

Noname manuscript No. (will be inserted by the editor)
---

---

# Evolution of a Predictive Internal Model in an Embodied and Situated Agent

Onofrio Gigliotta · Giovanni Pezzulo · Stefano Nolfi

Received: date / Accepted: date

**Abstract** We show how simulated robots evolved for the ability to display a context-dependent periodic behavior can spontaneously develop an *internal model* and rely on it to fulfill their task when sensory stimulation is temporarily unavailable. The analysis of some of the best evolved agents indicates that their internal model operates by anticipating sensory stimuli. More precisely, it anticipates functional properties of the next sensory state rather than the exact state that sensors will assume. The characteristics of the states that are anticipated and of the sensorimotor rules that determine how the agents react to the experienced states, however, ensure that they produce very similar behaviour during normal and blind phases in which sensory stimulation is available or is self-generated by the agent, respectively. Agents' internal models also ensure an effective transition during the phases in which agents' internal dynamics is decoupled and re-coupled with the sensorimotor flow. Our results suggest that internal models might have arisen for behavioral reasons, and successively exapted for other cognitive functions. Moreover, the obtained results suggest that self-generated internal states should not necessarily match in detail the corresponding sensory states and might rather encode more abstract and motor-oriented information.

**Keywords** Internal models · Evolutionary robotics · Prediction

## 1 Introduction

The idea that cognitive agents act on the basis of internal models can be considered foundational in cognitive science; consider for example Craik's (1943) *internal models*, Tolman's (1948) *cognitive maps*, or Johnson-Laird's (1983) *mental models*. The structure and functioning of internal models is however much more debated.

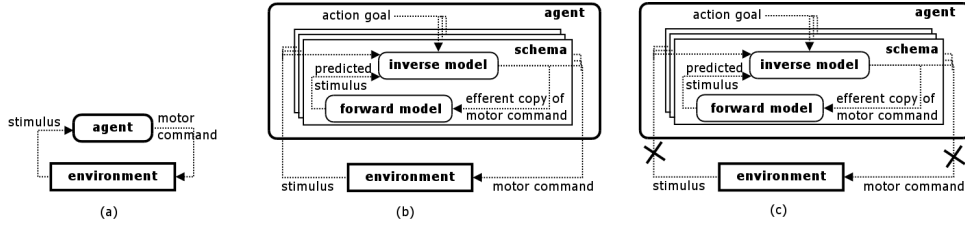
Traditional theories in cognitive science describe internal models as mental maps that support planning by means of symbolic representations, whose representational content is conceptual and is not tied to any sensorimotor modality (see Johnson-Laird, 1980 for a review). The de-emphasis of symbolic representations in cognitive science has determined a loss of attention for internal models in the traditional sense. The emphasis on the embodied and situated nature of intelligence and behavior has

---

O. Gigliotta  
Natural and Artificial Cognition Laboratory, University of Naples Federico II  
via Porta di Massa, 1 80133 Napoli (Italy)  
Tel.: +39-081-2535466  
Fax: +39-081-2535466  
E-mail: onofrio.gigliotta@istc.cnr.it

G. Pezzulo  
Istituto di Linguistica Computazionale "Antonio Zampolli"  
Via Giuseppe Moruzzi, 1 56124 Pisa (Italy)  
E-mail: giovanni.pezzulo@istc.cnr.it

S. Nolfi  
Institute of Cognitive Sciences and Technologies - CNR  
via San Martino della Battaglia, 44 00185 Rome (Italy)  
E-mail: stefano.nolfi@istc.cnr.it



**Fig. 1** Comparison between purely stimulus-response systems (b) and those endowed with anticipatory capabilities, which run an ‘internal loop’ on-line with action (b), or off-line (c).

lead instead to a non-representational and reactive view, whose most popular example in behavior-based robotics is the slogan that “the world is its own best model” (Brooks, 1991).

Recently, however, the idea of internal modeling is gaining consensus anew, as numerous researchers in cognitive psychology, neuroscience, and robotics have (re)integrated the ideas of internal modeling in an embodied view of cognition, thus essentially abandoning symbolic representations in favor of sensorimotor codes. For example Wolpert et al. (1995) claim that in many cases action is goal-directed and supported by internal models. Moreover, there is a vast literature that show how prediction and internal models play an important role in (visuo)motor control (Kawato, 1999; Miall and Wolpert, 1996) and in the stabilization of perception (von Holst and Mittelstaedt, 1950), allow to circumvent the problems caused by the delay with which sensory information reach the central nervous system (during the execution of fast reaching movements (Cerminara et al., 2009; Desmurget and Grafton, 2000) or during walking and posture control behavior (von Hofsten, 2004)), and allow to cancel self-produced stimuli so to focus the attention on relevant external stimuli (Webb, 2004). The idea of internal modeling is not confined to the motor domain but it is increasingly being used as a unifying concept that relates sensorimotor, cognitive and social abilities, including for example off-line motor planning, motor imagery, reasoning, imitation, mindreading, and cooperation (Frith et al., 2000; Grush, 2004; Jeannerod, 2006; Pezzulo, 2011; Schubotz, 2007; Wolpert et al., 2003). Recently, numerous empirical studies have focused on the neural correlates of internal models (Ebner and Pasalar, 2008; Imamizu et al., 2003), thus complementing previous theoretical and physiological analyses (Erlhagen, 2003; Kawato, 1999; Mehta and Schaal, 2002; Miall and Wolpert, 1996).

### 1.1 Structure and functioning of internal models

Internal models come in (at least) two varieties: *inverse models* and *forward models*. The former compute the motor command that produces a desired state in a certain condition, and the latter predict the sensory consequences of the execution of a motor command.

Fig. 1 highlights the differences between (a) a stimulus-response system, and (b) one endowed with (multiple pairs of) internal, forward and inverse models, which is inspired by the architecture for motor control described in (Wolpert et al., 1998). In the latter, the internal models (inverse and forward) realize an *inner loop*, which parallels actual sensorimotor interaction and mimics its input-output properties. Such loops can function on-line with action (b), or off-line (c), that is, *detached* from the current sensorimotor context. When this last condition holds, sensory inputs are substituted by predicted inputs, and motor outputs are inhibited.

This novel view of internal modeling, which incorporates control-theoretic ideas and an embodied view of cognition, is clearly synthesized in the *emulation theory of representation* proposed by Grush (2004, p. 1):

in addition to simply engaging with the body and environment, the brain constructs neural circuits that act as models of the body and environment. During overt sensorimotor engagement, *efference copies* (von Holst and Mittelstaedt, 1950) in parallel with the body and environment drive these models, in order to provide expectations of the sensory feedback, and to enhance and process sensory information. These models can also be run off-line in order to produce imagery, estimate outcomes of different actions, and evaluate and develop motor plans.

A related view is put forward by Hesslow (2002), who describes associative mechanisms as responsible of internal simulations of perception and behavior<sup>1</sup>. Both views emphasize that *anticipation* is a key element of internal models, and that predicting the effects of one's own actions can be used to improve motor control and can be at the basis of higher level cognitive skills (see also Pezzulo, 2011 for a discussion).

In this paper we focus our attention on the origins of internal models and on their use for on-line control of action in situations in which sensory information is temporarily unavailable. Other possible uses of internal models will be discussed in the concluding section and demanded to future investigations.

## 1.2 The origins of internal models

Why and how did internal models originate during the evolutionary history of living organisms? One possible answer to this question is that internal models might had arisen to improve motor control and might had been recruited to play additional functions (e.g. planning) later on (Pezzulo and Castelfranchi, 2007, 2009). More specifically, internal models might had arisen to cope with unreliability, and/or temporarily lack of sensory information and with the temporal delay with which sensory information reach the central nervous system. Indirect evidence in support to this hypothesis comes from evidences indicating that the human brain rely on prediction to deal with noise in motor control (Kording and Wolpert, 2006), and to compensate for transmission delays (Sheth et al., 2000).

Synthetic methodologies such us evolutionary robotics (Nolfi and Floreano, 2000), that allow to study how robots change evolutionarily while they attempt to adapt to a given task/environment, can provide complementary evidence. In particular, they can allow us to verify experimentally whether and in which conditions robots evolve an internal model, how such internal model is organized, and how it contributes to improve robots' performance.

## 1.3 Objectives of the study and relation to the state of the art

The primary goal of this paper is that to verify through a synthetic methodology whether internal models could spontaneously arise in living organisms for the sake of effective motor control. More specifically, in this paper we investigate whether artificial embodied agents, that are trained for the ability to exhibit a given behavioral skill (but are neither designed nor rewarded for prediction), develop and use an internal model that allow them to anticipate forthcoming stimuli to overcome the problems caused by the fact that sensory stimulation is temporarily missing.

The rationale behind our hypothesis is that, in the attempt to solve a problem that require to display a certain behavior during phases in which sensory stimuli are and are not available, the evolving robots will initially develop partial solutions that allow them to behave appropriately during the former but not during the latter phases by reacting appropriately to the experienced sensory stimuli. At this stage of the evolutionary process, agents might further improve their skill either by (1) developing a new behavioral strategy that allows them to behave appropriately also when sensory stimuli are not available, or (2) developing an ability to self-generate missing stimuli and react to both available and self-generated sensory stimuli on the basis of the previously developed behavioral strategy.

To verify whether this hypothesis holds, we ran and analyzed a series of experiments in which a population of robots is evolved for the ability to perform a simple behavior by keep producing it also when sensory stimulation is temporarily missing.

