This is a repository copy of *Understanding brain functional architecture through robotics*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/199900/

Version: Accepted Version

---

**Article:**
Prescott, T.J. orcid.org/0000-0003-4927-5390 and Wilson, S.P. orcid.org/0000-0001-8125-5133 (2023) Understanding brain functional architecture through robotics. Science Robotics, 8 (78). eadg6014. ISSN 2470-9476

https://doi.org/10.1126/scirobotics.adg6014

---

This is the author's version of the work. It is posted here by permission of the AAAS for personal use, not for redistribution. The definitive version was published in Science Robotics on Vol. 8, No. 78, 31st May 2023, DOI: 10.1126/scirobotics.adg6014.

**Reuse**
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

# Understanding Brain Functional Architecture Through Robotics

Tony J. Prescott[1*] Stuart P. Wilson[1,2]

[1]Department of Computer Science, University of Sheffield, UK.
[2]Department of Psychology, University of Sheffield, UK.

[*]To whom correspondence should be addressed; E-mail: t.j.prescott@shef.ac.uk

**Robotics is increasingly seen as a useful testbed for computational models of the brain functional architecture underlying animal behaviour. Here we provide an overview of past and current work, focusing on probabilistic and dynamical models, including approaches premised on the free energy principle, situating this endeavour in relation to evidence that the brain constitutes a layered control system. We argue that future neurobotic models should integrate multiple neurobiological constraints and be hybrid in nature.**

## Introduction

Twenty-five years ago, Kenji Doya suggested that the functional architecture of the mammalian brain is assembled around three specialized learning sub-systems (*1, 2*). Specifically, that modular circuits in the cerebellum implement supervised learning, those in the basal ganglia implement reinforcement learning, and those in the cerebral cortex implement unsupervised learning (see Figure 1). Since then, many detailed descriptions of the computations performed by a wide range of brain sub-systems have been developed and refined, together with theoretical proposals about how their biological substrates combine to generate adaptive behavior. This work, a synergy of effort from experimental and computational neuroscientists, control theorists, and cognitive scientists, has inspired a view of the brain as having a *hybrid* architecture in which different, partially modular brain sub-systems contribute distinct and complementary functionality (e.g. (*3–6*)). Further, a wide range of evidence supports the view that this brain architecture is also *layered*. That is, it is assembled from multiple, separable levels of control, where, at each level, the sensory apparatus is interfaced with the motor system (*7*). Broadly speaking, these controllers are arranged with fast but inflexible controllers at the bottom, and slow but flexible

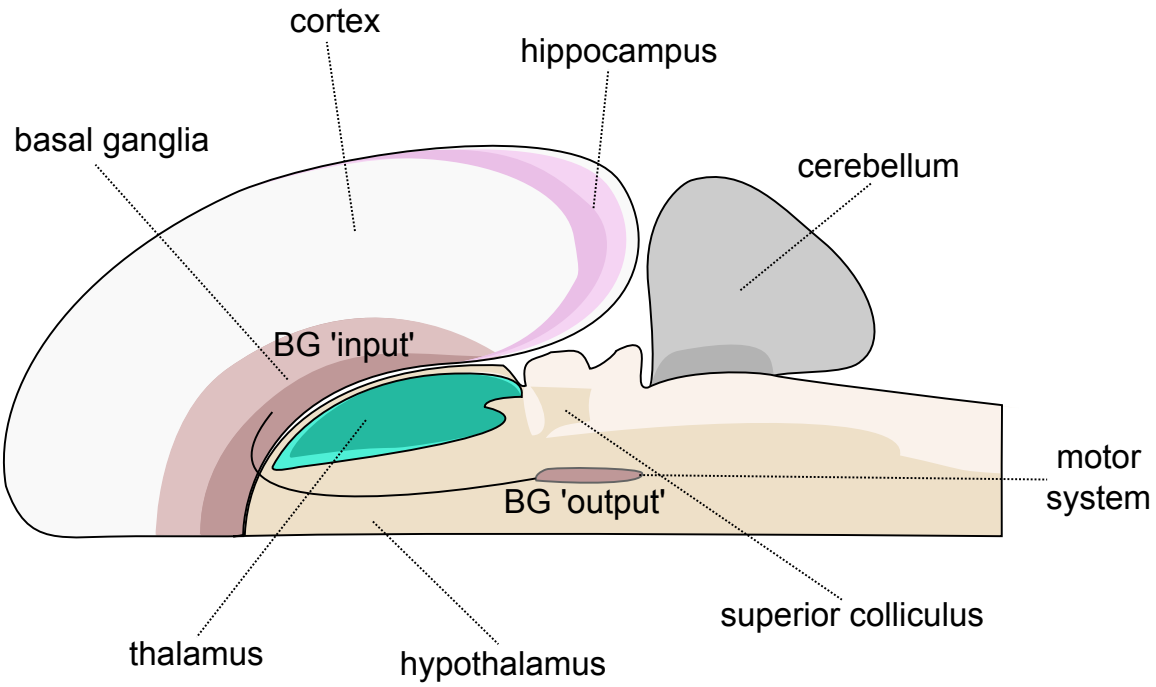


Figure 1: *Major functional divisions of the vertebrate brain.* 'Flatmap' of the rat brain, indicating the overall arrangement of several major functional regions. These have been the targets of some of the computational and neurorobotic modelling works discussed in this article.

ones at the top (*4, 7*). As originally proposed by John Hughlings Jackson (*8*), this architecture exhibits dissociations such that higher-level systems can be removed or damaged without fully compromising lower ones but not vice versa. This feature distinguishes layered control schemes from hierarchical ones and contributes to the robustness of the former (*7, 9*).

Neurorobotics, the effort to demonstrate the usefulness of brain-inspired control in embodied physical systems (see Figure 2), has challenged these proposals to be better specified, and to make their theoretical assumptions more visible. In so doing, it can provide a stringent test of the completeness of the underlying theories, particularly in terms of the purported role of target brain sub-systems in the real-time co-ordination of sensing with action (*10–13*). Robotic models have, for instance, added to our understanding of the role of the spinal cord/brainstem in motor pattern generation (*14, 15*), cerebellum in predictive control (*16–21*), of the basal ganglia as a key substrate for action selection and reinforcement learning (*22–25*), hippocampus as an attractor network that supports memory storage and retrieval (*26–30*), and the cerebral cortex as a locus for self-organising somatotopic maps (*31*), multisensory convergence (*32*), mental imagery (*33*) and meta-cognitive control (*34*).

More broadly, a growing literature is also offering an understanding of neural computation

