Embodied Inference: or "I think therefore I am, if I am what I think"

Karl Friston

Wellcome Trust Centre for Neuroimaging, London,
Institute of Neurology, University College of London, United Kingdom
k.friston@fil.ion.ucl.ac.uk

Introduction

This chapter considers situated and embodied cognition in terms of the free-energy principle. The free-energy formulation starts with the premise that biological agents must actively resist a natural tendency to disorder. It appeals to the idea that agents are essentially inference machines that model their sensorium to make predictions, which action then fulfils. The notion of an inference machine was articulated most clearly by Helmholtz[1] and developed in psychophysics by Gregory[2,3]. The basic premise is that agents, and in particular their brains, entail a model of how their sensory data are generated. Optimization of this model's parameters corresponds to perceptual inference and learning on a moment to moment basis; while optimization of the model *per se* rests on changes in the form or configuration of the phenotype at neurodevelopmental or evolutionary timescales. The free-energy formulation generalises the concept of agents as inference machines and considers each agent as a statistical model of its environmental niche (econiche). In brief, the free-energy principle takes the existence of agents as its starting point and concludes that each phenotype or agent *embodies* an optimal model of its econiche. This optimality is achieved by minimizing free-energy, which bounds the evidence for each agent (model), afforded by sensory interactions with the world. In this sense, each agent distils and embodies causal structure in its local environment. However, the key role of embodiment also emerges in a slightly deeper and more subtle argument: Not only does the agent embody the environment but the environment embodies the agent. This is true in the sense that the physical states of the agent (its internal milieu) are part of the environment. In other words, the statistical model entailed by each agent includes a model of itself as part of that environment. This model rests upon prior expectations about how environmental states unfold over time. Crucially, for an agent to exist, its model must include the prior expectation that its form and internal (embodied) states are contained within some invariant set. This is easy to see by considering the alternative: If the agent (model) entailed prior expectations that it will change irreversibly, then (as an optimal model of itself), it will cease to

exist in its present state. Therefore, if the agent (model) exists, it must *a priori* expect to occupy an invariant set of bounded states (cf., homeostasis). Heuristically, if I am a model of my environment and my environment includes me, then I model myself as existing. But I will only exist iff (sic) I am a veridical model of my environment. Put even more simply; "I think therefore I am,[4] iff I am what I think". This tautology is at the heart of the free-energy principle and celebrates the circular causality that underpins much of embodied cognition.

Under this view, each organism represents a hypothesis or model that contains a different set of prior expectations about the environment it inhabits. Interactions with the environment can be seen as hypothesis testing or model optimisation, using the free-energy as a measure of how good its model is. Phenotypes or species that attain a low free-energy (i.e., maximise the evidence for their model) represent optimal solutions in a free-energy or fitness landscape, where exchanges with the environment are consistent with their prior expectations. The characteristic of biological agents is that they *a priori* expect their physical states to possess key invariance properties. These priors are mandated by the very existence of agents and lead naturally to phenomena like homeostasis, and preclude surprising exchanges with the world. It can be seen that the role of prior expectations is crucial in this formulation: If each agent is a hypothesis that includes prior expectations, then these expectations must include the prior that the agent occupies an invariant (attracting) set of physical states. However, this is only a hypothesis, which the agent must test using sensory samples from the environment. Iff its hypothesis is correct, the agent will retain its priors and maintain its states within physiological bounds. This highlights the key role of priors and their intimate relationship to the structural form of phenotypes. It also suggests that simple prior expectations about homeostasis may be heritable and places the free-energy formulation (at least potentially) in an evolutionary setting. These arguments appeal to embodied cognition in that cognition and perception can be regarded as hypothesis testing about the environment in which the agent is situated, and which embodies the agent *per se*.

In summary, embodiment plays a fundamental and bilateral role in the free-energy formulation. On the one hand, agents embody (model) causal structure in the environment. On the other hand, the physical instantiation of this model is embodied in the environment. Only when the two are mutually compatible can the agent exist. This necessarily implies a low free-energy, which bounds the evidence for a model or hypothesis about an agent's milieu. As the long-term average of negative log-evidence is the entropy of an agent's sensory states, these low free-energy solutions implicitly resist a tendency to disorder and enable organisms to violate the second law of thermodynamics.

In what follows, we will go through these arguments in more detail

and try to connect them to established theories about perception, cognition and behaviour. This chapter comprises three sections. The first provides a heuristic overview of the free-energy formulation and its conceptual underpinnings. This formulation is inherently mathematical (drawing from statistical physics, dynamical systems and information theory). However, the basic ideas are intuitive and will be presented as such. For more technical readers, formal details (e.g., mathematical equations) can be found in the figures and their legends. In the second section, we examine these ideas in the light of existing theories about how the brain works. In the final section, we will look more closely at the role of prior expectations as policies for negotiating with the environment and try to link policies to itinerancy and related concepts from synergetics.

1. The Free-Energy Principle

In recent years, there has been growing interest in applying free-energy principles to the brain[5], not just in the neuroscience community, where it has caused some puzzlement[6] but from fields as far apart as psychotherapy[7] and social politics[8]. The free-energy principle has been described as a unified brain theory[9] and yet may have broader implications that speak to the way that any biological system interacts with its environment. This section describes the origin of the free-energy formulation, its underlying premises and the implications for how we represent and interact with our world.

The free-energy principle is a simple postulate that has complicated ramifications. It says that self-organising systems (like us) that are at equilibrium with their environment must minimise their free-energy[10]. This postulate is as simple and fundamental as Hamilton's law of Least Action and the celebrated H-theorems in statistical physics. The principle was originally formulated as a computational account of perception that borrows heavily from statistical physics and machine learning. However, it quickly became apparent that its explanatory scope included action and behaviour and was linked to our very existence: In brief, the free-energy principle takes well-known statistical ideas and applies them to deep problems in population (ensemble) dynamics and self-organisation. In applying these ideas, many aspects of our brains, how we perceive and the way we act become understandable as necessary and self-evident attributes of biological systems.

The principle is essentially a mathematical formulation of how adaptive systems (i.e., biological agents, like animals or brains) resist a natural tendency to disorder[11-14]. What follows is a non-mathematical treatment of its motivation and implications. We will see that although the motivation is quite straightforward, the implications are complicated and diverse. This diversity allows the principle to account for many aspects of brain

structure and function and lends it the potential to unify different perspectives on how the brain works. In the next section, we will see how the principle can be applied to neuronal systems, as viewed from these perspectives. This section is rather abstract and technical but the next section tries to unpack the basic idea in more familiar terms.

1.1 Resisting a tendency to disorder

The defining characteristic of biological systems is that they maintain their states and form in the face of a constantly changing environment[11-14]. From the point of view of the brain, the environment includes both the external and internal milieu. This maintenance of order is seen at many levels and distinguishes biological from other self-organising systems. Indeed, the physiology of biological systems can be reduced almost entirely to their homeostasis (the maintenance of physiological states within certain bounds[15]). More precisely, the repertoire of physiological and sensory states an organism can be in is limited, where those states define the organism's phenotype. Mathematically, this means that the probability distribution of the agent's (interoceptive and exteroceptive) sensory states must have low entropy. Low entropy just means that there is a high probability that a system will be in one of a small number of states, and a low probability that it will be in the remaining states. Entropy is also the average self-information. Self-information is the 'surprise' or improbability of something happening[16] or, more formally, its negative log-probability. Here, 'a fish out of water' would be in a surprising state (both emotionally and mathematically). Note that both entropy and surprise depend on the agent; what is surprising for one agent (e.g., being out of water) may not be surprising for another. Biological agents must therefore minimise the long-term average of surprise to ensure that their sensory entropy remains low. In other words, biological systems somehow manage to violate the Fluctuation Theorem, which says the entropy of (non-adaptive) systems can fall but the probability of entropy falling vanishes exponentially as the observation time increases[17].

In short, the long-term (distal) imperative, of maintaining states within physiological bounds, translates into a short-term (proximal) suppression of surprise. The sort of surprise we are talking about here is associated with unpredicted or shocking events (e.g., tripping and falling in the street or the death of a loved one). Surprise is not just about the current state (which cannot be changed) but also about the movement or transition from one state to another (which can). This motion can be very complicated and itinerant (wandering) provided it revisits a small set of states (called a global random attractor[18]) that are compatible with survival (e.g., driving a car within a small margin of error). It is this motion or these state-transitions that the free-energy principle optimises.

So far, all we have said is that biological agents must avoid surprises to ensure that their exchanges with the environment remain within bounds. But how do they do this? A system cannot know whether its sensations are surprising or avoid them even if it did know. This is where free-energy comes in: Free-energy is an upper bound on surprise, which means that if agents minimise free-energy they implicitly minimise surprise. Crucially, free-energy can be evaluated because it is a function of two things the agent has access to: its sensory states and a recognition density encoded by its internal states (e.g., neuronal activity and connection strengths). The recognition density is a probability distribution of putative environmental causes of sensory input; i.e., a probabilistic representation of what caused sensations. These causes can range from the presence of an object in the field of view that causes sensory impressions on the eye, to physiological states like blood pressure that cause interoceptive signals. The (variational) free-energy construct was introduced into statistical physics to convert difficult probability density integration problems into easier optimisation problems[19]. It is an information theoretic quantity (like surprise) as opposed to a thermodynamic energy. Variational free-energy has been exploited in machine learning and statistics to solve many inference and learning problems[20-22]. In this setting, surprise is called the (negative) model log-evidence (i.e., the log-probability of getting some sensory data, given it was generated by a particular model). In our case, the model is entailed by the agent. This means minimising surprise is the same as maximising the sensory evidence for a model or agent. In the present context, free-energy provides the answer to a fundamental question: How do self-organising adaptive systems avoid surprising states? They can do this by minimising their free-energy. So what does this involve?

1.2 Action and perception

In brief, agents can suppress free-energy by changing the two things free-energy depends on. They can change sensory input by acting on the world or they can change their recognition density by changing their internal states. This distinction maps nicely onto action and perception. One can understand this in more detail by considering three mathematically equivalent formulations of free-energy (see Fig. 1 and ref [5]; Supplementary material, for a more formal treatment). The free-energy bound on surprise is constructed by simply adding a non-negative term to surprise. This term is a function of the recognition density encoded by the agent's internal states. We will refer to this term as a posterior divergence. Creating the free-energy bound in this way leads to the first formulation:

1.2.1 Free-energy as posterior divergence plus surprise

The posterior divergence is a Kullback-Leibler divergence (cross entropy) and is just the difference between the recognition density and the posterior or conditional density on the causes of sensory signals. This conditional density represents the best possible guess about the true causes. The difference between the two densities is always non-negative and free-energy is therefore an upper bound on surprise. This is the clever part of the free-energy formulation; because minimising free-energy by changing the recognition density (without changing sensory data) reduces the difference, making the recognition density a good approximation to the conditional density and the free-energy a good approximation to surprise. The recognition density is specified by its sufficient statistics, which are the agent's internal states. This means an agent can reduce posterior divergence (i.e. free-energy) by changing its internal states. This is essentially perception and renders an agent's internal states representations of the causes of its sensations.

1.2.2 Free-energy as prior divergence minus accuracy

The second formulation expresses free-energy as prior divergence minus accuracy. In the model comparison literature, prior divergence is called 'complexity'. Complexity is the difference between the recognition density and the prior density on causes encoding beliefs about the state of the world before observing sensory data (this is also known as Bayesian surprise[23]). Accuracy is simply the surprise about sensations expected under the recognition density. This formulation shows that minimising free-energy by changing sensory data (without changing the recognition density) must increase the accuracy of an agent's predictions. In short, the agent will selectively sample the sensory inputs that it expects. This is known as active inference[24]. An intuitive example of this process (when it is raised into consciousness) would be feeling our way in darkness; anticipating what we might touch next and then trying to confirm those expectations. In short, agents can act on the world to minimise free-energy by increasing the accuracy of their predictions through selective sampling of the environment.

1.2.3 Free-energy as expected energy minus entropy

The final formulation expresses free-energy as an expected energy minus entropy. This formulation is important for three reasons. First, it connects the concept of free-energy as used in information theory with homologous concepts used in statistical thermodynamics. Second, it shows that the free-energy can be evaluated by an agent because the expected energy is the surprise about the joint occurrence of sensations and their perceived causes, while the entropy is simply the entropy of its recognition density. Third, it shows that free-energy rests upon a generative model of the

world; which is expressed in terms of the joint probability of a sensation and its causes occurring together. This means that an agent must have an implicit generative model of how causes conspire to produce sensory data. It is this model that defines both the nature of the agent and the quality of the free-energy bound on surprise.

