ORIGINAL PAPER

Biological Cybernetics

Action and behavior: a free-energy formulation

Karl J. Friston · Jean Daunizeau · James Kilner · Stefan J. Kiebel

Abstract We have previously tried to explain perceptual inference and learning under a free-energy principle that pursues Helmholtz's agenda to understand the brain in terms of energy minimization. It is fairly easy to show that making inferences about the causes of sensory data can be cast as the minimization of a free-energy bound on the likelihood of sensory inputs, given an internal model of how they were caused. In this article, we consider what would happen if the data themselves were sampled to minimize this bound. It transpires that the ensuing active sampling or inference is mandated by ergodic arguments based on the very existence of adaptive agents. Furthermore, it accounts for many aspects of motor behavior; from retinal stabilization to goal-seeking. In particular, it suggests that motor control can be understood as fulfilling prior expectations about proprioceptive sensations. This formulation can explain why adaptive behavior emerges in biological agents and suggests a simple alternative to optimal control theory. We illustrate these points using simulations of oculomotor control and then apply to same principles to cued and goal-directed movements. In short, the free-energy formulation may provide an alternative perspective on the motor control that places it in an intimate relationship with perception.

The free-energy principle is an attempt to explain the structure and function of the brain, starting from the fact that we exist: This fact places constraints on our interactions with the world, which have been studied for years in evolutionary biology and systems theory. However, recent advances in statistical physics and machine learning point to a simple scheme that enables biological systems to comply with these constraints. If one looks at the brain as implementing this scheme (minimizing a free-energy bound on disorder), then many aspects of its anatomy and physiology start to make sense. In this article, we show that free-energy can be reduced by selectively sampling sensory inputs. This leads to adaptive responses and provides a new view of how movement control might work in the brain. The main conclusion is that we only need to have expectations about the sensory consequences of moving in order to elicit movement. This means we that can replace the notion of *desired* movements with *expected* movements and understand action in terms of perceptual expectations.

K. J. Friston (✉) · J. Daunizeau · J. Kilner · S. J. Kiebel
The Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London, WC1N 3BG, UK
e-mail: k.friston@fil.ion.ucl.ac.uk

List of symbols

$\mathbf{\Psi} \supseteq \{\tilde{\mathbf{x}}, \tilde{\mathbf{v}}, \boldsymbol{\theta}, \boldsymbol{\gamma}\}$, $\Psi \supseteq \{\tilde{x}, \tilde{v}, \theta, \gamma\}$	Unknown causes of sensory input; variables in bold denote true values and those in italics denote variables assumed by the agent or model
$\tilde{x}(t) = [x, x', x'', \ldots]^T$, $\dot{\tilde{x}}(t) = f(\tilde{x}, \tilde{v}, \theta) + \tilde{w}$	Generalised hidden-states that act on an agent. These are time-varying quantities that include all high-order temporal derivatives; they represent a point in generalised coordinates of motion that encodes a path or trajectory
$\tilde{v}(t) = [v, v', v'', \ldots]^T$	Generalised forces or causal states that act on hidden states
$\tilde{s}(t) = g(\tilde{x}, \tilde{v}, \theta) + \tilde{z}$	Generalised sensory states caused by hidden states
$\theta \supseteq \{\theta_1, \theta_2, \ldots\}$	Parameters of the equations of motion and sensory mapping

$\gamma \supseteq \{\gamma^s, \gamma^x, \gamma^v\}$	Parameters of the precision of random fluctuations $\Pi(\gamma^i) : i \in s, x, v$
$\tilde{w}(t) = [w, w', w'', \ldots]^T$	Generalised random fluctuations of the motion of hidden states
$\tilde{z}(t) = [z, z', z'', \ldots]^T$	Generalised random fluctuations of sensory states
$\tilde{n}(t) = [n, n', n'', \ldots]^T$	Generalised random fluctuations of causal states
$\Pi^i := \Pi(\gamma^i) = \Sigma(\gamma^i)^{-1}$	Precisions or inverse covariances of generalised random fluctuations
$\mathbf{g}(\tilde{\mathbf{x}}, \tilde{\mathbf{v}}, \boldsymbol{\theta}), \mathbf{f}(\tilde{\mathbf{x}}, \tilde{\mathbf{v}}, \tilde{a}, \boldsymbol{\theta})$	Sensory mapping and equations of motion generating sensory states
$g(\tilde{x}, \tilde{v}, \theta), f(\tilde{x}, \tilde{v}, \theta)$	Sensory mapping and equations of motion modeling sensory states
$a(t)$	Policy: a scalar function of generalised sensory and internal states
$p(\tilde{\mathbf{x}}|m), p(\tilde{s}|m)$	Ensemble densities; the density of the hidden and sensory states of agents at equilibrium with their environment.
$D(q\|p) = \langle \ln(q/p) \rangle_q$	Kullback-Leibler divergence or cross-entropy between two densities
$\langle \rangle_q$	Expectation or mean of under the density q
m	Model or agent; entailing the form of a generative model
$H(X) = \langle \ln p(\tilde{\mathbf{x}}|m) \rangle_p$ $H(S) = \langle \ln p(\tilde{s}|m) \rangle_p$	Entropy of generalised hidden and sensory states
$-\ln p(\tilde{s}|m)$	Surprise or self-information of generalised sensory states
$F(\tilde{s}, \mu) \geq -\ln p(\tilde{s}|m)$	Free-energy bound on surprise
$q(\Psi|\mu)$	Recognition density on causes Ψ with sufficient statistics μ
$\mu = \{\tilde{\mu}(t), \mu_\theta, \mu_\gamma\}$ $\tilde{\mu} = \{\tilde{\mu}_x, \tilde{\mu}_v\}$	Conditional or posterior expectation of the causes Ψ; these are the sufficient statistics of the Gaussian recognition density
$\tilde{\eta}(t) = [\eta, \eta', \eta'', \ldots]^T$	Prior expectation of generalised causal states

$\xi_i = \Pi_i \tilde{\varepsilon}_i : i \in s, x, v$	Precision-weighted generalised prediction errors
$\tilde{\varepsilon} = \begin{bmatrix} \tilde{\varepsilon}_s = \tilde{s} - g(\mu) \\ \tilde{\varepsilon}_x = D\tilde{\mu}_x - f(\mu) \\ \tilde{\varepsilon}_v = \tilde{\mu}_v - \tilde{\eta} \end{bmatrix}$	Generalised prediction error on sensory states, the motion of hidden states and forces or causal states.

1 Introduction

This article looks at motor control from the point of view of perception; namely, the fitting or inversion of internal models of sensory data by the brain. Critically, the nature of this inversion lends itself to a relatively simple neural network implementation that shares many formal similarities with real cortical hierarchies in the brain. The idea that the brain uses hierarchical inference has been established for years (Mumford 1992; Rao and Ballard 1998; Friston 2005; Friston et al. 2006) and provides a nice explanation for the hierarchical organization of cortical systems. Critically, hierarchical inference can be formulated as a minimization of free-energy; where free-energy bounds the surprise inherent in sensory data, under a model of how those data were caused. This leads to the free-energy principle, which says that everything in the brain should change to minimize free-energy. We will see below that free-energy can be minimized by changing perceptual representations so that they approximate a posterior or conditional density on the causes of sensations. In short, the free-energy principle entails the Bayesian brain hypothesis (Knill and Pouget 2004; Ballard et al. 1983; Dayan et al. 1995; Lee and Mumford 2003; Rao and Ballard 1998; Friston 2005; Friston and Stephan 2007). However, the free-energy principle goes further than this. It suggests that our actions should also minimize free-energy (Friston et al. 2006): We are open systems in exchange with the environment; the environment acts on us to produce sensory impressions, and we act on the environment to change its states. This exchange rests upon sensory and effector organs (like photoreceptors and oculomotor muscles). If we change the environment or our relationship to it, then sensory input changes. Therefore, action can reduce free-energy by changing the sensory input predicted, while perception reduces free-energy by changing predictions. In this article, we focus in the implications of suppressing free-energy through action or behavior.

Traditionally, the optimization of behavior is formulated as maximizing value or expected reward (Rescorla and Wagner 1972; Sutton and Barto 1981). This theme is seen in cognitive psychology, in reinforcement learning models (Rescorla and Wagner 1972); in computational neuroscience and machine-learning as variants of dynamic programming, such as temporal difference learning (Sutton and Barto 1981;

Watkins and Dayan 1992; Friston et al. 1994; Daw and Doya 2006), and in behavioral economics as expected utility theory (Camerer 2003). In computational motor control (Wolpert and Miall 1996; Todorov and Jordan 2002; Todorov 2006; Shadmehr and Krakauer 2008), it appears in the form of optimal control theory. In all these treatments, the problem of optimizing behavior is reduced to optimizing value (or, conversely, minimizing expected loss or cost). Effectively, this prescribes an optimal control policy in terms of the value that would be expected by pursuing that policy.

Our studies suggest that maximizing value may represent a slight misdirection in explaining adaptive behavior, because the same behaviors emerge in agents that minimize free-energy (Friston et al. 2009). In brief, the minimization of free-energy provides a principled basis for understanding both action and perception, which replaces the optimal polices of control theory with prior expectations about the trajectory of an agent's states. In Sect. 2, we review the free-energy principle and active inference. In Sect. 3, we show how active inference can be used to model reflexive and intentional behavior. This section deals with visual and proprioceptive models to demonstrate the key role of prior expectations in prescribing movement. In Sect. 4, we consider the integration of visual and proprioceptive signals in finessing the control of cued reaching movements. Section 5 addresses how these prior expectations could be learned and illustrates the acquisition of goal-directed movements using the mountain-car problem (Sutton 1996; see also Friston et al. 2009). Section 6 revisits the learning of priors to prescribe autonomous behavior. We conclude by discussing the relationship between active inference and conventional treatments of computational motor control.