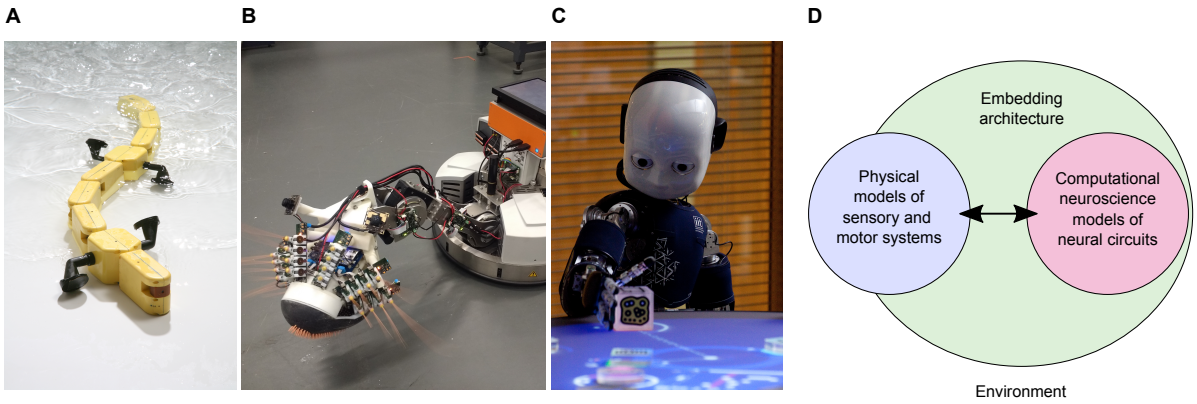


Figure 2: *Example neurobotic research platforms.* **A** *Salamandra Robotica*, an amphibious robot that can both swim and walk controlled by a model of oscillatory networks in the vertebrate spinal cord (*14, 15*). Image from Kostas Karakasiliotis, Biorobotics Laboratory, EPFL (with permission). **B** *WhiskEye*, a robot model of brain systems, including cerebellum, basal ganglia, superior colliculus and hippocampus, underlying navigation and active sensing in the mammalian vibrissal system (see (*41–44*)). Image from Martin Pearson, Bristol Robotics Laboratory (with permission). **C** *iCub*, a humanoid robot widely used to model human perception, cognition, motor control and social interaction (e.g. (*13, 29, 31–33*)). Image from the University of Sheffield (with permission). **D** The design of a neurorobot involves interfacing computational models of neural circuits with physical models of animal sensory and motor systems. Typically, the control systems also include components, that are not neurally-inspired, that mediate this interface (the "embedding architecture", see (*23*)). In a well-designed robot, that is suitably matched to its environment/niche, the morphology of the robot body can simplify some aspects of control (known as "morphological computation" (*11, 45*). The resulting physical model of the organism is able to operate in real-world environments serving as a strong test of the capacity of embedded neural models to generate adaptive behavior.

in both vertebrate and invertebrate nervous systems via neurorobotics (*11, 35–37*). These efforts range from embedding models of specific neural circuits implemented as networks of 'spiking' (leaky integrate and fire) artificial neurons, to those more concerned with computational principles implemented using more abstract network models. The latter are often formulated at a purely algorithmic level, frequently with an emphasis on matching wider system dynamics and behaviour. The approach, recently described with the phrase "cognitive neurorobotics" (*38*), will be our main focus here. This growing importance of robotic modelling in neuroscience evidences Rosenbleuth and Wiener's dictum that a suitably designed physical model can provide a practical and useful means of investigating scientific theories of complex biological systems (*39, 40*). Such a device can, in itself, stand as the most complete expression of a theory and as a tangible demonstration of its capacity to explain, predict and control (*36*).

The early history of AI and cognitive science was marked by a strong interest in unifying

theories of cognition, such as Soar (*46*) and ACT (Adaptive Control of Thought) (*47*), instantiated as working models of goal-oriented cognition known as cognitive architectures (for review see (*48*)). These theories sought to characterise the function of the mind/brain in terms of a relatively small number of core mechanisms. For instance, the Soar architecture cast the challenge of cognition as that of search in a problem space through recursive application of a limited set of operators (*46*). The quest for a unified theory has been re-ignited in recent decades by the identification of a single key principle – minimisation of free energy (*49*) – and its instantiation as a cognitive architecture through the framework of predictive processing (*50, 51*). The intriguing proposal is that what has appeared to be a hybrid brain architecture to a generation of theorists, may, after all, be explainable by a single underlying mechanism, recursively applied (*50*). Robotics is increasingly emerging as an important means for evaluating this hypothesis, and for determining whether it can fulfil its promise (*52*).

In this article, we review the foundations for this new framework in earlier theories of robot control, and look at their subsequent development through the incorporation of recurrent neural networks. A key observation is the convergence of approaches resulting in state-of-the-art bioplausible robotic control systems that combine elements of predictive processing and network attractor dynamics (*38, 53, 54*). We identify some promising insights that the framework has uncovered through robotic investigations. In doing so, we attempt to situate this endeavour, and wider research in neurorobotics, in relation to the evidence that the brain implements a form of layered control (*7*). From this perspective, we will also consider the view of living organisms as self-producing networks of processes that maintain themselves through constraint closure (*55–57*), and the implications of this view for the generality of the free energy principle (FEP). We conclude by considering what future theories of brain functional architecture, as instantiated in robotics, might look like.

# Selecting between control modules

The foundations for the current state-of-the-art trace back at least as far as the work of Doya's collaborators, and their now classic Mosaic model (MOSAIC; modular selection and identification for control) (*58*), which we examine first (see Figure 3).

The basic approach behind the Mosaic model (*58*) involves tuning pairs of forward and inverse models (e.g., artificial neural nets) based on how well their outputs match the future states (e.g., positions and velocities) of the robot. In each pair, the forward model registers the current state of the system and the current motor command and predicts its future state, and the inverse model registers desired changes in state and computes a motor command that should be appropriate to effect that change. As suggested by Wolpert & Kawato ( (*58*); see also (*60*)), the probability that each forward model best predicts the future state can be represented using a decreasing (Gaussian) function of the discrepancy between its prediction and the next measured state of the system — the prediction error — before being normalised to give a probability density over all forward/inverse pairs.

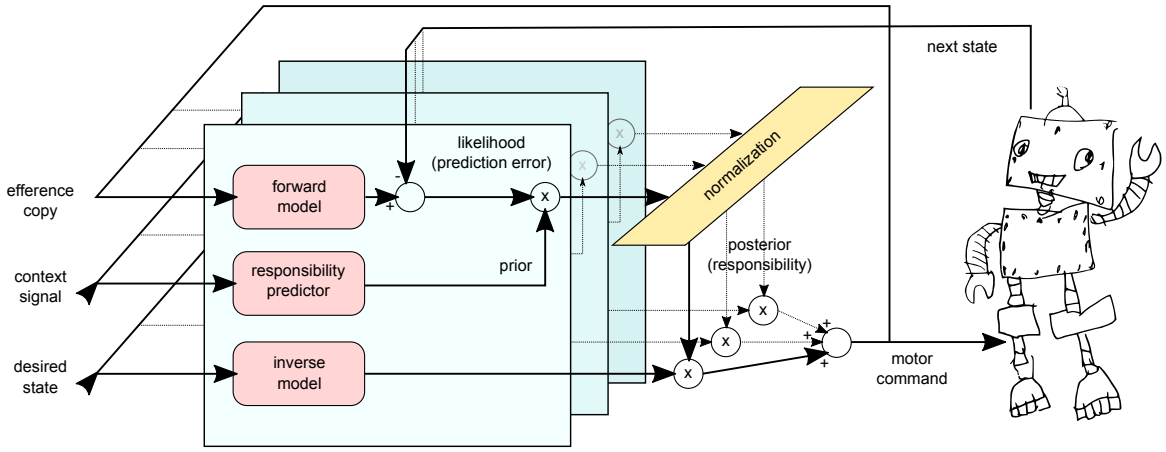


Figure 3: *Selecting between control modules*. Architecture diagram for the Mosaic model (*58*). Pairs of forward and inverse models are tuned according to how well their outputs match the future states of the robot. Proprioceptive prediction errors generated by each forward model are weighted by contextual signals from other sensors and then normalized across modules to create a distribution of 'responsibilities', which in turn weight the contribution of corresponding inverse models to the overall motor command. Prediction errors, responsibility predictors, and the normalized responsibilities can be considered as likelihood estimations, priors, and posteriors, and as such the model implements a form of the Bayesian belief updating that is used extensively in current predictive processing systems. Based on Haruno et al. (*59*), Figure 1., omitting notation and an additional feedback controller loop for clarity. 'Judy the Robot' drawing by Pippa Wilson.

The resulting distribution of 'responsibilities' then determines the proportion of the overall error to be allocated when tuning each model pair, as well as the contribution that a given inverse model should make to the overall motor command. The result is a (soft) selection amongst the actions requested by each inverse model, which should be appropriate for attaining more desirable future states. If a prediction of each responsibility, cued by contextual information from other sensors (visual, audio, etc.), is similarly tuned, the system can also be expected to bias requests for action in a contextually appropriate manner. For example, in a world where blue objects are heaviest, blue regions in a camera image should come to bias the selection of inverse models applying greater lifting forces.

