



LUND UNIVERSITY

The current status of the simulation theory of cognition

Hesslow, Germund

Published in:
Brain Research

DOI:
[10.1016/j.brainres.2011.06.026](https://doi.org/10.1016/j.brainres.2011.06.026)

2012

[Link to publication](#)

Citation for published version (APA):

Hesslow, G. (2012). The current status of the simulation theory of cognition. *Brain Research*, 1428(Online 27 June 2011), 71-79. <https://doi.org/10.1016/j.brainres.2011.06.026>

Total number of authors:
1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

The current status of the simulation theory of cognition
Germund Hesslow

Department of Experimental Medical Science, University of Lund, Sweden

BMC F 10, S 221 84 LUND, Sweden. Fax: +46 46 2224546 (office).

E-mail address: germund.hesslow@med.lu.se.

Keywords:

Simulation – Memory – Cognition – Consciousness –Thought - Anticipation

ABSTRACT

It is proposed that thinking is simulated interaction with the environment. Three assumptions underlie this 'simulation' theory of cognitive function. Firstly, behaviour can be simulated in the sense that we can activate motor structures, as during a normal overt action, but suppress its execution. Secondly, perception can be simulated by internal activation of sensory cortex in a way that resembles its normal activation during perception of external stimuli. The third assumption ('anticipation') is that both overt and simulated actions can elicit perceptual simulation of their most probable consequences. A large body of evidence, mainly from neuroimaging studies, that supports these assumptions, is reviewed briefly. The theory is ontologically parsimonious and does not rely on standard cognitivist constructs such as internal models or representations. It is argued that the simulation approach can explain the relations between motor, sensory and cognitive functions and the appearance of an inner world. It also unifies and explains important features of a wide variety of cognitive phenomena such as memory and cognitive maps. Novel findings from recent developments in memory research on the similarity of imaging and memory and on the role of both prefrontal cortex and sensory cortex in declarative memory and working memory are predicted by the theory and provide striking support for it.

Contents

1. Introduction
2. Simulation of behavior
3. Simulation of perception
4. Anticipation
5. Explanatory power of simulation theory
 - 5.1. Do we need cognitive maps?
 - 5.2. Declarative memory
 - 5.2.1. Recall of long-term memory
 - 5.2.2. Temporally extended actions and working memory
 - 5.3. Evolution and the role of motor structures in cognitive function
 - 5.4. Some problems of consciousness
6. Concluding remark

1. Introduction

Cognitive function, the ability to think thoughts, is usually considered to be a uniquely human function or at least restricted to higher primates. It has therefore been puzzling that there does not seem to be a brain structure specialised for cognitive function. The basic design of the brain is very similar in all mammals and even if some areas, such as the prefrontal cortex and cerebellar hemispheres, have become larger, they do not seem to contain any radically new type of circuitry. The simulation theory ([Hesslow, 1994, 2002](#)) is an attempt to explain the emergence of cognitive function in terms of basic perceptual and motor functions. In essence, the theory says that thinking, or at least some very important kinds of thinking, involves the same processes as interaction with the external environment, but where actions are covert and generate, via associative brain mechanisms, the sensory inputs that elicit further actions.

The theory has three components:

1) Simulation of behaviour. Pre-motor areas in the frontal lobes can be activated in a way that is similar to the way they are normally activated during movements, but where the chain of neural activity is interrupted before it results in overt behaviour.

2) Simulation of perception. Activity in sensory cortex that is similar to that which occurs during perception of external stimuli can be elicited from other parts of the brain. Imagining and recalling things seen, heard or felt is essentially the same kind of processes as actually seeing, hearing or feeling something.

3) Anticipation. We must also assume the existence of associative mechanisms in the brain, which enables both behavioural and perceptual activity to elicit activity in the sensory areas of the brain. A special, but important, case of this is that a simulated action can elicit perceptual activity that resembles the activity that would have occurred if the action had actually been performed. Such internally generated “simulated perceptions” can function as stimuli for new covert behaviour.

The simulation theory is a combination of ideas with a long history, going back to associationist philosophers such as [David Hume \(1739\)](#) and [Alexander Bain \(1868\)](#) and similar ideas have been suggested by several modern authors (too numerous to be credited here but see for instance [Grush \(2004\)](#) or [Decety and Grèzes \(2006\)](#)). Since the term ‘simulation’ has been used in several quite different theories, let me briefly point out a couple of distinctive features of the simulation theory proposed here.

Firstly, it is based on purely associative mechanisms and does not appeal to symbols, representations or internal models as has become common in theories of motor control ([Wolpert and Ghahramani, 2000](#)) and consciousness ([Holland and Goodman, 2003](#)). The simulation theory could even be integrated into a behaviourist framework ([Catania, 1998](#); [Donahoe and Palmer, 1994](#)). This does not mean that the theory is inconsistent with representations or internal models. Indeed, if an “internal model is simply a pattern of synaptic weights that happens to give the correct outputs”, the associations underlying anticipation could be said to explain how an internal model is constructed as suggested by [Holland and Goodman \(2003\)](#). However, since appeal to internal models, even in this ontologically harmless sense,

tends, like “dormitive virtues”, to suggest more explanatory power than it really has ([van Gelder, 1995](#)), I have tried to avoid this concept altogether. Secondly, it is sometimes suggested, implicitly or explicitly, that simulation (or emulation) is performed by a brain system, the cerebellum for instance, that is separate from those directly involved in movement and perception ([Ramnani, 2006](#); [Wolpert et al., 1998](#)). It is an essential part of the present argument that simulation is performed by the same neural mechanisms as those normally involved in movement and perception.