1.3 Generative models in the brain

We have just seen that one needs a generative model (denoted by $p(\tilde{s}, \vartheta \mid m)$ in the figures) of how the sensorium is caused to evaluate free-energy. These models combine the likelihood of getting some data, given their causes and prior beliefs about these causes. These models have to explain complicated dynamics on continuous states with hierarchical or deep causal structure. Many biological systems, including the brain, may use models with the form shown in Fig. 2. These are hierarchical dynamic models and provide a very general description of states in the world. They are general in the sense that they allow for cascades or hierarchies of nonlinear dynamics to influence each other. They comprise equations of motion and static nonlinear functions that mediate the influence of one hierarchical level on the next. Crucially, these equations include random fluctuations on the states and their motion, which play the role of observation noise at the sensory level and state-noise at higher levels. These random fluctuations induce uncertainty about states of the world and the parameters of the model. In these models, states are divided into *causal states*, which link states in different hierarchical levels and *hidden states*, which link states over time and lend the model memory. Gaussian assumptions about the random fluctuations furnish the likelihood and (empirical) priors on predicted motion that constitute a probabilistic generative model. These assumptions about random effects are encoded by their (unknown) precision, or inverse variance. See Fig. 2 for details. We will appeal to this sort of model below, when trying to understand how the brain complies with the free-energy principle, in terms of its architecture and dynamics.

In summary, the free-energy induces a probabilistic model of how sensory data are generated and a recognition density on the model's parameters (i.e., sensory causes). Free-energy can only be reduced by changing the recognition density to change conditional expectations about what is sampled or by changing sensory samples (i.e. sensory input) so that they conform to expectations. This corresponds to perception and action respectively. We will see later that minimising free-energy corresponds to minimising prediction errors. It then becomes almost self-evident that biological agents can suppress prediction errors by changing predictions (perception) or what is predicted (action): see Fig. 2. In the next section, we consider the implications of this formulation in light of some key theories about the brain.

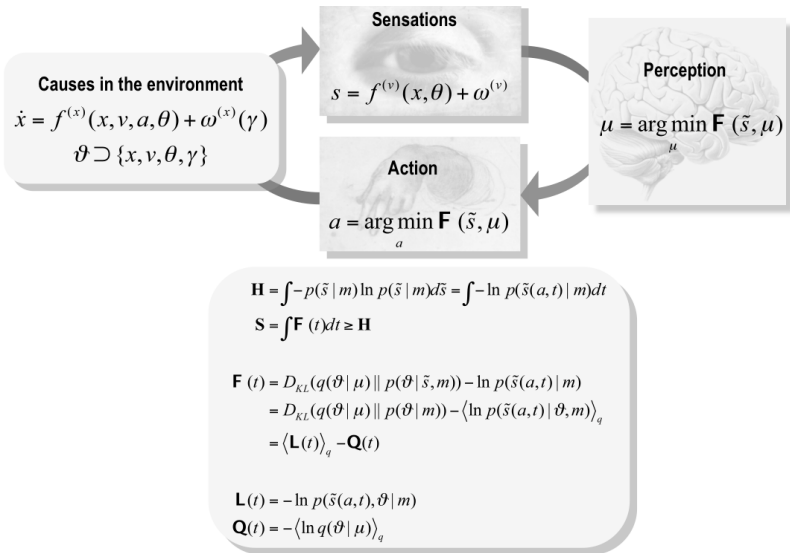
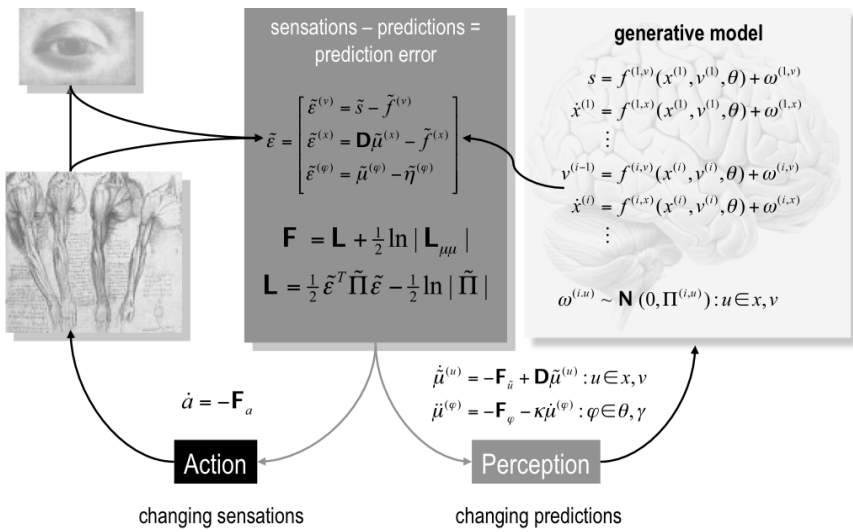


Fig. 1: The free-energy principle. This schematic shows the dependencies among the quantities that define the free-energy of an agent or brain, denoted by m. These include, its internal states $\mu(t)$, generalised sensory signals (i.e., position, velocity, acceleration etc.) $\tilde{s}(t) = [s, s', s'', ...]^T$ and action $a(t)$. The environment is described by equations, which specify the motion of its states $\{x(t), v(t)\}$. Both internal brain states and action minimise free-energy $F(\tilde{s}, \mu)$, which is a function of sensory input and the internal states. These states encode a recognition density $q(\vartheta \mid \mu)$ on the causes $\vartheta \supset \{x, v, \theta, \gamma\}$ of sensory input. These comprise states of the world $u(t) : u \in x, v$ and parameters $\varphi \in \theta, \gamma$ controlling the equations of motion and the amplitude of the random fluctuations $\omega^{(u)} : u \in x, v$ on the hidden states and sensory input. The lower panel provides the key equations behind the free-energy formulation. The first pair says that the path integral of free-energy (free-action; **S**) is an upper bound on the entropy of sensory states, **H**. This entropy is average surprise, which (under ergodic assumptions) is the long-term average or path integral of surprise. The free-energy per se $F(t) := F(\tilde{s}, \mu)$ is then expressed in three ways to show what its minimisation means. The first equality shows that optimising brain states, with respect to the internal states, makes the recognition density an approximate conditional density on the causes of sensory input. Furthermore, it shows that free-energy is an upper bound on surprise. This enables action to avoid surprising sensory encounters. The second equality shows that action can only reduce free-energy by selectively sampling data that are predicted by under the recognition density. The final equality expresses free-energy in terms of an expected energy $L(t)$ based on a generative model and the entropy $Q(t)$ of the recognition density. In this figure, $\langle \cdot \rangle_q$ denotes expectation or average, under the recognition density and $D_{KL}(\cdot \,||\, \cdot)$ is a non-negative Kullback-Leibler divergence (i.e., difference between two probability densities). In summary, free-energy rests on two probability densities; one that generates sensory samples and their causes, $p(\tilde{s}, \vartheta \mid m)$ and the recognition density, $q(\vartheta \mid \mu)$. The first is a probabilistic generative model, whose form is entailed by the agent or brain (denoted by m), while the second represents the best probabilistic estimate of the causes and is encoded by internal states. The free-energy principle states that all quantities that can change (sufficient statistics and action) minimise free-energy.

Fig. 2 (next page): Action and perception. This schematic illustrates the bilateral role of free-energy (i.e., prediction error) in driving action and perception: Action*:* Acting on the environment by minimising free-energy enforces a sampling of sensory data that is consistent with the current representation (i.e., changing sensations to minimise prediction error). This

is because free-energy is a mixture of complexity and accuracy (the second expression for free-energy in Fig. 1). Crucially, action can only affect accuracy. This means the brain will reconfigure its sensory epithelia to sample inputs that are predicted by its representations; in other words, to minimise prediction errors. The equation above action simply states that action performs a gradient decent on (i.e., minimises) free-energy (see ref [10] for details). Perception: Optimizing free-energy by changing the internal states that encode the recognition density makes it an approximate posterior or conditional density on the causes of sensations. This follows because free-energy is surprise plus a Kullback-Leibler divergence between the recognition and conditional densities (the first expression for free-energy in Fig. 1). Because this difference is non-negative, minimising free-energy makes the recognition density an approximate posterior probability. This means the agent implicitly infers or represents the causes of its sensory samples in a Bayes-optimal fashion. At the same time, the free-energy becomes a tight bound on surprise that is minimised through action. The equation above perception simply states that internal states perform a gradient decent on (i.e., minimise) free-energy. This gradient decent is in a moving frame of reference for generalised states and accumulates gradients over time for the parameters (see ref [5] for details). Prediction error: Prediction error is simply the difference between predicted and observed sensory states. The equations show that the free-energy comprises the expected energy L (*t*), which is effectively the (precision weighted) sum of squared error. This error contains the sensory prediction error and other differences that mediate empirical priors on the motion of hidden states and the parameters. The predictions rest on a generative model of how sensations are caused. These models have to explain complicated dynamics on continuous states with hierarchical or deep causal structure. An example of one such generic model is shown on the right. Generative model: Here $f^{(i,u)}: u \in x, v$ are continuous nonlinear functions of (hidden and causal) states, parameterised by $\boldsymbol{\theta} \subset \vartheta$ at the *i*-th level of a hierarchical dynamic model. The random fluctuations $\boldsymbol{\omega}^{(u)}: u \in x, v$ play the role of observation noise at the sensory level and state-noise at higher levels. Causal states $v^{(i)} \subset \vartheta$ link hierarchical levels, where the output of one level provides input to the next. Hidden states $x^{(i)} \subset \vartheta$ link dynamics over time and lend the model memory. Gaussian assumptions about the random fluctuations specify the likelihood and furnish empirical priors in terms of predicted motion. These assumptions are encoded by their or precision or inverse variance $\Pi^{(i,u)}: u \in x, v$, which depend on precision parameters $\gamma \subset \vartheta$. The associated message-passing scheme implementing perception is shown in the next figure. In this and subsequent figures, subscripts denote differentiation, *D* is a temporal derivative operator that acts on generalised states and κ is a large positive constant (see ref [48] for details).



THE BAYESIAN BRAIN AND OTHER THEORIES

This section attempts to place some key brain theories within the free-energy framework, in the hope of identifying common themes. It reprises and extends the review in ref [5]. We will consider a range of theories that derive from both the biological and physical sciences (e.g., neural Darwinism, information theory and optimal control). Crucially, one key theme runs throughout these theories; namely, optimization. Furthermore, if we look closely at what is optimized, the same quantity keeps emerging, namely value, expected reward, expected utility; or its complement: surprise, prediction-error or expected cost. We will see that this quantity is effectively free-energy.

2.1 The Bayesian brain hypothesis

The Bayesian brain hypothesis[25] uses Bayesian probability theory to formulate perception as a constructive process based on internal or generative models. The underlying idea is that the brain has a model of the world[1-3] that it tries to optimise using sensory inputs[28-33]. This idea is related to analysis by synthesis[27] and epistemological automata[26]. In other words, the brain is an inference machine that actively predicts and explains its sensations[1,3,30]. Central to this hypothesis is a probabilistic model that can generate predictions, against which sensory samples are tested to update beliefs about their causes. In Bayesian treatments, this generative model decomposes into the likelihood (the probability of sensory data, given their causes) and a prior (the *a priori* probability of those causes). Perception then becomes the inversion of the likelihood model (mapping from causes to sensations) to access the posterior probability of the causes, given sensory data (mapping from sensations to causes). This inversion is exactly the same as minimising the difference between the recognition and posterior densities (posterior divergence) to suppress free-energy. Indeed, the free-energy formulation was developed to finesse the difficult problem of exact inference by converting it into an easier optimisation problem[19-22]. This has furnished some powerful approximation techniques for model identification and comparison (e.g., variational Bayes or ensemble learning[34]). There are many interesting issues that attend the Bayesian brain hypothesis; we will focus on two.

The first is the form of the generative model and how it manifests in the brain. One criticism of Bayesian treatments is that they ignore the question of how prior beliefs, which are necessary for inference, are formed[32]. However, this criticism disappears under hierarchical generative models, in which the priors themselves are optimised[31,33]. In hierarchical models (cf, the right panel in Fig. 2), causes in one level of a model generate subordinate causes in a lower level, while the sensory data *per se*

are generated at the lowest level. Minimising the free-energy of representations effectively optimises empirical priors (i.e., the probability of causes at one level, given those in the level above). Crucially, because empirical priors are linked hierarchically, they are informed by sensory data, enabling the brain to optimise its prior expectations online. This optimisation makes every level in the hierarchy accountable to others, furnishing an internally consistent representation of sensory causes, at multiple levels of description. Not only do hierarchical models have a key role in statistics (e.g., random effects models and parametric empirical Bayes[35,36]), they may also be an important metaphor for the brain, given the hierarchical arrangement of its cortical sensory areas[37-39].

