodic-procedural memory are drawn from procedural memory and, again, operate associatively in their own right. Such procedural memories are not static but are dynamic and adapt as the action is executed (top right).



**FIGURE 7 |** The episodic elements of the joint episodic-procedural memory are drawn from episodic memory and therefore operate associatively in their own right. Furthermore, this provides a way to include other modalities of episodic memory (top right) including affective or hedonic memories.



**FIGURE 9 |** The network model of joint episodic-procedural memory facilitates prospection in three senses: (A) prospection by predicting the outcome of an action carried out in given perceptual circumstances, (B) prospection by predicting the action required to achieve a goal in given perceptual circumstances, and (C) abductive inference of the perceptual states that explains an outcome of a give action.

outcome of an action carried out in given perceptual circumstances, prospection by predicting the action required to achieve a goal in given perceptual circumstances, and abductive inference of the perceptual states that explains an outcome of a give action (see **Figure 9**).

Keeping episodic memory explicit in this framework preserves the flexibility for adaptive reconstruction and novel association. Since episodic memory operates by recombining imperfectly recalled past experience, this allows it to simulate new or unexpected events as outlined above.

There is, however, a potential problem in that the scope for exponential growth in association is significant. Something is needed to constrain this potential combinatorial explosion if such a joint episodic-procedural memory system is to be capable of

useful prospection through internal simulation. Because the associative links are exposed explicitly in the network organization, this framework for a joint episodic-procedural memory allows the internal simulation to be conditioned by current context, semantic memory, and the agent's value system by adjusting the associative links. Context and semantics constrain the combinatorial explosion of potential perception-action associations and allow effective action selection in the pursuit of goals, while the value system modulates the memory network to promote the agent's autonomy and cognitive development.

Finally, the approach being suggested here is an abstract schema and is therefore neutral regarding the final implementation of these episodic and procedural memories. These can be effected either as an emergent cognitive system, instantiating them sub-symbolically in a biologically inspired manner as associative networks [e.g., Hopfield nets such as in Mohan et al. (2014) or brain-based devices such as in Krichmar and Edelman (2005, 2006)]. Alternatively, they can be effected symbolically as more traditional AI systems. For example, episodic memory might be implemented using content-addressable image databases with traditional image indexing and recall algorithms, while procedural memory could be encapsulated in databases of motor-control scripts derived from experiential learning or from shared resources [e.g., Tenorth and Beetz (2009) and Tenorth et al. (2012, 2013)]. The traditional AI implementation, for the purpose of practical cognitive robotics, has a number of advantages.

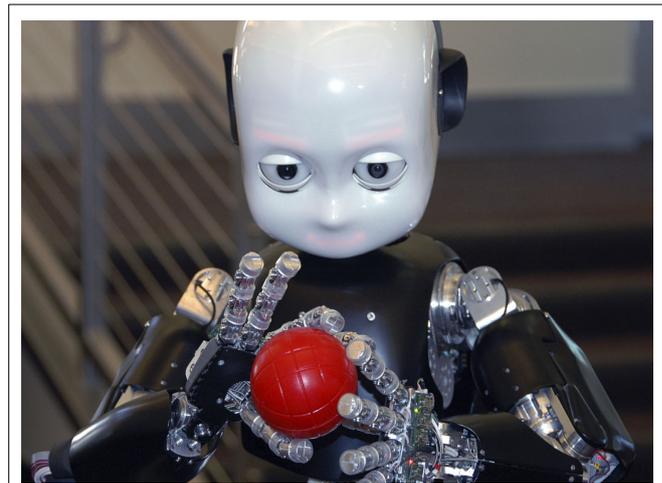
Although episodic memory will typically exploit by iconic representations, these representations are often augmented by symbolic tags when derived from on-line repositories. This symbolic tagging makes the integration of semantic knowledge much easier. The fact that both episodic memory and procedural memory are derived from experience, directly or indirectly, also finesses the symbol grounding problem (Harnad, 1990; Sloman). The traditional AI implementation also renders the knowledge contained in the memory inherently transferrable to other agents, provided their sensory systems are compatible and there is a known mapping – direct or indirect – between the embodiments of each agent, as described in Argall et al. (2009).