2 The free-energy principle

In this section, we try to establish the basic motivation for minimizing free-energy. This section rehearses material that we have used previously to understand perception. It is presented here with a special focus how action maintains a stationary relationship with the environment; and is developed more formally than in previous descriptions (e.g., Friston and Stephan 2007). The arguments for how perception decreases free-energy can be found in the neurobiological (Friston et al. 2006; Friston 2008) and technical (Friston et al. 2008) literature. These arguments are reviewed briefly but only to a depth that is sufficient to understand the simulations in subsequent sections.

What is free-energy? In statistics and machine learning, free-energy is an information theory quantity that bounds the evidence for a model of data (Hinton and von Camp 1993; MacKay 1995; Neal and Hinton 1998). Here, the data are sensory inputs, and the model is encoded by the brain. More precisely, free-energy is greater than the surprise (negative

log-probability) of some data, given a model of how those data were generated. In fact, under simplifying assumptions (see below), it is just the amount of prediction error. It is called free-energy because of formal similarities with thermodynamic free-energy in statistical physics; where energies are just negative log-probabilities (surprise) and free-energy is a bound on surprise. In what follows, we describe the nature of free-energy, and show why it is minimized by adaptive agents.

We start with the premise that adaptive agents or phenotypes must occupy a limited repertoire of physical states. For a phenotype to exist, it must possess defining characteristics or traits; both in terms of its morphology and exchange with the environment. These traits essentially limit the agent to a bounded region in the space of all states it could be in. Once outside these bounds, it ceases to possess that trait (cf., a fish out of water). This speaks to self-organized autopoietic interactions with the world that ensure these bounds are respected (cf., Maturana and Varela 1972). Later, we formalize this notion in terms of the entropy or average surprise associated with a probability distribution on the states an agent experiences. The basic idea is that adaptive agents must occupy a compact and bounded part of state–space and, therefore, avoid surprising states (cf., a fish out of water—sic). In terms of dynamical system theory, this set of states is a random attractor (Crauel and Flandoli 1994). Given this defining attribute of adaptive agents, we will look at how agents might minimize surprise and then consider what this means, in terms of their action and perception.

The free-energy principle rests on an ensemble density $p(\tilde{\mathbf{x}}|m)$ on generalized states, $\tilde{\mathbf{x}}(t) = [\mathbf{x}, \mathbf{x}', \mathbf{x}'', \ldots]^T$, which affect an agent, m. Generalized states cover position, velocity, acceleration, jerk, and so on (Friston 2008). This means that states include the position or configuration of the agent, its motion, and all influences acting on the agent: i.e., physical forces like gravity; thermodynamic states like ambient temperature, or physiological states such as hypoglycemia. Strictly speaking; the dimensionality of generalized states is infinite because the generalized motion of each state exists to infinite order. However, in practice one can ignore high-order temporal derivatives because their precision vanishes and they contain no useful information (i.e., their dispersion gets very large; see Friston 2008; Friston et al. 2008 for details). In the simulations below, we only used generalized states up to sixth order. In what follows, $\mathbf{x}(t)$ refers to a state vector and $\tilde{\mathbf{x}}(t)$ denotes the corresponding generalized state (i.e., the state and its generalized motion). Note that the density $p(\tilde{\mathbf{x}}|m)$ is conditioned on the agent or model. We will see later that the model entails formal constraints on the motion of an agent's states (i.e., its state-transitions or *policy*). This means the ensemble density is specific to each class of agent.

The ensemble density can be regarded as the probability of finding an agent in a particular state, when observed

on multiple occasions or, equivalently, the density of a large ensemble of agents at equilibrium with their environment. Critically, for an agent to exist, the ensemble density should have low entropy. This ensures that agents occupy a limited repertoire of states because a density with low entropy concentrates its mass in a small subset of state–space (i.e., its attractor). This places an important constraint on the states sampled by an agent; it means agents must somehow counter the dispersive effects of random forces, which increase entropy. This increase is a consequence of the fluctuation theorem (Evans 2003), which generalizes the second law of thermodynamics and says that the probability of entropy decreasing vanishes exponentially the longer a system is observed. Adaptive agents resist the implicit dispersion of their states through adaptive behaviors (like eating) that maintain their milieu within physiological bounds. In summary, biological agents must resist the natural tendency to disorder or increasing entropy; but how do they do this?

2.1 Active agents

At this point, we introduce the notion of active agents (Schweitzer 2003) that sense some states (with sensory organs) and can change others (with effector organs). We can describe this exchange with the environment in terms of sensory states $\tilde{s}(t)$ like the activity of photoreceptors, internal states $\tilde{\mu}(t)$ such as neuronal activity, and action $a(t)$ encoding peripheral motor control signals. Because agents only sense external states of the world through a probabilistic mapping, they are hidden from the agent. We will, therefore, call $\tilde{\mathbf{x}} \in X$ hidden states and assume that they evolve according to some complicated equations of motion, $\dot{\tilde{\mathbf{x}}} = \mathbf{f}(\tilde{\mathbf{x}}, a, \boldsymbol{\theta}) + \tilde{\mathbf{w}}$, where $\mathbf{w}$ are random fluctuations. We will see examples of this equation later, where $\tilde{\mathbf{x}}(t)$ represents the angular position and speed of a joint, and $\boldsymbol{\theta}$ encodes the elasticity and viscosity of muscles. We will use an unusual convention in which true hidden quantities are presented in bold and others in italics.

Control (e.g., of saccadic eye-movements) is represented by making the equations of motion a function of action. We will see below that action depends vicariously on sensory states and, therefore, constitutes a *policy* (from the point of view of optimum control theory). It is this policy that resists the increasing entropy of hidden states, because the only way that the agent can change hidden states is through action. We assume that agents do this by minimizing entropy until the movement of states toward the agent's attractor is balanced by dispersion due to random fluctuations. Sensation (e.g., proprioception) can be described with a probabilistic mapping $\tilde{s} = \mathbf{g}(\tilde{\mathbf{x}}, \boldsymbol{\theta}) + \tilde{\mathbf{z}}$ from hidden to sensory states $\tilde{s} \in S$, where $\mathbf{z}$ represents sensory noise. Later, we will see examples were $\mathbf{g}(\tilde{\mathbf{x}}, \boldsymbol{\theta})$ maps from the trajectory of an object in the visual field to the activity of photoreceptors. The equa-

tions of motion and sensory mapping are parameterized by (generally non-overlapping) subsets of some time-invariant parameters $\boldsymbol{\theta}$.

Under a sensory mapping, the entropy of the hidden states is bounded by the sensory entropy and a sensory mapping term, with equality in the absence of sensory noise (see Appendix 1).

$$
\begin{aligned}
H(X|m) &\leq H(S|m) - \int p(\tilde{\mathbf{x}}|m) \ln |\partial_{\tilde{\mathbf{x}}}\mathbf{g}| d\tilde{\mathbf{x}} \\
\tilde{\mathbf{s}} &= \mathbf{g}(\tilde{\mathbf{x}}, \boldsymbol{\theta}) + \tilde{\mathbf{z}} \\
\dot{\tilde{\mathbf{x}}} &= \mathbf{f}(\tilde{\mathbf{x}}, a, \boldsymbol{\theta}) + \tilde{\mathbf{w}}
\end{aligned}
\tag{1}
$$

where the respective entropies are

$$
\begin{aligned}
H(S|m) &= - \int p(\tilde{s}|m) \ln p(\tilde{s}|m) d\tilde{s} \\
H(X|m) &= - \int p(\tilde{\mathbf{x}}|m) \ln p(\tilde{\mathbf{x}}|m) d\tilde{\mathbf{x}}
\end{aligned}
\tag{2}
$$

The second term depends on the collective sensitivity $|\partial_{\tilde{\mathbf{x}}}\mathbf{g}| := \det(\partial_{\tilde{\mathbf{x}}}\mathbf{g})$ of sensory inputs to changes in hidden states, where $\partial_{\tilde{\mathbf{x}}}\mathbf{g}$ is the derivative of the sensory mapping with respect to the hidden states. For simplicity, we will assume that this sensitivity is uniformly high over the ranges of states considered. This appeals to the principle of maximum information transfer (Linsker 1990), which has been very useful in understanding the emergence of things like receptive fields (e.g., Olshausen and Field 1996). Under this simplifying assumption, the second term is constant, and it is sufficient to minimize sensory entropy to minimize the entropy of hidden states. This is important because the agent has access to sensory states but not hidden states. Note that agents cannot minimize sensory entropy by switching off sensory channels; they can only act by changing hidden states, not the sensory mapping per se.

Crucially, because the ensemble density on sensory states is at equilibrium, it can be interpreted as the proportion of time a single agent entertains these states (the sojourn time). This ergodic argument (Anosov 2001) means that the sensory entropy is the long-term average of the surprise $-\ln p(\tilde{s}|m)$ experienced by a particular agent:

$$
H(S|m) = \lim_{T \to \infty} -\frac{1}{T} \int_0^T dt \ln p(\tilde{s}(t)|m)
\tag{3}
$$

This means that in order to minimize sensory entropy, one needs to minimize surprise at all times (i.e., for all successive sensory samples).