The fact that biological organisms can overcome the problem caused by the temporarily lack of sensory information has been demonstrated, for example, in the experimental study carried out by (Lee and Thompson, 1982). In this work a group of blindfolded human subjects were asked to perform a series of task (e.g. walking to a given marked location, avoiding obstacles, and throwing objects toward different locations of the room) after having been asked to observe the room in which they were located and to direct their attention toward specific objects and markers. The fact that the subjects were able to accomplish these tasks rather well and almost as accurately with respect to a control situation in

---

<sup>1</sup> It is worth noting, however, that the *simulation hypothesis* is presented by Hesslow (2002) as an associative and non representational view, whilst the *emulation theory of representation* of Grush (2004) describes inner loops as representational. We will come back to this point in the final discussion.

which they were not blindfolded clearly indicates that they are able to compensate the lack of visual information through some form of internal process.

As far as we know, the first and only attempt to verify whether an artificial agent evolved for the ability to display a behavior (that crucially depends from the agent/environmental relation) can cope with the temporarily lack of sensory information is that reported in (Ziemke et al., 2002, 2005). In this work, the authors evolved a population of simulated wheeled robots for the ability to move along a square corridor in a normal and in a blind condition (during which sensory information was or was not available, respectively). The robots sensors included only a linear camera able to detect four visual landmarks located at the four corresponding edges of the corridor.

The analysis of the best individual reported in the paper indicates that during normal phases the robot accomplished the task by moving forward while slightly turning toward right, when the robot visually detects a landmark, and by turning right otherwise. The former behavior allows the robot to lose visual contact with the landmark toward the end of each corridor. The latter behavior allows the robot to negotiate a corner and to orient itself toward the following landmark. During blind phases the robot solved the problem by executing the same two behaviors (i.e. move forward while slightly turning toward the right and turn right). In this case, however, the arbitration between the two behaviors is not based on the current state of the vision sensors but rather on the state of the internal neurons that keep trace of time passed executing the current behavior. This information about time duration is used by the robot to switch between the first and the second behavior and vice versa after 30 and 5 time steps, respectively. This means that, in the case of these experiments, the evolved agents relies on two different strategies: (1) a reactive strategy, executed during normal phases, in which the action performed depends on the visual pattern currently perceived by the robot, and (2) a strategy, executed during blind phases, in which the action performed depends from the time spent by the robot executing the current action. The combination of these two strategies, however, allows the robot to effectively master the transition between normal and blind phases only occasionally. Indeed, in the majority of the cases, shortly after sensory states are no longer available, the robot crashes against walls as a consequence of the fact that it switches between the two actions too early or too lately. The fact that the problem concerns the transitions between normal and blind phases is demonstrated by the fact that during the blind phases in which the robot happens to make the first switch at the right time, it also manages to successfully master the next switches.

The analysis of the obtained results also demonstrated how, contrary to the authors' expectation, the evolved agents did not rely on an internal model and on an ability to internally generate the simulated experience of the stimuli that are temporarily missing. In other words, the lack of sensory information was not compensated by an ability to internally generate states that are identical or similar to those that would have been experienced during normal phases but rather through the generation of a different type of information (that encode the time spent by the robot executing the current action) and through the exhibition of two different strategies during normal and blind phases based on visual information and on time-duration information, respectively. Moreover, the combination of these two strategies, did not allow to appropriately master the transition between normal and blind phases in most of the cases thus leading to rather low performance during blind phases, on the average. In the experiment reported in this paper, instead, some of the evolved agents self-generate internal states that provide functionally equivalent information to that normally provided by sensors. Moreover, evolving agents display an ability to appropriately master the transition between normal and blind phases.

Another related work is constituted by the experiments reported in Beer and Gallagher (1992) and Beer (1995) in which the authors evolved the neural controller of a simulated hexapod robot for the ability to walk. For each leg, the agent is provided with two opposing muscles-like effectors that control the torque applied to the one degree of freedom of the leg, a foot effector that controls whether the foot is up or down, a sensor that encodes the current angular position of the leg along the frontal/rear arc, and two internal neurons. The authors compared the results obtained in a normal, a sensory-deprived, and a mixed condition in which the sensory feedback was always available, was never available, and was available only in part of the trials, respectively. However, they did not study the case in which sensory information was missing temporarily after a phase in which it was available as in the case of the work referred above. Agents evolved in a normal condition displayed a reactive pattern generator (i.e. a solution in which the state of the motors depends entirely on the current state of the sensors). Agents evolved in a sensory-deprived condition developed a central pattern generator solution (i.e. a solution in which the neural controller generated rhythmic movements without relying on any type of

sensory information). Agents evolved in the mixed condition displayed a controller that operates as a central pattern generator entrained by the rhythmic sensory feedback. As in the case of the controller evolved in the sensory-deprived condition, this controller produces rhythmic movements during the trials in which sensory information is not available. However, it also displays an ability to speed-up or slow-down the oscillatory behavior on the basis of the available sensory-feedback so to adapt the periodic behavior to the effects of the agent/environmental interaction. This in turn allows the agent to walk either with or without sensory-feedback and to exploit sensory information, when available, to improve the walking performance. The fact that the problem admits a sub-optimal but relatively good solution also without sensory feedback implies that it does not need to replace missing sensory states with some equivalent self-generated information when sensory information is not available. In this paper instead, we will study a problem in which the actions that should be produced by the agent crucially depend from the agent/environmental relation. In this situation agents are forced to find a way to compensate the lack of sensory information. This in turn, as we will see, allows the agents to display similar behavior during normal and blind phases, to achieve close to optimal performance also when sensory information is temporarily unavailable, and to appropriately master the transition between normal and blind phases.

A much larger body of research concerns artificial agents that have been trained to anticipate the next sensory states. For example, (Wolpert and Kawato, 1998) have studied whether the development of an internal model can improve motor control. The aim of these studies, however, is not to verify whether and how internal models could emerge or whether the ability to anticipate the next sensory states can be used to compensate the lack of sensory information, but rather that to investigate whether the development of an internal model can support the acquisition of other behavioural or cognitive skills such as the ability to self-localize in the environment (Nolfi and Tani, 1999), the development of a model of the robot’s own body (Bongard et al., 2006), the development of an ability to discriminate between free paths and dead-ends by “mentally simulating” the effect of alternative actions (Hoffmann, 2007; Moller and Schenck, 2008), the development of an ability to imitate (Demiris and Khadhour, 2005), the development of an ability to perform mental inferences (Oztop et al., 2005), and the development of an ability to manipulate objects (Nishimoto and Tani, 2009).

Another goal of this paper is to verify what type of internal models, if any, evolve. For example do the evolved internal models anticipate the forthcoming stimuli in detail or do they only anticipate certain properties of them ? Which is the time scale at which predictions are made? How accurate predictions are?

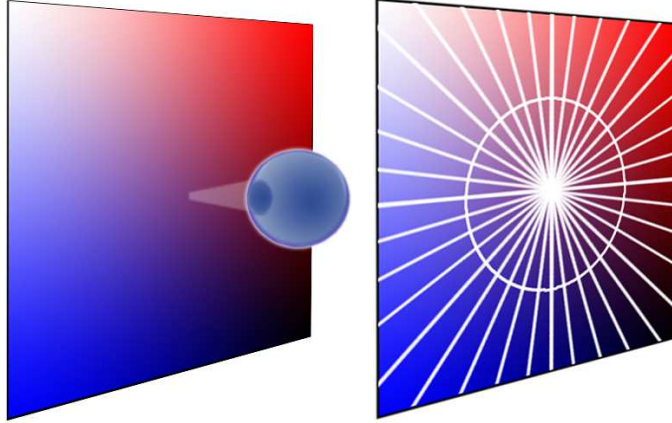
Finally, from a technological perspective, our study aims to develop a methodology that can be used to synthesize artificial embodied agents (robots) able to operate effectively in uncertain conditions.

## 2 Methodology and experimental scenario

To study the issues described in the previous section, we set up an experimental scenario in which an holonomic agent provided with an eye should develop an ability to move along a circular trajectory that allows it to observe a specific portion of the image displayed over a screen situated in front of the agent (fig. 2, left). More specifically, in each time step the agent should determine how much it has to move its eye in order to keep looking at different portions of the image located over a circular stripe of the image itself (fig. 2, right). Such behavior should be exhibited both during normal phases, in which the agent can infer the portion of the image that it is currently looking at on the basis of the color perceived, and during blind phases, in which sensory information is temporarily unavailable.

The characteristic of the agent/task/environment has been selected so to identify a scenario that is relatively simple for what concerns the sensorimotor system of the agent in which, however, the action to be performed by the agent crucially depends on the relative position of the eye with respect to the image (that can be inferred with limited precision from noisy sensory states during normal phases only). This implies that, differently from the case of the experiments reported in Beer and Gallagher (1992), the task does not admit optimal or close to optimal solutions based on control mechanism that can operate independently from the current agent/environmental relation.

The choice of a task that requires the execution of a periodic behavior (i.e. that involves the execution of a series of different actions, that allow the eye of the agent to move along a circular trajectory located on a specific position of the image, and that have to be repeated several times) has been made to identify a problem that admits qualitatively different solutions. More precisely, a problem in which the temporary



**Fig. 2** Left: The environment consists of a screen displaying an image composed by the combination of a blue and red gradients with an intensity that vary linearly along the left-right and bottom-up axis, respectively. Right: The image is virtually divided into 36 sector areas. The white circle indentifies the portion of image located at a distance of 130 pixels from the centre of the image along which the agent should keep moving its eye clockwise or anti-clockwise. The white lines drawn in the image represent a schematization and are not perceived by the agent

lack of sensory information can be solved either by developing and exploiting a form of memory that allows the agent to produce the right sequence of actions on the basis of a trace of previously executed actions stored in its internal states, or through the development of an ability to self-generate the missing sensory states by anticipating the sensory state that would have been experienced on the basis of previously experienced/self-generated sensory states. For analyses that show how the exhibition of periodic task such as human locomotion involve the ability to anticipate the next experienced sensory states and to perform anticipatory adjustments, see (McFadyen and Winter, 1991; McFadyen et al., 2001).