Allowing such a control system to calibrate itself in this way can be seen as a form of Bayesian belief updating, i.e., the iterative application of Bayes' rule to compute the next (posterior) distribution of responsibilities, using that computed on the previous iteration as the current prior. This process minimises the KL (Kullback–Leibler ; see (*61*)) divergence between the distribution of context-dependent predicted responsibilities (the prior), and that which maps appropriate motor commands to desirable states in the particular environment in which the robot learns. This minimisation process, we will see later on, is a key feature of approaches to control within the modern predictive processing framework. An important idea is that the acquired distribution of responsibilities constitutes a statistical model for effective motor control in the robot's particular environment (*16*).

To test these ideas, Haruno et al. (*59*) trained such a network to control a (simulated) robot arm to follow desired trajectories while holding objects with different physical properties paired with information (e.g., visual) about object identity. The network learnt to effectively switch between modules (forward/inverse pairs) appropriate for the control of each object substituted into the task. It was also able to adapt its control effectively for objects with novel physical parameters. Moreover, it generated sharp motor errors when objects were paired with incongruent object identity information. A notable feature of this study is the (optional) inclusion of an adaptable variance for each Gaussian function mapping prediction errors to responsibilities, resembling the idea of 'precision weighting' that features in later, more explicitly Bayesian, approaches to control (*62*).

The substrate for learning 'internal models' was identified initially with microcircuits of the cerebellum (*16*), in line with the anatomical decomposition of control by Doya (*1*). However, Doya and Kawato later worked together to extend the Mosaic framework to include responsibilities that are weighted by Sutton & Barto's classic temporal difference error signal (*63*). This work showed that reinforcement learning tasks, including learning the upswing phase in a cart-pole simulation, could be solved using the basic ingredients of modular decomposition and soft selection (*64*).

An approximation to the temporal-difference learning algorithm was later shown to be implementable within this framework, using a two-layered architecture broadly compatible with the anatomical structure of the basal ganglia (*3*). Here, one set of responsibility signals selects between forward models, a second selects amongst models driven to predict future levels of a reward signal, and a third learns to select amongst modules that combine the (responsibility

weighted) outputs of the first two using the temporal-difference learning formula. Accordingly, the outputs from the forward and reward models are interpreted in the reinforcement learning framework as estimates of the local value function and temporal difference errors, respectively. The output following selection at the top level is a control signal that, after adjustment of the responsibilities via Bayesian updating throughout, generates movements to maximise reward.

To summarise, the main ingredients for robot control under the Mosaic model are i) a separation of control into modules that compete for selection by minimising sensor-motor prediction errors, ii) a soft selection between modules based on responsibility weightings that are reconfigured through Bayesian updating, iii) incorporation of additional signals into responsibility weightings to generate context-dependent motor responses, iv) the potential to combine and reconfigure several such systems, e.g., to maximise reward signals during reinforcement learning tasks.

These ingredients have been re-used and elaborated in more general neurorobotic systems since; notably via the 'Hammer' model (*65*), as a theory of imitation learning inspired by the primate mirror neuron system. In this case, the approach involves selecting controllers based on the error each forward model would make if predicting the state of a demonstrator (e.g., a human) based on the command generated by its associated inverse model. A novel contribution is to suggest how the weighting of each controller (akin to its responsibility) may bias selection towards image features associated with successful predictions. This elaborates the basic operation of the responsibility predictors through a saliency map, into which the representations of multiple stimulus features (visual motion, colour etc.) may be projected, analogous to how activity patterns in multiple cortical feature maps may bias visual attention via interactions with the midbrain superior colliculus (*57*).

The Mosaic model thus provides a means of investigating how cerebellar and midbrain circuits might realise supervised and reinforcement learning, respectively, and how cortical learning may allow for the context-dependent elaboration of such computations. Next we will see that a parallel strand of robotics work from a different group shows how very similar ingredients might be combined for perceptual processing.

# Representing actions via attractor dynamics

At the same time as Wolpert and Kawato (*58*) were thinking about the role of cerebellar processing in controlling action, Tani and Nolfi (*66*) were thinking about the role of thalamo-cortical loops in perception along similar lines. They independently arrived at a similar method for weighting learning updates and control signals from an array of competing modules, via the more explicitly Bayesian formalism of the 'mixture of experts' learning algorithm developed by Jacobs et al. (*67*). An important difference is that the modules competing to respond to sensory-motor inputs were composed of neural networks with recurrent loops. These nets were of the kind shown by Elman (*68*) to be capable of extracting the statistical structure behind sequences of inputs using error backpropagation to refine predictions of the next input in the

sequence, i.e., to minimise prediction errors. A second key point is that the (responsibility weighted) outputs from each module were sampled at a slower timescale (every tenth iteration), before being presented as inputs to a second set of recurrent networks, at which the same responsibility-weighted soft-selection mechanism was applied. The second system was trained to predict the outputs of the first at the next timestep, i.e., to move the distribution of its outputs towards that of its inputs, and thus implicitly to approximate the dynamics achieved by Bayesian updating in the Mosaic model.

To test these ideas, a virtual robot was moved independently through a 2D environment consisting of two rooms, each containing a different set of obstacles (internal walls). After learning, the lower-level system was observed to abruptly select between its recurrent nets in lockstep with changes in sensory events that corresponded to specific movement patterns relative to the obstacle features, e.g., following walls or turning corners. Importantly, the second system learnt to select from its recurrent nets, that which produced the lowest prediction errors when the robot was in a given room, at the point when the robot moved between those rooms. Tani & Nolfi thus concluded that the system had learnt to simultaneously represent concepts corresponding to the sensor-motor 'flow' induced by obstacle features in the first layer, and the higher-order concepts of 'room A' or 'room B' in the second. The representation of these concepts manifest at each level as the selection of recurrent nets whose intrinsic attractor dynamics were evoked in the presence of the corresponding sensory-motor patterns. These self-organised dynamics provide a remarkable demonstration of how discrete symbols (selected attractors) for meaningful temporally extended sensor-motor events can emerge within a continuous-time system.

An important follow-up study by Yamashita & Tani (*69*) demonstrated the usefulness of attractor dynamics for representing aspects of motor control, and specifically the sequencing of actions across multiple timescales. They critiqued models based on the Mosaic architecture in terms of the conflict between segmentation and generalizability inherent in its modular design. As a more biologically motivated alternative they investigated the properties of highly recurrent networks comprising units with different intrinsic dynamical timescales. In Yamashita & Tani's model, a population of input/output units receiving robot state information (proprioception, and exteroception for object localisation) projected to the fast units (only), and the fast and slow units were otherwise similarly connected (see Figure 4).

Networks were trained by an experimenter moving the arms of a humanoid robot through sequences of actions, e.g., touch an object several times, move it left and right or up and down several times, clap the hands several times etc.. Transitions between these actions could take different paths with various lengths before the robot was to return to its initial pose. During training, the slow units were initialised to particular states at the beginning of each sequence, such that the potential for specific action sequences to be learnt could be investigated by resetting the initial state of the slow units accordingly. The network was trained offline, using a backpropagation algorithm to predict the next sensor-motor state. After training, the network was successful in reproducing the sequences of actions by controlling the robot via the input/output units.