At this point, most (astute) people say "but that means I should retire to a dark room and cover my ears." However, is not quite that simple: Recall that entropy is average surprise and surprise depends on the agent. This means that it might be surprising to find an agent in a dark room.

It helps here to imagine that each class of agent (e.g., species) has equations of motion that induce a random attractor in state–space (more formally, a compact random invariant set that attracts all trajectories; Crauel and Flandoli 1994). Entropy decreases with the volume or measure of this set. However, it is important not to confuse volume with shape or topological complexity. The attractor could be very complicated and space-filling but occupy a small volume (e.g., the states encountered while driving a car with a small margin of error). In this view, minimizing surprise ensures that states evolve according to the expected flow, $\mathbf{f}(\tilde{\mathbf{x}}, \tilde{a}, \boldsymbol{\theta})$ and remain in the attractor (e.g., within physiological bounds). In short, the ensemble and sensory entropy are lower bounded by a random attractor entailed by an agent's class. Agents try to attain this bound by minimizing surprise, while learning and natural selection may reduce the bound by optimizing the attractor per se. This means the notion of a "dark room" agent is a category error; in the sense that "dark room" or any fixed-point attractors are not invariant under real world dynamics and, therefore, cannot exist (e.g., at some point, you will get thirsty). Perhaps the most potent examples of surprise-reducing dynamics are found in physiology, which can be understood largely in terms of homeostatic mechanisms that avoid surprising trajectories of the internal milieu:

> The fixity of the milieu supposes a perfection of the organism such that the external variations are at each instant compensated for and equilibrated … All of the vital mechanisms, however varied they may be, have always one goal, to maintain the uniformity of the conditions of life in the internal environment … The stability of the internal environment is the condition for the free and independent life (Bernard 1974).

This is closely related to perspectives on motor control and sequence learning that "minimize deviations from the desired state, that is, to minimize disturbances of the homeostasis of the feedback loop." See Wörgötter and Porr (2005) for a fuller discussion. In summary, avoiding surprise is fundamental for survival and speaks to the basic need of organisms to maintain equilibrium within their environment. However, there is a problem:

2.2 Free-energy, action and perception

The problem faced by real agents is that they cannot quantify surprise, because this entails marginalizing over the unknown or hidden causes $\Psi \supset \{\tilde{\mathbf{x}}, \boldsymbol{\theta}\}$ of sensory input that include time-varying hidden states and time-invariant parameters

$$- \ln p(\tilde{s}|m) = - \ln \int p(\tilde{s}, \Psi|m) \mathrm{d}\Psi \qquad (4)$$

However, there is an alternative and elegant solution to minimizing surprise, which comes from theoretical physics (Feynman 1972) and machine learning (Hinton and von Camp 1993; MacKay 1995). This involves minimizing a free-energy bound on surprise that can be evaluated. Minimizing this bound implicitly minimizes surprise because the bound is always greater than surprise. The bound is induced by a recognition density $q(\Psi|\mu)$, which is parameterized by its sufficient statistics $\mu \supset \tilde{\mu}(t)$ (e.g., mean or covariance). These sufficient statistics are encoded by the internal states of the agent (e.g., by neuronal activity and connection strengths). The recognition density is a slightly mysterious construct because it is an *arbitrary* probability density specified by the internal states of the agent. Its role is to induce free-energy, which is a function of the internal states and sensory inputs. We will see below that when this density is optimized to minimize free-energy, it becomes the conditional density on the causes of sensory data. In Bayesian inference, this is known as the recognition density. In what follows, we summarize the key ideas behind a large body of study in statistics and machine learning referred to as *ensemble learning* or *variational Bayes*.

The free-energy bound is constructed by simply augmenting surprise with a non-negative divergence between the recognition density and conditional density $p(\Psi|\tilde{s}, m)$. This is described Appendix 2. The resulting expression for free-energy can be expressed in three ways (with the use of Bayes rule and simple rearrangements):

- Energy minus entropy
- Divergence plus surprise
- Complexity minus accuracy

Mathematically, these correspond to:

$$\begin{aligned} F &= - \langle \ln p(\tilde{s}, \Psi|m) \rangle_q + \langle \ln q(\Psi|\mu) \rangle_q \\ &= D(q(\Psi|\mu)||p(\Psi|\tilde{s}, m)) - \ln p(\tilde{s}|m) \qquad (5) \\ &= D(q(\Psi|\mu)||p(\Psi|m)) - \langle \ln p(\tilde{s}|\Psi, m) \rangle_q \end{aligned}$$

Here, $\langle \cdot \rangle_q$ means the expectation or mean under the density q and $D(\cdot||\cdot)$ is the cross-entropy or Kullback–Leibler divergence between two densities. The alternative formulations in Eq. 5 have some important implications: The first shows that free-energy is a scalar function of sensory data and the recognition density at any particular time. This can be evaluated by an agent; provided it has a probabilistic model of the environment. This model is usually expressed as the product of a likelihood and prior, $p(\tilde{s}, \Psi|m) = p(\tilde{s}|\Psi, m)p(\Psi|m)$. We will see examples of this model later (e.g., the likelihood model in Eq. 8).

The second formulation shows that minimizing the free-energy, by changing internal states (i.e., the sufficient statistics of $q(\Psi|\mu)$), reduces the divergence between the recognition and conditional densities. This makes the recognition density an approximate conditional density. This corresponds to Bayesian inference on the causes of sensory

signals and provides a principled account of perception; i.e., the Bayesian brain (Helmholtz 1860/1962; Barlow 1969; Ballard et al. 1983; Mumford 1992; Dayan et al. 1995; Rao and Ballard 1998; Lee and Mumford 2003; Knill and Pouget 2004; Kersten et al. 2004; Friston and Stephan 2007; Deneve 2008). Finally, it shows that free-energy is an upper bound on surprise because the divergence cannot be less than zero: Optimizing the recognition density makes the free-energy a tight bound on surprise; when the recognition and conditional densities coincide, free-energy is exactly surprise and perception is veridical.

The third equality shows that free-energy can be suppressed by action, through its effects on hidden states and the ensuing sensory signals. The key term here is the accuracy term, which is the expected surprise of sensory data under the recognition density (under Gaussian assumptions, this is just the amount of sensory prediction error). This means action will change the motion of sensory states so that they conform to conditional expectations. This minimizes surprise, provided perception makes free-energy a tight bound on surprise. In short, the free-energy principle prescribes an optimum perceptual representation and policy

learning or *sequential design* tries to harvest data to make recognition more precise using various criteria, based on how data inform model parameter estimates; "these criteria depend on the assumption that the hypothesis space is correct, which may prove to be their main weakness" (Mumford 1992). Under the free-energy formulation, this weakness becomes the strength; if the environment delivers data that are consistent with the agent's model or hypothesis space, then the agent must be at equilibrium with its environment.

2.3 Summary

In summary, the free-energy principle requires the internal states of an agent and its action to suppress free-energy. This corresponds to optimizing a probabilistic model of how sensations are caused, so that the resulting predictions can guide active sampling of sensory data. The requisite interplay between action and perception (i.e., active inference) engenders a policy that ensures the hidden states that act on the agent have low entropy. Put simply, if you search out things you expect, you will avoid surprises. This recapitulates the notion that "perception and behavior can interact synergistically, via the environment"

$$\mu(t)^* = \arg\min_{\mu} F(\tilde{s}(t), \mu(t)) \qquad \partial_\mu F(\tilde{s}, \mu) = 0 \Rightarrow \delta_\mu \int_0^T \mathrm{d}t\, F(\tilde{s}, \mu) = 0$$
$$a(t)^* = \arg\min_{a} F(\tilde{s}(t), \mu(t)) \quad \Rightarrow \quad \partial_a F(\tilde{s}, \mu) = 0 \Rightarrow \delta_a \int_0^T \mathrm{d}t\, F(\tilde{s}, \mu) = 0 \Bigg|_{\mu=\mu^*} \tag{6}$$

This policy reduces to sampling input that is expected under the recognition density (i.e., sampling selectively what one expects to experience). In other words, agents must necessarily (if implicitly) make inferences about the causes of their sensory signals and sample signals that are consistent with those inferences. Technically, minimizing free-energy ensures the variation of its path integral, with respect to small variations in action or internal representations, is zero (see Eq. 6). This time or path-integral is called Action (not to be confused with action). Under ergodic assumptions, Action is an upper bound on sensory entropy (see Eq. 3), which means optimal action and perception minimize a bound on sensory surprise and, implicitly, entropy.

The free-energy principle suggests that we should sample the world to ensure our predictions become a self-fulfilling prophecy. In this view, perception is enslaved by action to provide veridical predictions (more formally, to make the free-energy a tight bound on surprise) that guide active sampling of the sensorium. This active sampling should not be confused with *active learning* (Mumford 1992). Active

to optimize behavior (Verschure et al. 2003). Indeed, Distributed Adaptive Control (Verschure and Voegtlin 1998) relates closely to the free-energy formulation, because it provides an integrated solution to the acquisition of probabilistic models and policies. Active inference is an example of *self-referenced* learning (Maturana and Varela 1972; Porr and Wörgötter 2003) in which "the actions of the learner influence its own learning without any valuation process" (Porr and Wörgötter 2003).