The second issue is the form of the recognition density. This has to be encoded by physical attributes (i.e., internal states) of the brain, such as synaptic activity, efficacy and gain. In general, any density is encoded by its sufficient statistics (for example, the mean and variance of a Gaussian density). The way the brain encodes these statistics places important constraints on the sorts of schemes that underlie recognition. The differences between these schemes can usually be reduced to differences in the form of the recognition density. They range from free-form schemes, which use a vast number of sufficient statistics (e.g., particle filtering[31] and probabilistic population codes[40-43]) to simpler forms, which make stronger assumptions about the shape of the recognition density. These assumptions mean that the recognition density can be encoded with a small number of sufficient statistics. The simplest assumed form is Gaussian, which only requires the conditional mean or expectation. This is also known as the Laplace assumption [44], under which the free-energy reduces to the sum of squared prediction error at each level of the model (in fact, this assumption gives exact inference under some common models, such as factor analysis[45]). Minimising free-energy then corresponds to explaining away the prediction error (Fig. 2). This is known as predictive coding and has become a popular framework for understanding neuronal message-passing among different levels of sensory cortical hierarchies[46]. In this scheme, prediction error units compare conditional expectations with top-down predictions to elaborate a prediction error. This is passed forward to drive the units in the level above that encode conditional expectations and optimise top-down predictions to explain (i.e., reduce) prediction error in the level below. This just means countering excitatory bottom-up inputs to a prediction error neuron with inhibitory synaptic inputs that are driven by top-down predictions. See Fig. 3 and refs [47] and [48] for a detailed discussion. The reciprocal exchange of bottom-up prediction errors and top-down predictions proceeds until prediction error is minimised at all levels and conditional expectations are optimised. This scheme has been invoked to explain many features of early visual responses [46,49] and provides a plausible account of repetition suppression and

mismatch responses in electrophysiology[50]. Fig. 4 provides an example of perceptual categorisation that uses this scheme. Message-passing of this sort is consistent with known functional asymmetries in real cortical hierarchies [51], where forward connections (which convey prediction errors) are driving and backwards connections (that model the nonlinear generation of sensory input) show both driving and modulatory characteristics [52]. This asymmetric message-passing is also a characteristic feature of adaptive resonance theory[53,54], which shares formal similarities with predictive coding.

In summary, the theme underlying the Bayesian brain and predictive coding is that the brain is an inference engine that is trying to optimise probabilistic representations of what caused its sensory input. This optimisation can be finessed using a (variational free-energy) bound on surprise. In machine learning and statistics, surprise is known as the (negative) log-evidence or marginal likelihood of some data, given a model. In this sense, the free-energy principle subsumes the Bayesian brain hypothesis and can be implemented by the many schemes considered in this field. Almost invariably, these involve some form of message-passing or belief propagation among brain areas or units. We have focused on one of the simplest schemes, namely predictive coding, which lends itself to a neurobiologically plausible implementation. Furthermore, it allows us to connect to another principled approach to sensory processing, namely information theory:

2.2 The principle of efficient coding

This principle suggests that the brain optimises the mutual information (i.e., mutual predictability) between the sensorium and its internal representation, under constraints on the efficiency of those representations. This line of thinking was articulated by Barlow[55] in terms of a redundancy reduction principle (or principle of efficient coding) and formalised later in terms of the infomax principle[56]. It has been applied in machine learning[57], leading to things like independent component analysis[58], and in neurobiology, to understand the nature of neuronal responses[59-62]. This principle is extremely effective in predicting the empirical characteristics of classical receptive fields[59] and provides a formal explanation for sparse coding[61] and the segregation of processing streams in visual hierarchies[63]. It has been extended to cover dynamics and motion trajectories[64,65] and even used to infer the metabolic constraints on neuronal processing[66]. At its simplest, it says that neuronal activity should encode sensory information in an efficient and parsimonious fashion. It considers the mapping between one set of variables (sensory states) and another (variables representing those states). At first glance, this seems to preclude a probabilistic representation, because this would involve a mapping between sensory states and a probability density. However, the infomax principle can be

applied to the sufficient statistics of a recognition density. In this context, the infomax principle suggests that conditional expectations should afford an accurate but parsimonious prediction of sensory signals.

Crucially, the infomax principle is a special case of the free-energy principle, which arises when we ignore uncertainty in probabilistic representations (and when there is no action, see Fig. 5 and ref [5]; supplementary material for mathematical details). This is easy to see by noting that sensory signals are generated by causes. This means it is sufficient to represent the causes to predict these signals. More formally, the infomax principle can be understood in terms of the decomposition of free-energy into complexity and accuracy: Mutual information is optimised when conditional expectations maximise accuracy (or minimise prediction error), while efficiency is assured by minimising complexity (the prior divergence). This ensures that the generative model is not over-parameterised and leads to a parsimonious representation of sensory data that conforms to prior constraints on their causes. It is interesting that advanced model optimisation techniques use free-energy optimisation to eliminate redundant model parameters[67]. This might provide a nice explanation for synaptic pruning and homeostasis in the brain during neurodevelopment[68] and sleep[69].

The infomax principle pertains to a forward mapping from sensory input to representations. How does this relate to optimising generative models, which map from causes to sensory inputs? These perspectives can be reconciled by noting that all recognition schemes based on infomax can be cast as optimising the parameters of a generative model[70]. For example, in sparse coding models[61], the implicit priors posit independent causes that are sampled from a heavy tailed or sparse distribution[48]. The fact that these models predict empirically observed receptive fields so well, suggests that we are endowed with (or acquire) prior expectations that the causes of our sensations are largely independent and sparse.

Bayesian surprise was invoked recently to explain sampling in models of visual search and salience[23]. Bayesian surprise is the difference between the posterior and prior densities on the causes of sensory input and is formally identical to complexity. It is interesting because it appears to contradict the principle of efficient coding; in that maximising Bayesian surprise increases complexity. However, this apparent paradox is resolved easily by noting that any change to the posterior (or recognition) density that increases accuracy will incur a complexity cost and increase Bayesian surprise. However, under the free-energy formulation, Bayesian surprise *per se* is not optimised; it should be *minimised* in the absence of a recognisable stimulus. It might be interesting to test this prediction empirically.

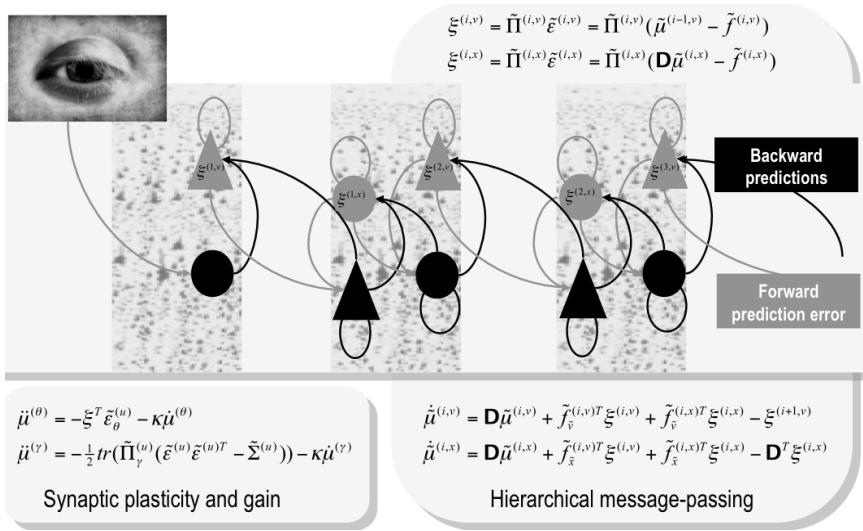
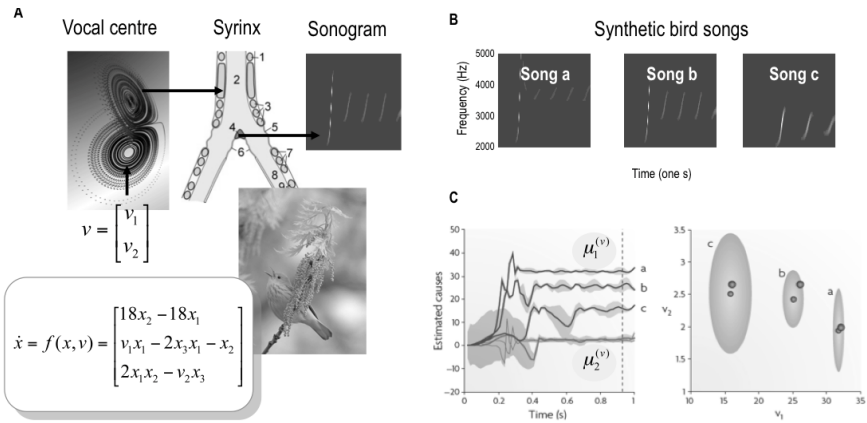


Fig. 3: Hierarchical message-passing in the brain. The schematic details a neuronal architecture that optimises the conditional expectations of causes in hierarchical models of sensory input of the sort illustrated in the previous figure. It shows the putative cells of origin of forward driving connections that convey prediction-error from a lower area to a higher area (grey arrows) and nonlinear backward connections (black arrows) that construct predictions [47]. These predictions try to explain away prediction-error in lower levels. In this scheme, the sources of forward and backward connections are superficial and deep pyramidal cells (triangles) respectively, where state-units are black and error-units are grey. The equations represent a gradient descent on free-energy using the generative model of the previous figure. Predictions and prediction-error: If we assume that neuronal activity encodes the conditional expectation of states, then recognition can be formulated as a gradient descent on free-energy. Under Gaussian assumptions, these recognition dynamics can be expressed compactly in terms of precision weighted prediction-errors $\xi^{(i,u)} : u \in x, v$ on the causal states and motion of hidden states. The ensuing equations suggest two neuronal populations that exchange messages; causal or hidden state-units encoding expected states and error-units encoding prediction-error. Under hierarchical models, error-units receive messages from the state-units in the same level and the level above; whereas state-units are driven by error-units in the same level and the level below. These provide bottom-up messages that drive conditional expectations $\mu^{(i,u)} : u \in x, v$ towards better predictions to explain away prediction-error. These top-down predictions correspond to $f^{(i,u)} : u \in x, v$. This scheme suggests the only connections that link levels are forward connections conveying prediction-error to state-units and reciprocal backward connections that mediate predictions. Note that the prediction errors that are passed forward are weighted by their precision. This tells us that precision may be encoded by the postsynaptic gain or sensitivity of error units, which also has to be optimised: Synaptic plasticity and gain: The corresponding equations for changes in the conditional expectation of the parameters of the model and the precisions of random fluctuations are related to formal models of associative plasticity and reinforcement learning: see refs [48] and [146] for further details.

Fig. 4 (next page): Birdsongs and perceptual categorisation. Left: The generative model of birdsong used in this simulation comprises a Lorenz attractor, whose shape is determined by two causal states (v_1 ,v_2). Two of the attractor's hidden states are used to modulate the amplitude and frequency of stimuli generated by a synthetic syrinx (an example is shown as a sonogram). The ensuing stimuli were then presented to a synthetic bird to see if it could recover the causal states (v_1 ,v_2) that categorise the chirp in a two-dimensional perceptual

space. This involves minimising free-energy by changing the internal representation ($\mu_1^{(v)}, \mu_2^{(v)}$) of the causes. Examples of this perceptual inference or categorisation are shown on the right. Right: Three simulated songs are shown (upper panels) in sonogram format. Each comprises a series of chirps whose frequency and number fall progressively (from a to c), as a causal state (known as the Raleigh number; v_1 on the left) is decreased. Lower left: This graph depicts the conditional expectations of the causal states, shown as a function of peristimulus time for the three songs. It shows that the causes are identified after about 600 milliseconds with high conditional precision (90% confidence intervals are shown in grey). Lower right: This shows the conditional density on the causes shortly before the end of peristimulus time (i.e., the dotted line in the left panel). The small dots correspond to conditional expectations and the grey areas correspond to the 90% conditional confidence regions. Note that these encompass the true values (large dots) used to generate the songs. These results illustrate the nature of perceptual categorisation under the inference scheme in Fig. 3: Here, recognition corresponds to mapping from a continuously changing and chaotic sensory input to a fixed point in perceptual space.