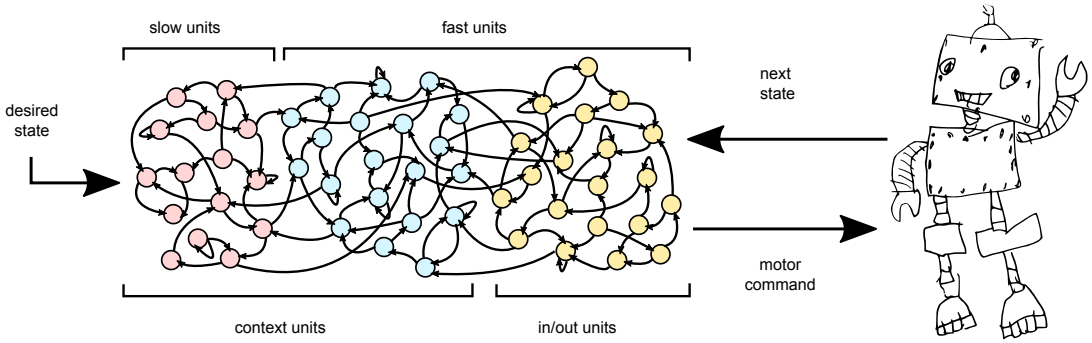


Figure 4: *Representing actions via attractor dynamics*. Architecture of the model by Yamashita & Tani (2008), in which distinct attractor dynamics in the fast units of a continuous time recurrent neural network come to represent distinct actions performed by a robot, and the dynamics of slower units come to represent configurations of actions. Inputs to the system are proprioceptive signals and visual information about the position of an object to be manipulated. Slow and fast context units are point neuron models with different intrinsic integration time constants. Every unit is connected to every other, including itself, with the exception of input units which do not have a direct connection to the slow context units (an illustrative subset of units and connections is shown for clarity). Activations of the output units are considered to be predictions about future proprioceptive states, and prediction errors are used to adjust the network connection weights using a backpropagation algorithm. Adapted from (*69*), Figure 3A.

During repetitive movements, similar activation dynamics were observed to repeat in the fast units, while activity patterns of the slow units were invariant except at the transitions between sub-tasks, complementing the findings of Tani & Nolfi for perception (*66*). Novel sequences of those sub-tasks could later be learnt when the weights of the fast units were frozen and only the units of the slow population were able to continue adapting. This suggests that modular segmentation of control can be achieved by the intrinsic self-organising properties of networks with multi-timescale dynamics, and hence that separation and selection between modules need not otherwise be prescribed by architectural constraints.

Nevertheless, the fundamental similarities between networks for robot perception (*66*) and control (*58, 69*) can be seen as early indications that perception and action may be two sides of the same coin, a key principle in the modern context of 'predictive processing', which we introduce next.

# Generating actions to avoid surprise

The 'predictive processing' framework (*50, 51*), also termed 'active inference', and the emerging 'physics of sentience' developed by Karl Friston (*70*), stems from a description of living systems in thermodynamic terms. Accordingly, an organism can be thought of as a particle, bouncing around an energy landscape, maintaining low entropy in defiance of the second law of thermodynamics. To persist in doing so – to self-organise – is to occupy a set of internal states that is separate from the set of external states of the environment, and to maintain a boundary between the two. This boundary, called a Markov blanket, consists of the states of the organism's sensors and effectors, and mediates the exchange of information between the organism and the environment. In these terms, perception is the flow of information from external to internal states via the states of the sensors, and action is the opposite flow, from internal to external states via the effectors.

From thermodynamic principles it follows, somewhat teleologically, that to self-organise is to maintain – by perception and action – a flow of information that, in both directions, reduces uncertainty about the underlying causes of external states. This uncertainty is bounded by a thermodynamic quantity called *free energy*, which is a measure of the inaccuracy of predictions about future sensory states, now and in the longer term ( (*49, 70*); see also (*61*)). Organisms thus defy the second law by minimising free energy, and in turn their uncertainty about future sensory states, by choosing actions least likely to have surprising consequences, and by adjusting predictions to make future sensory states less surprising. By doing so, internal states become implicit models for the causes of sensory states, insofar as the physiology of a fish or a bird may be considered a model of hydrodynamics or aerodynamics (see (*56*)). As actions in turn constrain (reduce the degrees of freedom of) sensory states, organisms implicitly construct models of their own dynamics, while modifying those dynamics to reinforce the model, in a cycle referred to as 'self evidencing' (see also (*71*)).

Implementations of robot control systems based on the FEP have recently been reviewed

by (*62*). These include applications for visually guided reaching (*72, 73*) and active vision (*74*), robot navigation (*75*), and self versus other discrimination ( (*76*); see also (*38, 77*)).

In the context of motor control, an interesting model was recently developed by Ahmadi and Tani (*53*). A key feature of their approach is to connect units of a recurrent network (comprising non-linear point neurons) with stochastic 'latent' units, whose activations are drawn, for example, from a Gaussian distribution with an adaptable mean and variance. This allows each latent unit to represent a prior distribution whose mean and variance is determined from activations of the recurrent units at the previous timestep, and a posterior distribution whose mean and variance are similarly determined by an adaptable parameter. As such, the (variational) free energy can be expressed at each latent unit as a sum of the KL divergence between the prior and posterior and the (negative log) prediction errors, potentially after backpropagation through several layers. The network weights, and the adaptable parameter, could then be adjusted to minimise the free energy using a stochastic gradient descent method. Importantly, this network can on one hand be used as an inference model by propagating state information from recurrent units to latent units. It can also be used as a generative model by projecting latent unit activations – i.e., samples from the distribution set by the adaptable parameter – to recurrent units and then, for example, to a PID controller.

Crucially, latent units allow signals to be generated from within a neural network and provide a simple interface between continuous-time and probability-based formalisms. This permits direct calculation and manipulation of statistical quantities, such as the free-energy, and therefore allows generative processes to be incorporated into potentially complex network architectures.

Ahmadi and Tani (*53*) set up two populations, one comprising neurons with fast dynamics, which projected to a second with slower dynamics, each with its own latent units. The model was interfaced with a humanoid robot and the fast units were presented with state information from its joint angles, and information about the hand location of a second robot through which identical movement patterns were generated by human manipulation. Movement patterns were sequences of actions, each containing repetitive elements similar to (*69*). The robot then registered prediction errors by comparison only with its 'view' of the other robot's hand movements. Calibrating the generative model by minimising its free energy enabled the robot to imitate the movements of the other, and even to generate anticipatory movements.

An interesting follow-up by Idei et al. (*54*) incorporated latent units into recurrent networks composed of neurons with multiple timescales as the basic description of a cortical 'area'. In this study, joint angles (proprioception) and hand coordinates (exteroception) of a robot arm were interfaced with a control architecture comprising two corresponding 'sensory areas', an 'association area', and an 'executive area', configured in a hierarchy from sensory to associative to executive. See Figure 5 for details.

During an initial learning phase the adaptive parameters for the posterior distribution of each latent unit, and the recurrent network weights that determine the corresponding priors, were adjusted to minimise free energy using a stochastic gradient descent method. The network was trained in this way while observing sequences of sensory states obtained from the robot as