In our experiments, the agents develop their skill through an evolutionary process, are provided with recurrent neural controllers, and are rewarded for the ability to exhibit the requested behavior only (and not for the ability to predict the next stimuli).

## 2.1 The agent, the environment, and the task

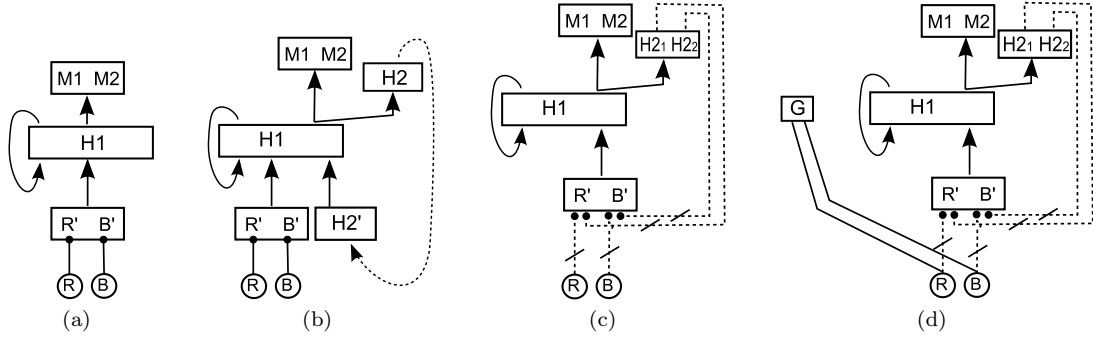
The agents consists of a simulated eye provided with a single photoreceptor located in front of a screen showing an 500x500 pixel image generated by the combination of a blue and red gradient ranging continuously from 0 to 255 along the left-right and the top-down dimensions, respectively (see fig. 2, left). Each time step, the photoreceptor detects the intensity (normalized within  $[0.0, 1.0]$ ) of the blue and red in the pixel corresponding to the current position of the eye. Noise is simulated by adding in each time step to the intensity of each color a random value with a uniform distribution within  $[-0.05, 0.05]$ . The agent is also provided with two motors that allow it to move left-right and/or top-down, with respect to its current position, up to a maximum of  $\pm 5$  pixels along each axis.

The task of the agent consists in navigating clockwise or anti-clockwise around the centre of the image at a distance of 130 pixels, so to keep moving along the different virtual sectors of the image shown in fig. 2, (left) and so to maintain a constant distance from the centre of the image.

## 2.2 Agent's neural controller

The agent's controller consists of an artificial neural network. We replicated the experiment four times by using four different neural architectures (see fig. 3).

In all cases, the network is provided with: two sensory neurons ( $B^1$  and  $R^1$ ) that encode the intensity of blue and red colour currently sensed by the photoreceptor of the eye ( $B$  and  $R$ ) perturbed by noise or the state of the H2 internal neurons at time  $t-1$  (see explanation below), eight internal neurons ( $H1$ ), and two motor neurons ( $M1$ ,  $M2$ ) that determine the amplitude of the eye movement with respect to its current position along the left-right and top-down dimension within a range of  $[-5,5]$  pixels. The



**Fig. 3** The four architectures of agents' neural controller used in four corresponding experiments. R and B encode the intensity of blue and red colour currently sensed by the photoreceptor of the eye.  $B^1$  and  $R^1$  encode the state of B and R or the state of H2 (see text for explanation). H1 and H2 are internal neurons. M1 and M2 are used to update the robot's position. G indicate a thresholded internal neurons that control whether the state of the sensory neurons  $B^1$  and  $R^1$  is updated on the basis of the state of the sensors (B and R) or on the basis of the state of the H2 internal neurons. Rectangles indicate blocks of neurons. Full arrows indicate connections between blocks of neurons. Dashed arrows indicate efferent copies.

H1 internal neurons receive connections from the sensory neurons and from themselves. Some of the architectures are also provided with two additional internal neurons (H2). The motor neurons (M1 and M2) and H2 internal neurons (when present) receive connection from the H1 internal neurons.

Architecture A is provided with the sensory neurons ( $B^1$  and  $R^1$ ) that are always set on the basis of sensors (B and R), eight internal neurons (H1), and two motor neurons (M1 and M2). Architecture B is also provided with a second layer of internal neurons (H2) and two additional neurons ( $H2^1$ ) that encode the state of H2 neurons at time  $t-1$ . The rationale between the use of this architecture is that it can facilitate the development of solutions that rely on sensory or on self-generated internal states during normal and blind phases, respectively. Architecture C is also provided with a built-in gating mechanism that ensures that the state of the sensory neurons ( $B^1$  and  $R^1$ ) is set on the basis of the state of the sensors (B and R) during normal phases and on the basis of the state of the two internal neurons (H2) at time  $t-1$  during blind phases. The rationale between the use of this architecture is that the presence of a mechanism that forces the agent to rely on sensory information while neglecting self-generated internal states and vice versa (depending on whether sensory information is or is not available, respectively) can facilitate the development of solutions that operate by self-generating the missing sensory information during blind phases. Architecture D is also provided with an additional internal neuron (G) that is used to determine at each time step whether the state of the sensory neurons ( $B^1$  and  $R^1$ ) is set on the basis of the sensors (B and R) or on the basis of the state of the two internal neurons (H2) at time  $t-1$ . The rationale between the use of this architecture is that to verifying whether leaving the agent free to determine when to use sensory or internal states can be advantageous, and whether the agent can successfully co-develop an ability to self-generate the missing sensory information and an ability to determine when to use sensory or self-generated states.

The sensory neurons are relay units that are activated on the basis of the intensity of the colours detected by the photoreceptor perturbed by noise.

Internal neurons are leaky integrators (i.e. neurons that hold a certain amount of the previous activation) and are updated according to the following equations:

$$A_j = \vartheta_j + \sum_{i,j=1}^N w_{ij} O_i \quad (1)$$

$$O_j = \tau_j O_{j(t-1)} + (1 - \tau_j)(1 + e^{-A_j})^{-1} \quad (2)$$

$$0 \leq \tau_j \leq 1 \quad (3)$$

With  $A_j$  being the activity of the  $j$ th neuron (or the state of the corresponding sensor in the case of sensory neurons),  $\vartheta_j$  the bias of the  $j$ th neuron,  $w_{ij}$  the weight from the  $i$ th to the  $j$ th neuron,  $O_i$  the output of the  $i$ th neuron,  $\tau_j$  the time constant of the  $j$ th neuron.

Motor neurons (M1 and M2) and H2 internal neurons are standard logistic neurons and are activated on the basis of following equation:

$$O_j = \frac{1}{1 + e^{-A_j}} \quad (4)$$

The G neuron is a binary neuron that is activated on the basis of the following equation ( $G_{net}$  is computed according to equation 1) :

$$G = \begin{cases} 1 & \text{if } \frac{1}{1+e^{-G_{net}}} > 0.5 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

The architecture of the neural network is fixed. The connection weights and biases and the time constant of the internal neurons are encoded in free parameters and adapted.

Notice that since the agent is not provided with proprio-sensors encoding information about the actual position of the eye with respect to the two axes and since the motor neurons encode the movements of the eye with respect to its current position, the agent's neural controller does not receive any indication on its absolute position or on its position on the image during blind phases beside the traces left by the sensory states experienced during the last normal phase.

### 2.3 Adaptive algorithm and training procedure

The free parameters of the agent's neural controller has been adapted through an evolutionary robotics method (Nolfi and Floreano, 2000). The initial population consists of 100 randomly generated genotypes which encode the free parameters of 100 corresponding individuals. Each parameter is encoded with 8 bits and is normalized in the interval [-5.0, +5.0] for the biases and the synaptic weights, and in the interval [0.0, 1.0] for the time constants. Each subsequent population is obtained by selecting the best 20 individuals of the previous population. Each selected individual is allowed to produce 5 offspring that are generated by duplicating the genotype of the reproducing individuals and by applying mutations (with 2% probability of flipping a bit) to each offspring.

Each individual is tested for 20 trials. At the beginning of each trial the eye is placed randomly in one of ten different positions distributed uniformly around the centre of the image at a distance of 130 pixels. The agent is then allowed to interact with the environment up to 4000 time steps. For each time step, the state of the agent's sensory neurons is updated on the basis of the current position of the eye, the state of the internal and motor neurons is updated, and the agent's eye is moved on the basis of the current state of the motor neurons. The agent experiences a sequence of phases in which sensory information is available (normal phases) or missing (blind phases). During the first half of each trial (i.e. during the first 2000 time steps) the agent has always access to the sensory stimulation coming from the environment (normal phase). During the next half of the trial, instead, the agent experiences an alternation of blind and normal phases with the same length. The length of these phases varies from 1 to 21 time steps during the 20 corresponding trials.

The performance (fitness) of the individuals during a trial is computed on the basis of the following equation that rewards individuals for the ability to visit subsequent sectors of the image (clockwise or anti-clockwise) and for the ability to keep a distance as close as possible to 130 pixels from the centre of the image. In particular, for each new sector visited, the fitness  $F$  of the individuals is updated by adding  $dF$ :

$$dF = \begin{cases} \frac{1 - \frac{|D_t - 130|}{100}}{36} & \text{if } 30 \leq D_t \leq 230 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where  $D_t$  represents the distance between the point of the image observed by the agent at time  $t$  from the center of the image, and 36 corresponds to the number of sectors. In order to not force the system to maximize its fitness by moving as fast as possible, the fitness scored during the first 2000 time step is truncated to 30.