In machine learning and statistics, the free-energy is used as a bound on the log-evidence for a model. This means, given some data, one can optimize the model to maximize its evidence; i.e., find the model that minimizes the surprise of some fixed data. The free-energy principle turns this optimization around and considers that the model (agent) is fixed and the aim is to minimize surprise by re-sampling the data. This minimization is mandated by the fact that average surprise is entropy and agents must have a low entropy ensemble density. This argument rests on noting that the ensemble density is formally identical to the marginal likelihood in statistics. In the analysis of time-series data, one optimizes the model

with respect to the path integral of free-energy over time (see Friston 2008): Minimizing free-energy at each point in time provides a stationary solution for Action (by the fundamental lemma of variational calculus): this is basically Hamilton's principle of stationary Action. All that we are doing here is noting that Action is a bound on the path-integral of surprise and, under ergodic assumptions, entropy. This means finding a minimum for Action implicitly finds a minimum for sensory entropy. It is interesting to note that a tendency to disorder or the second law of thermodynamics (which applies only to closed systems) can be resisted by appealing to the more general tendency of (open) systems to reduce their free-energy (Gontar 2000; Evans 2003). A related treatment of self-organization in non-equilibrium systems can be found in synergetics; where "patterns become functional because they consume in a most efficient manner the gradients which cause their evolution" (Tschacher and Haken 2007). Here, these gradients might be regarded as surprise. This section has tried to establish the theoretical motivation for free-energy minimization. We now turn to how this might be implemented in the brain.

3 Active inference

In this section, we unpack the theory above to establish the nature of processes underlying action and perception. In the previous section, we saw that optimal action and perception could be cast as minimizing free-energy and that free-energy is defined in relation to a probabilistic generative model of the world. In what follows, we will examine particular forms of this model, which enable free-energy to be expressed in terms of prediction errors (the mismatch between observed sensory signals and those predicted by the model). This leads to a formulation of action and perception as the suppression of prediction errors (or surprise). By assuming the minimization in Eq. 6 is implemented using gradient decent, we can then write down differential equations that prescribe the dynamics of action and recognition as functions of prediction error.

In order to see how active inference works, we must first define an environment and the agent's model of it. We will assume that both can be cast as dynamical systems with additive random effects. For the environment, we have equations of motion and a sensory mapping that describe the processes generating sensory data

$$
\begin{aligned}
\tilde{s} &= \mathbf{g}(\tilde{\mathbf{x}}, \tilde{\mathbf{v}}, \boldsymbol{\theta}) + \tilde{\mathbf{z}} \\
\dot{\tilde{\mathbf{x}}} &= \mathbf{f}(\tilde{\mathbf{x}}, \tilde{\mathbf{v}}, a, \boldsymbol{\theta}) + \tilde{\mathbf{w}}
\end{aligned}
\tag{7}
$$

This describes how sensory data are actually generated, while the agent assumes that these data are generated by the following system:

$$
\begin{aligned}
\tilde{s} &= g(\tilde{x}, \tilde{v}, \theta) + \tilde{z} \\
\dot{\tilde{x}} &= f(\tilde{x}, \tilde{v}, \theta) + \tilde{w} \\
\tilde{v} &= \tilde{\eta} + \tilde{n}
\end{aligned}
\tag{8}
$$

Equation 8 is just a way of expressing a generative model in terms of equations that have a formal homology with the true generative process (Eq. 7). In other words, Eq. 8 represents a parameterized model of sensory input, which becomes a probabilistic model, if the agent makes some assumptions about the probability distribution of the random fluctuations (see below). Note that we have partitioned the hidden states into those that are autonomous $\tilde{\mathbf{v}}(t)$ and those that are not $\tilde{\mathbf{x}}(t)$ (cf. Eq. 1). The autonomous hidden states (e.g., exogenous forces) allow us to model exogenous perturbations of the agent's state. In previous treatments (e.g., Friston et al. 2006; Friston 2005), we have called these *casual states*. Furthermore, these forces are modeled as mixture of some deterministic prior $\tilde{\eta}(t)$ and random fluctuations $\tilde{n}(t)$.

Equations 7 and 8 pertain to the same sensory states and play a similar role. However, the first is the true stochastic process generating states, and the second is a probabilistic model of this process. The equations of motion and sensory mapping may or may not have the same form, which means the agent's model of the world may be different from the real process generating sensory data. Furthermore, we allow for a difference between the true unknowns $\boldsymbol{\Psi} \supset \{\tilde{\mathbf{x}}, \tilde{\mathbf{v}}, \boldsymbol{\theta}\}$ and those assumed by the model $\Psi \supset \{\tilde{x}, \tilde{v}, \theta\}$.

Gaussian assumptions about the random fluctuations $\{\tilde{z}, \tilde{w}, \tilde{n}\}$ in Eq. 8 furnish three things; a likelihood model, $p(\tilde{s}|\Psi) = N(g, \Sigma(\gamma^s))$, empirical priors on the generalized motion $p(\tilde{x}'|\tilde{v}, \Psi) = N(f, \Sigma(\gamma^x))$, and priors on the forces, $p(\tilde{v}) = N(\tilde{\eta}, \Sigma(\gamma^v))$. Here, the inverse variances or precisions $\Psi \supset \gamma^i : i \in s, x, v$ determine the covariance $\Sigma(\gamma^i)$ of the generalized fluctuations. Note that the equations of motion for the true states depend on action, whereas the generative model has no notion of action; it just produces predictions that action tries to fulfill. However, the generative model contains a prior on exogenous forces, which, as we will see later, can be used to model the effects of action. Equation 8 is a simple hierarchical dynamic model in generalized coordinates of motion. This sort of model is considered in detail in Friston (2008). Recall that the equations generating data (Eq. 7) and those of the generative model (Eq. 8) do not have to be the same; it is this discrepancy that action tries to cancel.

Given a probabilistic generative model, the free-energy can now be minimized: This minimization obliges the agent to infer the states of the world and learn the unknown parameters responsible for its motion by optimizing the sufficient statistics of its recognition density. This is perceptual inference and learning and can be implemented in a biologically plausible fashion as described in Friston (2008). In brief, this scheme assumes that a mean-field

approximation: $q(\Psi|\mu) = q(\tilde{x}, \tilde{v}|\tilde{\mu})q(\theta|\mu^\theta)q(\gamma|\mu^\gamma)$ with Gaussian marginals, whose sufficient statistics are expectations and covariances. Under this Gaussian or Laplace approximation, it is sufficient to optimize the expectations because the covariances can be derived as a function of the expectations. From now on, we assume that the sufficient statistics are the conditional expectations of states, parameters, and precisions, $\mu = \{\tilde{\mu}, \mu_\theta, \mu_\gamma\}$, where the conditional expectations of states $\tilde{\mu}(t) = \{\tilde{\mu}_x, \tilde{\mu}_v\}$ are in generalized coordinates of motion. Using these (mean-field and Laplace) approximations, we can formulate Eq. 6 as a gradient descent that describes perceptual inference, learning, and action (see Friston et al. 2008; Friston 2008 for details):

$$\dot{\tilde{\mu}}_x = D\tilde{\mu}_x - \frac{\partial F}{\partial \tilde{\mu}_x} \approx D\tilde{\mu}_x - \frac{\partial \tilde{\varepsilon}^T}{\partial \tilde{\mu}_x}\xi$$
$$\dot{\tilde{\mu}}_v = D\tilde{\mu}_v - \frac{\partial F}{\partial \tilde{\mu}_v} \approx D\tilde{\mu}_v - \frac{\partial \tilde{\varepsilon}^T}{\partial \tilde{\mu}_v}\xi$$

$$\ddot{\mu}_\theta = -\frac{\partial F}{\partial \mu_\theta} \approx -\frac{\partial \tilde{\varepsilon}^T}{\partial \mu_\theta}\xi \qquad (9)$$
$$\dot{a} = -\frac{\partial F}{\partial a} = -\frac{\partial \tilde{\varepsilon}_s^T}{\partial a}\xi_s$$

$$F \approx \frac{1}{2}\tilde{\varepsilon}^T \Pi \tilde{\varepsilon} - \frac{1}{2}\ln|\Pi|$$
$$\xi = \Pi\tilde{\varepsilon}$$

This equation is used in all the simulations of active inference described later. It has a simple form because the free-energy in Eq. 5 reduces (effectively) to a sum of squared errors because of our Gaussian assumptions about the random fluctuations. The approximate equalities in Eq. 9 are due to the fact that we are neglecting mean-field terms (and retain only terms that depend on conditional expectations). The derivative operator D is a block matrix with identity matrices along the first diagonal such that $\tilde{x}' := [x', x'', \dots]^T = D[x, x', \dots]^T = D\tilde{x}$. The first pair of equations prescribes recognition dynamics, in terms of how expected states change over time. The second terms in these equations are free-energy gradients. The first terms reflect the fact that we are working in generalized coordinates; they ensure that $\dot{\tilde{\mu}} = D\tilde{\mu} \Rightarrow \partial_{\tilde{\mu}}F = 0$ when free-energy is minimized and its gradient is zero (i.e., they ensure the motion of the expectation is the expected motion). This scheme can be implemented in a biologically plausible manner by assuming that the activity of one population of neurons encodes the conditional expectations $\tilde{\mu}(t)$ and another population encodes precision-weighed prediction errors

$$\tilde{\varepsilon} = \begin{bmatrix} \tilde{\varepsilon}_s = \tilde{s} - g(\mu) \\ \tilde{\varepsilon}_x = D\tilde{\mu}_x - f(\mu) \\ \tilde{\varepsilon}_v = \tilde{\mu}_v - \tilde{\eta} \end{bmatrix} \qquad (10)$$

The prediction errors $\tilde{\varepsilon}(t)$ are just the difference between sensory samples (resp. expected motion of hidden states) and the predictions afforded by the sensory mapping (resp. equations of motion) evaluated with the expected states and parameters. The matrix $\Pi(\mu_\gamma)$ is the expected precision (inverse covariance) of the random effects and effectively boosts prediction errors in proportion to their precision (see Friston 2008 for more details). Note that under the Laplace assumption, minimizing free-energy corresponds to minimizing prediction error. This sort of scheme is also called *predictive coding* (Rao and Ballard 1998). Figure 2a shows a schematic of the neuronal circuits that might implement predictive coding of this sort. These circuits comprise neuronal state-units or populations (white) that encode sensory input and the conditional expectations of hidden states causing that input. Error-units are shown in red and encode the prediction error for their corresponding state-unit. The dotted lines are connections that are intrinsic to the model, and solid lines represent coupling between action-units (black) and effectors. The red lines convey prediction errors to the state (and action) units, while the black lines pass predictions to the error-units (and effectors). See, Friston (2008), for a fuller description of the implicit message-passing scheme.