## An Example Joint Episodic-Procedural Memory for Overt Attention

The iCub is a 53 degree-of-freedom humanoid robot (see **Figure 10**) that was designed to be an open-systems platform for research in cognitive development (Sandini et al., 2007; Tsagarakis et al., 2007; Metta et al., 2010). It is approximately 1 m tall, weighs 22 kg, has visual, vestibular, auditory, and haptic sensors, and is capable of dexterous manipulation. To date, iCubs have been delivered to over 20 research laboratories in Europe and one in the U.S.A.<sup>6</sup>

The original iCub cognitive architecture (Sandini et al., 2007; Vernon et al., 2010) focused on gaze-modulated goal-directed reaching and locomotion. Episodic memory and procedural memory were designed to effect internal simulation in order to provide capabilities for prediction and model construction bootstrapped by learned affordances. Motivations encapsulated in the system's affective state addressed curiosity and experimentation, both of which are exploratory motives, triggered by exogenous and endogenous factors, respectively. This distinction between the exogenous and the endogenous was reflected in the overt attention system that could be triggered by both external and internal events. A simple process of homeostatic self-regulation governed by the affective state provided elementary action selection. Finally, all the various components of the cognitive architecture operated concurrently so that a sequence of states representing cognitive behavior emerges from the interaction of many separate parallel processes rather than being dictated by some pre-programmed state-machine.

In the variant of the iCub cognitive architecture presented here, the separate episodic and procedural memories have been replaced by a simple proof-of-principle joint episodic-procedural memory (see **Figure 11**). This is the focus of the current article and the specific objective is to investigate how a joint episodic-procedural memory can be used for representation, development, and adaptation of scan-path patterns that result from overt and covert attention. This particular model of attention uses an information-theoretic saliency map (Bruce and Tsotsos, 2009) with an overt attention system comprising (1) the winner-take-all process effected by a selective tuning model to identify a single focus of attention (Tsotsos et al., 1995; Tsotsos, 2006, 2011), (2) an Inhibition-Of-Return (IOR) mechanism to attenuate the attention



**FIGURE 10 |** The iCub humanoid robot: an open-systems platform for research in cognitive development.

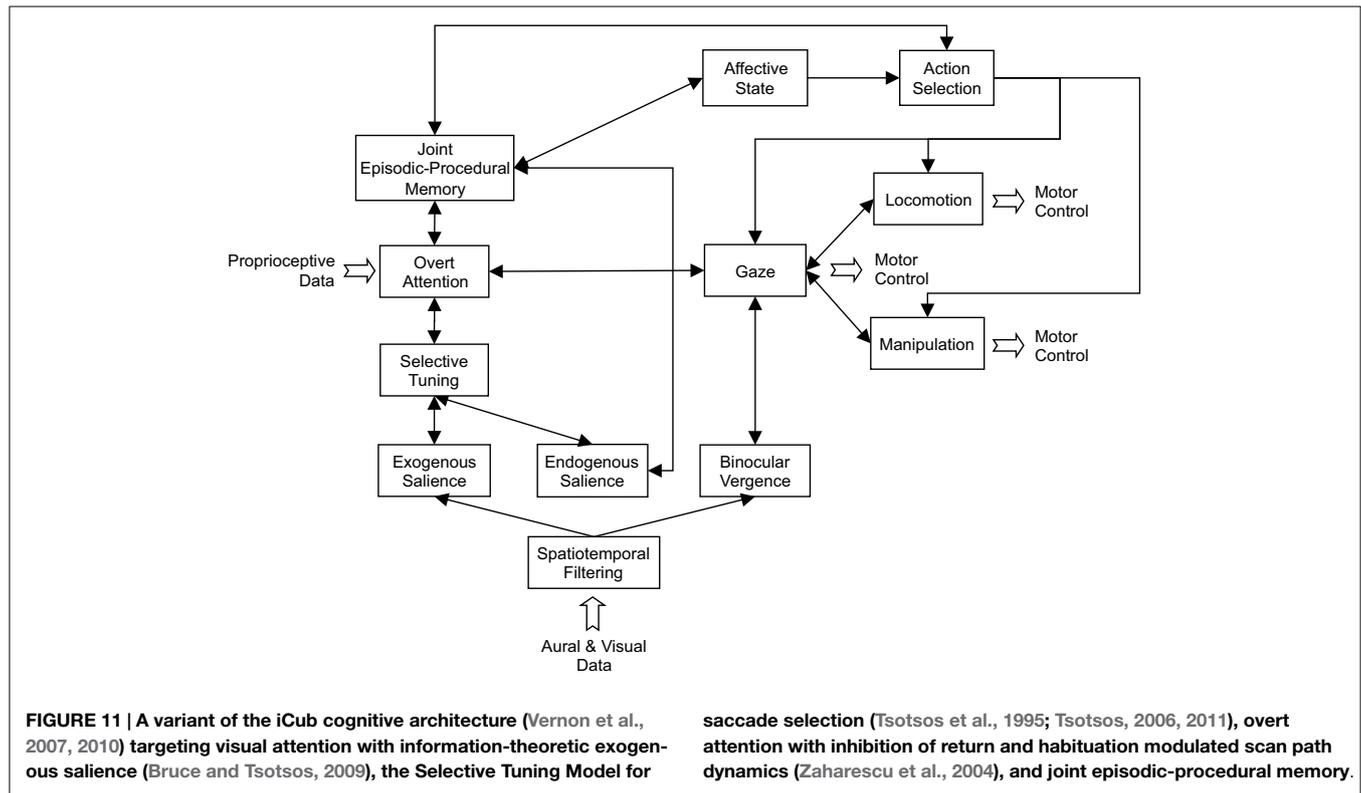
value of previous winning locations so that new regions become the focus of attention, and (3) a habituation process to reduce the salience of the current focus of attention with time thereby ensuring that attention is fixated on a given point for a limited period (Zaharescu et al., 2004). Fixation points are represented using retinotopic images rather than conventional rectangular regularly sampled images. The retinotopic images are constructed using a scale and rotation-invariant log-polar transform (Braccini et al., 1981; Berton, 2006; Berton et al., 2006; Traver and Bernardino, 2010) to map the Cartesian camera image data to a non-uniformly sampled image that reflects the foveated sampling in the primate retina. The resultant scan path patterns are captured in an elementary joint episodic-procedural memory: the episodes are retinotopic log-polar images of the fixation points and the actions are the saccade angles.