Trials last up to 4000 time steps but are terminated as soon as the agent's eye looks to a point located outside the 500x500 pixels area or looks back to a recently visited sector. The total performance of an individual is calculated by averaging the performance obtained during the 20 trials.



**Table 1** Distribution of solutions for different type of architectures. F1, F2 and F3 indicate the family of the solution. A,B,C, and D indicate the type of architecture. Each replication of the experiment has been categorized on the basis of the behavior exhibited by the best evolved individual.

Architectures	F1	F2	F3
A	40	0	0
B	40	0	0
C	23	13	4
D	37	2	1

The evolutionary process is continued for 1600 generations. The evolutionary experiment has been replicated 40 times starting with different, randomly generated, initial populations.

### 3 Results

By analysing the behaviour and the average performance displayed by evolved individuals we observed that most of them display optimal or close to optimal performance during normal phases as well as a good ability to handle short blind phases. However, only some of the individuals provided with the architecture C and D also demonstrate an ability to handle relatively long blind phases and to generalize their skill to blind phases lasting 42 and 84 time steps (i.e. two or four times the maximum duration experienced during the evolutionary process).

The analyses conducted and reported in the next sections indicate that the solutions adopted by evolving individuals can be categorized in three different families.

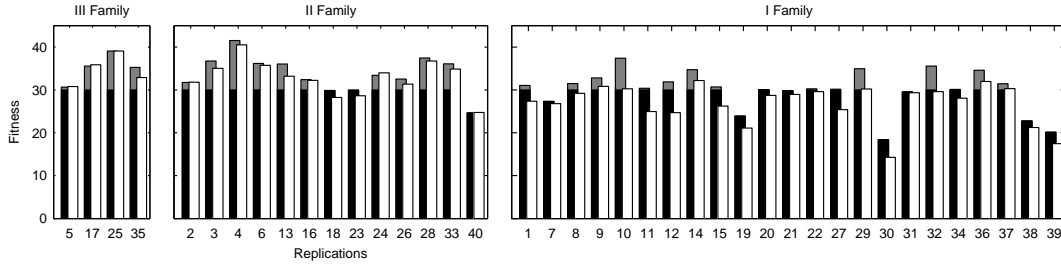
The *first family* includes individuals that display an ability to keep rotating around the centre of the image during normal phases but not during blind phases. Indeed, in these individuals, the activation state of the motor neurons immediately or quickly converges on a fixed state, after the transition between a normal and a blind phase, which leads to the production of a linear trajectory. In other words they are unable to maintain a self-sustained internal dynamics. In some cases, depending on the relative position of the agents eye at the beginning of the blind phase, this simple strategy allow the individuals to handle successfully short blind phases. However, it does not allow to master long blind phases.

On the contrary, the individuals of the *second* and *third family* show an ability to keep varying their motor actions also during blind phases. As we will see, many of the individuals belonging to the second and third family are able to master long blind phase and to generalize their skill to blind phases lasting 42 and 84 time steps. What differentiates the latter two families is the fact that in the second family the agents exhibit different behaviors during normal and blind phases, while in the third family the agents produce similar behavior in the two phases (and we will argue that, to do so, agents develop internal models; see below).

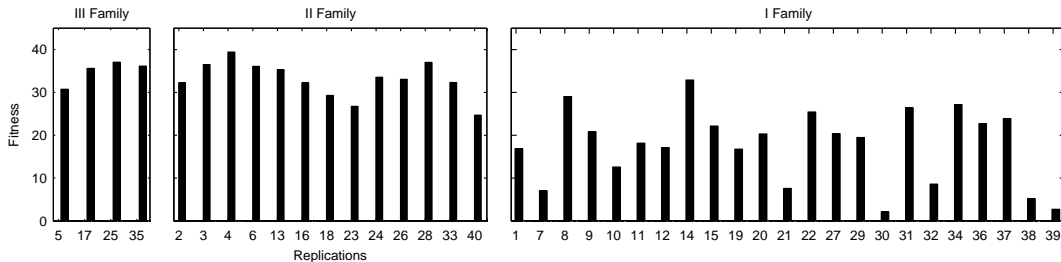
To discriminate between family 2 and 3 we used two measures of behavior similarity: a) the average euclidean distance between the state of the motors over time recorded during normal and blind phases ( $d_{motorsfft}$ ) and b) the average correlation between fast fourier transformations of the state of the motors over time ( $\rho_{motorsfft}$ ) recorded in normal and blind phases (for more information see below). The individuals were assigned to family 3 if  $\rho_{motorsfft} < 0.7$  and  $d_{motorsfft} < 0.4$ , and to family 2 otherwise.

The fact that only some of the individuals with the architecture C and D display solutions belonging to families 2 and 3 (see table 1 indicates that the presence of the gating mechanism (that updates  $B^1$  and  $R^1$  on the basis of B and R or H2 during normal and blind phases, respectively) represents a necessary prerequisites for the development of effective solutions. Moreover, the fact that the number of replication that lead to family 2 and 3 solutions is significantly higher in the case of the architecture C with respect to architecture D (17/40 and 3/40, respectively) indicates that the need to also develop an ability to determine when to rely on sensory or self-generated internal states makes the task harder.

The analysis of the performance of the best evolved individuals of the experiment with architecture C (fig. 4) indicates that most of the replications lead to optimal or close to optimal performance during normal phases as well as a good ability to handle blind phases, on the average. Moreover some of them (fig. 5 and fig. 6) also demonstrate an ability to generalize their skill to blind phases lasting 42 and 84 time steps (i.e. two or four times the maximum duration experienced during the evolutionary process).



**Fig. 4** Performance of the best individual of each replication of the experiment with architecture C. The black and white histograms display the average performance obtained during the first and the second 2000-cycles parts of each trial in which agents experience a normal only and an alternation of blind and normal phases, respectively. Average performance calculated over 20 trials. Grey histograms on-top of black histograms indicate the average performance that would have been obtained during the first part without truncation to 30. See text for explanation.



**Fig. 5** Performance of the best individual of each replication of the experiment with architecture C tested in a control condition in which normal and blind phases lasting 42 time steps alternate from the beginning of the trial. Average performance calculated over 20 trials. Note that a few replications (and in particular 4, 6, 24, 28 and 33 for the second family, and 17 and 25 for the third family) display an ability to generalize their skill in this condition.

Fig. 6 also shows that only the solutions belonging to the second and third families display close to optimal performance (i.e. performance  $> 30$ ) when tested on blind phases lasting 84 times steps.

Notice that the differences in performance are not only due to the ability to keep moving by producing a circular trajectory during blind phases but also to the speed with which the agents circle around the centre of the image and to the extent to which they succeed in keeping a distance of 130 pixels from the centre of the image.

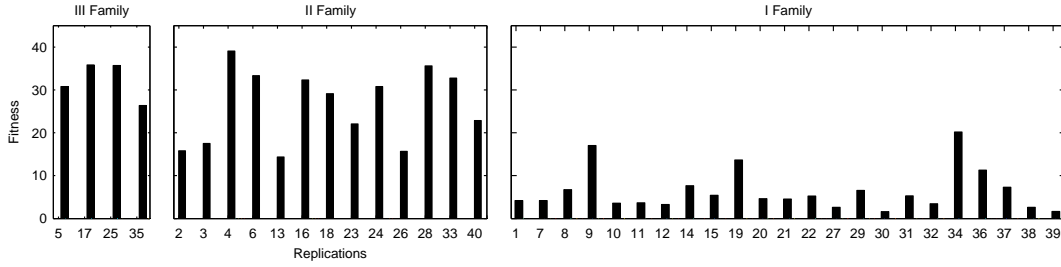
The analysis of the performance of the best evolved individuals with architecture D (Fig. 7) displays qualitatively similar results. In this cases, however, the number of replications that lead to sub-optimal family-1 solutions is much larger. The analysis of the best individual of the single replication that lead to a family-3 solution indicates that it relies on the sensory information while neglecting self-generated states depending on whether the state of the sensors is or not above 0.0. In other words, they evolved a set of weights that play the same role of the built-in gating mechanisms included in architecture C.

In the following section we analyze in more details the best solutions of each family in the case of the experiment with architecture C and we describe in more detail how the solutions belonging to family 2 and 3 differ.

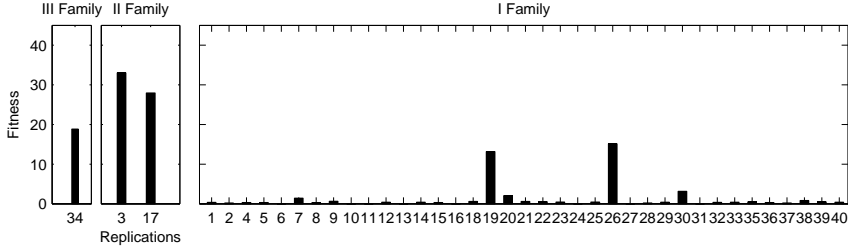
### 3.1 Analysis of the third family of solutions (of architecture C), and its internal modeling strategy

In this section, we analyze in detail the best individual of replication 17, which belongs to the third family ( $d_{motorsfft} = 0.12$ ,  $\rho_{motorsfft} = 0.98$ ). As we will see, this and the other individuals belonging to this family possess and rely on an internal model that anticipates relevant characteristics of the forthcoming sensory stimulus, and allows the agents to display similar effective behaviors during normal and blind phases as well as during the transition between the two.