Perceptual inference and categorisation

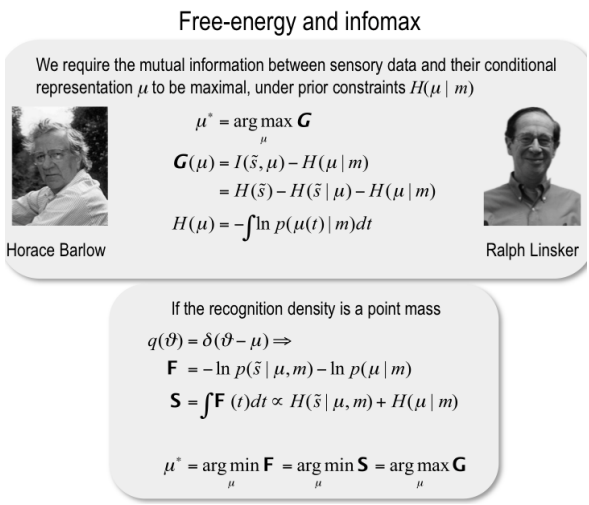


Fig. 5: Free-energy and infomax. This schematic provides the key equalities that show the infomax principle is a special case of the free-energy principle that obtains when we discount uncertainty and represent sensory data with point estimates of their causes. Alternatively, the free-energy is a generalization of the infomax principle that covers probability densities on the unknown causes of data. Horace Barlow and Ralph Linsker are two of the key people behind the principle of efficient coding and infomax.

In summary, the principle of efficient coding says the brain should optimise the mutual information between its sensory signals and some parsimonious neuronal representations. This is the same as optimising the parameters of a generative model to maximise the accuracy of predictions (i.e., to minimise prediction error), under complexity constraints. Both are mandated by the free-energy principle, which can be regarded as a probabilistic generalisation of the Infomax principle (see Fig. 5). We now turn to more biologically inspired ideas about brain function that focus on neuronal dynamics and plasticity. This takes us deeper into neurobiological mechanisms and implementation of theoretical principles above.

2.3 The cell assembly and correlation theory

The cell assembly theory was proposed by Hebb[71] and entails Hebbian — or associative — plasticity, which is a cornerstone of neural network theory and of the empirical study of use-dependent or experience-dependent plasticity[72]. There have been several elaborations of this theory; for example, the correlation theory of von der Malsburg[73,74] and formal refinements to Hebbian plasticity *per se*[75]. The cell assembly theory posits the formation of groups of interconnected neurons through a strengthening of their synaptic connections that depends on correlated pre- and post-synaptic activity; i.e., 'cells that fire together wire together'. This enables the brain to distil statistical regularities from the sensorium. The correlation theory considers the selective enabling of synaptic efficacy and their plasticity (cf. meta-plasticity[76]) by fast synchronous activity induced by different perceptual attributes of the same object (e.g., a red bus in motion). This resolves a putative deficiency of classical plasticity, which cannot ascribe a pre-synaptic input to a particular cause (i.e., redness of the bus)[73]. The correlation theory underpins theoretical treatments of synchronised brain activity and its role in associating or binding attributes to specific objects or causes[74,77]. Another important field that rests upon associative plasticity is the use of attractor networks as models of memory formation and retrieval[78-80]. So how do correlations and associative plasticity figure in the free-energy formulation?

Hitherto, we have considered only inference on states of the world that cause sensory signals, where conditional expectations about states are encoded by synaptic activity. However, the causes covered by the recognition density are not restricted to time-varying states (e.g., the motion of an object in the visual field); they also include time-invariant regularities that endow the world with causal structure (e.g., objects fall with constant acceleration). These regularities are parameters of the generative model and have to be inferred by the brain. The conditional expectations of these parameters may be encoded by synaptic efficacy (these expectations are $\mu^{(\theta)}$ in Fig. 3). Inference on parameters corresponds to optimising connection strengths in the brain; i.e., plasticity that underlines learning. So what

form would this learning take? It transpires that a gradient descent on free-energy (i.e., changing connections to reduce free-energy) is formally identical to Hebbian plasticity[33,48] (see Fig. 3). This is because the parameters of the generative model determine how expected states (synaptic activity) are mixed to form predictions. Put simply, when the pre-synaptic predictions and post-synaptic prediction-errors are highly correlated, the connection strength increases, so that predictions can suppress prediction errors more efficiently. Fig. 6 shows a simple example of this sort of sensory learning, using an oddball paradigm to elicit repetition suppression.

In summary, the formation of cell assemblies reflects the encoding of causal regularities. This is just a restatement of cell assembly theory in the context of a specific implementation (predictive coding) of the free-energy principle. It should be acknowledged that the learning rule in predictive coding is really a delta rule, which rests on Hebbian mechanisms; however, Hebb's wider notions of cell assemblies were formulated from a non-statistical perspective. Modern reformulations suggest that both inference on states (i.e., perception) and inference on parameters (i.e., learning) minimise free-energy (i.e., minimise prediction error) and serve to bound surprising exchanges with the world. So what about synchronisation and the selective enabling of synapses?

2.4 Biased competition and attention

To understand what is represented by the modulation of synaptic efficacy — or synaptic gain — we have to consider a third sort of cause in the environment; namely, the amplitude of random fluctuations. Causal regularities encoded by synaptic efficacy control the deterministic evolution of states in the world. However, stochastic or random fluctuations in these states play an important part in generating sensory data. Their amplitude is usually parameterized as precision (i.e., inverse variance) that encodes the reliability of prediction errors. Precision is important, especially in hierarchical schemes, where it controls the relative influence of bottom-up prediction errors and top-down predictions. So how is precision encoded in the brain? In predictive coding, expected precision modulates the amplitude of prediction errors (these expectations are $\mu^{(\gamma)}$ in Fig. 3), so that prediction errors with high precision have a greater impact on units encoding conditional expectations. This means that precision corresponds to the synaptic gain of prediction error units. The most obvious candidates for controlling gain (and implicitly encoding precision) are classical neuromodulators like dopamine and acetylcholine, which provides a nice link to theories of attention and uncertainty[81-83]. Another candidate is fast synchronised pre-synaptic input that lowers effective post-synaptic membrane time constants and increases synchronous gain[84]. This fits comfortably with the correlation theory and speaks to recent ideas about the role of synchronous activity in mediating attentional gain[85,86].

In summary, the optimisation of expected precision in terms of synaptic gain links attention and uncertainty in perception (through balancing top-down and bottom-up effects on inference) to synaptic gain and synchronisation. This link is central to theories of attentional gain and biased competition[86-91], particularly in the context of neuromodulation[92,93]. Clearly, these arguments are heuristic but show how different perspectives can be linked by examining mechanistic theories of neuronal dynamics and plasticity under a unifying framework. Fig. 7 provides a summary of the various neuronal processes that may correspond to optimising conditional expectations about states, parameters and precisions; namely, optimising synaptic activity, efficacy and gain respectively. In cognitive terms, these processes map nicely onto perceptual inference, learning and attention. The theories considered so far have dealt only with perception. However, from the point of view of the free-energy principle, perception just makes free-energy a good proxy for surprise. To actually reduce surprise we need to act. In the next section, we retain a focus on cell assemblies but move to the selection and reinforcement of stimulus-response links.

2.5 Neural Darwinism and value-learning

In the theory of neuronal group selection[94], the emergence of neuronal assemblies or groups is considered in the light of selective pressure. The theory has four elements: Epigenetic mechanisms create a primary repertoire of neuronal connections, which are refined by experience-dependent plasticity to produce a secondary repertoire of neuronal groups. These are selected and maintained through reentrant signalling (the recursive exchange of signals among neuronal groups). As in cell assembly theory, plasticity rests on correlated pre and post-synaptic activity but here it is modulated by value. Value is signalled by ascending neuromodulatory transmitter systems and controls which neuronal groups are selected and which are not. The beauty of neural Darwinism is that it nests selective processes within each other. In other words, it eschews a single unit of selection and exploits the notion of meta-selection (the selection of selective mechanisms; e.g. ref [95]). In this context, value confers adaptive fitness by selecting neuronal groups that meditate adaptive stimulus-stimulus associations and stimulus-response links. The capacity of value to do this is assured by natural selection; in the sense that neuronal systems reporting value are themselves subject to selective pressure.

This theory, particularly value-dependent learning[96], has deep connections with reinforcement learning and related approaches in engineering such as dynamic programming and temporal difference models[97,98] (see below). This is because neuronal systems detecting valuable states reinforce connections to themselves, thereby enabling the brain to label a sensory state as valuable iff it leads to another valuable

state. This ensures that agents move through a succession of states that have acquired value to access states (rewards) with genetically specified (innate) value. In short, the brain maximises value, which may be reflected in the discharge of dedicated neuronal systems (e.g., dopaminergic systems[98-102]). So how does this relate to the optimisation of free-energy?

The answer is simple: value is inversely proportional to surprise, in the sense that the probability that a phenotype is in a particular state increases with the value of that state. More formally $V = -\Gamma \ln p(\tilde{s} \mid m)$, where Γ encodes the amplitude of random fluctuations (see ref [5]; supplementary material). This means the adaptive fitness of a phenotype is the negative surprise averaged over all the states it experiences, which is simply its negative entropy. Indeed, the whole point of minimising free-energy (and implicitly entropy) is to ensure agents spend most of their time in a small number of valuable states. In short, that free-energy is (a bound on) the complement of value and its long-term average is (a bound) on the complement of adaptive fitness. But how do agents know what is valuable? In other words, how does one generation tell the next which states have value (i.e., are unsurprising). Value or surprise is determined by the agent's generative model and its implicit expectations — these specify the value of sensory states and, crucially, are heritable. This means prior expectations that are specified epigenetically can prescribe an attractive state. In turn, this enables natural selection to optimise prior expectations and ensure they are consistent with the agent's phenotype. Put simply, valuable states are just states the agent expects to frequent. These expectations are constrained by the form of its generative model, which is specified genetically and fulfilled behaviourally, under active inference. It is important to appreciate that prior expectations include not just what will be sampled from the world but how the world sampled. This means natural selection may equip agents with the prior expectation they will explore their environment, until attractive states are encountered. We will look at this more closely in the next section, where priors on motion through state-space are cast in terms of policies in reinforcement learning.

In summary, neuronal group selection rests on value, which depends on prior expectations about what agents expect to encounter. These expectations are sensitive to selective pressure at an evolutionary timescale and are fulfilled as action minimises free-energy. Both Neural Darwinism and the free-energy principle try to understand somatic changes in an individual in the context of evolution: Neuronal Darwinism appeals to selective processes, while the free-energy formulation considers the optimisation of ensemble or population dynamics in terms of entropy and surprise. The key theme that emerges here is that (heritable) prior expectations can label things as innately valuable (unsurprising); but how does labelling states

lead to adaptive behaviour? In the final section, we return to reinforcement learning and related formulations of action that try to explain adaptive behaviour in terms of policies and cost-functions.

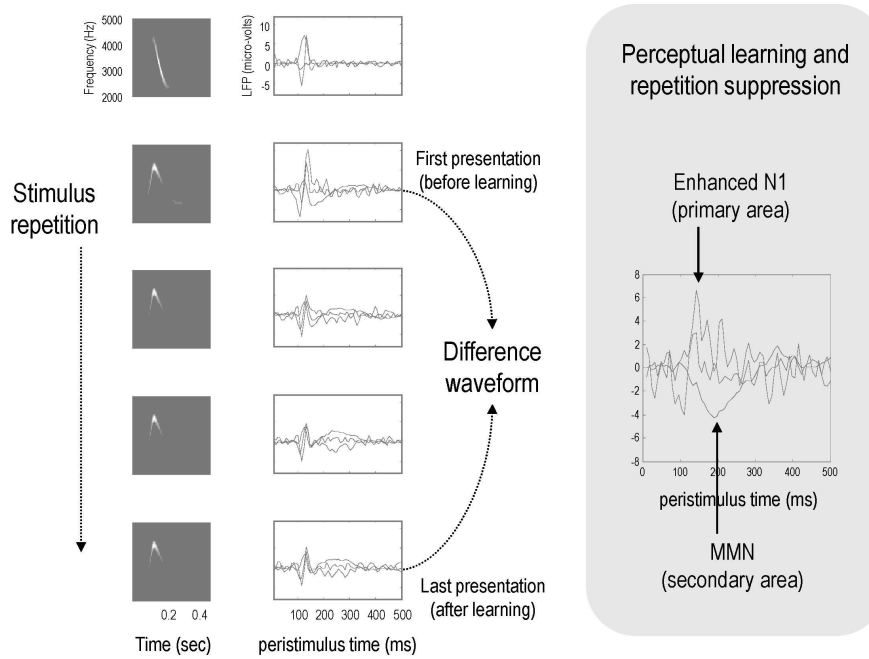


Fig. 6: A demonstration of perceptual learning. This figure shows the results of a simulated roving oddball paradigm, in which a stimulus is changed sporadically to elicit an oddball (i.e., deviant) response. The stimuli used here are chirps of the same sort as those used in Fig. 4. Left panels: The left column shows the percepts elicited in sonogram format. These are simply the predictions of sensory input, based on their inferred causes (i.e., the expectations about hidden states). The right column shows the evolution of prediction error at the first (dotted lines) and second (solid line) levels of a simple linear convolution model (in which a causal state produces time-dependent amplitude and frequency modulations). The results are shown for one learned chirp (top graph) and the first four responses to a new chirp (lower graphs). The new chirp was generated by changing the parameters of the underlying equations of motion. It can be seen that following the first oddball stimulus, the prediction errors show repetition suppression (i.e., the amplitudes of the traces get smaller). This is due to learning the model parameters over trials (see synaptic plasticity and gain in Fig. 3). Of particular interest is the difference in responses to the first and last presentations of the new stimulus: these correspond to the deviant and standard responses, respectively. Right panel: This shows the difference between standard and oddball responses, with an enhanced negativity at the first level early in peristimulus time (dotted lines for inferred amplitude and frequency), and a later negativity at the higher or second level (solid line for the causal state). These differences could correspond to phenomena like enhanced N1 effects and the mismatch negativity (MMN) found in empirical difference waveforms. Note that superficial pyramidal cells (see Fig. 3) dominate event related potentials and that these cells may encode prediction error[47,146].