The episodic memory in the iCub cognitive architecture is a simple associatively recalled memory of autobiographical events. It is a form on one-shot learning and does not generalize multiple instances of an observed event. In the current implementation, the episodic memory provides a purely visual iconic memory of landmark appearance using scale- and rotation-invariant<sup>7</sup> retinotopic log-polar images as the landmark representation (Braccini et al., 1981; Berton, 2006; Berton et al., 2006; Traver and Bernardino, 2010) with image recognition being effected using color histogram intersection (Swain and Ballard, 1990, 1991). In essence, the iCub episodic memory implements a form of content-addressable memory which is populated by log-polar landmark images acquired under the control of the iCub's covert and overt attention sub-system.

Procedural memory maintains a very simple repository of elementary actions. The current implementation comprises gaze motor commands in a body-centered frame of reference and symbolic tags denoting one of five possible associated actions (reach, push, grasp, locomote, or wait). These are just placeholders for

<sup>6</sup>For more information on the iCub robot see <http://www.icub.org>.

<sup>7</sup>The rotation invariance of log-polar images is restricted to roll: rotation about the camera's principal axis.



more flexible and adaptive gaze-directed motor control schemes [e.g., Lukic et al. (2012)] to be implemented later.

The joint episodic-procedural memory itself is a network of associations between motor events and pairs of sensory events. In this variant of the iCub cognitive architecture, a sensory event is a visual landmark which has been acquired by the iCub and stored in the episodic memory. A motor event is a gaze saccade with an optional reaching, grasping, or locomotion movement. Thus, joint episodic-procedural memory can be viewed as a directed network with two types of nodes, one representing sensory patterns – retinotopic log-polar images of the fixation points – and the other representing motor patterns – the saccade motor commands. A path through the network traverses alternately sensory and motor nodes and any clique in this memory network effectively captures a causal relationship between a sensory state, a motor state, and a subsequent sensory state (or a sequence of such associations). An extended path in this memory captures the scan path pattern of the robot as it pays attention to its visual environment (see **Figure 12**).