Thirdly, the expression ‘simulation theory’ has sometimes been used for the idea that understanding other minds rests on our ability to simulate the situation of another person ([Gordon, 1986](#)). The simulation theory defended in the present paper is a different idea, but the mechanism it suggests is likely to be operative when we understand other minds ([Decety and Chaminade, 2003](#); [Decety and Grèzes, 2006](#)). In the rest of this paper I will attempt to a) review direct evidence for the three assumptions of the simulation theory, b) show that it explains and unifies a wide range of cognitive phenomena, thereby providing further indirect evidence and c) show that it provides plausible answers to some of the classical philosophical problems of consciousness.

2. Simulation of behavior

It is an old idea, expressed for instance by [Bain \(1868\)](#) that thinking is essentially a ‘weak’ form of behaviour that might also be called “preparatory”, “incipient” or “covert”. Recent developments in neuroscience allow us to formulate the idea more precisely and also to see that there is quite strong evidence for it.

Complex actions can be described as hierarchically organised, such that a general goal, and a corresponding action, consist of more specific sub-goals and component responses down to the simple muscle contractions. The action of making a sandwich consists of components such as slicing the bread, spreading the butter and placing a slice of ham on the bread. Slicing the bread in turn consists of a complicated combination of elementary movements down to the contractions of the single muscle cells.

This hierarchy corresponds to a hierarchical organisation of the frontal lobes ([Fuster, 2004, 2008](#); [Koechlin and Jubault, 2006](#)). Actions are generated in the frontal lobes in response to input from the sensory cortex and motivational structures. Cells in the primary motor cortex signal to single muscles or muscle groups, for instance that a finger should be bent or a hand raised. These cells form the lowest layer of the hierarchy. In order to generate a more complex movement, such as grasping and lifting an object, many muscles and hence many cell groups in the primary motor cortex must be activated in a particular temporal sequence. The command signal for this sequence comes from more anterior areas in the frontal lobes with assistance from the cerebellum and basal ganglia. Put simply, the more anterior we go, the more global the behavioural command signals become and the higher up in the neural hierarchy we get.

Saying that behaviour can be simulated here means nothing more than that the signal flow from the prefrontal cortex via the pre-motor areas may occur even if it is interrupted before it activates the primary motor cortex and results in overt behaviour. A simulated action is thus essentially a suppressed or unfinished action. There is now an impressive body of evidence to support this hypothesis. Since it has been extensively reviewed elsewhere ([Bonnet et al., 1997](#); [Decety, 1996](#); [Jeannerod, 1994](#); [Jeannerod and Frak, 1999](#); [Kosslyn et al., 2001](#); [Moulton and Kosslyn, 2009](#)), I will only give a brief summary here. Many early behavioural experiments have demonstrated parallels between simulated and actual movements. For instance, temporal features of simulated movements correspond to those of actual movements, such as walking blindfolded in familiar room ([Decety et al., 1989](#)) or performing 'mental' rotation of three-dimensional objects ([Cooper et al., 1973](#); [Shepard and Metzler, 1971](#)). These experiments can be interpreted in many ways and at best provide suggestive evidence for a similarity between "mental" and physical movements.

More compelling evidence has been obtained with imaging techniques. More than 30 years ago, [Ingvar and Philipsson \(1977\)](#) showed that subjects instructed to simulate and physically perform hand movements had increased neural activity in the anterior parts of the frontal lobes, whilst only the second group activated the primary motor cortex. Although some subtle differences between imagined and executed movements have been found ([Deiber et al., 1998](#)), later studies using PET and fMRI have also found activation of pre-motor and supplementary motor areas during imagined movements ([Decety et al., 1994](#); [Lotze et al., 1999](#); [Rao et al., 1993](#)). In one recent study of pianists it was shown that playing and imagining playing a particular piece of music would activate the same areas of frontal and parietal cortex ([Meister et al., 2004](#)).

A striking piece of evidence for the same thing is the recent case study of a patient with bilateral parietal cortex lesions ([Schwoebel et al., 2002](#)). When the patient was asked to only imagine a hand movement, the movement was nevertheless performed although the subject was unaware of this. This case clearly suggests that an imagined movement is the same as an overt one, except that the overt performance is suppressed in the former case. Another type of evidence for this is finding that when a subject imagines a movement, the thresholds for exciting the relevant pyramidal tract neurons with transcranial magnetic stimulation is decreased ([Fadiga et al., 1999](#)). Imagining the movement thus seems to activate, though weakly, even the final motor output from the cortex. It is important to realise that simulated actions, like overt ones, may be quite "abstract". When imagining that I am drawing a triangle or travelling abroad, I do not need to imagine all the component movements. I can imagine going to Paris in a single step so to speak, and perceptual consequences, such as seeing the Eiffel tower, will immediately appear.

3. Simulation of perception

The idea that the sensory parts of the brain can be activated from "within", that is, without any input from the sense organs was suggested by Hume and was quite

common in the 19th century. [James \(1890\)](#) wrote the “commonly received idea is that [imagination] is only a milder degree of the same process which took place when the thing now imagined was sensibly perceived”. This hypothesis was advanced without any real empirical support, but today there is extensive evidence for it from cognitive psychology and neuroscience. This evidence has been reviewed extensively elsewhere ([Farah, 1988](#); [Hurley, 2008](#); [Kosslyn et al., 2001](#); [Moulton and Kosslyn, 2009](#)) and few examples will suffice here. Many behavioural studies, for instance the famous mental rotation experiments of Shepard and Metzler ([Shepard and Metzler, 1971](#)) have been interpreted as evidence that images have visual properties and therefore probably utilise the same mechanisms as the visual system ([Kosslyn, 1994](#)). The experiments are hard to evaluate, however, and alternative interpretations have been suggested ([Pylyshyn, 1984, 2003](#)).