The solutions to the third line in Eq. 9 are the optimum parameters that may be encoded in the brain with synaptic efficiency (Friston 2008). These are second-order differential equations because these expectations optimize Action (the path-integral of free-energy), due to the fact we know a priori, that they do not change with time. The resulting changes in synaptic efficacy are formally identical to Hebbian or associative plasticity (under some simplifying assumptions; see Friston 2008 for details). The final equation describes action as a gradient descent on free-energy. The only way action can affect free-energy is through changing the motion of sensory signals. This means action must suppress sensory prediction errors, $\tilde{\varepsilon}_s = \tilde{s}(a) - g(\mu)$. Equation 9 embodies a nice convergence of action and perception: perception tries to suppress prediction error by adjusting expectations to furnish better predictions of signals, while action tries to fulfil these predictions by changing those signals.

3.1 Motor control and prediction errors

Optimal behavior is a powerful guide for understanding motor control. The notion of forward or generative models has been established in this field for decades (Wolpert et al. 1995). Some have even used forward motor models to motivate generative models in perception (Kawato et al. 1993). Usually, forward models in the motor control literature are used to finesse control, given motor commands prescribing a desired movement trajectory (see Sect. 7). The free-energy formulation offers a simpler view; prediction errors are not there to finesse motor signals; *they are the motor signals.*

In other words, there are no top-down motor commands; just prior expectations about proprioceptive input that are fulfilled at the spinal level, given a compliant environment and motor plant (see Fig. 1). In this view, the central nervous system is not divided into motor and sensory systems but is one perceptual inference machine that provides predictions of optimal action, in terms of its expected consequences. This formulation also specifies what movement kinetics should optimize; the only thing that action can affect is the prediction error at the sensory level. This means action can only suppress the weighted sensory prediction error variance, $\tilde{\varepsilon}_s^T \xi_s = \tilde{\varepsilon}_s^T \Pi_s \tilde{\varepsilon}_s$. In generalized coordinates, these include prediction errors on displacement, velocity, acceleration, jerk, etc. This ensures smooth trajectories, because the variance of higher temporal derivatives measures the smoothness of trajectories. The relative weighting of these error variances is controlled by the optimized precision. This mixture of weighted error variance is related formally to phenomenological cost functions proposed to account for movement trajectories (cf. Todorov and Jordan 1998; Nakano et al. 1999; Berret et al. 2008).

3.2 Summary

In this section, we have seen that minimizing free-energy with respect to neuronal activity encoding the recognition density can be formulated as a gradient decent. This leads to recognition dynamics that suppress prediction error, making them formally equivalent to predictive coding. Policies are prescribed in exactly the same way; however, action can only affect prediction error at the sensory level. This suggests that sensory prediction errors generate motor command signals and predicts an intimate anatomical relationship between primary motor and sensory systems. This is because motor commands must emanate from sensory systems responsible for processing the proprioceptive consequences of movement. Empirically, this is seen at cortical, subcortical, and spinal levels. For example, the primary motor and sensory cortex lie next to each other in the central sulcus and are strongly interconnected (Huffman and Krubitzer 2001). At a subcortical level, the superior colliculus provides a site of convergence for sensory information (through direct projections from the retina) and predictions (from visual, parietal, and frontal cortex to the intermediate and deep layers). Neuronal discharges in deep layers, that initiate saccades, define motor-fields that coincide with visual receptive fields in superficial layers (Andersen 1989). Finally, at the spinal level, the ventral (motor) and dorsal horn (proprioceptive) are juxtaposed (see Fig. 1).

In summary, under active inference, perception tries to explain away prediction errors by changing predictions, while action tries to explain them away by changing the signals being predicted. This formulation suggests that motor control may not be the selection of an action but an emergent

property of predictive coding, where proprioceptive prediction errors are quenched by movement. In the next section, we use the recognition dynamics above and see what sorts of behavior emerge under active inference.
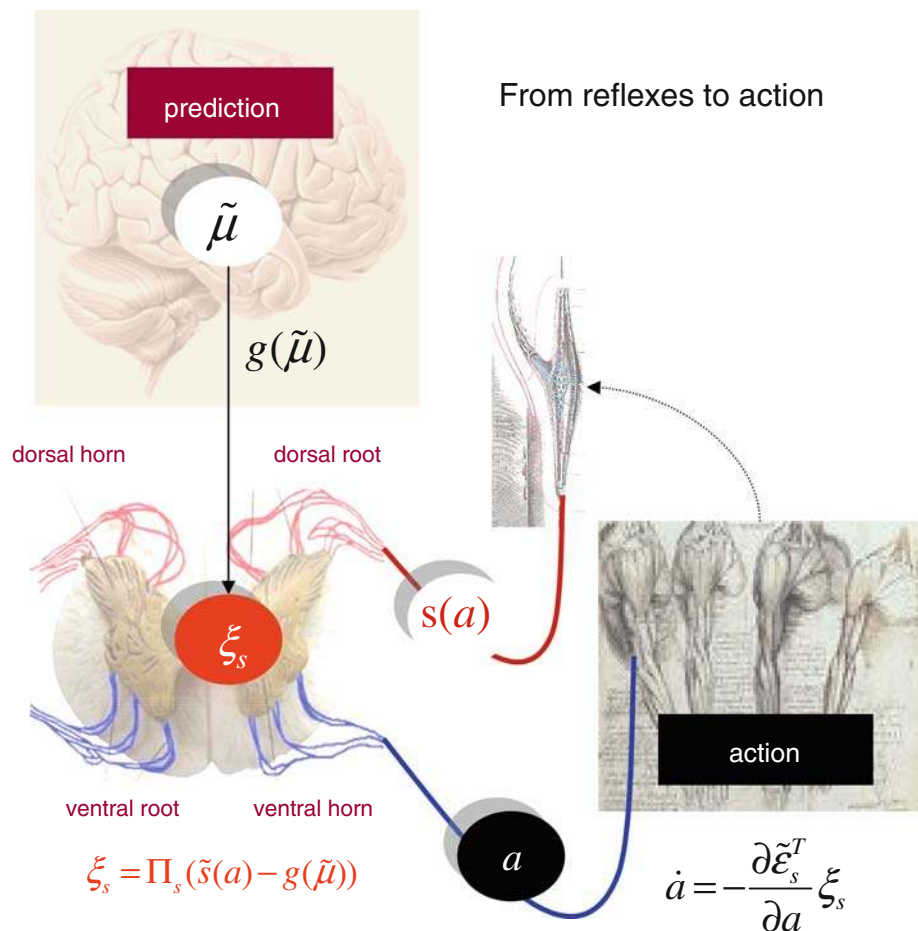
4 Simulations

4.1 Oculomotor control

In this section, we illustrate active inference using a series of simple simulations of oculomotor control. These simulations assume that the agent has already learned the parameters that optimize its inference about sensory states. We use these simulations to focus on the nature of active inference and show how it can lead to phenomena like retinal stabilization, orientating responses, and pursuit movements. We then use the same model to illustrate how active inference can lead to intentional behavior that is prescribed by prior expectations about sensory input.