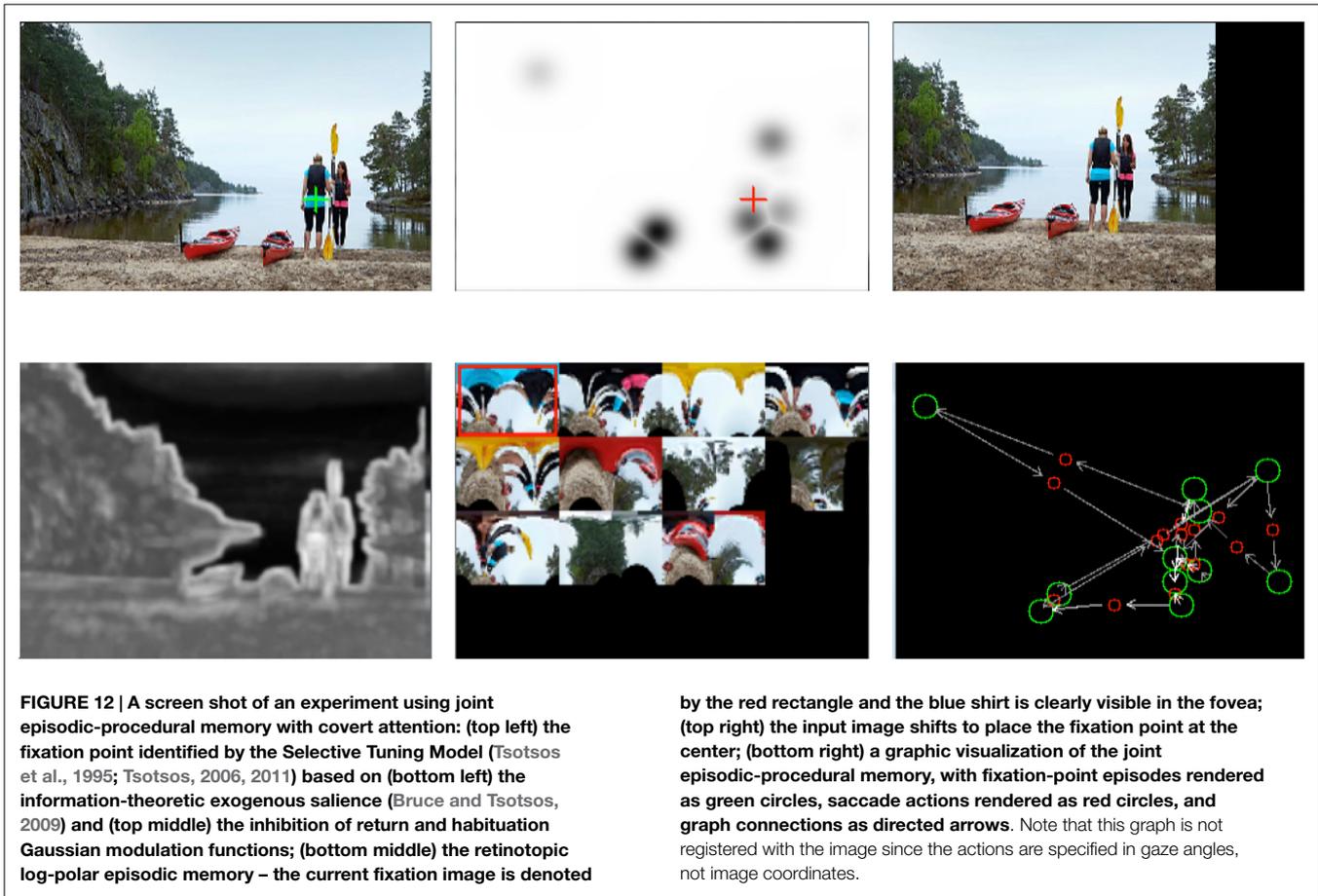
The key feature of this form of joint episodic-procedural memory representation of the attention pattern of the robot is that it lends itself to development: modulation or dynamically reconfiguration of the connectivity of this network – which is learned from experience – so that its prospective capacity increases as new memories are added as a result of the agent's interaction with its environment. Various forms of adaptive reconfiguration are currently being examined, some based on small world networks (Watts and Strogatz, 1998; Newman, 2000; Bohland and Minai, 2001; Kleinberg, 2006; Telesford et al., 2011) and others based

on information theoretic models that dynamically modulate the pathways in flow networks (Ulanowicz, 2000).

## Conclusion

While action and prospection are intimately linked, most research on prospection has tended to focus on the constructive role of episodic memory (Tulving, 1972, 1984; Seligman et al., 2013), i.e., the so-called episodic future thinking (Atance and O'Neill, 2001), often achieved through internal simulation, i.e., the mental construction of an imagined alternative perspectives (Buckner and Carroll, 2007) and simulated embodied interaction (Svensson et al., 2007). Although hedonic affective experience has been addressed to some extent (Gilbert and Wilson, 2007; Lowe and Ziemke, 2011), procedural memory has been neglected in modeling prospective capacities. When it is included, it usually takes the form of distinct forward models that predict the sensory outcome of a given motor command (Hesslow, 2002, 2012; Shanahan, 2006) and inverse models that determine the action required to produce a given goal perception (Demiris and Khadhouri, 2006). Ideo-motor theory (Stock and Stock, 2004; Iacoboni, 2009) is an exception to this. It assumes that perception and action share a common representational framework and that action is the causal result of internally generated goals. Such a joint representation provides greater flexibility in prospection through both inductive inference and abductive inference.

With few exceptions, such as the Theory of Event Coding (Hommel et al., 2001) and object-action complexes



(Krüger et al., 2011), joint perceptuo-motor representations have received little attention and none have addressed integration of hedonic affective experience. Our conjecture is that an internal simulation capability founded on ideo-motor theory and joint representations, and drawing on recent progress in the modeling-related mirror neuron system (Gallese et al., 1996; Rizzolatti et al., 1996; Rizzolatti and Craighero, 2004; Thill et al., 2013), provides a viable way to approach the integration of procedural and episodic memory as a joint perceptuo-motor system. Our specific contention is that it is helpful to conceive of this joint episodic-procedural memory – for goal-directed internal simulation – as a network of associations between elements of both episodic and procedural memories. This perspective is neutral regarding the final implementation of these episodic and procedural memories and it can facilitate both emergent and cognitivist AI approaches.