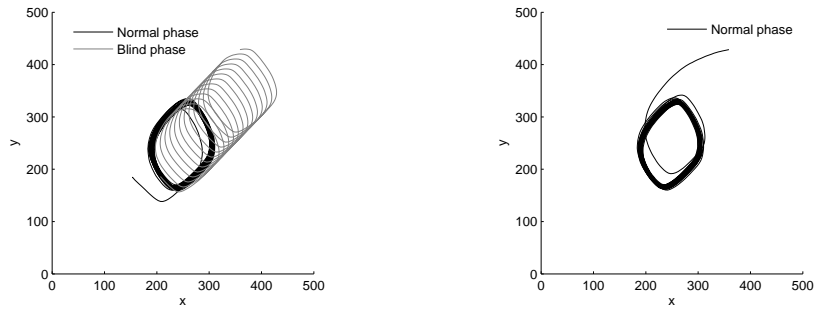
A first aspect that we can observe is that this agent displays very similar (and close to optimal) behaviour during normal and blind phases. Fig. 8 (left) shows the behaviour displayed by the agent, which is initially provided with external stimuli, and successively (after the first half of the trial) experiences a blind phase lasting 1000 time steps. We can see how, during the blind phase, the agent keeps producing



**Fig. 6** The same as fig. 5, but for phases lasting 84 time steps. Again, note that a few replications (and in particular 4, 6, 24, 28 and 33 for the second family, and 17 and 25 for the third family) display an ability to generalize their skill to blind phases lasting 84 time steps.



**Fig. 7** Performance of the best individual of each replication of the experiment with architecture D tested in a control condition in which normal and blind phases lasting 84 time steps alternate from the beginning of the trial. Average performance calculated over 20 trials.

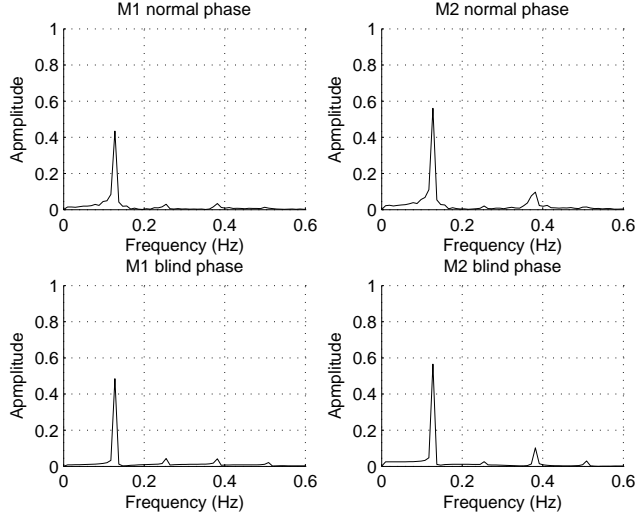


**Fig. 8** Behaviour displayed by the best individual of replication number 17 during a test in which the blind phase last 1000 time steps. Left, trajectory produced during the initial normal phase lasting 2000 time steps and during the succeeding blind phase lasting 1000 time steps (dark and light lines, respectively). Right: trajectory produced during the last 1000 time step in which the agent has access again to stimulation coming from the external environment.

the same quasi-circular trajectory as in the normal phase, while slowly drifting toward the top-right part of the image (for an example of how the behavior vary during the two phases in an individual belonging to family 2, see fig. 14). Moreover, by observing the trajectory produced by the agent during the successive normal phase (fig. 8, right), we can see how the agent manages to quickly recover from the drift as soon as the sensory stimulation return available. In other words, the agent is able to (1) de-couple from external stimuli when they are unavailable, by keep producing a close to optimal behavior, and (2) re-couple with external stimuli when they are available again.

The fact that the agent manages to keep producing a rather similar motor behaviour during normal and blind phases is also confirmed by the comparison of how the state of the motors (M1 and M2) varies over time during normal and blind phases (see fig. 11 and fig. 13, right) and by the comparison of the fourier transform of the sequence of motor states produced during normal and blind phases (fig. 9).

Overall, these analyses demonstrate that the individual solves the problem by displaying almost identical behaviour during normal and blind phases. The fact that the behaviours exhibited by the individual during the two phases are so similar allows the agent to master successfully also relatively long blind phases. On the other hand, the lack of sensory information can be tolerated only for a limited



**Fig. 9** Fourier transform of M1 and M2 over time (calculated using a standard fast fourier transform algorithm) for the best individual of replication 17, belonging to the third family. The top and bottom figures correspond to the data collected during a normal and a blind phase, respectively, lasting 1000 time steps.

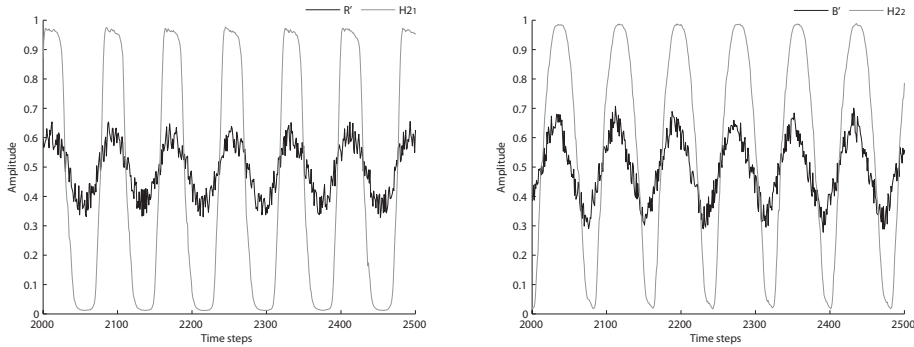
amount of time since small differences between the behaviour produced in normal and blind conditions tend to cumulate over time during blind phases. Finally, the analyses reported above indicate that the individual is able to tolerate quite well the transition between the normal and blind conditions, irrespective of the relative position of eye when the transition occurs, and to quickly recover from drift occurring during blind phases as soon as sensory information returns available. This means that the internal model does not only successfully compensate for a lack of external input, but also allows a timely ‘switch’ from the presence to the absence of external stimuli, and vice versa.

### 3.1.1 Prediction

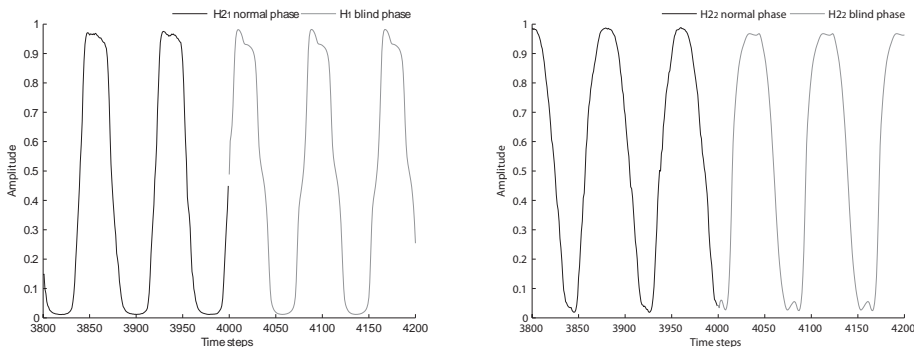
As discussed in sec. 1, internal models operate by predicting the state that the sensors of the agent will assume after the execution of a given action. To verify whether the agent predicts at time  $t$  the state that the sensory neurons will assume at time  $t+1$  we recorded and plotted the state of H2 neurons at time  $t$  and the state of neurons (R and B) at time  $t+1$  during normal and blind phases (Fig. 10). The obtained results indicate that during normal phases the state of the H2 neurons differ significantly from the state of the sensory neurons. More specifically, the self-generated states vary within the full range  $[0.0, 1.0]$  while the sensory states experienced in normal condition vary within a much smaller range  $[0.35, 0.65]$ . Notice also how, beside the range, the shape of the oscillatory curve varies in the two cases. From this analysis we can conclude that the agent does not predict the exact state that the sensory neurons would assume at time  $t+1$  in a normal condition. However, further analyses reported below indicate that it anticipates more abstract properties of how sensory states vary over time, which are sufficient to maintain the same behavior during normal and blind phases.

This hypothesis is confirmed by the data reported in (Fig. 11) that show how the state of H2 neurons vary over time during a normal and a blind phase. The comparison indicates that they vary in a rather similar way independently from the fact that the sensory neurons are feed with actual data collected from the environment or with self-generated data (i.e. the state of H2 neurons themselves), despite the two types of data differ significantly. This means that the agent self-generates stimuli that provide the regularities that are necessary to produce the same behaviour elicited by real stimuli even if the self-generated and real stimuli differ significantly. This is, in brief, a *motor-oriented encoding*, which is functional to enable an appropriate robot’s behavior, rather than a mere replication of the sensations.

Theoretical considerations lead us to hypothesize that, to generate appropriate behavior, and behavior that is similar in the two conditions (during normal and blind phases), internal models should capture the essential regularities of the stimuli, while, at the same time, discarding other unnecessary parts. This constitutes a form of *abstraction* (see sec. 4 for a discussion of this point). One abstract property of the



**Fig. 10** State of the sensors and of H2 neurons during a normal phase. Left: the dark and light lines indicate the state of neurons R and of the first H2 neuron, respectively. Right: the dark and left light indicate the state of neurons B and of the second H2 neuron, respectively.