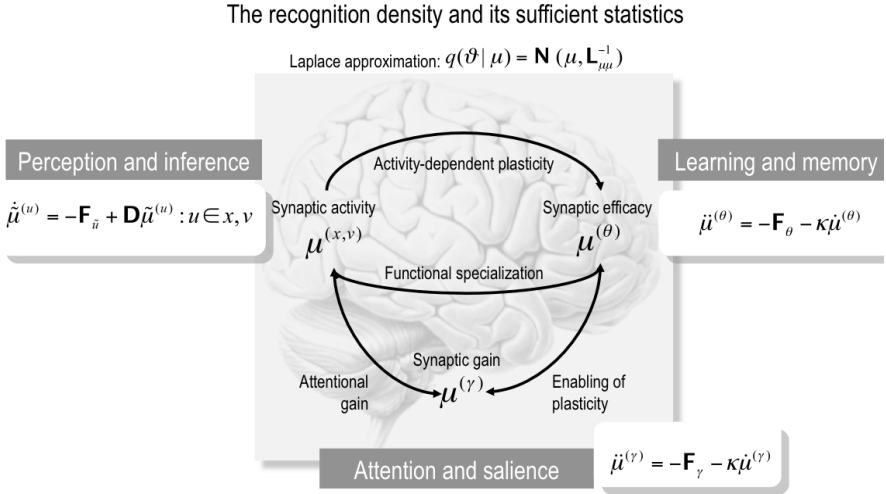


Fig. 7: The recognition density and its sufficient statistics. This schematic maps free-energy optimisation of the recognition density to putative processes in the brain: Under the Laplace assumption, the sufficient statistics of the recognition density (encoded by internal states) reduce to the conditional expectations (i.e., means). This is because the conditional precision is the curvature of the energy evaluated at the mean. Optimizing the conditional means of states of the world may correspond to optimising synaptic activity that mediates hierarchical message passing. Optimising the conditional means of parameters encoding causal structure may be implemented by associative mechanisms implementing synaptic plasticity and, finally, optimizing the conditional precisions may correspond to optimising synaptic gain (see Fig. 3).

POLICIES AND PRIORS

So far, we have established a fundamental role for generative models in furnishing a free-energy bound on surprise (or the value of attracting states an agent occupies). We have considered general (hierarchical and dynamic) forms for this model that prescribe predictions about how an agent will move through its state-space: in other words the state-transitions it expects. This expected motion corresponds to a *policy* that action is enslaved to pursue. However, we have not considered the form of this policy; i.e., the form of the equations of motion. In this section, we will look at universal forms for policies that define an agent's generative model. Because policies are framed in terms of equations of motion they manifest as (empirical) priors on the state-transitions an agent expects to make. This means that policies and priors are the same thing (under active inference) and both rest on the form of generative models embodied by agents. We first consider universal forms based on optimal control theory and reinforcement learning. These policies use an explicit representation of value to guide motion, under simplifying assumptions about state-transitions. Although useful heuristics these policies do not generalise to

dynamical settings. This is because they only lead to fixed (low-cost) states (i.e., fixed-point attractors). Although this is fine for plant control in engineering or psychology experiments with paradigmatic end-points, fixed-point policies are not viable solutions for real agents (unless they aspire to be petrified or dead). We will then move on to wandering or itinerant policies that lead to invariant sets of attracting states. Itinerant policies may offer universal policies and implicitly, universal forms for generative models.

From the previous section, policies (equations of motion in Fig. 2) have to satisfy constraints that are hereditable. In other words, they have to be elaborated given only sparsely encoded information about what states are innately attractive or costly, given the nature of the agent's phenotype. We will accommodate this with the notion of cost-functions. Cost-functions can be thought of as standing in for the genetic specification of attractive states but they also allow us to connect to another important perspective on policies from engineering and behavioural economics:

3.1 Optimal control and Game Theory

Value is central to theories of brain function that are based on reinforcement learning and optimum control. The basic notion that underpins these treatments is that the brain optimises value, which is expected reward or utility (or its complement, expected loss or cost). This is seen in behavioural psychology as reinforcement learning[103], in computational neuroscience and machine-learning, as variants of dynamic programming such as temporal difference learning[104-106], and in economics, as expected utility theory[107]. The notion of an expected reward or cost is crucial here; it is the cost expected over future states, given a particular policy that prescribes action or choices. A policy specifies the states an agent will move to from any given state (or motion through state-space in continuous time). This policy has to access sparse rewarding states given only a cost-function, which labels states as costly or not. The problem of optimising the policy is formalised in optimal control theory as the Bellman equation and its variants [104], which expresses value as a function of the optimal policy and a cost-function. If one can solve the Bellman equation, one can associate each sensory state with a value and optimise the policy by ensuring the next state is the most valuable of the available states. In general, it is impossible to solve the Bellman equation exactly but a number of approximations exist, ranging from simple Rescorla-Wagner models[103] to more comprehensive formulations like Q-learning[105]. Cost also has a key role in Bayesian decision theory, where optimal decisions minimise expected cost, not over time but in the context of uncertainty about outcomes; this is central to optimal decision (game) theory and behavioural economics[107-109].

So what does free-energy bring to the table? If value is inversely proportional to surprise (see above), then free-energy is (an upper bound on) expected future cost. This makes sense, because optimal control theory assumes that action minimises expected cost, whereas the free-energy principle states that it minimises free-energy. Furthermore, the dynamical perspective provides a mechanistic insight into how policies are specified in the brain: Under the Principle of Optimality[104] cost is the rate of change of value, which depends on changes in sensory states. This suggests that optimal policies can be prescribed by prior expectations about the motion of sensory states. Put simply, if priors induce a fixed-point attractor, when the states arrive at the fixed point, value will stop changing and cost will be minimised. A simple example is shown in Fig. 8, in which a cued arm movement is simulated using only prior expectations that the arm will be drawn to a fixed point (the target). This figure illustrates how computational motor control[110-114] can be formulated in terms of priors and the suppression of sensory prediction errors[115]. More generally, it shows how rewards and goals can be considered as prior expectations that action is obliged to fulfil[24] (see also ref [116]).

However, fixed-point policies based on maximising value (minimising surprise) explicitly are flawed in two respects. First, they lead to fixed-point attractors, which are not viable solutions for agents immersed in environments with autonomous and dissipative dynamics. The second and slightly more subtle problem with optimal control and its ethological variants is that they assume the existence of a policy (flow though state-space) that always increases value. Mathematically, this assumes value is 'Lyapunov function' of the policy. Unfortunately, these policies do not necessarily exist. Technically, value is proportional to (log) eigensolution to the Fokker-Planck equation describing the density dynamics of an infinite number of agents pursuing the same policy under random fluctuations. This eigensolution is the equilibrium density and is a function of the policy. However, this does not imply that the policy or flow always increases value: According to the Helmholtz decomposition (also known as the fundamental lemma of vector calculus) flow can always be decomposed into two components: an irrotational (curl-free) flow and a solenoidal (divergence-free) flow. When these components are orthogonal it is relatively easy to show that value is a Lyapunov function of the flow. However, there is no lemma or requirement for this orthogonality to exist and the Principle of Optimality[104] is not guaranteed. In summary, although value can (in principle) be derived from the policy, the policy cannot (in general) be derived from the flow. So where does that leave us in a search for universal policies? We turn for an answer to itinerant policies that are emerging as a new perspective on behaviour and purposeful self-organisation.

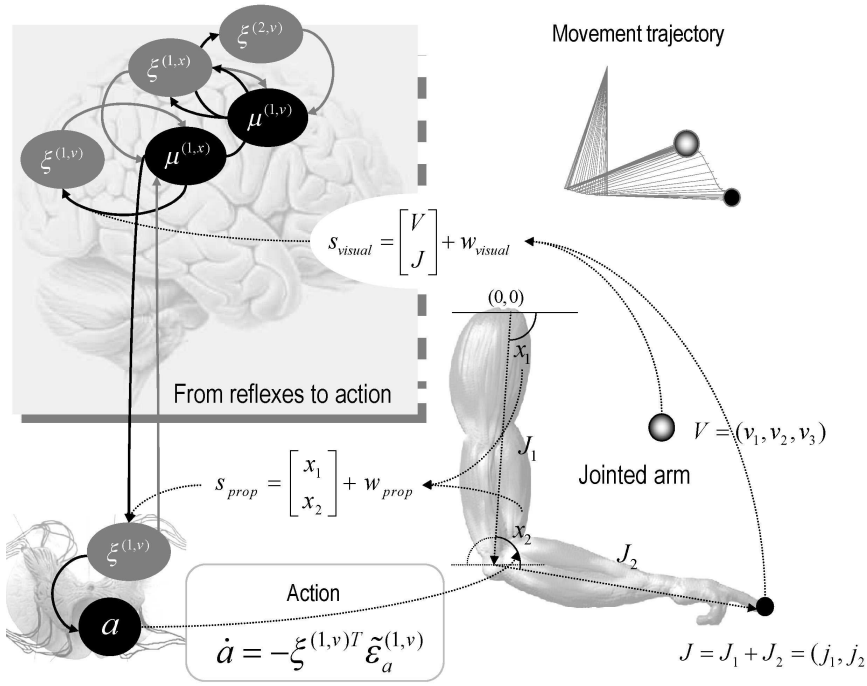


Fig. 8: A demonstration of cued-reaching movements. Lower right: motor plant, comprising a two-jointed arm with two hidden states, each of which corresponds to the angular position of joints. The position of the finger (black circle) is the sum of the vectors describing the location of each joint. Here, causal states in the world are the position and brightness of the target (grey sphere). The arm obeys Newtonian mechanics, specified in terms of angular inertia and friction. Left: The brain senses hidden states directly in terms of proprioceptive input (s_{prop}) that signals the angular positions (x_1,x_2) of the joints and indirectly, through seeing the location of the finger in space (j_1,j_2). In addition, the agent senses the target location (v_1,v_2) and brightness (v_3) through visual input (s_{visual}). Sensory prediction errors are passed to higher brain levels to optimise the conditional expectations of hidden states (i.e., the angular position of the joints) and causal (i.e., target) states. The ensuing predictions are sent back to suppress sensory prediction errors. At the same time, sensory prediction errors are also trying to suppress themselves by changing sensory input through action. The grey and black lines denote reciprocal message-passing among neuronal populations that encode prediction error and conditional expectations; this architecture is the same as that depicted in Fig. 3. The descending black line represents motor control signals (predictions) from sensory state-units. The agent's generative model includes priors on the motion of hidden states that effectively engage an invisible spring between the finger and target (when the target is illuminated). This induces a prior expectation that the finger will be drawn to the target, when cued appropriately. Insert (upper right): The ensuing movement trajectory caused by action. The black circles indicate the initial and final positions of the finger, which reaches the target (grey ball) quickly and smoothly.