We argue that such a framework meets several challenges in cognitive robotics such as the need to accommodate modal and

modal episodic data and extended perceptuo-motor sequences, as well as mechanisms for conditioning the association dynamics with external constraints derived from semantic declarative knowledge, current context, and affective value signals. It also addresses the need to integrate the episodic and procedural knowledge gathered by robots as they operate of their physical environment with information extracted from web-based knowledge bases. This is particularly important if the power of indirect knowledge (acquired by interpreting third-party descriptions) is to be harnessed in the development of robot skills.

## Acknowledgments

This work was supported in part by the European Commission, Project 611391 DREAM: Development of Robot-enhanced Therapy for Children with Autism Spectrum Disorders ([www.dream2020.eu](http://www.dream2020.eu)).

## References

- Argall, B. D., Chernova, S., Veloso, M., and Browning, B. (2009). A survey of robot learning from demonstration. *Rob. Auton. Syst.* 57, 469–483. doi:10.1016/j.robot.2008.10.024
- Atance, C. M., and O'Neill, D. K. (2001). Episodic future thinking. *Trends Cogn. Sci.* 5, 533–539. doi:10.1016/S1364-6613(00)01804-0
- Baars, B. J. (1998). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Baars, B. J. (2002). The conscious access hypothesis: origins and recent evidence. *Trends Cogn. Sci.* 6, 47–52. doi:10.1016/S1364-6613(00)01819-2
- Bennett, A. (1996). Do animals have cognitive maps. *J. Exp. Biol.* 199, 219–224.
- Berton, F. (2006). "A brief introduction to log-polar mapping," in *Technical Report, LIRA-Lab*, (University of Genova).