More compelling support for the idea, that imaging utilises the same mechanisms as perception, has been obtained by measuring activity in various parts of the brain when subjects imagine a stimulus. In one of the earliest of these experiments, the electrical activity was measured in both the visual cortex and in the cortical area receiving tactile information from the arm whilst subjects imagined various visual and tactile stimuli. When the task was to imagine light flashes, nervous activity increased specifically in the visual cortex. When the subjects imagined someone touching their arm instead, activity increased in the part of somatosensory cortex that receives input from the arm ([Davidson and Schwartz, 1977](#)). Experiments using modern imaging techniques have subsequently confirmed that the primary visual cortex is strongly activated when we imagine a visual stimulus or recall a visual memory ([Kosslyn et al., 1993](#); [Le Bihan et al., 1993](#)) and that auditory cortex is activated when sounds are imagined ([Schurmann et al., 2002](#); [Zatorre et al., 1996](#)).

The parallels actually go quite far. Perceiving different objects, such as faces, houses and chairs, will elicit slightly different activation patterns. These differences turn up also when subjects are imagining the same objects ([Ishai et al., 2000](#)). It is a critical assumption of the theory that perceptual simulations can function as stimuli. This is supported by the fact that perceptual simulations can have physiological effects that resemble those of actual perception. Anyone can generate a certain degree of anxiety by thinking about a traumatic experience. Subjects who imagine leg exercise increase their heart rate and respiration rate in proportion to the imagined effort ([Decety et al., 1991](#)). There is also evidence that emotional responses involved in feelings of sympathy are elicited by a simulation mechanism ([Decety and Chaminade, 2003](#); [Decety and Grèzes, 2006](#)).

Perceptual simulation may also be able to induce motor learning. If a subject tries to point to a target that is seen through laterally displacing prisms, there will be an initial pointing error. With repeated pointing the nervous system quickly adapts and the errors disappear. In a series of rigorously controlled experiments, [Finke \(1979\)](#) has shown that merely imagining the pointing errors is sufficient for this adaptation to occur.

4. Anticipation

A widely held view in psychology has been that a Pavlovian conditioning process can generate associations between different sensory modalities and there is evidence supporting this. For instance, when a visual word stimulus has been paired with a sound, later presentation of the visual stimulus elicits activity in the auditory cortex (Nyberg et al., 2000). Although such associations are no doubt important, I would like to focus here on a different source of perceptual simulation, namely actions generated in the forebrain. What we perceive is not only determined by the external world but also by our own behaviour. Visual input changes when we move our heads. Tactile stimulation is generated in the feet by walking and in the hands by manipulating objects. The sensory consequences of actions are to a large extent predictable and it would be extremely wasteful if an organism did not make use of this in anticipating the consequences of its behaviour. This was clearly understood by writers like Hume (1739) and Bain. The latter wrote extensively about the “bond of association ... between our own actions and the sensible effects that follow from them” (Bain, 1868, p. 427). Suppose that we have an organism that frequently performs the response R1 in the situation S1 and that this regularly results in the new situation S2, which in turn usually elicits R2. The implication of Bain's suggestion, illustrated in Fig. 1, is that the regularity, R1 leads to S2, can be learned, so that performing the neural activity r1 in the presence of s1, elicits the “expected” activity s2. In other words, the preparatory stages of a behavioural response can elicit sensory activity that resembles the activity that would normally be caused by the completed overt behaviour. Notice that the proposal here is not only that the final motor command from the motor cortex, what is sometimes called an “efference copy”, is signalled to the sensory cortex as might be suggested by Fig. 1. It is a crucial point that all preparatory stages, also the early global command signals (“going to Paris”), can elicit sensory simulation. Once the mechanism of anticipation is in place, there is nothing to prevent the appearance of long chains of simulated responses and perceptions. A simulated action in the frontal lobe generates a simulated perception of its probable consequences in the sensory cortex. This activity may serve as a stimulus for new action and so on. Such a chain of simulated stimuli and actions can in principle continue indefinitely or until it is terminated by some external stimulus.

It is easy to overlook the dramatic improvement in the purposefulness of behaviour that this mechanism would afford an organism. The obvious advantage is that it will enable it to respond in advance of various dangers that its own behaviour might cause. Less obvious is that it enables the organism to test out the consequences of an action or a course of action in advance.

It is also easy to underestimate the creative potential of simulating chains of behaviour and perception. I can imagine doing many things that I have never done before by using similar behaviour as building blocks. For instance, I have never flown an airplane but I have sat in front of instrument panels, pulled levers, looked out of airplane windows and seen similar things on film. I have never been another person, but I can imagine the situation of someone else and use the ensuing simulated behaviours and sensory consequences help to predict what that person would do or experience. Thus, the theory does not limit sensory simulations to those that have been experienced in exactly the same form and following exactly the same behaviour

as previously performed. There is not much direct evidence for the anticipation mechanism suggested here, but there is plenty of indirect evidence. To begin with, there is a plausible neural substrate for such a mechanism in the form of an extensive fibre projection from the frontal lobe to all parts of sensory cortex ([Pandya et al., 1985](#)). The functions of these pathways are not well understood, but there is physiological evidence from monkeys that neurons in polysensory cortex can be modulated by movement ([Hietanen and Perrett, 1996](#)).

Some interesting evidence for movement-elicited perceptual simulation comes from psychological experiments, which suggest that imagining movement has sensory consequences and is a crucial mechanism in many forms of problem solving. For instance [Wexler et al. \(1998\)](#) found that manual rotation of a visual object interfered with mental rotation in a way that suggested that the perceived mental rotation was controlled by the simulated movement. [Schwartz and Black \(1999\)](#) reported that subjects who were asked to predict at what level of tilting a water-filled glass would spill over, they did better when they imagined actually moving the glass. Clinical observations also support the anticipation hypothesis. Patients suffering from hemiplegia after brain damage are sometimes unaware of their handicap. When trying to move their arms, some of these patients apparently feel their arms moving although no movement occurs ([Feinberg et al., 2000](#); [Heilman et al., 1998](#)).