3.2 Itinerant policies

This subsection considers attractive states that are not fixed-points but bounded sets that arise from itinerant (wandering or searching) dynamics. This speaks to optimising space-filling attractors that ensure low-cost equilibria. The importance of itinerancy has been articulated many times in the past (see ref [117]), particularly from the perspective of computation and autonomy (with a focus on Milnor attractors[118]). It has also been considered formally in relation to cognition (with a focus on attractor relics, ghosts or ruins[119]) and implicitly in ethology[120]. The ethological perspective is useful here because it suggests that some species are equipped with prior expectations that they will engage in exploratory or social play. For example, 'rough and tumble play' may be a fundamental form of play comprising a unique set of behaviours that can be distinguished from aggression and other childhood activities. Tani et al. [121] consider itinerant dynamics in terms of bifurcation parameters that generate multiple goal-directed actions on the behavioural side, and optimization of the same parameters when recognizing actions. They provide a series of elegant robotic simulations to show generalization by learning with this scheme. See also ref [122] for interesting simulations of itinerant exploration, using just prediction errors on sensory samples over time.

Although there may not be a universal form for itinerant policies, the principles upon which they are based may be universal. One principle (which we focus on here) is the vitiation or destruction of costly attractors. This idea appears in several guises and has found important applications in a number of domains. For example, it is closely related to the notion of autopoiesis and self-organisation in situated (embodied) cognition[123]. It is formally related to the destruction of gradients in synergetic treatments of intentionality[124]. Mathematically, it is finding a powerful application to universal optimisation schemes[125] and in models of perceptual categorization[126]. The dynamical phenomena, upon which these schemes rest, involve an itinerant wandering through state-space along heteroclinic channels (orbits connecting different fixed-points). Crucially, these attracting sets are weak (Milnor) attractors or attractor ruins that expel the state until it finds the next weak attractor or ruin. The result is a sequence of transitions through state-space that, in some instances, can be stable and repeating. The resulting stable heteroclinic channels have already been proposed as a metaphor for neuronal dynamics and underlying cognitive processing[127]. Furthermore, the notion of Milnor or ruined attractors underlies much of the technical and cognitive literature on itinerant dynamics. For example, Tyukin et al. [126] can explain "a range of phenomena in biological vision, such as mental rotation, visual search, and the presence of multiple time scales in adaptation" using the concept of weakly attracting sets.

To illustrate itinerant policies we will focus on the simplest of examples: An examination of the density dynamics, upon which the free-energy principle is based, suggests it is sufficient to keep moving until an *a priori* attractor is encountered (see ref [5]; supplementary material). This entails destroying unexpected (costly) fixed-points in the environment by making them unstable (like shifting to a new position when sitting uncomfortably). Mathematically, this reduces to adopting a policy that ensures a positive divergence in costly states (intuitively, this is like moving through a liquid with negative viscosity). Fig. 9 illustrates a solution to the classical mountain car problem using a simple prior that induces this sort of policy. This prior is on the motion of (i.e., changes in) states and enforces exploration until an attractive state is found. Priors of this sort may provide a principled way to understand the exploration-exploitation trade-off[128-130] and related issues in evolutionary biology [131]. The implicit use of priors to induce dynamical instability (i.e., autovitiation) also provides a key connection to dynamical systems theory approaches to the brain that emphasise the importance of itinerant dynamics, metastability, self-organised criticality and winner-less competition[127,132-139], which play a key role in synergetic and autopoietic accounts of adaptive behaviour[13,122,124].

The mountain car example (Fig. 9) provides a fairly abstract example of a very simple (if effective) itinerant policy. It may help to consider formally related policies in simple organisms whose genetic and cellular mechanisms are well understood: The bacterium Escherichia coli (E. coli) is an organism of choice for unravelling biochemical pathways, deciphering the genetic code and studying the molecular biology of behaviour[140,141]. E. coli is propelled in aqueous media by long thin helical filaments, each driven by a reversible rotary engine at its base. As peritrichous bacteria they alternately swim and tumble (thrash about with little forward progress), elaborating a random walk; with relatively straight swims interrupted by tumbles that reorient the bacterium. Bacteria such as E. coli cannot choose the direction in which they swim and are unable to swim in a straight line for more than a few seconds due to rotational diffusion. Given these limitations, it is remarkable that they can direct their motion to high concentrations of attractants (i.e., chemotaxis). If the bacterium senses that it is moving in the right direction (towards an attractant), it will keep swimming in a straight line for a longer time before tumbling. If it is moving in the wrong direction, it will tumble sooner and try a new direction. In short, by selective modulation of tumbling frequency, these bacteria show chemotaxis[140]. This is a nice example of an itinerant policy based on the prior expectation (endowed by natural selection) that the organism will only change its motion through state-space when it encounters unexpected (costly) generalised states (here, a decease in the concentration of attractants).

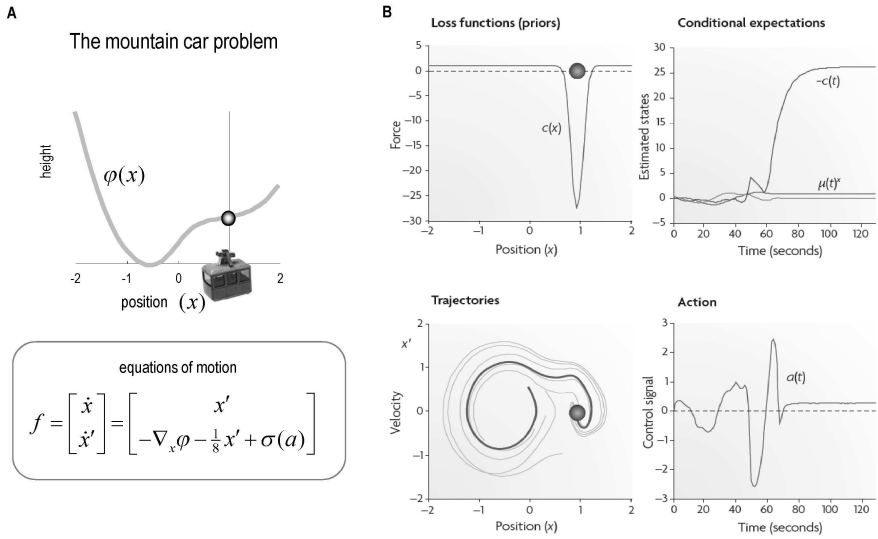


Fig. 9: Active inference and behavior. Solving the mountain car problem with prior expectations. This example shows how paradoxical but adaptive behaviour (e.g. moving away from a target to secure it later) emerges from simple priors on the motion of hidden states in the world. Panel A shows the landscape or potential energy function (with a minimum at $x = -0.5$) that exerts forces on the car. The car is shown at the target position on the hill at $x = 1$, indicated by the grey ball. The equations of motion of the car are shown below the figure. Crucially, at $x = 0$ the force on the car cannot be overcome by the agent, because a squashing function $-1 \leq \sigma(a) \leq 1$ is applied to action to prevent it being greater than one. This means that the agent can only access the target by starting halfway up the left hill to gain enough momentum to carry it up the other side. Panel B: The results of active inference under priors that destabilise fixed points outside the target domain. The priors are encoded in a cost-function $c(x)$ (lower left), which drives a hidden state corresponding to friction. When friction is negative the car expects to go faster (see ref [5], Supplementary material for details). The inferred hidden states (upper right: position, velocity and negative dissipation) show that the car explores its landscape until it encounters the target. At this point friction increases dramatically to prevent the car from escaping (i.e., falling down the hill). The ensuing trajectory is shown on the upper left. The paler lines provide exemplar trajectories from other trials, with different starting positions. In the real world friction is constant. However, the car expects friction to change with position, thus enforcing exploration or exploitation. These expectations are fulfilled by action (lower right).

In summary, the predictions afforded by generative models of the world oblige action to pursue policies specified in terms of equations of motion through state-space. Fixed-point policies, of the sort found in optimal control and decision (game) theory, start with the notion of cost or utility and try to construct value-functions of states, whose gradients guide the flow. Conversely, the free-energy formulation starts with (a bound on) the value of states, which is specified (via flow) by priors on the motion of hidden environmental states. These priors can incorporate cost-functions to vitiate costly states, leading to itinerant policies. In this view, the problem of finding sparse rewards in the environment is na-

ture's solution to the problem of how to minimise the entropy (average surprise or free-energy) of an agent's states; by ensuring they occupy a limited invariant set of attracting (i.e., rewarding) states. These dynamics rest on the complementary self-construction (autopoiesis) and destruction (autovitiation) of attracting sets, which are mandated by the existence of agents that are at equilibrium with their environment.

DISCUSSION

Although contrived to highlight commonalities, the material reviewed in this chapter suggests that many global theories of brain function can be united under a Helmholtzian perspective on the brain as generative model of the world it inhabits[1,27,2,30]. Notable examples include the integration of the Bayesian brain and computational motor control, the objective functions shared by predictive coding and the infomax principle, hierarchical inference and theories of attention (e.g., biased competition), the embedding of perception in natural selection and the link between optimum control (i.e., reinforcement learning and dynamic programming) and more exotic phenomena in dynamical systems theory (i.e., attractors, winner-less competition and itinerancy). The constant theme in all these theories is that the brain optimises a (free-energy) bound on surprise or its complement, value. This manifests as perception (so as to change predictions), or action (so as to change the sensations that are predicted). Crucially, these predictions depend on prior expectations (that furnish policies), which are optimised at different (somatic and evolutionary) time scales and define what is valuable. See Fig. 10 for a schematic summary of free-energy optimisation at different scales.

What does the free-energy principle portend for the future? If its main contribution is to integrate established theories, then the answer is probably "not a lot". On the other hand, it may provide a framework in which current debates could be resolved; e.g., does dopamine encode reward prediction error or surprise[142,143]. This is particularly important for understanding things like addiction, Parkinson's disease and schizophrenia. Indeed the free-energy formulation has already been used to explain the positive symptoms of schizophrenia (i.e., hallucinations and delusions), in terms of false inference[144]. The free-energy formulation may also provide some new approaches to old problems that might call for a reappraisal of conventional notions (particularly in reinforcement learning and motor control; see the previous section). If the arguments underlying the free-energy principle hold, then the real challenge is to understand how it manifests in the brain. This speaks to a greater appreciation of hierarchical message-passing[47] and the functional role of specific neurons and microcircuits; and the dynamics they support (e.g., what is the relationship between predictive coding, attention and dynamic coordination in the

brain?[145]). Beyond neuroscience, many exciting applications in engineering, robotics, embodied cognition and evolutionary biology suggest themselves; although fanciful, it is not difficult to imagine building little free-energy machines that garner and model sensory information (like our children) to maximise the evidence for their own existence.

CONCLUSION

The free-energy principle rests on a fundamental imperative for biological systems; namely, to select exchanges with the environment that ensure their physical states constitute an invariant bounded set. This precludes phase-transitions and underwrites the system's (agent's) longevity. In the introduction, we summarised this as "I think therefore I am, iff I am what I think". In other words, I model myself as embodied in my environment and harvest sensory evidence for that model. If I am what I model, then confirmatory evidence will be available. If I am not, then I will experience things that are incompatible with my (hypothetical) existence. And, after a short period, will cease to exist in my present form.

The implicit duality between 'being' and 'thinking' is not the Cartesian duality that preoccupies philosophers. It is a pragmatic duality between physical states and a probabilistic representation they entail. These entailed constructs are the generative and recognition densities in Figure 1. The free-energy is a functional (function of a function) of these densities and is therefore a function of their sufficient statistics (internal states). In principle, it should be possible to infer the functional form of the free-energy given the action and internal states of any organism. In short, the densities are well-defined (if not necessarily unique) mathematical constructs that are paired with (entailed by) the physical states of an agent. These constructs can be quantified and studied empirically. A simple example here is the duality between neuronal activity as a physiological process and as a conditional expectation about a hidden state of the world. This exemplifies one functional form for the free-energy. To establish that this is the right form, one would need to show that it is minimised by action and perception.

The free-energy perspective does not mean that we get up in the morning and set about minimising our free-energy; any more than E. coli are purposefully trying to minimise prediction error when tumbling through their milieu. We are saying that if biological systems attain equilibrium with their environment, their internal states must entail a generative model of their world, whose free-energy is minimised by action and perception. This is true whether you are an E. coli or an evangelist. Because free-energy is a function of sensations and internal states it is, in essence, an attribute of an embodied inference.