- Berton, F., Sandini, G., and Metta, G. (2006). Anthropomorphic visual sensors. *Encycl. Sens.* 10, 1–16. doi:10.3389/fnsys.2014.00109
- Bickhard, M. H. (2000). *Autonomy, function, and representation*. Available at: <http://www.lehigh.edu/~mh0/autfuncprep.html>
- Bohland, J. W., and Minai, A. A. (2001). Efficient associative memory using small-world architecture. *Neurocomputing* 3, 489–496. doi:10.1016/S0925-2312(01)00378-2
- Braccini, C., Gambardella, G., and Sandini, G. (1981). A signal theory approach to the space and frequency variant filtering performed by the human visual system. *Signal Processing* 3, 231–240. doi:10.1016/0165-1684(81)90038-4
- Bratman, M. E. (1998). “Intention and personal policies,” in *Philosophical Perspectives*, ed. J. E. Tomberlin (Oxford: Blackwell), 3.
- Bruce, N. D. B., and Tsotsos, J. K. (2009). Saliency, attention, and visual search: an information theoretic approach. *J. Vis.* 9, 1–24. doi:10.1167/9.3.5
- Buckner, R. L., and Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57. doi:10.1016/j.tics.2006.11.004
- Demiris, Y., Aziz-Zahdeh, L., and Bonaiuto, J. (2014). Information processing in the mirror neuron system in primates and machines. *Neuroinformatics* 12, 63–91. doi:10.1007/s12021-013-9200-7
- Demiris, Y., and Khadhour, B. (2006). Hierarchical attentive multiple models for execution and recognition (HAMMER). *Rob. Auton. Syst.* 54, 361–369. doi:10.1016/j.robot.2006.02.003
- Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., and Tanila, H. (1999). The hippocampus, memory, and place cells: is it spatial memory or a memory space? *Neuron* 23, 209–226. doi:10.1016/S0896-6273(00)80773-4
- Flanagan, J. R., and Johansson, R. S. (2003). Action plans used in action observation. *Nature* 424, 769–771. doi:10.1038/nature01861
- Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609. doi:10.1093/brain/119.2.593
- Gallistel, C. R. (1989). Animal cognition: the representation of space, time and number. *Annu. Rev. Psychol.* 40, 155–189. doi:10.1146/annurev.ps.40.020189.001103
- Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge, MA: MIT Press.
- Gaussier, P., Revel, A., Banquet, J. P., and Babeau, V. (2002). From view cells and place cells to cognitive map learning: processing stages of the hippocampal system. *Biol. Cybern.* 86, 15–28. doi:10.1007/s004220100269
- Gilbert, D. T., and Wilson, T. D. (2007). Propection: experiencing the future. *Science* 317, 1351–1354. doi:10.1126/science.1144161
- Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behav. Brain Sci.* 27, 377–442. doi:10.1017/S0140525X04000093
- Harnad, S. (1990). The symbol grounding problem. *Physica D* 42, 335–346. doi:10.1016/0167-2789(90)90087-6
- Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* 6, 242–247. doi:10.1016/S1364-6613(02)01913-7
- Hesslow, G. (2012). The current status of the simulation theory of cognition. *Brain Res.* 1428, 71–79. doi:10.1016/j.brainres.2011.06.026
- Hommel, B., Müseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–937. doi:10.1017/S0140525X01000103
- Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annu. Rev. Psychol.* 60, 653–670. doi:10.1146/annurev.psych.60.110707.163604
- Johansson, R. S., Westling, G., Bäckström, A., and Flanagan, J. R. (2001). Eye-hand coordination in object manipulation. *J. Neurosci.* 21, 6917–6932.
- Kleinberg, J. (2006). “Complex networks and decentralized search algorithms,” in *Proceedings of the International Congress of Mathematicians (ICM)*. Madrid, 1019–1044.
- Krichmar, J. L., and Edelman, G. M. (2005). Brain-based devices for the study of nervous systems and the development of intelligent machines. *Artif. Life* 11, 63–77. doi:10.1162/1064546053278946
- Krichmar, J. L., and Edelman, G. M. (2006). “Principles underlying the construction of brain-based devices,” in *Proceedings of AISB '06 - Adaptation in Artificial and Biological Systems*, Volume 2 of *Symposium on Grand Challenge 5: Architecture of Brain and Mind*, eds T. Kovacs, and J. A. R. Marshall (Bristol: University of Bristol), 37–42.
- Krüger, N., Geib, C., Piater, J., Petrick, R., Steedman, M., Wörgötter, F., et al. (2011). Object-action complexes: grounded abstractions of sensory-motor processes. *Rob. Auton. Syst.* 59, 740–757. doi:10.1016/j.robot.2011.05.009