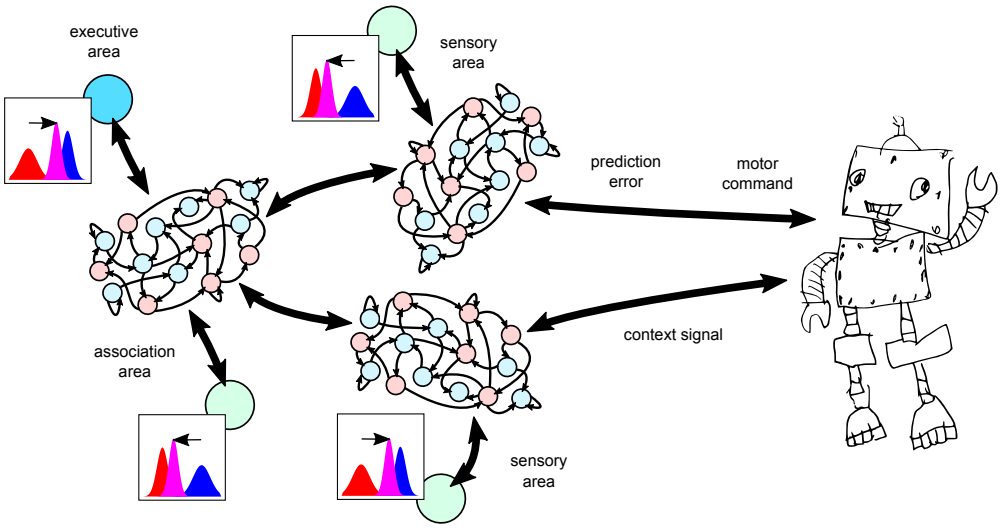


Figure 5: *Generating actions to avoid surprise*. Architecture diagram based on the study of Idei et al. (*54*), which demonstrates the emergence of different free energy regimes when proprioceptive and exteroceptive signals provide congruent or incongruent information. The executive area consists of a single latent unit (large blue circle), the association area contains 15 recurrent units (eight fast, blue, and seven slow, red) and three latent units (large green circle, one shown), and the two sensory areas each contain 15 recurrent units and one latent unit. Each recurrent unit in the association area receives weighted inputs from the latent unit in the executive area, from the three latent units in the association area, and from the activation of all recurrent units in the association area at the previous timestep. In a given sensory area, each recurrent unit receives weighted inputs from the three latent units in the association area, from the latent unit in the corresponding sensory area, and from the activation of all recurrent units in that sensory area at the previous timestep. A subset of the recurrent units in each sensory area are assigned as output units and a weighted sum of their activations for each sensor value is interpreted as a prediction of its future state. The predicted proprioceptive states are used as the target values for a PID controller to determine the robot's movements, and the proprioceptive and exteroceptive predictions are continually compared to the actual future states to produce prediction errors. Inset graphs indicate shifts in the latent variables (pink) to minimize free energy.

it was manually moved by a human. During testing the weights were fixed and only the latent (posterior) distributions were updated, and the robot generated movement patterns resembling those from the training set.

Interestingly, half way through testing, the exteroceptive inputs were substituted for novel examples from the training set, making the two sensory modalities uncorrelated. At this point the network consistently transitioned from one dynamically stable regime to another. Initially, the prior distribution in the sensory areas had a high precision (small variance relative to that of the posterior) causing large shifts in the posterior that attenuated responses to prediction errors in the sensory areas. However, when the inputs were uncorrelated the posterior distribution in the executive area shifted, rendering the prior distribution in the sensory and association areas relatively low in precision. This caused smaller shifts in the sensory posteriors and an amplification of sensory responses to prediction errors.

The authors interpret these two regimes, which were associated with distinct levels of free energy, in terms of an ability to discriminate between causal structures attributable to the self versus others (*54*). They see the emergent amplification versus attenuation of sensory responses in either dynamical regime as a potential explanation for the observation that we cannot tickle (i.e., surprise) ourselves (see also (*78*)).

Echoing earlier findings by (*66*) and (*69*), a key insight from this work is that while discrete changes in the activity of the top-level executive area correspond to higher-order changes in control strategy, they emerge spontaneously from the overall network dynamics. That is, under the general requirement to avoid surprise. This has important implications for understanding the emergence of 'executive' function in the brain, i.e., as a network-level effect that requires no implicit homunculus to be 'behind the wheel'.

The success of these two uses of multi-timescale dynamics, either between neurons (*53*) or within neurons (*54*) grouped by a latent variable, raises the intriguing possibility for a control scheme in which free energy minimization may be implemented at multiple levels of network 'embedment'. More generally, the work of Tani and colleagues since the 1990s, using predictive processing techniques to orchestrate the dynamics of multi-scale recurrent nets for robotic control, has illuminated several promising points of contact between predictive processing and enactionist views of control and cognition. We will consider some of this work further below, following a brief consideration of some architectural properties of vertebrate brains.

# The brain as a layered hourglass

Current formulations of the FEP most readily permit the engineering of control architectures within a hierarchical design based on homogenous processing elements and (bi-directional) message passing. These algorithmic accounts, when related to the brain, have been most closely matched to intra-cortical and thalamocortical circuitry (*51, 79, 80*). However, it is not clear whether the models we have reviewed so far have specifically utilised, or at least particularly benefited from, the incorporation of such strict hierarchical constraints into the design of their

architectures. Indeed, some of the works considered here have demonstrated that the emergence of functional hierarchies in control systems need not be isomorphic with the strict hierarchical constraints that might otherwise be used to engineer control architectures (e.g. (*69*)). An important alternative to hierarchical control is heterarchical control. Bechtel and Bich (*81*) argue that strict hierarchies, wherein many low-level controllers are combined by fewer at a higher level and potentially by one at the top of a pyramid, are seldom, if ever, realised in biological systems. Instead, biological control is realised through heterarchies, often involving an inverted pyramid structure, with a smaller number of lower-level controllers being reused and repurposed by many more higher-level controllers.

The evolutionary neurobiology of the vertebrate brain indicates a highly-conserved architecture at the macro-level, in which all major sub-systems, including the cortex/pallium are present in all vertebrate classes including ancestral jawless fish (*82*). As previously noted, this architecture has a layered design (illustrated in Figure 6a), incorporating elements of both hierarchy and heterarchy (*7, 57*), whilst providing a robustness against damage and extremes (through dissociation) that is also a requirement for many modern engineered systems (*9*). Layered control is also a characteristic of several brain-inspired robot control systems such as Distributed Adaptive Control (*4, 13*), and the ArmarX architecture (*83*).

Not only is the overall layered architecture of the brain conserved, but also that of core systems along the midline of the brain involved in functional integration, specifically, the basal ganglia and medial reticular formation (*7, 84, 86*). The nuclei that compose the basal ganglia connect brain sub-systems that can command behaviour with sensorimotor sub-systems that implement real-time control. Collectively the basal ganglia hold a veto over all voluntary movement, acting to ensure that the animal does not simultaneously attempt incompatible actions (*86*). This design constrains the manner in which sensorimotor systems can be accessed while increasing the flexibility of the system to add new behavioral capabilities and to control diverse sensory and musculoskeletal systems in different animal forms. A similar principle of "constraints that deconstrain" has been described for engineered systems that implement such an 'hourglass' architecture (see (*9*) and Figure 6b). The presence of an hourglass architecture for action selection within an overall layered control system provides a robust but flexible way of assembling a brain-inspired control architecture which we have demonstrated in a variety of animal-like robot platforms (*11, 41, 87*) and in a robotic model of human gaze control (*85*).

# Cognition through constraint closure in self-producing systems

For Bechtel and Bich (*81*), even very simple biological systems may be considered to embody cognition. They suggest that, to some extent, an allosteric enzyme 'decides' the rate of a chemical reaction based on sensing the presence of a particular molecule, and in doing so helps maintain the functional identity of the organism in return for the matter and energy required for its construction and repair. Accordingly, control mechanisms are special kinds of constraints, whose effect on the system (to reduce its degrees of freedom) is not permanent, and even the