Another line of evidence comes from recent work on eye movements as a tool for studying mental imagery ([Johansson et al., 2006](#)). During the last decade, a large number of studies have been published that show that recalling a visual memory is associated with eye movements that reflect the structure of the recalled scene ([Brandt and Stark, 1997](#); [Laeng and Teodorescu, 2002](#)). If subjects are asked to memorise a scene with a number of different objects and then asked to recall it whilst looking at a blank screen, the eyes tend to move as if they were scanning the scene and as if the subjects were sequentially focusing on the various objects. Eye movements also reflect the position of objects when subjects imagine a novel scene based on a spoken description ([Johansson et al., 2006](#)). These results support, or at least are consistent with, the hypothesis that perceptual simulation is driven by (covert or overt) behaviour. Given the similarity between imagining and recalling memories (see below), all the evidence that supports a role for behavioural simulation in memory recall also indirectly supports a similar role in imagination. Some evidence could be taken to suggest that frontal lobe activity suppresses sensory activity rather than elicits it. It is known, for instance that cortical responses to stimuli that are generated by self-initiated movements, such as speech or tickling, are suppressed ([Blakemore et al., 1998](#); [Martikainen et al., 2005](#)). However, in these cases the cortical response to the consequences of a movement comes after completion of the movement and does not contradict the assumption of an earlier anticipatory simulated perception.

5. Explanatory power of simulation theory

The simulation theory can offer unified view of cognitive function and provide novel treatments of a number of issues in cognitive science. It is not possible to deal with all such applications within this review, but a few examples will be discussed in [Section 5.1–5.4](#).

5.1. Do we need cognitive maps?

An idea introduced by Tolman and still popular in cognitive science is that both animals and humans form “cognitive maps” of their environment ([Tolman, 1948](#)). In a classic experiment, [Tolman and Gleitman \(1949\)](#) let a rat freely explore a T-maze with a dark goal box in the left arm of the maze and a light goal box to the right. Both boxes contained food. The rat was then placed in dark chamber, similar to the left goal box, and subjected to anxiety producing electrical foot shocks. When the rat was later placed in the T-maze, it went directly to the right goal box and never entered the left arm of the maze, in spite of the fact that it had never been punished for doing so. It seemed as if the rat had access to a picture or a map of the maze and could infer that an unpleasant experience awaited it in the left goal box. But there is a simpler explanation for this experiment. When the rat reaches the choice point, it will sometimes begin walking to the left. When this behaviour is still at the preparatory stage, it elicits the usual perceptual consequences, learned during the initial exploration phase of the experiment, that is, the sight of the dark goal box. This in turn elicits conditioned anxiety, which suppresses completion of the initiated behaviour. This is a fairly concrete and unproblematic sense in which the rat “anticipates” the consequences of walking to the left.

It is actually doubtful that human navigation depends on cognitive maps. [Moeser \(2009\)](#) studied student nurses who had learned to find their way in a complex hospital building . Even after traversing the building for 2 years, the nurses had failed to form “survey”-type cognitive maps, yet could navigate quite effectively in the building. It is an essential aspect of the map metaphor, that knowledge of routes is symmetric, that is, if I know the way from A to B, this automatically entails knowledge of the route from B to A. However, there is evidence that knowledge of routes is learned asymmetrically, so that a route can be followed in one direction but not in the other ([Kuipers, 1982](#)). Even if this does not exclude that humans may under some circumstances develop and use cognitive maps, it does suggest that some simpler mechanism exists that can achieve similar results.

[Chrisley \(1990\)](#) has shown how a robot using a connectionist network can learn “to predict what sensations it would have if it were to move in a particular way”, that is the anticipation mechanism described above. The robot can use this knowledge to generate the equivalent of a “cognitive map”. Since this is based on a simple connectionist principle, very similar to what I have called anticipation, it is perfectly compatible with the view developed here. Thus, although Chrisley appeals to “models of the environment” and “cognitive maps”, what he actually shows is that these notions can be treated metaphorically and that a map can be derived from a set of predictive associations. If an organism has acquired a set of associations such as “movement m1 during sensation a will be followed by sensation b” and “m2 during b will be followed by c”, it can simulate the sequence m1 plus m2 and predict the final outcome. The organism can then behave as if it was consulting a map. Connectionist

models based on the novel Associative Self-Organizing Map ([Johnsson et al., 2011](#)) that capture this idea have been explored in [Johnsson et al. \(2010\)](#). That the anticipation mechanism can actually enable robots to navigate in a simple environment has recently been demonstrated in robot experiments ([Jirenhed et al., 2001](#); [Ziemke et al., 2005](#)).

5.2. Declarative memory

5.2.1. Recall of long-term memory

There is no clear distinction between imagery and recall of memory and both can be interpreted as perceptual simulation. If I imagine that I am walking around in a familiar city, I am also recalling memories of the city. Even when we are imagining things that we have never experienced, we are using remembered experiences as building blocks (which is an important source of our creative ability). The difference between imagery and memory recall has less to do with the properties of the simulated perceptions than with the behaviour that elicits them. Two important predictions follow from this. Firstly, we should expect that the frontal lobes are instrumental in eliciting memories. We recall what a particular place looks like by walking around and looking, that is, by performing the appropriate (possibly abstract) behaviour. We may recall Eiffel tower by “going to Paris”. Secondly, we should expect similarities in the neural activities involved in imagination and memory recall. Both predictions have been amply confirmed during the last decade.