4.1.1 Retinal stabilization and oculomotor reflexes

Here, we use a linear state–space model and a nonlinear sensory mapping to generate high-dimensional sensory data. We use the same model as a generative model for active inference and explore the consequences of action under perturbation. In this model, a single force perturbs the dynamics of two hidden states to produce a damped transient. These states control the location of object (a Gabor patch) in retinotopic space. For simplicity, these states cause the eye to move (as opposed to the object). Sensory signals are sampled by photoreceptors on a sparse grid rectangular. The equations of motion are linear and parameterized in such a way that the perturbing force causes displacement in one direction. This displacement is then communicated to the other direction, to induce elliptical motion of the visual target. The sensory mapping, on the other hand, is a highly nonlinear function of the hidden states (direction of gaze or stimulus location) because the visual stimulus has a nonlinear contrast profile. We will use this system to ask what would happen under different prior expectations about stimulus movement. This generative process and model are given by Eqs. 7 and 8 where

$$\mathbf{g}(\mathbf{x})_{ij} = \upsilon \left(i - \tfrac{3}{4}\mathbf{x}_1 - \tfrac{N}{2} \right) \upsilon \left(j - \tfrac{3}{4}\mathbf{x}_2 - \tfrac{N}{2} \right)$$

$$\upsilon(x) = \exp\left(-\tfrac{1}{4}x^2\right) \cos\left(\tfrac{2}{3}\pi x\right)$$

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \mathbf{v} + a \\ 0 \end{bmatrix} - \tfrac{1}{8} \begin{bmatrix} 1 & -4 \\ 2 & 1 \end{bmatrix} \mathbf{x} \qquad (11)$$

$$g(x)_{ij} = \mathbf{g}(x)_{ij}$$

$$f(x) = \begin{bmatrix} v \\ 0 \end{bmatrix} - \tfrac{1}{8} \begin{bmatrix} 1 & -4 \\ 2 & 1 \end{bmatrix} x$$

Fig. 1 Schematic detailing the components of the motor reflex arc and how it might minimize sensory prediction error (free-energy) through action. The basic idea here is that the central nervous system provides predictions about the expected (proprioceptive) consequences of movements, which action is enslaved to fulfill. In the context of sensorimotor reflexes, motor signals emitted by ventral efferents of the spinal cord try to suppress prediction errors induced by an unexpected firing of stretch-receptors. However, from the point of view of the motor neurons, this prediction error could also have been mediated by a top-down suppression of predicted stretch-receptor discharge rates. In short, sensory prediction error controls movements and top-down predictions control sensory prediction error

These equations, which specify the nature of the environment and model, are used to simulate environmental dynamics and the agent's responses. This entails integrating the stochastic differential equation 7 (to evaluate the environment's states) and the ordinary differential equation 9 (to evaluate the agent's states). See Appendix 3 for details of the integration scheme. Note that these equations are specified in terms of states. The corresponding equations for generalized states are derived easily using the chain rule (see Friston 2008; Eqs. 1 and 2).

In this example, a force or cause $\mathbf{v}(t)$ excites changes in the hidden states through the equations of motion, $\mathbf{f}(\mathbf{x})$. The hidden states (stimulus location in two dimensions) are then mapped to sensory data and delivered to the agent, using state and observation noise with log-precisions (inverse variances) of four and eight respectively. A log-precision of four means that the standard deviation is $\exp(-4/2) = 0.135$. In all the simulations below, we used relatively smooth noise for both the generative process and model. This smoothness is controlled by the precision of its higher derivatives, which were chosen to give a Gaussian autocorrelation function with a standard deviation of one half of a simulated time bin (usually 1/32 s). In this example, the sensory data are a nonlinear

function of stimulus location, generating $6 \times 6 = 36$ channels of data. These functions can be thought of as receptive fields that map the retinotopic location of the stimulus to the pattern of excitation, $\mathbf{g}(\mathbf{x})_{ij}$ in an $N \times N$ array of sensory receptors (see Fig. 2), where $i, j \in \{1, 2, \ldots N\} : N = 6$.

The agent then infers the hidden states (i.e., stimulus location) by optimizing the sufficient statistics of its recognition density, as described in the previous section (Eq. 9). The only extra thing that we need to consider here is action, $a(t)$, which is part of the generative process but not part of the generative model. The only way that action can change free-energy is by re-sampling sensory data to reduce sensory prediction error. From Eq. 9, we can see that this just depends on sensory prediction error and the changes in sensory signals with action. Appendix 3 provides more details about the ensuing dynamics for the interested reader.

Here, we have deliberately chosen to put action in opposition to the force, so that the agent can oppose perturbations to its hidden states. By manipulating its prior expectations on these perturbations, $v(t)$, we can elicit very different behaviors from the agent. This is because the action is used to explain away prediction error, and prediction error depends upon the agent's prior expectations. For example, if we
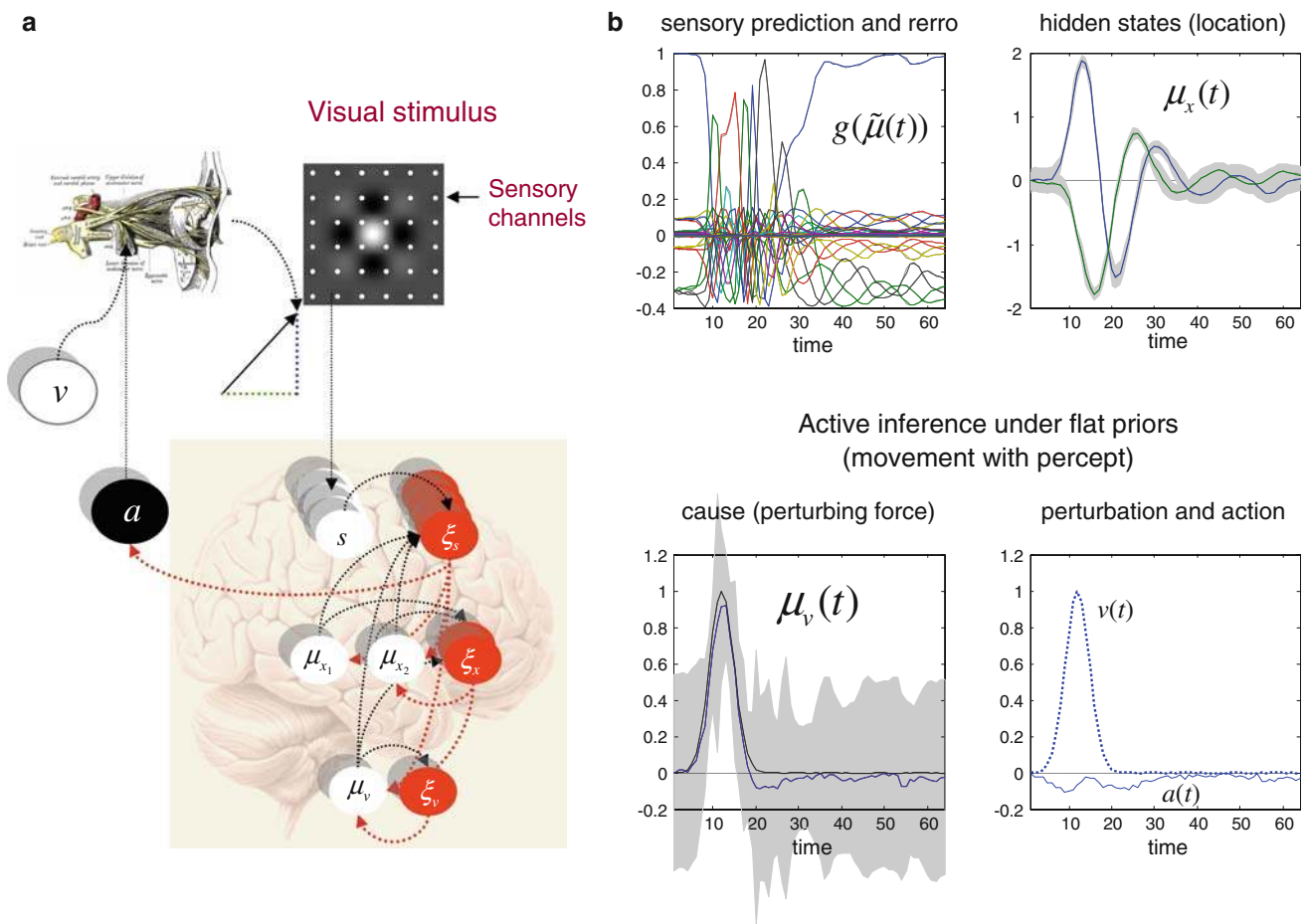
Fig. 2 a Schematic showing the relationship of the generative model (*lower panel*) to the generative process modeling oculomotor sampling of the world. The schematic of the generative model used by the brain comprises state-units or populations (*white*) that encode sensory input and the conditional expectations of hidden states causing that input. Error-units are shown in *red* and encode the prediction error for their corresponding state-unit. The *dotted lines* are connections that are intrinsic to the model, and *solid lines* represent coupling between action-units (*black*) and effectors. The *red lines* convey prediction errors to the state (and action) units, while the *black lines* pass predictions to the error-units (and effectors). Please see Friston et al. (2008) for a fuller description of the scheme implied by these circuit diagrams. In this example, action causes eye movements which change the visual information sampled by an array of photoreceptors (*top panel*). The locations of these receptors are shown by the white dots, superimposed on the Gabor patch-like stimulus. Here, the states of the world are the direction of gaze in two dimensions. **b** *Upper left* These are the predicted sensory states and corresponding prediction error (*red*

lines); they represent the predicted sensory input over the 36 photo-receptor channels induced by motion of the Gabor patch on the *left*. This rather complicated set of transients is the prediction based on a simple elliptical movement of the stimulus described by the expected hidden states. *Upper right* The expected hidden states causing sensory predictions. These can be thought of as vertical and horizontal displacements in retinotopic space. The *colored lines* correspond to conditional expectations, and the *dashed gray lines* to true values (superimposed in this example). The *gray regions* correspond to 90% confidence intervals based on the conditional covariance. *Lower left* The true (*dotted line*) and conditional expectation (*solid line*) of the exogenous cause; a Gaussian bump function of peristimulus time. In this example, the prior was zero with a large prior variance. The conditional expectation conforms roughly to the true perturbation. *Lower right* Action (*solid line*) and true cause (*dotted line*). In this instance, action does not change very much because it is not needed to explain away prediction error. This is because a priori the agent expects perturbations of unspecified amplitude

induce a damped elliptical motion in the target, the agent will sample the stimulus differently, depending upon whether it expects the stimulus to move or not.