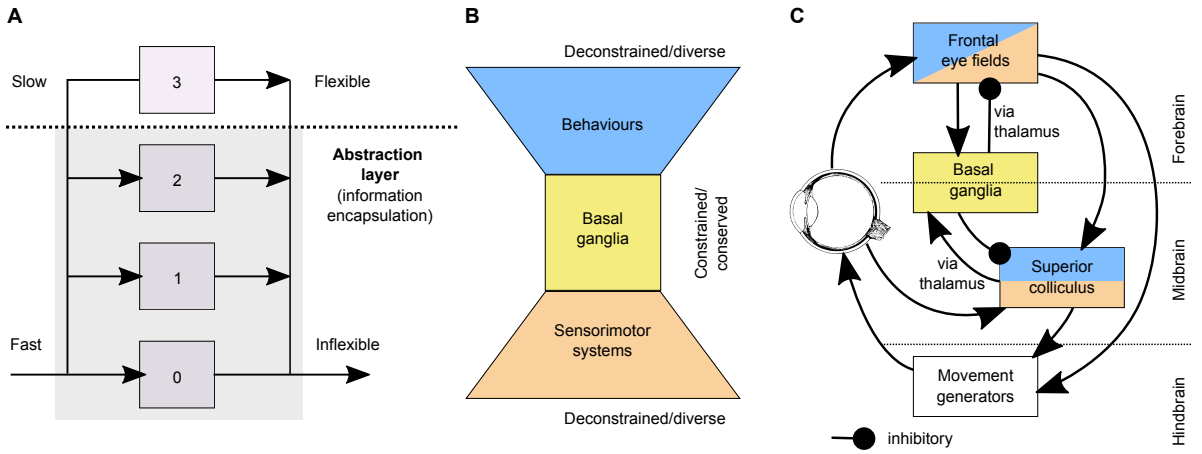


Figure 6: *Layered and hourglass architectures*. **A** Layered architectures exhibit dissociations, such that the lower levels still operate, and exhibit some sort of behavioural competence, in the absence through damage (or removal) of the higher layers but not vice versa. Layers may implement different solutions to similar problems, typically these solutions are fast but inflexible in lower layers, and slower but more flexible in higher layers. Layered control systems are not strict control hierarchies. The interface between layers may lead to information encapsulation, that is processing in lower layers may not be fully accessible at higher levels. This can be described as giving rise to an abstraction layer, in computer science terms, and can be related to the idea of a Markov Blanket (see text). **B** In an hourglass (or bowtie) architecture conserved processes and protocols at the interfaces between layers can operate as "constraints that deconstrain" (*9*). For example, the highly conserved intrinsic circuitry of the vertebrate basal ganglia (BG) (*22, 84*), which holds a veto over all voluntary movement, and connects brain sub-systems that can command behaviour with sensorimotor sub-systems that implement real-time control, may operate as a form of hourglass architecture, constraining the manner in which sensorimotor systems can be accessed but increasing flexibility in the capacity of the system to add new behavioral capabilities. **C** Layered and hourglass control in the mammalian oculomotor system. A forebrain system involving the frontal eye fields and the midbrain superior colliculus (SC) are both capable of generating signals that direct eye movements. The forebrain system operates, in part, by modifying the behaviour of the SC, but the SC is able to generate eye movements in the absence of cortex (blindsight). The overall system can therefore be described as a layered architecture. Both systems project into the basal ganglia (striatum) and receive inhibitory output from the basal ganglia (substantia nigra). The basal ganglia thus instantiates a form of hourglass architecture gating the capacity of both systems to access motor outputs (see (*85*)).

simplest organisms embody multiple such mechanisms. Moreover, "control mechanisms in living organisms are not the product of such [hierarchical] design; they have arisen through evolutionary processes in which variants that improve, or at least do not significantly diminish, the ability of organisms to maintain themselves and leave offspring become fixed as a result of selection or drift. In such an evolutionary process, one should expect new control mechanisms to be added independently of one another in an opportunistic manner" (*81*) (p5). The sub-systems underlying mammalian thermoregulation show this characteristic – a layered architecture of parallel and partially independent sub-systems that are not fully integrated within the brain, but interact through feedback loops involving the body and the environment (*57*). Rather than directly co-operating to maintain temperature at a set-point, in the classic manner of feedback control, they operate synergistically to maintain a dynamic 'balance point' within the organism's adaptive range.

Here the physics of sentience meets an alternative thermodynamic treatment of living systems, according to which a prerequisite for the emergence of a biological whole is 'constraint closure' (*55, 57, 81, 88–90*). Processes that constrain a system become integral to its dynamics if, for each in a set of such constraints, the constraint both depends on another (its dynamics are slower) and is generative for another (its dynamics are faster). For the theoretical biologist Stuart Kauffman, constraint closure systems meet the criteria for evolvable life (*91*). Each new constraint that is incorporated into the dynamics of the system can be identified with a new biological function, insofar as it also becomes necessary for maintaining the system dynamics, and each yields further opportunities for more thermodynamic work to be done and new functionality to arise. The role of a given constraint in maintaining the system is then, by definition, its function – e.g., the function of a heart is to pump blood only insofar as pumping blood is a necessary condition for maintaining the whole. So when new opportunities for maintaining the whole arise, new constraints can be incorporated and new functionality can emerge. We have recently explored this principle in the context of the functional architecture of the mammalian brain, emphasising how control loops can extend beyond the body and operate across multiple – evolutionary, developmental, and behavioural – timescales (see (*57*)).

A prerequisite for constraint closure is that a path must exist through which all constraints can affect all others (operational closure), as has been claimed for collectively free-energy minimising systems (*56*). According to this interpretation of the FEP, as developed by Kirchhoff et al. (*56*), organisms are better thought of as systems of interacting particles, each wrapped up in its own Markov blanket, interacting with the others so as to collectively maintain a 'superordinate' blanket that partitions the dynamics of the collective system from that of the environment (*56*). Cells maintain blankets around interacting molecular processes, tissues blanket interacting cells, organs blanket interacting tissues, organisms blanket interacting organs, and indeed ecosystems maintain blankets around interacting organisms. From this viewpoint, an important concept is that to maintain a superordinate Markov blanket, and thus to be self-evidencing at the level of the organism (or ecosystem), the component parts of the system must be coupled and their internal states must be conditionally independent — a condition that is met when the characteristic timescales of the internal dynamical processes are distinct (*56*).

Whilst Kirchhoff et al. (*56*) see the potential for convergence between the frameworks derived from free energy and constraint closure principles, via Maturana and Varela's notion of autopoesis (*92*), this supposed point of connection has been challenged from the perspective of enaction. As we have seen, the FEP follows from a description of organisms as self-organising systems, which maintain their structural integrity and persist over time, i.e., they maintain, under ergodic assumptions, a (non-equilibrium) steady state. But crystals have this property. Enaction instead describes living systems as networks of processes that maintain themselves (self-distinction) by continually regenerating the networks of processes that they maintain (self-production): Crystals do not have this property. From the enactionist view, a key feature of living systems is that instead of avoiding phase transitions (i.e., by minimizing surprise) they undergo discrete, often irreversible phase transitions, on evolutionary, developmental, and behavioural timescales, triggered by interactions with the environment (*93*). As such, this view emphasises the historical nature of living systems. Crucial to understanding the qualitative changes in states available to living systems, which occur through changes to the body and via adaptations to the environment etc., is to understand the sequence of phase transitions that the system may undergo. Put simply, the brains of humans and mice, of infants and adults, of experts and novices, may all minimise free energy, but the FEP does not illuminate the differences between them.

# The future of neurorobotics

An important aspect of the FEP, evidenced by some of the robotics work reviewed in this paper, is that it provides a specific formal methodology for calibrating a given control network (to minimise surprise), which has led to impressive demonstrations of robot control and provided new insights into how perception and action can interact in cognitive systems (*62*). The work of Tani and colleagues (*38, 53*), using predictive processing techniques to orchestrate the dynamics of multi-scale recurrent nets for robotic control, has further illuminated promising points of contact between the FEP and enactionist views. This work has demonstrated that recurrent networks of processing units with distinct intrinsic timescales, and with generative properties derived from the incorporation of latent variables, can be coordinated by free energy minimisation to yield functional systems within which distinct processing levels become identifiable (e.g., (*54*)).

As a guiding principle, it is important to recognise the distinction between a formal model of a physical system and the system itself (*94*). Whilst it can be tempting to see the physical system as isomorphic to the formal model, i.e., as embedding some version of it, in reality organisms as well as robots are faced with many constraints that are not fully captured by general formal models. These include efficiency, evolvability, learnability, and the capacity of the system to self-assemble (see (*95*) and comments about historicity above). Whilst theorists may seek unifying principles, the biologist is also required to understand at the level of mechanism and to respect constraints deriving from development. Similarly, the roboticist is required to

design at the level of implementation, and to respect constraints deriving from the process of construction. In this context it is notable that the FEP, and the predictive processing approach more broadly, is not prescriptive in terms of specifying architectural and design constraints for robust and efficient control and cognition *a priori*. Indeed, leading proponents assert that predictive processing is not realised exclusively within the brain but within the whole organism and its environment (*96, 97*).