**Fig. 11** State of H2 neurons over time. The dark and light lines indicate the state of the neuron during a normal and blind phase, respectively. Left: data for the first H2 neuron. Right: data for the second H2 neuron.

sensory flow that could be anticipated by the agent is constituted by the phase with which sensory states vary over time. To verify whether this hypothesis is true, we carried out a cross-correlation analysis between the first H2 neuron and R, and the second H2 neuron and B. The analysis of the data collected during a normal phase (fig. 12, left) indicates that the first H2 neuron at time  $t$  oscillates in phase with R at time  $t+1$  while the oscillation of the second H2 neuron at time  $t$  is delayed with respect to that of B at time  $t+1$ . The analysis of the data collected during a blind phase (fig. 12, right) indicates that first and the second H2 neurons (at time  $t$ ) are in phase with R and B (at time  $t+1$ ), respectively.

The presence of a phase correspondence between the state of H2 neurons at time  $t$  and the state of the neurons R and B at time  $t+1$ , during blind phases, confirms the hypothesis that the agents' neural controller anticipates the phase with which the value of sensory states oscillate over time. This means that the agent's internal model is predictive, or, in other terms, that it is an *internal forward model*, which encodes transitions from current state and action to future state (i.e.:  $s, a \rightarrow s_{t+1}$ )<sup>2</sup>

The analysis of the other replications of the experiment belonging to the same family of solutions provided qualitatively similar results. More precisely, during blind phases the cross-correlation analysis indicates that the variations of H2 at time  $t$  are in phase with variations of R and B at time  $t+1$  in all the four replications belonging to the third family. During normal phases, instead, the two H2 neurons oscillate with the same period but with different phases. Usually, the phase of one of the output neuron anticipates that of the corresponding sensor neuron of [1-5] time steps while the phase of the other output neuron follow of [1-5] time steps that of the corresponding sensory neuron.

Finally, to understand how self-generated sensory states differ from sensory information gathered directly from the external environment we computed (see Fig. 13, left) how the point of the image

<sup>2</sup> As shown in fig. 1 (b), internal models encode a transition from sensory states and actions to future (predicted) sensory states ( $s, a \rightarrow s_{t+1}$ ). This is typically done by giving the forward model an efference copy of the last motor command. Rather, in our implementation state and action information is available through the connections between *H1* and *H2*.

fixated by the eye would have varied over time if the state of the H2 neurons would correspond to real experienced sensory states (i.e. to real perceived colours). As can be seen, the relative positions that correspond to the self-generated sensory states are squeezed toward the border of the image with respect to the position really assumed by the eye (Fig. 8). Interestingly, the average discrepancy between the virtual and the actual position of the eye is low for what concerns the angular information (i.e. the information which is crucial in order to produce the required exploratory behaviour) while is much higher for what concerns the distance information. Indeed, by testing the agent for 20 trials with normal and blind phases lasting 42 time steps we observed that during blind phases the average discrepancy for what concerns the angle with respect to the centre of the image is 14.25 degrees (standard deviation = 8.69 degrees). The average discrepancy for what concerns the distance with respect to the centre of the image instead is 205.40 pixels (standard deviation = 33.59 pixels). This implies that the dynamics of the H2 neurons are transformed, with respect to the dynamic of the sensory neuron R and B, so to preserve angular information that is crucial for the accomplishment of the task. This is, again, an essential aspect of the motor-oriented encoding.

Overall these results indicate that the individuals belonging to the third family rely on an internal model that anticipates how the state of the sensors varies as a result of the execution of the current planned action. Rather than anticipating the exact value that the sensors would have assumed, however, the internal model anticipates more abstract properties of the sensory flow, namely its phase variation as well as information that provide an indication of the current angular position of the eye with respect to the centre of the image. Here “abstractions” are sensory abstractions (e.g., the phase with which sensory states vary over time rather than the exact value of the stimuli), which capture the regularities that make the agent able to act in similar ways with and without sensory stimuli (this is why we consider this encoding *motor-oriented* rather than being aimed to reconstruct the details of the sensory stimulations). This, in turn, might imply that anticipation can constitute a route to sensory abstraction, and to perceptual capabilities that goes beyond the mere reconstruction of sensory stimuli.

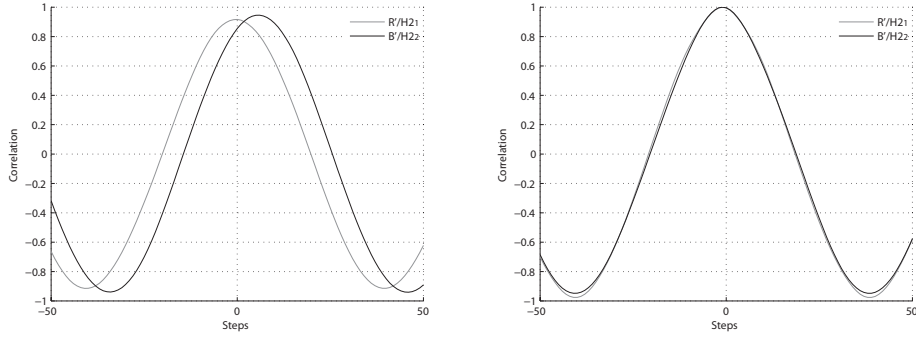
There are several reasons that might explain while evolved agents anticipate abstract properties of the forthcoming stimuli rather than the actual stimuli. A first reason is that predicting the detailed characteristics of the next stimuli is more difficult than predicting only selected properties of them (Schmidhuber and Prelinger, 1993; Nolfi and Tani, 1999). A second reason is that predicting only the functionally relevant properties is more parsimonious. A third reason is that predicting something different from the actual sensory state allows to encode information to be predicted in a way that simplifies the use of this information for triggering the appropriate motor action; we consider it a *motor-oriented* encoding. Indeed, the stimuli self-generated by our evolved robots are easier to discriminate, with respect to their natural counterpart, since they vary more sharply within a larger range (fig. 10). Overall this means that only the essential characteristics of the stimuli are anticipated in a form that simplifies their use, while at the same time it permits to (re)generate appropriate behavior.

### 3.2 Analysis of the second family of solutions (Architecture C)

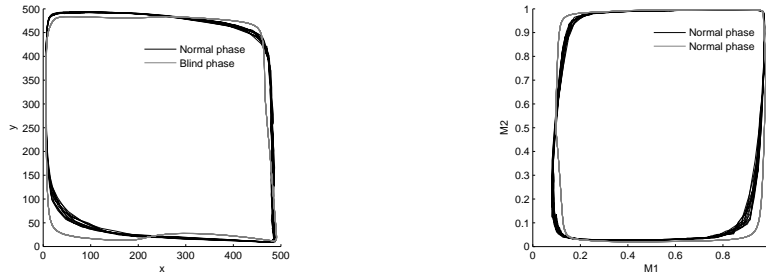
We now describe the characteristics of the agents belonging to the second family of solutions. As stated above, also some of the individuals belonging to the second family display the ability to master relatively long blind phases through the generation of a self-sustained activity. Moreover, the analysis of the correlation over time of the state of the two additional motor neurons and the state of the two corresponding sensory neurons indicates that, also most of the individuals belonging to the second family anticipate the phase with which sensory state vary over time (more precisely phase anticipation is observed in the best evolved individuals of 10 out of 13 replications of the experiment belonging to the second family). Finally, also in the case of second family, self-generated states experienced during blind phases differ significantly from the real sensory states experienced during normal phases.

Differently from the third family, however, the individuals belonging to the second family display rather different motor behaviours during normal and blind phases. This can be clearly seen, for example, by comparing the motor trajectories (Fig. 14, left) and the Fourier transform analysis of the motor actions (Fig. 14, right) during normal and blind phases in the case of the best individual of replication 4.

While in the individuals belonging to the third family the problem caused by the lack of sensory information during blind phases is solved through the self-generation of states that are functionally



**Fig. 12** Correlation over time between the states of the H2 neurons and the states of the two corresponding sensors. Analysis performed on the data collected for 1000 time steps during a normal and a blind phase (left and right pictures, respectively). During blind phases, the state that the sensors would had assumed if the agent were not blind were used. The light and dark line indicate the correlation of the first or of the second sensor and H2 neuron, respectively. The analysis has been repeated by shifting the output sample window of  $k$  steps, where  $k$  varies from  $-50$  to  $50$ . The position of the peak along the x-axis indicates the extent of the anticipation (for value below 0) or of the delay (for value above 0). Note that the peak at 0 indicates that the phase of variation of the additional output neurons anticipates the phase of variation of the sensory neurons of 1 time step.



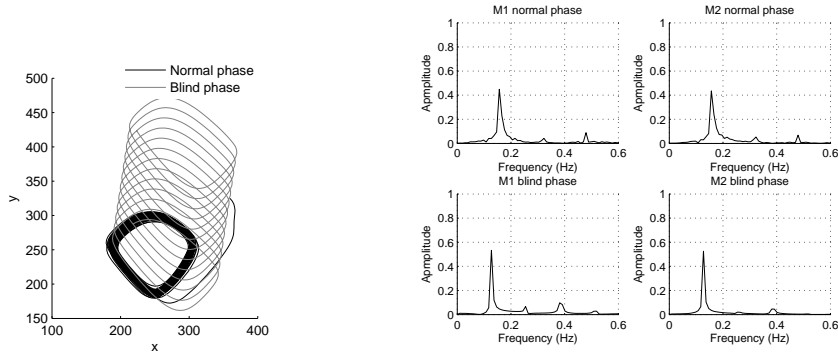
**Fig. 13** Left: The graph shows how the position of the eye on the image, corresponding to the state of the H2 neurons, varies during a normal and blind phases lasting 1000 time steps (black and gray lines respectively). X and Y axis indicate the position along the two corresponding axis obtained by normalizing the state of the H2 neurons within the  $[0, 500]$  range, i.e. within the range of variation of the red and blue gradients over the  $500 \times 500$  pixel image. Right: the graph shows how the state of the motor neurons varies over time during a normal and blind phases lasting 1000 time steps.

equivalent to the missing sensory states (i.e. that trigger the same motor actions during normal and blind phases), in the individuals belonging to the second family the problem is solved through the generation of a self-sustained activity and through the development of two different behavioural strategies that lead to the exhibition of two qualitatively different solutions for normal and blind phases. In other words, self-generated stimuli play different roles in the third and second families, and are used to generate the same behaviour in the former but not the latter.