The idea that declarative memory is a kind of reactivation of the sensory activity that occurred at the time the memory was encoded ([Fuster, 1997](#)) is now quite common and supported by direct evidence ([Nyberg et al., 2000](#); [Slotnick, 2004](#)). The prefrontal cortex plays a crucial role in activating declarative memories, a prediction for which there is also extensive and compelling evidence ([Badre and Wagner, 2007](#)). [Kent and Lamberts \(2008\)](#) have recently summarised the evidence for the thesis that memory retrieval is elicited by mental simulation. One of the sources of evidence they point to is the eye movement research mentioned above in [Section 4](#). It is noteworthy that not only does this work support the ability of eye movements to elicit imagery in general. The correlation between eye movement paths during encoding and during recall in turn correlated with accuracy in a memory test ([Laeng and Teodorescu, 2002](#)).

There is also now extensive evidence for the claim that imagining and recalling memories are similar neuronal processes. For instance, amnesic patients with hippocampal damage have a markedly impaired ability to imagine new experiences ([Hassabis et al., 2007b](#)). Imaging studies have confirmed a large overlap in the brain areas activated by imagining and recalling memories ([Addis et al., 2007](#); [Hassabis et al., 2007a](#); [Schacter et al., 2008](#); [Szpunar et al., 2007](#)). The overlap is not restricted to sensory cortex but also other areas such as the prefrontal cortex, the hippocampus and the precuneus ([Cavanna and Trimble, 2008](#)).

5.2.2. Temporally extended actions and working memory

In a typical test of working memory, a monkey observes an experimenter place a food morsel in one of two food wells. Identical cards cover the wells and a screen is lowered in front of the monkey. After several seconds, the screen is raised and the animal is allowed to select one of the food wells. The usual interpretation is that the delay forces the monkey to keep the “information” about the location of the food in its “working memory” which is usually thought of as a specialised structure located in the prefrontal cortex (Goldman-Rakic et al., 1999). According to what has been called the “standard model” (Postle, 2006), working memory depends on specialised neural circuits in the prefrontal cortex that can act as a temporary buffer system. But there is a more parsimonious alternative account that views working memory as an emergent effect of the frontal lobe action system in combination with associated perceptual simulations in sensory cortex.

Many actions can be extended in time. I can hold a cup of coffee in my hand, keep on walking or just stand still for long periods. There is no reason why simulated actions could not be similarly extended. If so, the sensory activity generated by a sustained action would also be extended in time. When the monkey sees the food being covered, it prepares an appropriate action. As long as it is preparing to lift the card, it will “see” the food and hence “know” where it is hidden. Working memory may thus be construed as an emergent property of a set of associations between frontal and sensory cortices (for a more detailed exposition and defence of this view, see Postle (2006)). No special storage device is necessary to understand this phenomenon. The simulation account explains why the prefrontal cortex should be critical for working memory and also why working memory tasks should involve sensory cortex, as a growing body of recent evidence suggests (Fuster, 1997; Postle, 2006; Ruchkin et al., 2003). Rather than being a “buffer store” for “information”, the “Prefrontal cortex provides the pointer system for maintaining activation in the appropriate posterior processing systems” (Ruchkin et al., 2003).

5.3. Evolution and the role of motor structures in cognitive function

It is often assumed that human cognitive abilities depend on specialised neural circuits, the most well-known example being the idea of an in-born universal grammar. It is difficult to see how such circuits could arise without fairly extensive changes in the brain and large evolutionary leaps. Conscious thought seems to be performed by brain structures, which existed long before anything approaching human intellectual function had appeared on the evolutionary scene. The motor cortex, the sensory cortex, the basal ganglia, cerebellum, hippocampus etc. are present in all mammals. These structures vary in their relative size and some variations in the micro-circuitry also exist, but the basic design of the brain is the same in humans as in monkeys, rats and cows. Thus, cognitive function must be explained in terms of neural circuits, which were designed by evolution to enable animals to move around and find shelter, mates and food. There may of course be quantitative differences in various respects such as in the speed, precision or storage capacity of associative mechanisms or differences in motivation and consequently in

how effectively various kinds of behaviour are learned. These are matters of degree, however, not of principle or of basic mechanism. If one accepts this argument, it is a strong point of the simulation theory that it explains the appearance of cognitive functions without postulating any evolutionarily novel mechanisms.

The simulation theory also makes sense of the accumulating data showing that “motor” structures such as the cerebellum and the basal ganglia appear to be involved in cognitive tasks. The cerebellum is activated during imagined movements (Decety et al., 1990; Ryding et al., 1993), in the Tower of London task (Baker et al., 1996) and during mental rotation (Parsons et al., 1995, 1997). Cerebellar lesions seem to cause various forms of cognitive impairments (Dominey et al., 1995; Schmahmann et al., 1997). We saw above that imagining movement errors can elicit cerebellar learning. These findings should not surprise us.

Since simulated movements can be appropriately timed and coordinated, they will need the same kind of assistance from the cerebellum as overt movement and various sequences of frontal lobe activity in simulated movements should activate the cerebellum in the same way regardless of whether the movement is completed or remains covert. On the perceptual side, if simulated error perception is sufficiently similar to actual error perception, it should be able to send the same signals to the cerebellum and induce cerebellar learning.

Similarly, the basal ganglia, particularly the striatum, are also activated during various cognitive tasks such as the Tower of London (Baker et al., 1996; Dagher et al., 1999) and performance in this task is impaired or altered in patients with Parkinson's disease (Dagher et al., 2001; Morris et al., 1988).

5.4. Some problems of consciousness

There are several problems of consciousness and many of them currently beyond the reach of neuroscience. Nevertheless, the simulation mechanism can provide plausible answers to a couple of these problems (Hesslow, 1994; Hesslow and Jirenhed, 2007).