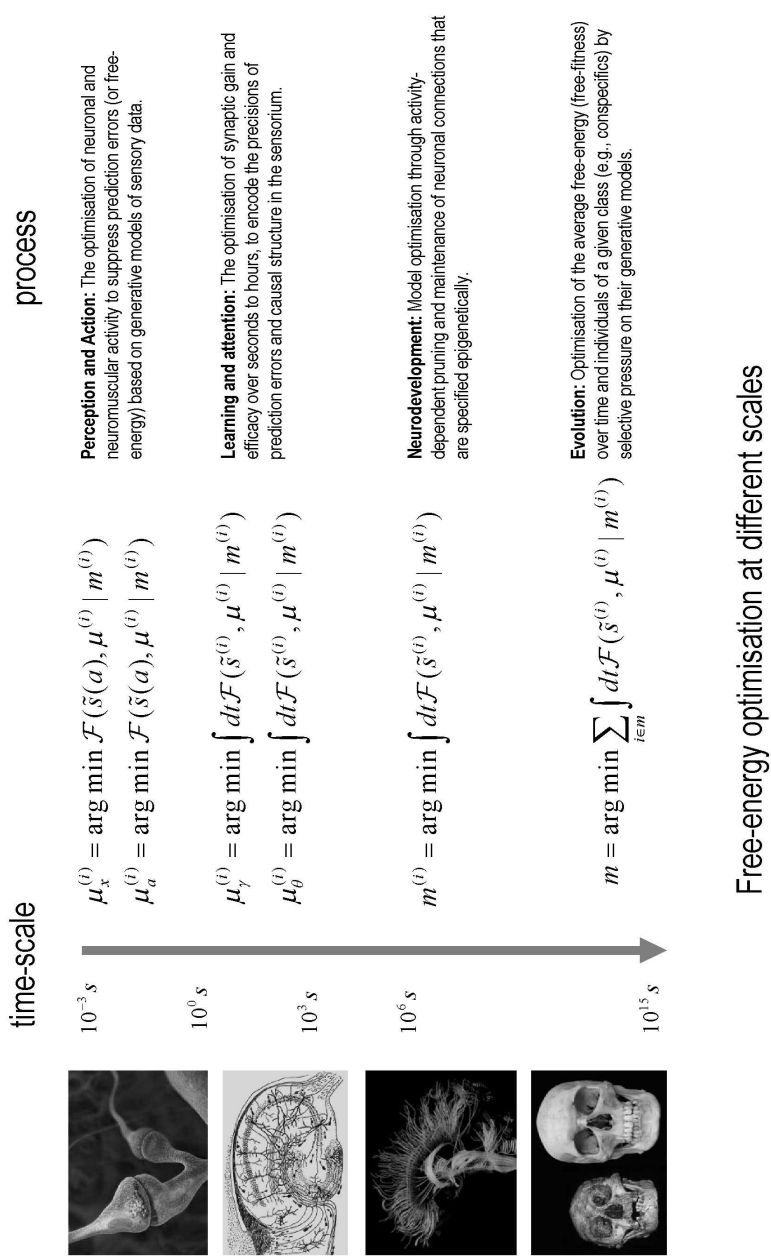


Fig. 10: Optimising free-energy over different time-scales. This schematic summarises the various time-scales over which minimisation of free-energy can be considered as optimising the state (perception), configuration (action), connectivity (learning and attention), anatomy (neuro-development) and the phenotype (evolution) of an agent. Here, $F(\tilde{s}, \mu^{(i)} \mid m^{(i)})$ is the free-energy of the sensory data (and its temporal derivatives) $\tilde{s}(a)$ and states μ of an agent $m^{(i)} \in m$ that belongs to class m, while action a determines the sampling of sensory data. The physical states of the phenotype μ encode an implicit recognition density. In the brain, these representations could correspond to synaptic activity, gain and strength.

Acknowledgments

This work was funded by the Wellcome Trust. I would like to thank my colleagues at the Wellcome trust Centre for Neuroimaging, The Institute of Cognitive Neuroscience and the Gatsby Computational Neuroscience Unit for invaluable collaborations and discussions. I would also like to thank the organisers and participants of the 15th Herbstakademie / Fall Academy: Embodied Cognition and Embodied Communication for insights and discussions.

References

1. Helmholtz H. von (1866). Concerning the perceptions in general. In *Treatise on physiological optics*, vol. III, 3rd edn (translated by J. P. C. Southall 1925 Opt. Soc. Am. Section 26, reprinted New York: Dover, 1962).
2. Gregory R. L. (1968). Perceptual illusions and brain models. *Proc R Soc Lond B, 171*, 179-96.
3. Gregory R. L. (1980). Perceptions as hypotheses. Phil. *Trans R Soc Lond B, 290*, 181-97.
4. Descartes R. (1637/1960). *Discourse on Method and Meditations*. L. J. Lafleur (transl.). New York: The Liberal Arts Press.
5. Friston K. (2010). The free-energy principle: a unified brain theory? *Nat Rev Neurosci, 11*, 127-38.
6. Thornton C. (2010). Some puzzles relating to the free-energy principle: comment on Friston. *Trends Cogn Sci, 14*, 53-4; author reply 54-5.
7. Carhart-Harris R. L. & Friston K. J. (2010). The default-mode, ego-functions and free-energy: A neurobiological account of Freudian ideas. *Brain, 133*, 1265-83.
8. Grist M. (2010). *Changing the Subject.* RSA. www.thesocialbrain.wordpress.com (pp. 74-80).
9. Huang G. (2008). Is this a Unified Theory of the Brain? *New Scientist*, *2658*, 30-3.
10. Friston K., Kilner J., & Harrison L. (2006). A free-energy principle for the brain. *J Physiol Paris, 100*, 70-87.
11. Ashby W. R. (1947). Principles of the self-organising dynamic system. *J Gen Psychology*, *37*, 125-8.
12. Nicolis G. & Prigogine I. (1977). *Self-organisation in non-equilibrium systems* (p. 24). New York: John Wiley.
13. Haken H. (1983). *Synergetics: An introduction. Non-equilibrium phase transition and self-organisation in physics, chemistry and biology* (3rd ed). Berlin: Springer.
14. Kauffman S. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford: Oxford University Press.
15. Bernard C. (1974). *Lectures on the phenomena common to animals and plants.* H. E. Hoff, R. Guillemin, & L. Guillemin (transl.). Springfield: Charles C Thomas.
16. Applebaum D. (2008). *Probability and Information: An Integrated Approach*. Cambridge: Cambridge University Press.
17. Evans D. J. (2003). A non-equilibrium free-energy theorem for deterministic systems. *Molecular Physics, 101*, 1551-4.
18. Crauel H. & Flandoli F. (1994). Attractors for random dynamical systems. *Probab Theory Relat Fields, 100*, 365-93.
19. Feynman R. P. (1972). *Statistical mechanics.* Reading: Benjamin.
20. Hinton G. E. & van Camp D. (1993). Keeping neural networks simple by minimising the description length of weights. *Proceedings of COLT-93* (pp. 5-13).

21. MacKay D. J. C. (1995). Free-energy minimisation algorithm for decoding and cryptoanalysis. *Electronics Letters, 31*, 445-7.
22. Neal R. M. & Hinton G. E. (1998). A view of the EM algorithm that justifies incremental, sparse, and other variants. In M. I. Jordan (ed.). *Learning in Graphical Models* (pp. 355-68). Dordrecht: Kluwer Academic Publishers.
23. Itti L. & Baldi P. (2009). Bayesian Surprise Attracts Human Attention. *Vision Res., 49,* 1295-306.
24. Friston K., Daunizeau J., & Kiebel S. (2009). Active inference or reinforcement learning? *PLoS-ONE, 4,* e6421.
25. Knill D. C. & Pouget A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci*, 27, 712-9.
26. MacKay D. M. (1956). The epistemological problem for automata. In C. E. Shannon & J. McCarthy (eds.), *Automata Studies (*pp. 235-51*)*. Princeton: Princeton University Press.
27. Neisser U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.
28. Ballard D. H., Hinton G. E., & Sejnowski T. J. (1983). Parallel visual computation. *Nature*, *306*, 21-6.
29. Kawato M., Hayakawa H., & Inui T. (1993). A forward-inverse optics model of reciprocal connections between visual areas. Network. *Computation in Neural Systems, 4*, 415-22.
30. Dayan P., Hinton G. E., & Neal R. M. (1995). The Helmholtz machine. *Neural Computation, 7*, 889-904.
31. Lee T. S. & Mumford D. (2003). Hierarchical Bayesian inference in the visual cortex. *J Opt Soc Am Opt Image Sc Vis, 20,* 1434-48.
32. Kersten D., Mamassian P., & Yuille A. (2004). Object perception as Bayesian inference. *Annu Rev Psychol*, *55*, 271-304.
33. Friston K. J. (2005). A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci, 360,* 815-36.
34. Beal M. J. (2003). *Variational Algorithms for Approximate Bayesian Inference.* PhD. Thesis, Gatsby Computational Neuroscience Unit, University College London.
35. Efron B. & Morris C. (1973). Stein's estimation rule and its competitors — an empirical Bayes approach. *J Am Statist Assoc, 68,* 117-30.
36. Kass R. E. & Steffey D. (1989). Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). *J Am Stat Assoc, 407*, 717-26.
37. Zeki S. & Shipp S. (1988). The functional logic of cortical connections. *Nature, 335,* 311-7.
38. Felleman D. J. & Van Essen D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*, 1-47.
39. Mesulam M. M. (1998). From sensation to cognition. *Brain, 121*, 1013-52.
40. Sanger T. (1996). Probability density estimation for the interpretation of neural population codes. *J Neurophysiol, 76*, 2790-3.
41. Zemel R., Dayan P., & Pouget A. (1998). Probabilistic interpretation of population code. *Neural Computat, 10*, 403-30.
42. Paulin M. G. (2005). Evolution of the cerebellum as a neuronal machine for Bayesian state estimation. *J Neural Eng*, *2*, 219-34.
43. Ma W. J., Beck J. M., Latham P. E., & Pouget A. (2006). Bayesian inference with probabilistic population codes. *Nat Neurosci*, *9*, 1432-8.
44. Friston K., Mattout J., Trujillo-Barreto N., Ashburner J., & Penny W. (2007). Variational free-energy and the Laplace approximation. *Neuroimage, 34,* 220-34.
45. Roweis S. & Ghahramani Z. (1999). A unifying review of linear Gaussian models. *Neural Computation, 11*, 305-45

46. Rao R. P. & Ballard D. H. (1998). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive field effects. *Nature Neuroscience, 2*, 79-87.
47. Mumford D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol Cybern, 66*, 241-51.
48. Friston K. (2008). Hierarchical models in the brain. *PLoS Comput Biol, 4*, e1000211.
49. Murray S. O., Kersten D., Olshausen B. A., Schrater P., & Woods D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proc Natl Acad Sci USA, 99*, 15164-9.
50. Garrido M. I., Kilner J. M., Kiebel S. J., & Friston K. J. (2009). Dynamic causal modeling of the response to frequency deviants. *J Neurophysiol*, *101*, 2620-31.
51. Sherman S. M. & Guillery R. W. (1998). On the actions that one nerve cell can have on another: distinguishing "drivers" from "modulators". *Proc Natl Acad Sci USA, 95*, 7121-6.
52. Angelucci A. & Bressloff P. C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Prog Brain Res*, *154*, 93-120.
53. Grossberg S. (2007). Towards a unified theory of neocortex: laminar cortical circuits for vision and cognition. *Prog Brain Res, 165*, 79-104.
54. Grossberg S. & Versace M. (2008). Spikes, synchrony, and attentive learning by laminar thalamocortical circuits. *Brain Res*, *1218*, 278-312.
55. Barlow H. (1961). Possible principles underlying the transformations of sensory messages. In W. Rosenblith (ed.). *Sensory Communication* (pp. 217-234). Cambridge: MIT Press.
56. Linsker R. (1990). Perceptual neural organisation: some approaches based on network models and information theory. *Annu Rev Neurosci, 13*, 257-81.
57. Oja E. (1989). Neural networks, principle components, and subspaces. *Int J Neural Systems, 1*, 61-8.
58. Bell A. J. & Sejnowski T. J. (1995). An information maximisation approach to blind separation and blind de-convolution. *Neural computation, 7*, 1129-59.
59. Atick J. J. & Redlich A. N. (1992). What does the retina know about natural scenes? *Neural Computation, 4*, 196-210.
60. Optican L. & Richmond B. J. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior cortex. II Information theoretic analysis. *J Neurophysiol*, *57*, 132-46.
61. Olshausen B. A. & Field D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*, 607-9.
62. Simoncelli E. P. & Olshausen B. A. (2001). Natural image statistics and neural representation. *Annu Rev Neurosci*, 24, 1193-216.
63. Friston K. J. (2000). The labile brain. III. Transients and spatio-temporal receptive fields. *Philos Trans R Soc Lond B Biol Sci, 355*, 253-65.
64. Bialek W., Nemenman I., Tishby N. (2001). Predictability, complexity, and learning. *Neural Comput*, *13*, 2409-63.
65. Lewen G. D., Bialek W., & de Ruyter van Steveninck R. R. (2001). Neural coding of naturalistic motion stimuli. *Network*, *12*, 317-29.
66. Laughlin S. B. (2001). Efficiency and complexity in neural coding. *Novartis Found Symp, 239*, 177-87
67. Tipping M. E. (2001). Sparse Bayesian learning and the Relevance Vector Machine. *J Machine Learning Research*, *1*, 211-44.
68. Paus T., Keshavan M., & Giedd J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nat Rev Neurosci, 9*, 947-57.