Before concluding this section is important to clarify that, although we believe that the discrimination of solutions in three classes is useful to understand the mechanisms that allow individuals to solve the task, there are no clear-cut boundaries between different families. Indeed, for individuals that display similar, but not very similar, behaviour during normal and blind phases, the categorization into family two or three is somewhat arbitrary. Similarly, individuals displaying short term self-sustained dynamics which allow them to master successfully short blind phases can be categorized between family one and two.

## 4 Discussion

The obtained results demonstrate how some of the evolved robots manage to solve their task also when they are temporarily 'blindfolded' and to appropriately master the transition between normal and blind phases. However, only some of these robots — those belonging to the third family — solve the problem



**Fig. 14** Left: Behaviour displayed by the best individual of the replication number 4 ( $d_{motorsfft} = 0.70$ ,  $\rho_{motorsfft} = 0.21$ ) during a test in which normal and blind phase last 1000 time steps. Right: Fourier transform of M1 and M2 over time for the best individual of replication 4, belonging to the second family. The top and bottom figures correspond to the data collected during a normal and a blind phase, respectively, lasting 1000 time steps.

by producing the same behaviour during normal and blind phases. We argue that robots belonging to the third family rely on an *internal model* that has the following characteristics:

- It is autonomously developed depending on the demands of the agent-environment interactions rather than externally designed. Indeed, individuals have not been rewarded for the ability to predict the state of forthcoming stimuli.
- It is primarily driven by its own dynamic properties, and can be triggered by both external and internal, self-generated inputs. It is self-sustained, i.e. it can (re)generate similar dynamics and behavior by using self-produced rather than external stimuli.
- It generates and relies on abstract motor-oriented states rather than on states that mimic in details the characteristics of sensory stimuli.
- It has anticipatory nature, since it correlates with future stimuli more than with past or present stimuli, and can be self-sustained by using *predicted* sensory stimuli instead of ‘real’ sensory stimuli. Note that the anticipatory nature, defined as the ability to generate future stimuli on the basis of previously experience stimuli, is not the only way to solve the adaptive task. Indeed, robots could rely on alternative strategies that, for example, determine the appropriate action on the basis of the previously executed actions, as in the case of the study reported in (Ziemke et al., 2002, 2005).

The solutions belonging to the second family, instead, lack at least one of the features that we considered essential for a full-fledged internal model: the ability to support *the same* (or similar) behavior interchangeably both in the presence of external and self-generated stimuli. Although the lack of this feature does not necessarily lead to a lower performance in the experimental setting considered in this paper, it other cases in might prevent a flexible reuse of their self-generated stimuli. For example, it might prevent the possibility to effectively fuse sensed and self-generated stimuli to reduce the unreliability of sensory information (as is done in models based on Kalman filters). Similarly it might prevent the possibility to use external, goal stimuli as *reference signals* for inferring intentional action, an aspect that is essential in closed-loop and ideomotor theories (Adams, 1971; Hommel et al., 2001). Furthermore, it prevents using predictions for cancelation of self-produced stimuli (Blakemore et al., 1998).

## 5 Conclusions

Theoretical studies suggest that internal models could have originated in living organisms for the sake of adaptive behavior, not for cognition, and were than exapted for advanced cognitive and social operations (Pezzulo and Castelfranchi, 2007, 2009). Unfortunately, little effort has been devoted to the verification of this idea—something that is admittedly very complicated by using empirical means, but is more feasible by using the methodology of *evolutionary robotics* (Nolfi and Floreano, 2000), which we adopted in this study.

The central hypothesis that motivated our design methodology is that a (temporary) deprivation of external stimuli and the possibility to substitute the missing sensory information with self-generated



internal states can create the adaptive conditions for the development of an internal model in an embodied and situated agent even in absence of any explicit reward for prediction. The rationale behind this hypothesis is that the agent will first develop a strategy which is effective when sensory stimulation is available. The development of such strategy might then channel the adaptive process toward the development of an ability to self-generate sensory information, which is temporarily not available, rather than toward the development of a two different motor strategies able to operate effectively when sensory information is or is not available, respectively.

The obtained results demonstrated that indeed, a population of agents evolved for the ability to perform an exploration task display an ability to solve the task during both normal and blind phases (i.e. during phases in which sensory stimulation is temporarily not available). Moreover the analysis of the strategy adopted by the best evolved individuals demonstrate how, in some cases, they solve their task through the development of an internal model that has anticipatory aspects, can be (temporarily) detached from the current sensorimotor flow, and endogenously reactivated by self-generated signals. It is worth mentioning that, as testified by the presence of three families of solutions, and by the results of the study reported in (Ziemke et al., 2002, 2005), the arose of internal modeling strategy (in the sense we have described here) is far from trivial, and it is not simply dictated by our particular architectural choices. In other terms, certain architectural constraints, such as the availability of a gating/regulatory mechanisms that allows the agent to neglect sensory or self-generated states during blind and normal phases, could indeed be necessary conditions for the arose of internal models, but they are certainly not sufficient.

Given the similarities between the specific problems solved by living organisms during natural evolutions, and by the artificial agents described here, our results support the hypothesis that internal models might had originated for motor control and might have been exapted for additional cognitive and social functions later on. Moreover, we have demonstrated that an ability to anticipate can evolve even if the agents are not directly rewarded for predicting the forthcoming stimuli, providing that there are compelling environmental conditions.

At a general level of description, this study can be situated within a new tendency in cognitive science that take an ecological perspective in which cognitive processes are integrated with behavioural and ultimately sensorimotor processes. Within this area, some studies tend to frame the issue at a theoretical level (Clark and Grush, 1999; Keijzer, 2001; Grush, 2004; Pezzulo and Castelfranchi, 2007, 2009), while other studies (Tani, 1996, 2003; Beer, 2003; Nolfi, 2005; Tani et al., 2008; Moller and Schenck, 2008; Gigliotta and Nolfi, 2008; Gigliotta et al., 2010; Johnsson et al., 2009), including the present article, attempt to come up with complete and detailed models implemented in embodied agents that are situated in an external environment with which they interact.

In that respect, it is important to notice that the realization and the analysis of models implemented in embodied and situated agents can support or disconfirm alternative theoretical models and/or can highlight aspects that need to be incorporated in these abstract models. For instance, the results described in this paper support the hypothesis that internal models might had arisen for compensating the problems caused by the temporary unavailability of sensory information and the hypothesis that internal model can operate by producing a one-step prediction at the level of raw sensory data. However they disconfirm the assumption, often made implicitly, that anticipation should be realized by generating states that encode the anticipated information in the same way in which it is encoded in sensors. Moreover, our results demonstrate that the presence of a learning mechanism that drive the learning process on the basis of the prediction error (i.e. the difference between predicted and actual sensory states) does not necessarily represent a prerequisite for the development of an anticipatory capability.

*The possible representational role of internal modeling* As remarked in sec. 1, some researchers attribute a *representational role* to internal modeling, and most notably Grush (2004) in his ‘emulation theory of representation’. Unfortunately, the term ‘representation’ is inflated in cognitive science. It is worth noting, however, that the internal models that we observe in the third family of solutions met all the criteria for representationality that are accepted in a variety of approaches. First, not only they correlate with external events, but they *causally* determine behavior. As many philosophers have argued, causal role is one important criterion that distinguish representations from mere correlations (Dretske, 1981; Millikan, 2004). Second, the internal models go beyond causal power in that they also allow ‘internal’ manipulations, in the sense that they permit closing the sensorimotor loop internally, or without receiving external stimuli. This fact is compliant with the foundational idea in cognitive science, which can be

traced back at least to (Piaget, 1954), that the ability to act on internal representations rather than (or before than) in the external reality is one of the hallmarks of cognition and autonomy.

In addition, one intriguing property of the evolved internal models is that not only they facilitate motor execution and control when *coupled* with the external environment but, under certain conditions (e.g., when the robot is blindfold), they can self-sustain, so that they afford action *in absence* of external stimuli, or when the agent is *detached* from the external environment. Crucially, this dual property of coupling and detachment makes internal models ideal candidates to support many more on-line and off-line processes. Several researches have recently argued that off-line reuse and reenactment of internal models is the ideal candidate for the passage from sensorimotor to more advanced cognitive capabilities. Some examples of ‘cognitive’ reuses of internal models that have been recently advanced are off-line motor planning by *chaining* multiple predictions and inhibiting external inputs and motor commands (see e.g., Jeannerod, 2006; Nishimoto and Tani, 2009; Pezzulo, 2008), as well as advanced social skills such as mindreading and imitation (see e.g., Hurley, 2008; Oztop et al., 2005; Wolpert et al., 2003).