One of the classical puzzles is the nature of mental objects, the ‘mind-body problem’ in philosophy. How is it possible that there can exist objects in the mind that lack physical properties? What could they be made of, if they are not physical? One answer is that they do not need to exist in order to be visible. I can see people in a television set or in a mirror. The reason is not that there are ‘mental’ objects in the TV or in the mirror, but that these things can generate light patterns on my retina that closely resemble those generated by real people. Similarly, parts of the brain can elicit activity patterns in my visual cortex that resemble those normally caused by impulses from the eyes. It should not be puzzling therefore, that we can see things without having to assume that there must be something there to be seen. In his attempt to identify what he terms “the hard problem” of consciousness, David Chalmers (1995) wrote “It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises. Why should

physical processing give rise to a rich inner life at all?" Other formulations suggest that Chalmers means more than this, but finding out the physiological mechanisms that generate our inner world is certainly also one of the important problems of consciousness.

It should by now be obvious that the appearance of an inner world is an unavoidable consequence of the simulation process. If covert behaviour can generate perceptual activity, which resembles the activity generated by perception of the external world, then one will be able to do something that resembles acting and perceiving the consequences of actions without actually interacting with the external world.

6. Concluding remark

In addition to the scientific evidence for the three constituent hypotheses of the simulation theory, there are some general considerations that speak strongly in its favour. One of these is that the theory is ontologically parsimonious; it is formulated in simple cause and effect terms and does not rely on controversial constructs like 'internal models' or 'representations'. It also makes evolutionary sense. The associative mechanisms assumed by the theory are very likely present in all mammals. As noted above the theory can make sense of many cognitive phenomena and provide a plausible account of their evolution without assuming any dramatic leaps or the appearance of specialised cognitive circuitry.

Acknowledgments

This work was supported by grants from the Swedish Research Council, the Linnaeus Centre for Cognition, Communication and Learning at Lund University and the Söderberg and Åhlen foundations.

REFERENCES

- Addis, D.R., Wong, A.T., Schacter, D.L., 2007. Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia* 45, 1363–1377.
- Badre, D., Wagner, A.D., 2007. Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* 45, 2883–2901.
- Bain, A., 1868. *The Senses and the Intellect*. Longmans, Green & Co., London.
- Baker, S.C., Rogers, A.D., Owen, A.M., Frith, C.D., Dolan, R.J., Frackowiak, R.S., Robbins, T.W., 1996. Neural systems engaged by planning: a PET study of the Tower of London task. *Neuropsychologia* 34, 515–526.
- Blakemore, S.-J., Wolpert, D.M., Frith, C.D., 1998. Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640.
- Bonnet, M., Decety, J., Requin, J., Jeannerod, M., 1997. Mental simulation of an action modulates the excitability of spinal reflex pathways in man. *Cogn. Brain Res.* 5, 221–228.
- Brandt, S.A., Stark, L.W., 1997. Spontaneous eye movements during visual imagery reflect the content of the visual scene. *J. Cogn. Neurosci.* 9, 27–38.

Catania, C., 1998. Learning. Prentice-Hall, Upper Saddle River. Cavanna, A.E., Trimble, M.R., 2008. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129, 564–583. Chalmers, D.J., 1995. Facing up to the problem of consciousness. *J. Conscious. Stud.* 2, 200–219. Chrisley, R., 1990. Cognitive map construction and use: a parallel distributed processing approach. In: Touretzky, D., Elman, J., Hinton, G., Sejnowski, T. (Eds.), *Connectionist Models: Proceedings of the 1990 Summer School*. Morgan Kaufman, San Mateo, CA, pp. 287–302. Cooper, L.A., Shepard, R.N., 1973. Chronometric Studies of the rotation of mental images. In: Chase, W.G. (Ed.), *Visual Information Processing*. Academic Press, New York. Dagher, A., Owen, A.M., Boecker, H., Brooks, D.J., 1999. Mapping the network for planning: a correlational PET activation study with the Tower of London task. *Brain* 122, 1973–1987. Dagher, A., Owen, A.M., Boecker, H., Brooks, D.J., 2001. The role of the striatum and hippocampus in planning: a PET activation study in Parkinson's disease. *Brain* 124, 1020–1032. Davidson, R.J., Schwartz, G.E., 1977. Brain mechanisms subserving self-generated imagery: electrophysiological specificity and patterning. *Psychophysiology* 14, 598–601. Decety, J., 1996. The neurophysiological basis of motor imagery. *Behav. Brain Res.* 77, 45–52. Decety, J., Chaminade, T., 2003. Neural correlates of feeling sympathy. *Neuropsychologia* 41, 127–138. Decety, J., Grèzes, 2006. The power of simulation: imagining one's own and other's behavior. *Brain Res.* 1079, 4–14. Decety, J., Jeannerod, M., Prablanc, C., 1989. The timing of mentally represented actions. *Behav. Brain Res.* 34, 35–42. Decety, J., Sjöholm, H., Ryding, E., Stenberg, G., Ingvar, D.H., 1990. The cerebellum participates in cognitive activity: tomographic measurements of regional cerebral blood flow. *Brain Res.* 535, 313–317. Decety, J., Jeannerod, M., Germain, M., Pastene, J., 1991. Vegetative response during imagined movement is proportional to mental effort. *Behav. Brain Res.* 42, 1–5. Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., Mazziotta, J.C., Fazio, F., 1994. Mapping motor representations with positron emission tomography. *Nature* 371, 600–602. Deiber, M.P., Ibanez, V., Honda, M., Sadato, N., Raman, R., Hallett, M., 1998. Cerebral processes related to visuomotor imagery and generation of simple finger movement studied with positron emission tomography. *Neuroimage*. 7, 73–85. Dominey, P.F., Decety, J., Broussolle, E., Chazot, G., Jeannerod, M., 1995. Motor imagery of a lateralized sequential task is asymmetrically slowed in hemi-Parkinson patients. *Neuropsychologia* 33, 727–741. Donahoe, J.W., Palmer, D.C., 1994. Learning and Complex Behavior. Allyn & Bacon, Needham Heights. Fadiga, L., Buccino, G., Craighero, L., Fogassi, L., Gallese, V., Pavesi, G., 1999. Corticospinal excitability is specifically modulated by motor imagery. A magnetic stimulation study. *Neuropsychologia* 37, 147–158. Farah, M.J., 1988. Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychol. Rev.* 95, 307–317. Feinberg, T.E., Roane, D.M., Ali, J., 2000. Illusory limb movements in anosognosia for hemiplegia. *J. Neurol. Neurosurg. Psychiatry* 68, 511–513. Finke, R.A., 1979. The functional equivalence of mental images and errors of movement. *Cogn. Psychol.* 11, 235–264. Fuster, J.M., 1997. Network memory. *Trends Neurosci.* 20, 451–459. Fuster, J.M., 2004. Upper processing stages of the perception-action cycle. *Trends Cogn. Sci.* 8, 143–145. Fuster, J.M., 2008. *The Prefrontal Cortex*. Elsevier, Amsterdam. Goldman-Rakic, P.S., 1999. Working Memory, Neural Basis of. In: Wilson, R.A., Keil, F.C. (Eds.), *The MIT Encyclopedia of the Cognitive Sciences*. MIT Press, Cambridge, Mass., pp. 890–894. Gordon, R., 1986. Folk psychology as simulation. *Mind Lang.* 1, 158–171. Grush, R., 2004. The emulation theory of representation: motor control, imagery and perception. *Behav. Brain Sci.* 27, 377–