69. Gilestro G. F., Tononi G., & Cirelli C. (2009). Widespread changes in synaptic markers as a function of sleep and wakefulness in Drosophila. *Science, 324*, 109-12.
70. Roweis S. & Ghahramani Z. (1999). A unifying review of linear Gaussian models. *Neural Computation, 11*, 305-45.
71. Hebb D. O. (1949) *The organization of behaviour*. New York: Wiley.
72. Paulsen O. & Sejnowski T. J. (2000). Natural patterns of activity and long-term synaptic plasticity. *Current opinion in neurobiology, 10*, 172–9.
73. von der Malsburg C. (1981). *The Correlation Theory of Brain Function*. Internal Report 81-2, Dept. of Neurobiology, Max-Planck-Institute for Biophysical Chemistry, Gottingen.
74. Singer W. & Gray C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annu Rev Neurosci, 18*, 555-86.
75. Bienenstock E. L., Cooper L. N., & Munro P. W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J Neurosci, 2*, 32-48.
76. Abraham W. C. & Bear M. F. (1996). Metaplasticity: the plasticity of synaptic plasticity. *Trends Neurosci, 19*, 126-30.
77. Pareti G. & De Palma A. (2004). Does the brain oscillate? The dispute on neuronal synchronization. *Neurol Sci*, *5*, 41-7.
78. Leutgeb S., Leutgeb J. K., Moser M. B., & Moser E. I. (2005). Place cells, spatial maps and the population code for memory. *Curr Opin Neurobiol, 15*, 738-46.
79. Durstewitz D. & Seamans J. K. (2006). Beyond bistability: biophysics and temporal dynamics of working memory. *Neuroscience, 139*, 119-33.
80. Anishchenko A. & Treves A. (2006). Autoassociative memory retrieval and spontaneous activity bumps in small-world networks of integrate-and-fire neurons. *J Physiol Paris, 100*, 225-36.
81. Abbott L. F., Varela J. A., Sen K., & Nelson S. B. (1997). Synaptic depression and cortical gain control. *Science, 275*, 220-4.
82. Yu A. J. & Dayan P. (2005). Uncertainty, neuromodulation and attention. *Neuron, 46*, 681-92.
83. Doya K. (2002). Metalearning and neuromodulation. *Neural Netw, 15*, 495-506.
84. Chawla D., Lumer E. D., & Friston K. J. (1999). The relationship between synchronization among neuronal populations and their mean activity levels. *Neural Comput, 11*, 1389-411.
85. Fries P., Womelsdorf T., Oostenveld R., & Desimone R. (2008). The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *J Neurosci, 28*, 4823-35.
86. Womelsdorf T. & Fries P. (2006). Neuronal coherence during selective attentional processing and sensory-motor integration. *J Physiol Paris, 100*, 182-93.
87. Desimone R. (1996). Neural mechanisms for visual memory and their role in attention. *Proc Natl Acad Sci USA, 93*, 13494-9.
88. Treisman A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London Series B*, 353, 1295-306.
89. Maunsell J. H. & Treue S. (2006). Feature-based attention in visual cortex. *Trends in Neuroscience, 29*, 317-22.
90. Spratling M. W. (2008). Predictive-coding as a model of biased competition in visual attention. *Vision Research, 48*, 1391-408.
91. Reynolds J. H. & Heeger D. J. (2009). The normalization model of attention. *Neuron, 61*, 168-85.
92. Schroeder C. E., Mehta A. D., & Foxe J. J. (2001). Determinants and mechanisms of attentional modulation of neural processing. *Front Biosci, 6*, D672-84.

93. Hirayama J., Yoshimoto J., & Ishii S. (2004). Bayesian representation learning in the cortex regulated by acetylcholine. *Neural Netw*, *17*, 1391-400.
94. Edelman G. M. (1993). Neural Darwinism: selection and reentrant signaling in higher brain function. *Neuron, 10*, 115-25.
95. Knobloch F. (2001). Altruism and the hypothesis of meta-selection in human evolution. *J Am Acad Psychoanal*, *29*, 339-54.
96. Friston K. J., Tononi G., Reeke G. N. Jr, Sporns O., & Edelman G. M. (1994). Value-dependent selection in the brain: simulation in a synthetic neural model. *Neuroscience, 59*, 229-43.
97. Sutton R. S. & Barto A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychol Rev*, *88*, 135-70.
98. Montague P. R., Dayan P., Person C., & Sejnowski T. J. (1995). Bee foraging in uncertain environments using predictive Hebbian learning. *Nature*, *377*, 725-8.
99. Schultz W. (1998). Predictive reward signal of dopamine neurons. *J Neurophysiol, 80*, 1-27.
100. Daw N. D. & Doya K. (2006). The computational neurobiology of learning and reward. *Curr Opin Neurobiol*, *16*, 199-204.
101. Redgrave P. & Gurney K. (2006). The short-latency dopamine signal: A role in discovering novel actions? *Nature Reviews Neuroscience, 7*, 967-75.
102. Berridge K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl), 191*, 391-431.
103. Rescorla R. A. & Wagner A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64-99). New York: Appleton Century Crofts.
104. Bellman R (1952). On the Theory of Dynamic Programming. *Proceedings of the National Academy*, *38*, 716-9.
105. Watkins C. J. C. H. & Dayan P. (1992). Q-learning. *Machine Learning*, 8, 279-92.
106. Todorov E. (2006). Linearly-solvable Markov decision problems. In Scholkopf et al. (eds). *Advances in Neural Information Processing Systems 19* (pp. 1369-76). Cambridge: MIT Press.
107. Camerer C. F. (2003). Behavioural studies of strategic thinking in games. *Trends Cogn Sci*, *7*, 225-31.
108. Smith J. M. & Price G. R. (1973). The logic of animal conflict, *Nature, 246*, 15-8.
109. Nash J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences of the United States of America, 36*, 48-9.
110. Wolpert D. M. & Miall R. C. (1996). Forward Models for Physiological Motor Control. *Neural Netw, 9*, 1265-79.
111. Todorov E. & Jordan M. I. (1998). Smoothness maximization along a predefined path accurately predicts the speed profiles of complex arm movements. *J Neurophysiol, 80*, 696-714.
112. Tseng Y. W., Diedrichsen J., Krakauer J. W., Shadmehr R., & Bastian A. J. (2007). Sensory prediction-errors drive cerebellum-dependent adaptation of reaching. *J Neurophysiol, 98*, 54-62.
113. Bays P. M. & Wolpert D. M. (2007). Computational principles of sensorimotor control that minimize uncertainty and variability. *J Physiol*, *578*, 387-96.
114. Shadmehr R. & Krakauer J. W. (2008). A computational neuroanatomy for motor control. *Exp Brain Res*, *185*, 359-81.
115. Friston K. J., Daunizeau J., Kilner J., & Kiebel S. J. (2010). Action and behaviour: A free-energy formulation. *Biol Cybern*, 102, 227-60.
116. Verschure P. F., Voegtlin T., & Douglas R. J. (2003). Environmentally mediated synergy between perception and behaviour in mobile robots. *Nature*, *425*, 620-4.

117. Nara S. (2003). Can potentially useful dynamics to solve complex problems emerge from constrained chaos and/or chaotic itinerancy? *Chaos*, *13*, 1110-21.
118. van Leeuwen C. (2008). Chaos breeds autonomy: connectionist design between bias and baby-sitting. *Cogn Process*, *9*, 83-92.
119. Gros C. (2009). Cognitive computation with autonomously active neural networks: an emerging field. *Cognitive Computation*, *1*, 77-99.
120. Panksepp J., Siviy S., & Normansell L. (1984). The psychobiology of play: theoretical and methodological perspectives. *Neurosci Biobehav Rev.*, *8*, 465-92.
121. Tani J., Ito M., & Sugita Y. (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using RNNPB. *Neural Networks*, *17*, 1273-89.
122. Herrmann J. M., Pawelzik K., & Geisel T. (1999). Self-localization of autonomous robots by hidden representations. *Autonomous Robots*, 7, 31-40.
123. Maturana H. R. & Varela F. (1972). De máquinas y seres vivos. Santiago, Chile: Editorial Universitaria. English version: Autopoiesis: the organization of the living. In H. R. Maturana & F. G. Varela (eds.) (1980). *Autopoiesis and Cognition*. Dordrecht, Netherlands: Reidel
124. Tschacher W. & Haken H. (2007). Intentionality in non-equilibrium systems? The functional aspects of self-organised pattern formation. *New Ideas in Psychology*, *25*, 1-15.
125. Tyukin I., van Leeuwen C., & Prokhorov D. (2003). Parameter estimation of sigmoid superpositions: dynamical system approach. *Neural Comput.*, *15*, 2419-55.
126. Tyukin I., Tyukina T., & van Leeuwen C. (2009). Invariant template matching in systems with spatiotemporal coding: A matter of instability. *Neural Netw*, *22*, 425-49.
127. Rabinovich M., Huerta R., & Laurent G. (2008). Neuroscience. Transient dynamics for neural processing. *Science*, *321*, 48-50.
128. Cohen J. D., McClure S. M., & Yu A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos Trans R Soc Lond B Biol Sci*, *362*, 933-42.
129. Ishii S., Yoshida W., & Yoshimoto J. (2002). Control of exploitation-exploration meta-parameter in reinforcement learning. *Neural Netw*, *15*, 665-87.
130. Usher M., Cohen J. D., Servan-Schreiber D., Rajkowski J., & Aston-Jones G. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science, 283,* 549-54.
131. Voigt C. A., Kauffman S., & Wang Z. G. (2000). Rational evolutionary design: the theory of in vitro protein evolution. *Adv Protein Chem, 55,* 79-160.
132. Freeman W. J. (1994). Characterization of state transitions in spatially distributed, chaotic, nonlinear, dynamical systems in cerebral cortex. *Integr Physiol Behav Sci, 29,* 294-306.
133. Tsuda I. (2001). Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. *Behav Brain Sci, 24,* 793-810.
134. Jirsa V. K., Friedrich R., Haken H., & Kelso J. A. (1994). A theoretical model of phase transitions in the human brain. *Biol Cybern, 71*, 27-35.
135. Breakspear M. & Stam C. J. (2005). Dynamics of a neural system with a multiscale architecture. *Philos Trans R Soc Lond B Biol Sci, 360,* 1051-74.
136. Bressler S. L & Tognoli E. (2006). Operational principles of neurocognitive networks. Int *J Psychophysiol, 60,* 139-48.
137. Werner G. (2007). Brain dynamics across levels of organization. *J Physiol Paris, 101*, 273-9.

138. Pasquale V., Massobrio P., Bologna L. L., Chiappalone M., & Martinoia S. (2008). Self-organization and neuronal avalanches in networks of dissociated cortical neurons. *Neuroscience*, *153*, 1354-69.
139. Kitzbichler M. G., Smith M. L., Christensen S. R., & Bullmore E. (2009). Broadband criticality of human brain network synchronization. *PLoS Comput Biol*, *5*, e1000314.
140. Ordaq G. W. & Fields R. B. (1977). A biochemical mechanism for bacterial chemotaxis. *Journal of Theoretical Biology*, *68*, 491-500.
141. Berg H. C. (2004). *E. coli in Motion*. Series: *Biological and Medical Physics, Biomedical Engineering XI*. New York: Aip Press.
142. Fiorillo C. D., Tobler P. N., & Schultz W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*, 1898-902.
143. Niv Y., Duff M. O., & Dayan P. (2005). Dopamine, uncertainty and TD learning. *Behav Brain Funct*, *1*, 6.
144. Fletcher P. C. & Frith C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci*, *10*, 48-58.
145. Phillips W. A. & Silverstein S. M. (2003). Convergence of biological and psychological perspectives on cognitive coordination in schizophrenia. *Behav Brain Sci*, *26*, 65-82.
146. Friston K. & Kiebel S. (2009). Cortical circuits for perceptual inference. *Neural Netw*, *22*, 1093-104.