Although our experiments did not address directly any of these abilities, and acting blindfold is, admittedly, only one example of internal models ‘reuse’, it is worth noting that it is not qualitatively different from more sophisticated forms of reuse such as motor planning or motor imagery. Indeed, even in our experiments it is required that the agent ‘*closes*’ its (*sensorimotor*) loop internally, in the internal model, instead than externally (in the world, via its actions and the feedback it receives); see (Grush, 2004; Hesslow, 2002; Jeannerod, 2006; Pezzulo, 2011; Pezzulo and Castelfranchi, 2009) for further discussions on the off-line reuse of internal modeling for higher-level cognitive abilities, and (Pezzulo, 2008) for a preliminary study on this topic.

Finally, one last point that we would like make is the fact that the model proposed in this paper might also represent an effective methodology for an engineering perspective, i.e. from the point of view of developing autonomous robots than can operate reliably in uncontrolled conditions. Indeed, the ability to overcome the problems caused by the fact that sensory information is temporarily unavailable (for example, due to visual occlusions, incompleteness and/or noise) represents a crucial prerequisite for the possibility to tackle a large variety of potentially useful applications.

*Acknowledgments* The research leading to these results has received funding from the Europeans Community 7<sup>th</sup> Framework Programme under grant agreements ITALK (ICT-214668), Goal-Leaders (ICT-270108), and HUMANOBS (ICT-231453). The authors thank Domenico Parisi for his suggestions and insightful comments on a preliminary version of this paper.

## References

- Adams, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior*, 3:111–149.
- Beer, R. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72:173–215.
- Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11:209–243.
- Beer, R. D. and Gallagher, J. C. (1992). Evolving dynamical neural networks for adaptive behavior. *Adapt. Behav.*, 1(1):91–122.
- Blakemore, S.-J., Wolpert, D. M., and Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7):635–640.
- Bongard, J., Zykov, V., and Lipson, H. (2006). Resilient machines through continuous self-modeling. *Science*, 314(5802):1118–1121.
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47(47):139–159.
- Cerminara, N. L., Apps, R., and Marple-Horvat, D. E. (2009). An internal model of a moving visual target in the lateral cerebellum. *J Physiol*, 587(Pt 2):429–442.
- Clark, A. and Grush, R. (1999). Towards a cognitive robotics. *Adaptive Behavior*, 7(1):5–16.
- Craik, K. (1943). *The Nature of Explanation*. Cambridge University Press, Cambridge.
- Demiris, Y. and Khadhour, B. (2005). Hierarchical attentive multiple models for execution and recognition (hammer). *Robotics and Autonomous Systems Journal*, 54:361–369.
- Desmurget, M. and Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends Cogn. Sci.*, 4:423–431.

- Dretske, F. (1981). *Knowledge and the flow of information*. MIT Press, Cambridge MA.
- Ebner, T. J. and Pasalar, S. (2008). Cerebellum predicts the future motor state. *Cerebellum*, 7(4):583–588.
- Erlhagen, W. (2003). Internal models for visual perception. *Biological Cybernetics*, 88:409–417.
- Frith, C. D., Blakemore, S. J., and Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philos Trans R Soc Lond B Biol Sci*, 355(1404):1771–1788.
- Gigliotta, O. and Nolfi, S. (2008). On the coupling between agent internal and agent/ environmental dynamics: Development of spatial representations in evolving autonomous robots. *Adaptive Behavior - Animals, Animats, Software Agents, Robots, Adaptive Systems*, 16(2-3):148–165.
- Gigliotta, O., Pezzulo, G., and Nolfi, S. (2010). Emergence of an internal model in evolving robots subjected to sensory deprivation. In *Proceedings of SAB 2010*.
- Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27(3):377–96.
- Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, 6:242–247.
- Hoffmann, H. (2007). Perception through visuomotor anticipation in a mobile robot. *Neural Networks*, 20:22–33.
- Hommel, B., Musseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (tec): a framework for perception and action planning. *Behavioral and Brain Science*, 24(5):849–78.
- Hurley, S. (2008). The shared circuits model (scm): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences*, 31:1–22.
- Imamizu, H., Kuroda, T., Miyauchi, S., Yoshioka, T., and Kawato, M. (2003). Modular organization of internal models of tools in the human cerebellum. *Proc Natl Acad Sci U S A*, 100(9):5461–5466.
- Jeannerod, M. (2006). *Motor Cognition*. Oxford University Press.
- Johnson-Laird, P. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge University Press and Harvard University Press, Cambridge.
- Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, 4:71–115.
- Johnsson, M., Balkenius, C., and Hesslow, G. . . (2009). Neural network architecture for crossmodal activation and perceptual sequences. In *Papers from the AAAI Fall Symposium (Biologically Inspired Cognitive Architectures) 2009*, pages 85–86, Arlington, Virginia, USA.
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9:718–27.
- Keijzer, F. (2001). *Representation and behavior*. MIT Press, Cambridge, MA.
- Kording, K. and Wolpert, D. (2006). Bayesian decision theory in sensorimotor control. *Trends Cogn. Sci.*, 10:319–326.
- Lee, D. N. and Thompson, J. A. (1982). Vision in action: the control of locomotion. In Ingle, D. J., Goodale, M. A., and Manfield, R. J. W., editors, *Analysis of Visual Behavior*, pages 411–433. MIT, Cambridge.
- McFadyen, B. and Winter, D. (1991). Anticipatory locomotor adjustments during obstructed human walking. *Neuroscience Research Communications in Mathematical Physics*, 9(1):37–44.
- McFadyen, B. J., Malouin, F., and Dumas, F. (2001). Anticipatory locomotor control for obstacle avoidance in mid-childhood aged children. *Gait Posture*, 13(1):7–16.
- Mehta, B. and Schaal, S. (2002). Forward models in visuomotor control. *Journal of Neurophysiology*, 88:942–53.
- Miall, R. C. and Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, 9(8):1265–1279.
- Millikan, R. G. (2004). *Varieties Of Meaning*. Mit Press.
- Moller, R. and Schenck, W. (2008). Bootstrapping cognition from behavior—a computerized thought experiment. *Cognitive Science*, 32(3):504–542.
- Nishimoto, R. and Tani, J. (2009). Development process of functional hierarchy for actions and motor imagery: A constructivist view from synthetic neuro-robotics study. *Psychological research*.
- Nolfi, S. (2005). Categories formation in self-organizing embodied agents. In Cohen, H. and Lefebvre, C., editors, *Handbook of Categorization in Cognitive Science*. Elsevier, Amsterdam.
- Nolfi, S. and Floreano, D. (2000). *Evolutionary Robotics*. MIT Press.
- Nolfi, S. and Tani, J. (1999). Extracting regularities in space and time through a cascade of prediction networks: The case of a mobile robot navigating in a structured environment. *Connection Science*,

- 11(2):125–148.
- Oztop, E., Wolpert, D., and Kawato, M. (2005). Mental state inference using visual control parameters. *Cognitive Brain Research*, 22:129–151.
- Pezzulo, G. (2008). A study of off-line uses of anticipation. In Asada, M., Tani, J., Hallam, J., and Meyer, J.-A., editors, *Proceedings of SAB 2008*, volume LNAI 5040, pages 372–382. Springer.
- Pezzulo, G. (2011). Grounding procedural and declarative knowledge in sensorimotor anticipation. *Mind and Language*, 26(1):78–114.
- Pezzulo, G. and Castelfranchi, C. (2007). The symbol detachment problem. *Cognitive Processing*, 8(2):115–131.
- Pezzulo, G. and Castelfranchi, C. (2009). Thinking as the control of imagination: a conceptual framework for goal-directed systems. *Psychological Research*, 73(4):559–577.
- Piaget, J. (1954). *The Construction of Reality in the Child*. Ballentine.
- Schmidhuber, J. and Prelinger, D. (1993). Discovering predictable classifications. *Neural Computation*, 5(4):625–635.
- Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, 11(5):211–218.
- Sheth, B. R., Nijhawan, R., and Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nat Neurosci*, 3(5):489–495.
- Tani, J. (1996). Model-based learning for mobile robot navigation from the dynamical systems perspective. *IEEE Transactions on Systems, Man, and Cybernetics*, 26:421–436.
- Tani, J. (2003). Learning to generate articulated behavior through the bottom-up and the top-down interaction processes. *Neural Netw*, 16(1):11–23.
- Tani, J., Nishimoto, R., and Paine, R. W. (2008). Achieving "organic compositionality" through self-organization: Reviews on brain-inspired robotics experiments. *Neural Netw.*, 21(4):584–603.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55:189–208.
- von Hofsten, C. (2004). An action perspective on motor development. *Trends in Cognitive Science*, 8(6):266–272.
- von Holst, E. and Mittelstaedt, H. (1950). The reafference principle: Interaction between the central nervous system and the periphery. *Naturwissenschaften*, 37:464–76.
- Webb, B. (2004). Neural mechanisms for prediction: do insects have forward models? *Trends Neurosci*, 27(5):278–282.
- Wolpert, D., Miall, C., and Kawato, M. (1998). Internal models in the cerebellum. *Trends Cogn Sci*, 2:338–347.
- Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philos Trans R Soc Lond B Biol Sci*, 358(1431):593–602.
- Wolpert, D. M., Gharamani, Z., and Jordan, M. (1995). An internal model for sensorimotor integration. *Science*, 269:1179–1182.
- Wolpert, D. M. and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11(7-8):1317–1329.
- Ziemke, T., Jirenghed, D.-A., and Hesslow, G. (2002). Blind adaptive behavior based on internal simulation of perception. Technical Report HS-IDA-TR-02-001, Department of Computer Science (School of Humanities & Informatics), University of Skovde, Sweden.
- Ziemke, T., Jirenghed, D.-A., and Hesslow, G. (2005). Internal simulation of perception: a minimal neuro-robotic model. *Neurocomputing*, 68:85–104.