For instance, Clark (*96*) considers a brain-centric version of the predictive processing approach that builds "a kind of neuronally-encoded rich inner recapitulation of an observer-independent reality" (p. 12). He constrasts this with a version better aligned with embodied cognition (and his own 'extended mind') approaches, which is considerably more frugal in its use of internal models. Generative models, in this view, should be as simple as possible to get the job done and should make use of approximations and heuristics. These can include tuning of body morphology to the environment or task (*45*) and outsourcing some of the work of tracking causal relationships to the cognitive niche as restructured by the organism and its conspecifics (*97*). Clark also emphasises how predictive models constructed over slow timescales, can modify and co-ordinate "fast, knowledge-sparse modes of response" (p. 14). This account again tallies with notions of layered control, which typically feature fast but relatively inflexible sub-systems at lower levels, and of constraint closure across different time-scales. For instance, in (*57*) we review evidence that the primate superior colliculus instantiates simple visual filters, present at birth, that can detect species-relevant stimuli such as the faces of conspecifics. These can be viewed as a form of hardwired prior that provides cortical systems with a rich supply of useful data and scaffolds learning of critical cortically-centered skills such as face recognition. The collicular filters themselves are shaped over evolutionary time-scales adapting with the changing needs and capabilities of cortical processing (*98*).

This instrumental view of FEP (see also (*99, 100*)) suggests that homogeneity of processing units is not a prerequisite for cognitive architectures consistent with this general principle. Indeed, the concept of Markov blankets also permits a considerable degree of flexibility in terms of realising the implementation of a predictive processing framework. Specifically, the core property of a Markov blanket is to relax constraints on what forms of processing occur on the other side of the boundary, as long as appropriate signals are provided at the interface. A similar notion in computer science is that of the abstraction boundary which conceals some of the intrinsic processing of a sub-system from external influence, and is also a characteristic of layered architectures in animals and robots (see Figure 6a).

As noted in the introduction, the functionality of various brain sub-systems has been successfully captured within a variety of formalisms. These include formalisms derived from reinforcement learning, adaptive control, and dynamical systems, consistent with Doya's proposal (*1*) that different parts of the brain realise different kinds of learning device. By extending earlier insights, such as those of the Mosaic model (*58*), predictive processing has galvanised recent neurorobotics research, highlighting, from the top-down, formal relationships between different approaches to control. Nevertheless, in our view, predictive processing should not be seen as a singular pathway to exploring natural design principles via robotics. A broader grounding

in biological principles, including those arising from constraint closure and enactivism, and in the history of the emergence of biological intelligence, will be needed (see also (*36*)). As a route to understanding the brain through physical modelling, we consider that a hybrid, layered architecture, including components that minimize free energy, may provide the most promising path. This approach could lead to neurorobotic platforms capable of generating complex life-like behaviours that surprise us... and possibly even themselves!

# References


1. K. Doya, *Neural Networks* **12**, 961 (1999).
2. K. Doya, *Current Opinion in Neurobiology* **10**, 732 (2000).
3. M. Kawato, K. Samejima, *Current Opinion in Neurobiology* **17**, 205 (2007).
4. P. F. Verschure, C. M. Pennartz, G. Pezzulo, *Philosophical Transactions of the Royal Society B: Biological Sciences* **369** (2014).
5. D. Hassabis, D. Kumaran, C. Summerfield, M. Botvinick, *Neuron* **95**, 245 (2017).
6. J. Merel, M. Botvinick, G. Wayne, *Nature Communications* **10**, 1 (2019).
7. T. J. Prescott, P. Redgrave, K. Gurney, *Adaptive Behavior* **7**, 99 (1999).
8. J. Jackson, *Selected Writings of John Hughlings Jackson*, J. Taylor, ed. (Staples Press, London, 1884/1958), vol. 2, book section 2, pp. 1–10.
9. J. C. Doyle, M. Csete, *Proceedings of the National Academy of Sciences of the United States of America* **108**, 15624 (2011).
10. D. Floreano, A. J. Ijspeert, S. Schaal, *Current Biology* **24**, R910 (2014).
11. T. J. Prescott, J. L. Ayers, F. Grasso, P. F. M. J. Verschure, *From Neuron to Cognition via Computational Neuroscience* pp. 483–512 (2016).
12. J. L. Krichmar, *Frontiers in Neurorobotics* **12** (2018).
13. C. Moulin-Frier, *et al.*, *IEEE Transactions on Cognitive and Developmental Systems* **10**, 1005 (2018).
14. A. J. Ijspeert, A. Crespi, D. Ryczko, J.-M. Cabelguen, *Science* **315**, 1416 (2007).
15. A. Ijspeert, P. Ramdya, *Science Robotics* (2023).
16. M. Kawato, *Current Opinion in Neurobiology* **9**, 718 (1999).

17. P. F. Verschure, M. Mintz, *Neurocomputing* **38-40**, 1019 (2001).

18. J. L. McKinstry, G. M. Edelman, J. L. Krichmar, *Proceedings of the National Academy of Sciences* **103**, 3387 (2006).

19. P. Dean, J. Porrill, C. F. Ekerot, H. Jörntell, *Nature Reviews Neuroscience* **11**, 30 (2010).

20. I. Abadia, F. Naveros, J. A. Garrido, E. Ros, N. R. Luque, *IEEE Transactions on Cybernetics* **51**, 2476 (2021).

21. E. D. Wilson, *et al.*, *Journal of The Royal Society Interface* **18**, 20200750 (2021).

22. K. Gurney, T. J. Prescott, J. R. Wickens, P. Redgrave, *Trends in Neurosciences* **27**, 453 (2004).

23. T. J. Prescott, F. M. M. González, K. Gurney, M. D. Humphries, P. Redgrave, *Neural Networks* **19**, 31 (2006).

24. M. Khamassi, B. Girard, A. Berthoz, A. Guillot, *Proceedings of the Eighth International Conference on Intelligent Autonomous Systems IAS-8, Amsterdam, The Netherlands, 10-13 March 2004.* pp. 10–13 (2004).

25. M. Khamassi, S. Lallée, P. Enel, E. Procyk, P. F. Dominey, *Frontiers in Neurorobotics* **5**, 1 (2011).

26. N. Burgess, J. G. Donnett, K. J. Jeffery, J. O–keefe, *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **352**, 1535 (1997).

27. M. Milford, G. Wyeth, D. Prasser, *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004* **2004**, 403 (2004).

28. A. Arleo, W. Gerstner, *Biological Cybernetics* **83**, 287 (2000).

29. T. J. Prescott, D. Camilleri, U. Martinez-Hernandez, A. Damianou, N. D. Lawrence, *Philosophical Transactions of the Royal Society B: Biological Sciences* **374** (2019).

30. M. T. Whelan, A. Jimenez-Rodriguez, T. J. Prescott, E. Vasilaki, *Bioinspiration & Biomimetics* **18**, 015007 (2023).

31. M. Hoffmann, Z. Straka, I. Farkas, M. Vavrecka, G. Metta, *IEEE Transactions on Cognitive and Developmental Systems* **10**, 163 (2018).

32. S. Lallee, P. F. Dominey, *Adaptive Behavior* **21**, 274 (2013).

33. K. Seepanomwan, D. Caligiore, A. Cangelosi, G. Baldassarre, *Neural Networks* **72**, 31 (2015).

34. M. Khamassi, P. Enel, P. F. Dominey, E. Procyk, *Medial prefrontal cortex and the adaptive regulation of reinforcement learning parameters* (Elsevier B.V., 2013), vol. 202, pp. 441–464.

35. B. Webb, *Nature* **417**, 359 (2002).

36. T. J. Prescott, N. Lepora, P. F. M. J. Verschure, eds., *The Handbook of Living Machines: Research in Biomimetic and Biohybrid Systems* (OUP, Oxford, UK, 2018).