442. Hassabis, D., Kumaran, D., Maguire, E.A., 2007a. Using imagination to understand the neural basis of episodic memory. *J. Neurosci.* 27, 14365–14374.

Hassabis, D., Kumaran, D., Vann, S.D., Maguire, E.A., 2007b. Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl Acad. Sci.* 104, 1726–1731.

Heilman, K.M., Barrett, A.M., Adali, T., 1998. Possible mechanisms of anosognosia: a defect in self-awareness. *Philos. Trans. R. Soc. Lond. B* 353, 1903–1909.

Hesslow, G., 1994. Will neuroscience explain consciousness? *J. Theor. Biol.* 171, 29–39.

Hesslow, G., 2002. Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* 6, 242–247.

Hesslow, G., Jirnhed, D.-A., 2007. The inner world of a simple robot. *J. Conscious. Stud.* 14, 85–96.

Hietanen, J.K., Perrett, D.I., 1996. Motion sensitive cells in the macaque superior temporal polysensory area: response discrimination between self-generated and externally pattern motion. *Behav. Brain Res.* 76, 155–167.

Holland, O., Goodman, R., 2003. Robots with internal models — a route to machine consciousness? *J. Conscious. Stud.* 10, 77–109.

Hume, D., 1739. *A Treatise of Human Nature*. Oxford University Press, Oxford.

Hurley, S.L., 2008. The shared circuits model (SCM): how control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behav. Brain Sci.* 31, 1–58.

Ingvar, D.H., Philipsson, L., 1977. Distribution of the cerebral blood flow in the dominant hemisphere during motor ideation and motor performance. *Ann. Neurol.* 2, 230–237.

Ishai, A., Ungerleider, L., Haxby, J., 2000. Distributed neural systems for the generation of visual images. *Neuron* 28, 979–990.

James, W., 1890. *Principles of Psychology*. Macmillan. (republished by Dover 1950).

Jeannerod, M., 1994. The representing brain: neural correlates of motor intention and imagery. *Behav. Brain Sci.* 17, 187–245.

Jeannerod, M., Frak, V., 1999. Mental imaging of motor activity in humans. *Curr. Opin. Neurobiol.* 9, 735–739.

Jirnhed, D.A., Hesslow, G., Ziemke, T., 2001. Exploring internal simulation of perception in a mobile robot. *Lund University Cognitive Studies* 107–113.

Johansson, R., Holsanova, J., Holmqvist, K., 2006. Pictures and spoken descriptions elicit similar eye movements during mental imagery. Both in light and in complete darkness. *Cogn. Sci.* 30, 1053–1079.

Johnsson, M., Gil, D., Balkenius, C., Hesslow, G., 2010. Supervised Architectures for Internal Simulation of Perceptions and Actions. In: Hernandez, C., Gomez, J., Sanz, R. (Eds.), *Proceedings of Brain Inspired Cognitive Systems (BICS)*. Madrid.

Johnsson, M., Martinsson, M., Gil, D., Hesslow, G., 2011. Associative Self-OrganizingMap. In: Mwasiagi, J.I. (Ed.), *SelfOrganisingMaps— Applications and Novel Algorithm Design*. InTech, pp. 603–626.

Kent, C., Lamberts, K., 2008. The encoding-retrieval relationship: retrieval as mental simulation. *Trends Cogn. Sci.* 12, 92–98.

Koechlin, E., Jubault, T., 2006. Broca's area and the hierarchical organization of human behavior. *Neuron* 50, 963–974.

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

Kosslyn, S.M., Alpert, N.M., Thompson, W.L., Maljkovic, V., Weise, S.W., Chabris, C.F., Hamilton, S.E., Rauch, S.L., Buonanno, F.S., 1993. Visual mental imagery activates topographically organized visual cortex: PET investigations. *J. Cogn. Neurosci.* 5, 263–287.

Kosslyn, S.M., Ganis, G., Thompson, W.L., 2001. Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.

Kuipers, B., 1982. The “map in the head” metaphor. *Environ. Behav.* 14, 202–220.

Laeng, B., Teodorescu, D.S., 2002. Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cogn. Sci.* 26, 207–231.

Le Bihan, D., Turner, R., Zeffiro, T.A., Cuénod, C.A., Jezzard, P., Bonnerod, V., 1993. Activation of human primary visual cortex during visual recall: a magnetic resonance imaging study. *Proc. Natl Acad. Sci.* 90, 11802–11805.