- Lowe, R., and Ziemke, T. (2011). The feeling of action tendencies: on the emotional regulation of goal-directed behaviours. *Front. Psychol.* 2:1–24. doi:10.3389/fpsyg.2011.00346
- Lukic, L., Santos-Victor, J., and Billard, A. (2012). “Learning coupled dynamical systems from human demonstration for robotic eye-arm-hand coordination,” in *Proceedings of the IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, (Osaka, Japan), 552–559.
- McNamara, T. P., and Shelton, A. L. (2003). Cognitive maps and the hippocampus. *Trends Cogn. Sci.* 7, 333–335. doi:10.1016/S1364-6613(03)00167-0
- McNaughton, B., Battaglia, F., Jensen, O., Moser, E., and Moser, M. (2006). Path integration and the neural basis of the ‘cognitive map’. *Nat. Rev.* 7, 663–678. doi:10.1038/nrn1932
- Merrick, K. E. (2010). A comparative study of value systems for self-motivated exploration and learning by robots. *IEEE Trans. Auto. Mental Dev.* 2, 119–131. doi:10.1109/TAMD.2010.2051435
- Metta, G., Natale, L., Nori, F., Sandini, G., Vernon, D., Fadiga, L., et al. (2010). The iCub humanoid robot: an open-systems platform for research in cognitive development. *Neural Netw.* 23, 1125–1134. doi:10.1016/j.neunet.2010.08.010
- Mohan, V., Sandini, G., and Morasso, P. (2014). A neural framework for organization and flexible utilization of episodic memory in cumulatively learning baby humanoids. *Neural Comput.* 26, 1–43. doi:10.1162/NECO\_a\_00664
- Morse, A., Lowe, R., and Ziemke, T. (2008). “Towards an enactive cognitive architecture,” in *Proceedings of the First International Conference on Cognitive Systems*, (Karlsruhe, Germany).
- Moulton, S. T., and Kosslyn, S. M. (2009). Imagining predictions: mental imagery as mental emulation. *Philos. Trans. R. Soc. B* 364, 1273–1280. doi:10.1098/rstb.2008.0314
- Newman, M. E. J. (2000). Models of the small world: a review. *J. Stat. Phys.* 101, 819–841. doi:10.1023/A:1026485807148
- O’Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Exp. Neurol.* 51, 78–109. doi:10.1016/0014-4886(76)90055-8
- O’Keefe, J., and Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford: Clarendon.
- Ondobaka, S., and Bekkering, H. (2012). Hierarchy of idea-guided action and perception-guided movement. *Front. Cogn.* 3:1–5. doi:10.3389/fpsyg.2012.00579
- Østby, Y., Walhovd, K. B., Tamnes, C. K., Grydeland, H., Westlye, L. G., and Fjell, A. M. (2012). Mental time travel and default-mode network functional connectivity in the developing brain. *PNAS* 109, 16800–16804. doi:10.1073/pnas.1210627109
- Rizzolatti, G., and Craighero, L. (2004). The mirror neuron system. *Annu. Rev. Physiol.* 27, 169–192. doi:10.1146/annurev.neuro.27.070203.144230
- Rizzolatti, G., and Fadiga, L. (1998). “Grasping objects and grasping action meanings: the dual role of monkey rostroventral premotor cortex (area F5),” in *Sensory Guidance of Movement, Novartis Foundation Symposium 218*, eds G. R. Bock, and J. A. Goode (Chichester: John Wiley and Sons), 81–103.
- Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141. doi:10.1016/0926-6410(95)00038-0
- Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson’s University Library.
- Sandini, G., Metta, G., and Vernon, D. (2007). “The iCub cognitive humanoid robot: An open-system research platform for enactive cognition,” in *50 Years of AI*, Vol. 4850, eds M. Lungarella, F. Iida, J. C. Bongard, and R. Pfeifer (Heidelberg: Springer), 359–370.
- Schacter, D. L., and Addis, D. R. (2007a). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philos. Trans. R. Soc. B* 362, 773–786. doi:10.1098/rstb.2007.2087
- Schacter, D. L., and Addis, D. R. (2007b). Constructive memory – the ghosts of past and future: a memory that works by piecing together bits of the past may be better suited to simulating future events than one that is a store of perfect records. *Nature* 445, 27. doi:10.1038/445027a
- Schacter, D. L., Addis, D. R., and Buckner, R. L. (2008). Episodic simulation of future events: concepts, data, and applications. *Ann. N. Y. Acad. Sci.* 1124, 39–60. doi:10.1196/annals.1440.001
- Seligman, M. E. P., Railton, P., Baumeister, R. F., and Sripada, C. (2013). Navigating into the future or driven by the past. *Perspect. Psychol. Sci.* 8, 119–141. doi:10.1177/1745691612474317