We first examined the responses of the agent under flat or uninformative priors on the force. These priors were implemented by making the agent expect large amplitude random fluctuations (with a log-precision of $\ln \mu_\gamma^v = -16$) about the prior mean $\eta(t) = 0$. This simply means that, a priori,

forces can be very large or small and the prior distribution is very wide (i.e. flat). Under these priors, we would expect the force and consequent motion to be perceived, because there are no constraints on whether perturbations should be large or small. Consequently, we would expect the perceived location to change as it would in the absence of active sampling. The results in Fig. 2 show this to be the case. Here, we perturbed the system with a Gaussian bump function of peristimulus
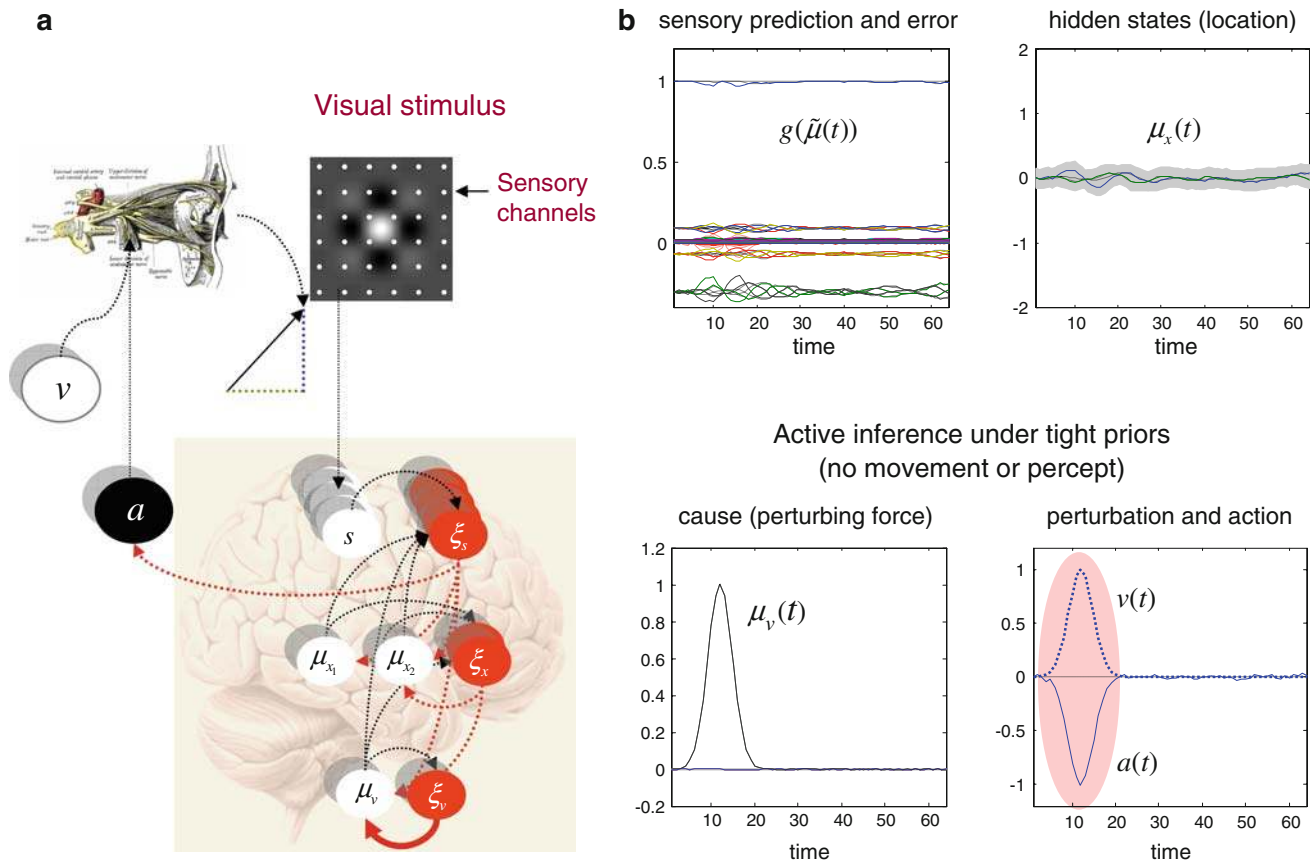
a

b   sensory prediction and error   hidden states (location)

Fig. 3 As per the previous figure but now using informative priors on the cause. These priors are induced by increasing the conditional log-precision of random fluctuations in the cause; indicated by *thick red line* in the *lower panel* of **a**. The ensuing priors drive action to explain away prediction error. In this instance, action mirrors the real cause (*circled region* in the *lower right panel* of **b**, so that their respective influences cancel and there is an attenuation of perceived movement. Note that the prediction error under these informative priors is larger (*red lines*). It is this prediction error that drives action. **a** This shows the relationship

of the generative model (*lower panel*) to the generative process modeling oculomotor sampling of the world. **b** *Upper left* These are the predicted sensory states and corresponding prediction error (*red lines*); *upper right* The expected hidden states causing sensory predictions: The *colored lines* correspond to conditional expectations, and the *gray regions* correspond to 90% confidence intervals. *Lower left* The true (*dotted line*) and conditional expectation (*solid line*) of the exogenous cause. *Lower right* Action (*solid line*) and true cause (*dotted line*)

time. It can be seen that the conditional expectation and true location are roughly the same and that very little action has been emitted. The results of this simulation should be compared with those in Fig. 3. Here, we placed tight priors ($\ln \mu_\gamma^v = 16$) on the force, so that the agent did not expect any movement of the stimulus that, essentially, it is foveating. When the stimulus actually moves, the consequent prediction error is explained away by action, and the excursion of the stimulus is countered almost immediately. The dynamics of action can be seen in the lower right panel and, as anticipated, mirror the exogenous force so that it is opposed almost exactly. The retinotopic trajectories with flat and informative priors on exogenous forces are plotted in Fig. 4 for comparison. These show that the excursion of the target has been completely suppressed by precise expectations (of no perturbations). Note that the only difference between the simulations in Figs. 2 and 3 was that we increased the precision of

the priors on the force, as indicated by the thicker red line at the bottom of the schematic in Fig. 3. We will return to the key role of precision in controlling behavior in the next section.

There are several interesting interpretations of this simple demonstration: If we consider it in the light of vision, then action corresponds to saccadic or reflexive eye movements. Note that the agent can only move its eye in one direction, although the motor plant entailed by the state equations of the generative process couple the resulting movements in both directions. Despite this, it is able to stabilize the image of the stimulus under active sampling. In other words, the image is stabilized in accordance with the prior expectation of no movement. This may provide a nice metaphor for retinotopic stabilization through visual and vestibular oculomotor reflexes (Tatler and Wade 2003). Alternatively, we could consider the action as causing a shift in the sensory frame
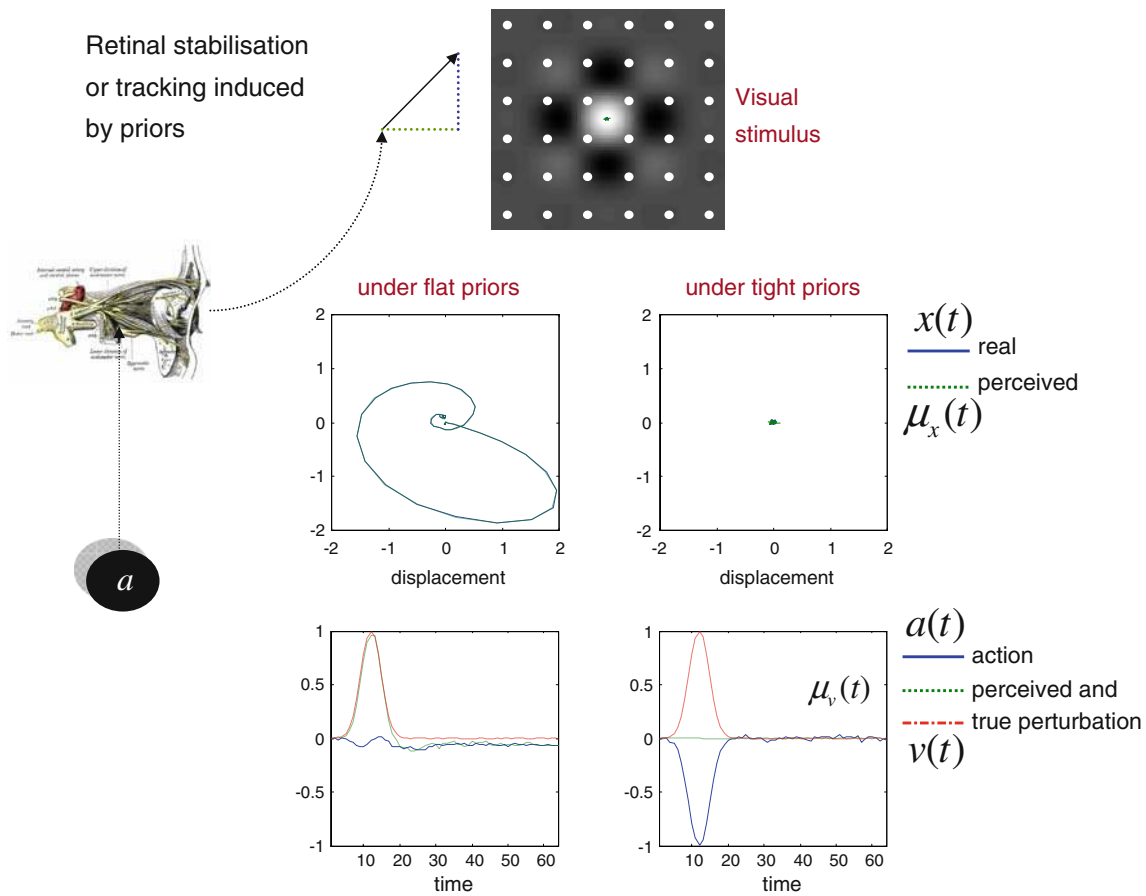
Fig. 4 The stimuli and trajectories with uninformative (Fig. 2) and tight (Fig. 3) priors: The *upper panel* details the stimulus or target in sensory space. *Middle panels*: These are the displacements of the stimulus perceived under flat and informative priors. The displacement under tight priors is re-plotted in retinotopic space in the *top panel*. Note the shrinkage of the perceived excursion toward the prior expectation of zero under tight priors (*middle right*). *Lower panels*: Action and causes with (*right*) and without (*left*) tight priors. Here, action (*solid blue line*) is engaged by informative priors to attenuate the perceived excursion of the stimulus (*dotted green line*)