Lotze, M., Montoya, P., Erb, M., Hüsmann, E., Flor, H., Klose, U., Birbaumer, N., Grodd, W.,

1999. Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *J. Cogn. Neurosci.* 11, 491–501. Martikainen, M.H., Kaneko, K.i., Hari, R., 2005. Suppressed responses to self-triggered sounds in the human auditory cortex. *Cereb. Cortex* 15, 299–302. Meister, I.G., Krings, T., Foltys, H., Müller, M., Töpfer, R., Thron, A., 2004. Playing piano in the mind — an fMRI study on music imagery and performance in pianists. *Cogn. Brain Res.* 19, 219–228. Moeser, S.D., 2009. Cognitive mapping in a complex building. *Environ. Behav.* 20, 21–49. Morris, R.G., Downes, J.J., Sahakian, B.J., Evenden, J.L., Heald, A., Robbins, T.W., 1988. Planning and spatial working memory in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 51, 757–766. Moulton, S.T., Kosslyn, S.M., 2009. Imagining predictions: mental imagery as mental emulation. *Philos. Trans. R. Soc. Lond. B* 364, 1273–1280. Nyberg, L., Habib, R., McIntosh, Aron, Tulving, E., 2000. Reactivation of encoding-related brain activity during memory retrieval. *Proc. Natl Acad. Sci.* 97, 11120–11124. Pandya, D.N., Yeterian, E.H., 1985. Architecture and connections of cortical association areas. In: Peters, A., Jones, E.G (Eds.), *Cerebral Cortex*, Vol. 4. Plenum Press, New York, pp. 3–61. Parsons, L.M., Fox, P.T., Downs, J.H., Glass, T., Hirsch, T.B., Martin, C.C., Jerabek, P.A., Lancaster, J.L., 1995. Use of implicit motor imagery for visual shape discrimination as revealed by PET. *Nature* 375, 54–58. Parsons, L.M., Fox, P.T., Schmahmann, J.D., 1997. Sensory and cognitive functions. In: Schmahmann, J.D. (Ed.), *The Cerebellum and Cognition*. J.D. Schmahmann, ed. Academic Press, New York, pp. 255–271. Postle, B.R., 2006. Working memory as an emergent property of the mind and brain. *Neuroscience* 139, 23–38. Pylyshyn, Z.W., 1984. *Computation and Cognition*. MIT Press, Cambridge, MA. Pylyshyn, Z.W., 2003. *Mental Imagery: In search of a theory*. *Behav. Brain Sci.* 25, 157–182. Ramnani, N., 2006. The primate cortico-cerebellar system: anatomy and function. *Nat. Rev. Neurosci.* 7, 511–522. Rao, S.M., Binder, J.R., Bandettini, P.A., Hammeke, T.A., Yetkin, F.Z., Jesmanowicz, A., Lisk, L.M., Morris, G.L., Mueller, W.M., Estikowski, L.D., Wong, E.C., Haughton, V.M., Hyde, J.S., 1993. Functional magnetic resonance imaging of complex human movements. *Neurology* 43, 2311–2318. Ruchkin, D.S., Grafman, J., Cameron, K., Berndt, R.S., 2003. Working memory retention systems: a state of activated long-term memory. *Behav. Brain Sci.* 26, 709–777. Ryding, E., Decety, J., Sjöholm, H., Stenberg, G., Ingvar, D.H., 1993. Motor imagery activates the cerebellum regionally: a SPECT rCBF study with 99m Tc-HMPAO. *Cogn. Brain Res.* 1, 94–99. Schacter, D.L., Addis, D.R., Buckner, R.L., 2008. Episodic simulation of future events: concepts, data, and applications. *Ann. N. Y. Acad. Sci.* 1124, 39–60.

Schmahmann, J.D., et al., 1997. *The Cerebellum and Cognition*. Academic Press, New York.

Schurmann, M., Raij, T., Fujiki, N., Hari, R., 2002. Mind's ear in a musician: where and when in the brain. *Neuroimage* 16, 434–440.

Schwartz, D.L., Black, T., 1999. Inferences through imagined actions. Knowing by simulated doing. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 116–136.

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

- Shepard, R.N., Metzler, J., 1971. Mental rotation of three-dimensional objects. *Science* 171, 701–703.
- Slotnick, S.D., 2004. Visual memory and visual perception recruit common neural substrates. *Behav. Cogn. Neurosci. Rev.* 3, 207–221.
- Szpunar, K.K., Watson, J.M., McDermott, K.B., 2007. Neural substrates of envisioning the future. *Proc. Natl Acad. Sci.* 104, 642–647.
- Tolman, E.C., 1948. Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.
- Tolman, E.C., Gleitman, H., 1949. Studies in learning and motivation: I. Equal reinforcements in both end-boxes, followed by shock in one end-box. *J. Exp. Psychol.* 39, 810–819.
- van Gelder, T.J., 1995. What might cognition be, if not computation? *J. Philos.* 91, 345–381.
- Wexler, M., Kosslyn, S.M., Berthoz, A., 1998. Motor processes in mental rotation. *Cognition* 68, 77–94.
- Wolpert, D.M., Ghahramani, Z., 2000. Computational principles of movement neuroscience. *Nat. Rev. Neurosci.* 3, 1212–1217 Suppl.
- Wolpert, D.M., Miall, R.C., Kawato, M., 1998. Internal models in the cerebellum. *Trends Cogn. Sci.* 2, 338–347.
- Zatorre, R.J., Halpern, A.R., Perry, D.W., Meyer, E., Evans, A.C., 1996. Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J. Cogn. Neurosci.* 8, 29–46.
- Ziemke, T., Jirnhed, D.-A., Hesslow, G., 2005. Internal simulation of perception: a minimal neuro-robotic model. *Neurocomputing* 28, 85–104.