37. J. Ayers, *Biological Intelligence for Biomimetic Robots* (MIT Press, Cambridge, MA, 2023).

38. J. Tani, J. White, *Adaptive Behavior* (2020).

39. A. Rosenblueth, N. Wiener, *Philosophy of Science* **12(4)**, 316 (1945).

40. B. Mitchinson, M. J. Pearson, A. G. Pipe, T. J. Prescott, *Neuromorphic and Brain-Based Robots* pp. 23–57 (2011).

41. T. J. Prescott, M. J. Pearson, B. Mitchinson, J. C. W. Sullivan, A. G. Pipe, *IEEE Robotics I& Automation Magazine* **16**, 42 (2009).

42. M. J. Pearson, B. Mitchinson, J. C. Sullivan, A. G. Pipe, T. J. Prescott, *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 3085 (2011).

43. T. J. Prescott, *et al.*, *The robot vibrissal system: Understanding mammalian sensorimotor co-ordination through biomimetics* (Springer New York, 2015), book section 10, pp. 213–240.

44. M. J. Pearson, *et al.*, *Frontiers in Robotics and AI* **8** (2021).

45. R. Pfeifer, J. Bongard, *How the Body Shapes the Way We Think: A New View of Intelligence* (MIT Press, Cambridge, MA, 2006).

46. J. E. Laird, A. Newell, P. S. Rosenbloom, *Artificial Intelligence* **33**, 1 (1987).

47. J. R. Anderson, *The architecture of of Cognition* (Harvard University Press, 1983).

48. P. Langley, J. E. Laird, S. Rogers, *Cognitive Systems Research* **10**, 141 (2009).

49. K. Friston, *Trends in Cognitive Sciences* **13**, 293 (2009).

50. A. Clark, *Behavioral and Brain Sciences* **36**, 181 (2013).

51. J. Hohwy, *The predictive mind* (Oxford University Press, 2013).

52. L. D. Costa, P. Lanillos, N. Sajid, K. Friston, S. Khan, *Entropy* **24** (2022).

53. A. Ahmadi, J. Tani, *Neural Computation* **31**, 2025 (2019).

54. H. Idei, W. Ohata, Y. Yamashita, T. Ogata, J. Tani, *Scientific Reports* **12** (2022).

55. M. Montévil, M. Mossio, *Journal of Theoretical Biology* **372**, 179 (2015).

56. M. Kirchhoff, T. Parr, E. Palacios, K. Friston, J. Kiverstein, *Journal of the Royal Society Interface* **15** (2018).

57. S. Wilson, T. Prescott, *Philosophical Transactions of the Royal Society B* **377**, 20200519 (2021).

58. D. Wolpert, M. Kawato, *Neural Networks* **11**, 1317 (1998).

59. M. Haruno, D. M. Wolpert, M. Haruno, D. Wolpert, M. Kawato, *Neural Computation* **13**, 2201 (2001).

60. D. Wolpert, Z. Ghahramani, *Nat. Neurosci.* **3**, 1212 (2000).

61. S. F. Mann, R. Pain, M. D. Kirchhoff, *Biology and Philosophy* **37** (2022).

62. A. Ciria, G. Schillaci, G. Pezzulo, V. V. Hafner, B. Lara, *Neural Computation* **33**, 1402 (2021).

63. R. S. Sutton, *Machine Learning* **3**, 9 (1988).

64. K. Doya, K. Samejima, K.-i. Katagiri, M. Kawato, *Neural Computation* **14**, 1347 (2002).

65. Y. Demiris, B. Khadhouri, *Robotics and Autonomous Systems* **54**, 361 (2006).

66. J. Tani, S. Nolfi, *Neural Networks* **12**, 1131 (1999).

67. R. A. Jacobs, M. I. Jordan, S. J. Nowlan, G. E. Hinton, *Neural Computation* **3**, 79 (1991).

68. J. L. Elman, *Cognitive Science* **14**, 179 (1990).

69. Y. Yamashita, J. Tani, *PLoS Computational Biology* **4** (2008).

70. K. Friston, *arxiv* (2019).

71. M. Lungarella, O. Sporns, *Proceedings of 2005 4th IEEE International Conference on Development and Learning* **2005**, 25 (2005).

72. L. Pio-Lopez, A. Nizard, K. Friston, G. Pezzulo, *Journal of The Royal Society Interface* **13**, 20160616 (2016).

73. G. Oliver, P. Lanillos, G. Cheng, *arxiv* (2019).

74. T. V. de Maele, T. Verbelen, O. Çatal, C. D. Boom, B. Dhoedt, *Frontiers in Neurorobotics* **15** (2021).

75. O. Çatal, T. Verbelen, T. V. de Maele, B. Dhoedt, A. Safron, *Neural Networks* **142**, 192 (2021).

76. P. Lanillos, J. Pages, G. Cheng, *arxiv* (2020).

77. J. Hwang, J. Kim, A. Ahmadi, M. Choi, J. Tani, *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **50**, 1918 (2020).

78. S. Blakemore, D. Wolpert, F. CD, *Nat. Neurosci.* **1**, 635 (1998).

79. K. S. Walsh, D. P. McGovern, A. Clark, R. G. O'Connell, *Annals of the New York Academy of Sciences* **1464**, 242 (2020).

80. M. Miller, A. Clark, *Synthese* **195**, 2559 (2018).

81. W. Bechtel, L. Bich, *Philosophical Transactions of the Royal Society B: Biological Sciences* **376** (2021).

82. A. B. Butler, W. Hodos, *Comparative vertebrate neuroanatomy* (Wiley-Liss, New York, 1996).

83. F. Peller-Konrad, *et al.*, Conceptual design of the memory system of the robot cognitive architecture armarx (2022).

84. S. Grillner, B. Robertson, *Current Biology* **26**, R1088 (2016).

85. A. J. Cope, J. M. Chambers, T. J. Prescott, K. N. Gurney, *bioRxiv* (2017).

86. P. Redgrave, T. J. Prescott, K. Gurney, *Neuroscience* **89**, 1009 (1999).

87. B. Mitchinson, T. J. Prescott, *Biomimetics and Biohydrid Systems V)* (2016), vol. 9793, pp. 179–191.

88. M. Montévil, M. Mossio, A. Pocheville, G. Longo, *Progress in Biophysics and Molecular Biology* **122**, 36 (2016).

89. M. Mossio, M. Montévil, G. Longo, *Progress in Biophysics and Molecular Biology* **122**, 24 (2016).

90. J. M. Brooke, S. S. James, A. Jimenez-Rodriguez, S. P. Wilson, *PLOS Computational Biology* **18**, 1 (2022).

91. S. A. Kauffman, *A world beyond physics: The emergence and evolution of life* (Oxford University Press, 2019).

92. H. Maturana, F. Varela, *Autopoiesis and Cognition: The Realization of the Living* (D Reidel Publishing Company, 1980).

93. E. Di Paolo, E. Thompson, R. Beer, *Philosophy and the Mind Sciences* **3** (2022).

94. R. Rosen, *Life Itself a Comprehensive Inquiry Into the Nature, Origin, and Fabrication of Life* (Columbia University Press, 2005).

95. J. C. T. H. Hallam, C. A. M. Malcolm, *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences* **349**, 29 (1994).

96. A. Clark, *The Southern Journal of Philosophy* **53**, 3 (2015).

97. A. Constant, A. Clark, M. Kirchhoff, K. J. Friston, *Mind & Language* **37**, 373 (2022).

98. L. Zhaoping, *Current Opinion in Neurobiology* **40**, 94 (2016).

99. J. Bruineberg, J. Kiverstein, E. Rietveld, *Synthese* **195**, 2417 (2018).

100. M. J. D. Ramstead, K. J. Friston, I. Hipólito, *Entropy* **22** (2020).

# Acknowledgements

Funding: Horizon 2020 FET Flagship Human Brain Project SGA3, grant no. 945539 (T.J.P.); UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee for the EIC Pathfinder CAVAA project, grant no. 101071178 (T.J.P.); the James S. McDonnell Foundation, grant no. 220020516 (S.P.W.). Author contributions: T.J.P. and S.P.W. contributed equally. Competing interests: T.J.P. is a director and shareholder in two UK robotics companies Consequential Robotics Ltd. and Bow Ltd.