- Shanahan, M. P. (2005a). "Cognition, action selection, and inner rehearsal," in *Proceedings IJCAI Workshop on Modelling Natural Action Selection*, 92–99.
- Shanahan, M. P. (2005b). "Emotion, and imagination: a brain-inspired architecture for cognitive robotics," in *Proceedings AISB 2005 Symposium on Next Generation Approaches to Machine Consciousness*, 26–35.
- Shanahan, M. P. (2006). A cognitive architecture that combines internal simulation with a global workspace. *Conscious. Cogn.* 15, 433–449. doi:10.1016/j.concog.2005.11.005
- Shanahan, M. P., and Baars, B. (2005). Applying global workspace theory to the frame problem. *Cognition* 98, 157–176. doi:10.1016/j.cognition.2004.11.007
- Sloman, A. Available at: <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/talks/models.pdf>.
- Squire, L. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiol. Learn. Mem.* 82, 171–177. doi:10.1016/j.nlm.2004.06.005
- Stachenfeld, K., Botvinick, M., and Gershman, S. J. (2014). "Design principles of the hippocampal cognitive map," in *Advances in Neural Information Processing Systems*. eds Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (New York: Curran Associates, Inc.), 2528–2536.
- Sterling, P. (2004). "Principles of allostasis," in *Allostasis, Homeostasis, and the Costs of Adaptation*, ed. J. Schulkin (Cambridge: Cambridge University Press), 17–64.
- Sterling, P. (2012). Allostasis: a model of predictive regulation. *Physiol. Behav.* 106, 5–15. doi:10.1016/j.physbeh.2011.06.004
- Stock, A., and Stock, C. (2004). A short history of ideomotor action. *Psychol. Res.* 68, 176–188. doi:10.1007/s00426-003-0154-5
- Svensson, H., Lindblom, J., and Ziemke, T. (2007). "Making sense of embodied cognition: simulation theories of shared neural mechanisms for sensorimotor and cognitive processes," in *Body, Language and Mind*, Vol. 1, eds T. Ziemke, J. Zlatev, and R. M. Frank (Berlin: Mouton de Gruyter), 241–269.
- Svensson, H., Morse, A. F., and Ziemke, T. (2009). "Representation as internal simulation: a minimalistic robotic model," in *Proceedings of the Thirty-first Annual Conference of the Cognitive Science Society*, eds N. Taatgen, and H. van Rijn (Austin, TX: Cognitive Science Society), 2890–2895.
- Svensson, H., Thill, S., and Ziemke, T. (2013). Dreaming of electric sheep? Exploring the functions of dream-like mechanisms in the development of mental imagery simulations. *Adapt. Behav.* 21, 222–238. doi:10.1177/1059712313491295
- Swain, M., and Ballard, D. (1991). Color indexing. *Int. J. Comput. Vis.* 7, 11–32. doi:10.1007/BF00130487
- Swain, M. J., and Ballard, D. H. (1990). "Indexing via colour histograms," in *International Conference on Computer Vision – ICCV90*, 390–393.
- Szpunar, K. K. (2010). Episodic future thought: an emerging concept. *Perspect. Psychol. Sci.* 5, 142–162. doi:10.1177/1745691610362350
- Telesford, Q. K., Joyce, K. E., Hayasaka, S., Burdette, J. H., and Laurienti, P. J. (2011). The ubiquity of small-world networks. *Brain Connect.* 1, 367–375. doi:10.1089/brain.2011.0038
- Tenorth, M., and Beetz, M. (2009). "KnowRob – knowledge processing for autonomous personal robots," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4261–4266.
- Tenorth, M., Perzylo, A. C., Lafrenz, R., and Beetz, M. (2012). "The roboearth language: representing and exchanging knowledge about actions, objects, and environments," in *IEEE International Conference on Robotics and Automation*, 1284–1289.
- Tenorth, M., Perzylo, A. C., Lafrenz, R., and Beetz, M. (2013). Representation and exchange of knowledge about actions, objects, and environments in the roboearth framework. *IEEE Trans. Auto. Sci. Eng.* 10, 643–651. doi:10.1109/TASE.2013.2244883
- Thill, S., Caligiore, D., Borghi, A. M., Ziemke, T., and Baldassarre, G. (2013). Theories and computational models of affordance and mirror systems: an integrative review. *Neurosci. Biobehav. Rev.* 37, 491–521. doi:10.1016/j.neubiorev.2013.01.012
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208. doi:10.1037/h0061626
- Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–735. doi:10.1017/S0140525X05000129
- Traver, V. J., and Bernardino, A. (2010). A review of log-polar imaging for visual perception in robotics. *Rob. Auton. Syst.* 58, 378–398. doi:10.1016/j.robot.2009.10.002
- Tsagarakis, N. G., Metta, G., Sandini, G., Vernon, D., Beira, R., Santos-Victor, J., et al. (2007). iCub – the design and realisation of an open humanoid platform for cognitive and neuroscience research. *Int. J. Adv. Robot.* 21, 1151–1175. doi:10.1163/156855307781389419
- Tsotsos, J. K. (2006). "Cognitive vision need attention to link sensing with recognition," in *Cognitive Vision Systems: Sampling the Spectrum of Approaches*, Vol. 3948, eds H. I. Christensen, and H.-H. Nagel (Heidelberg: Springer), 25–36.
- Tsotsos, J. K. (2011). *A Computational Perspective on Visual Attention*. Cambridge, MA: MIT Press.
- Tsotsos, J. K., Culhane, S., Wai, W., Lai, Y., David, N., and Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artif. Intell.* 78, 507–547. doi:10.1016/0004-3702(95)00025-9
- Tulving, E. (1972). "Episodic and semantic memory," in *Organization of Memory*, eds E. Tulving, and W. Donaldson (New York: Academic Press), 381–403.
- Tulving, E. (1983). *Elements of Episodic Memory*. Oxford: Oxford University Press.
- Tulving, E. (1984). Précis of elements of episodic memory. *Behav. Brain Sci.* 7, 223–268. doi:10.1017/S0140525X0004440X
- Ulanowicz, R. E. (2000). *Growth and Development; Ecosystems Phenomenology*. Lincoln: Excel Press.
- Vernon, D. (2014). *Artificial Cognitive Systems – A Primer*. Cambridge, MA: MIT Press.
- Vernon, D., Metta, G., and Sandini, G. (2007). A survey of artificial cognitive systems: implications for the autonomous development of mental capabilities in computational agents. *IEEE Trans. Evol. Comput.* 11, 151–180. doi:10.1109/TEVC.2006.890274
- Vernon, D., von Hofsten, D., and Fadiga, L. (2010). *A Roadmap for Cognitive Development in Humanoid Robots, Cognitive Systems Monographs (COSMOS)*, Vol. 11, (Berlin: Springer).
- von Hofsten, C. (2004). An action perspective on motor development. *Trends Cogn. Sci.* 8, 266–272. doi:10.1016/j.tics.2004.04.002
- Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of small world networks. *Nature* 393, 440–442. doi:10.1038/30918
- Wintermute, S. (2012). Imagery in cognitive architecture: representation and control at multiple levels of abstraction. *Cogn. Syst. Res.* 1, 1–29. doi:10.1016/j.cogsys.2012.02.001
- Wood, R., Baxter, P., and Belpaeme, T. (2012). A review of long-term memory in natural and synthetic systems. *Adapt. Behav.* 20, 81–103. doi:10.1177/1059712311421219
- Zaharescu, A., Rothenstein, A. L., and Tsotsos, J. K. (2004). "Towards a biologically plausible active visual search model," in *Proceedings of the Second International Workshop on Attention and Performance in Computational Vision, WAPCV*, Vol. 3368, eds L. Paletta, J. K. Tsotsos, E. Rome, and G. Humphreys (Berlin: Springer), 133–147.
- Ziemke, T., and Lowe, R. (2009). On the role of emotion in embodied cognitive architectures: from organisms to robots. *Cogn. Comput.* 1, 104–117. doi:10.1007/s12559-009-9012-0

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Vernon, Beetz and Sandini. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.