of reference that enables the agent to track a moving target. This alternative perspective speaks to orientating or tracking behavior that enables the perception of a moving target as if it were stationary. Although the target may be moving, the full extent of this movement is not registered by perception. This is because the prediction error, under tight priors on the perturbation, is explained away by action, and action is not part of the perceptual model. In other words, from the point of view of the agent, the target is not moving. This suggests that orientating responses might be understood as active sampling of the sensorium, which conforms to the prior expectation that targets do not move. This is a simple but generic example of the sort of behavior that emerges with action under the free-energy principle.

4.1.2 Action and intended movements

In the previous simulations, we used an exogenous force with flat priors to elicit behavior that rendered a moving target stationary. What would happen if we switched the cause and priors around? In other words, what would happen if we took the real cause and used it as a prior expectation but did not actually move the stimulus? (i.e., if we swapped $\eta(t)$ and $\mathbf{v}(t)$ in Eq. 11). In this instance, there would be the same prediction error (with opposite sign) and an active response. In this case, the prior expectation is that the target is moving and active sampling will ensure that expectation is met. This means the agent should saccade away from the target and then return, once the prior $\eta(t)$ falls back to zero. The results in Fig. 5 show an example of this, using a model in which we replaced the nonlinear sensory mapping to simulated photoreceptors with a linear mapping to four oculomotor proprioceptors (see the schematic on the left). In this case, action reproduces the prior expectation on the cause to suppress prediction error (because they play the same role in the generative process; see Eq. 11). The result is apparent motion of the stimulus, shown in the upper panels of Fig. 6. This should be contrasted with the corresponding action in
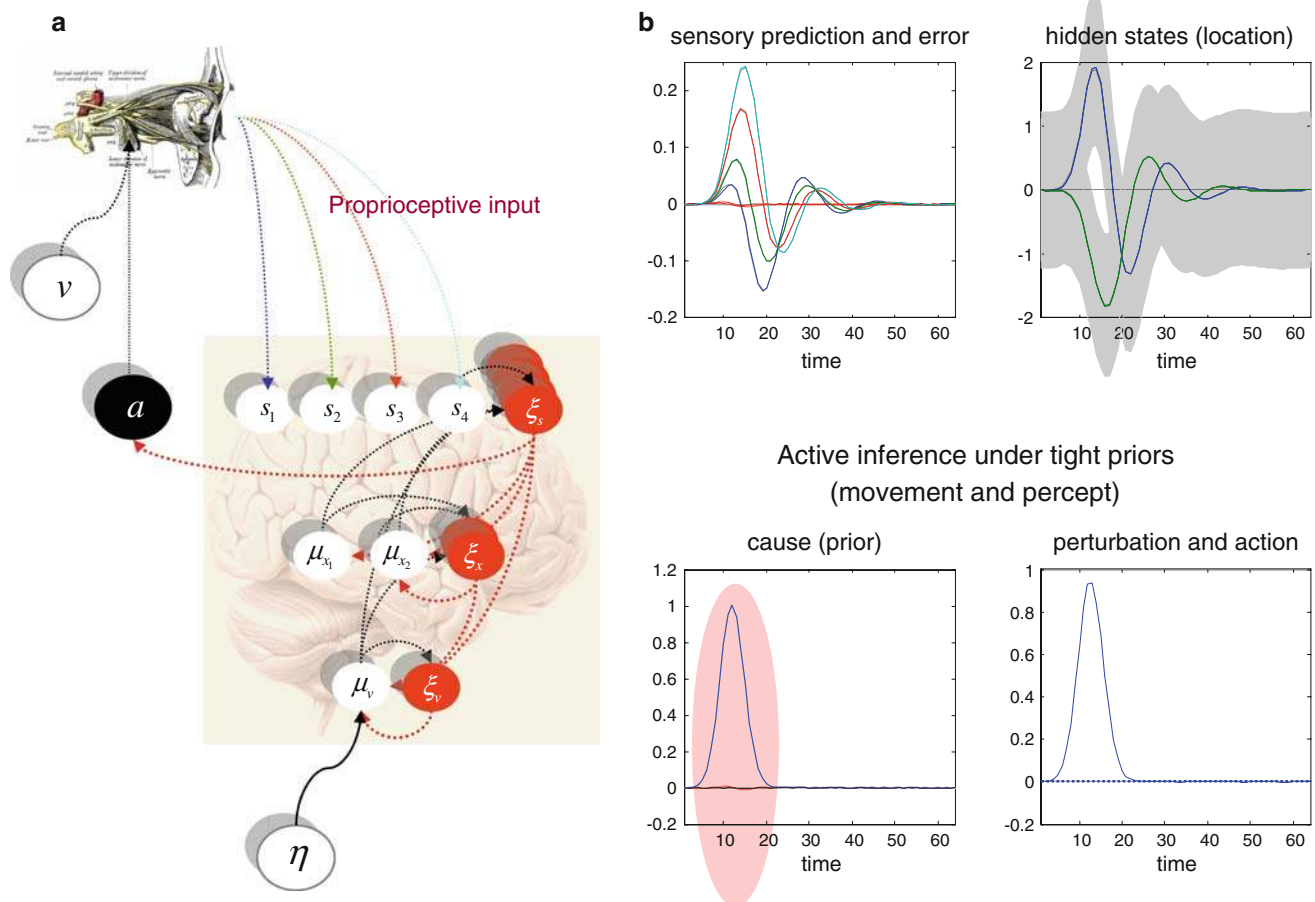
Fig. 5 As per Fig. 3 but here the true cause and prior expectation have been switched: i.e., the prior expectation is a Gaussian bump function, and the true cause is zero. In this instance, action is in the reverse direction to fulfill the prior expectation, causing a saccade away from the target and then back again. The resulting trajectory in retinotopic space is shown in the next figure. In this example, sensory data were generated by proprioceptors in the oculomotor system, as opposed to simulated photoreceptors: This is shown schematically in panel (**a**), which shows the relationship of the generative model (*lower panel*) with the generative process modeling oculomotor sampling of the world. **b** *Upper left*: These are the predicted sensory states and corresponding prediction error (*red lines*); *upper right*: The expected hidden states causing sensory predictions. The *colored lines* correspond to conditional expectations and the *gray regions* correspond to 90% confidence intervals. *Lower left*: The true (*dotted line*) and conditional expectation (*solid line*) of the exogenous cause. *Lower right*: Action (*solid line*) and true cause (*dotted line*)

the Fig. 4 where action and forces mirror each other, thereby canceling to mimic the flat prior.

This example suggests that movements can be induced by prior expectations of their sensory consequences. In other words, we only have to expect the sensory consequences of a movement to make it happen. Prior expectations induce predictions about what we should be sensing and action automatically moves sensory epithelia to fulfill these predictions. We are not aware of the movement per se, but register its sensory consequences that are encoded by prior expectations. In this example, the priors were prescribed by us. In more realistic hierarchical models, one would imagine that movement-inducing expectations would arise as empirical priors during inference on a more expansive set of multimodal sensory inputs.

This form of movement specification is a little counter-intuitive, but has some appealing and biologically plausible properties. For example, if we only have to specify the consequences of an intended or expected movement, then the actual action will be robust to variations in the true generative process or unexpected perturbations to that process. In other words, specifying a movement in terms of what we want to see, as opposed to what we want to do, automatically makes behavior much more robust. In order to illustrate this, we added an exogenous but late perturbation to the target stimulus during the intended movement. This is a simple metaphor for empirical studies of movement trajectories under perturbation (e.g., Paulignan et al. 1991; Fourneret and Jeannerod 1998; Liu and Todorov 2007). The resulting inference and action in Fig. 6 (middle panels) show that action (blue line)

Fig. 6 These plots show the trajectory of a perceived stimulus location (*left panels*), under movement induced by priors and consequent action (*right panels*). *Upper row*: This corresponds to a movement elicited by priors, where the prior expectation is formally identical to the true cause of the previous simulations (see Fig. 3). *Middle row*: These are the results of a simulation where a small exogenous perturbing force was applied late in the movement trajectory (*arrow in right panel*). The impact on the trajectory is almost imperceptible (*arrow in left panel*). In this instance, action has compensated for the unanticipated perturbation. *Lower row*: Here, we doubled the sensitivity of the motor plant to command signals or action. Note that action has automatically compensated for this change and that the ensuing trajectory, as perceived, has not deviated from its prior expectation
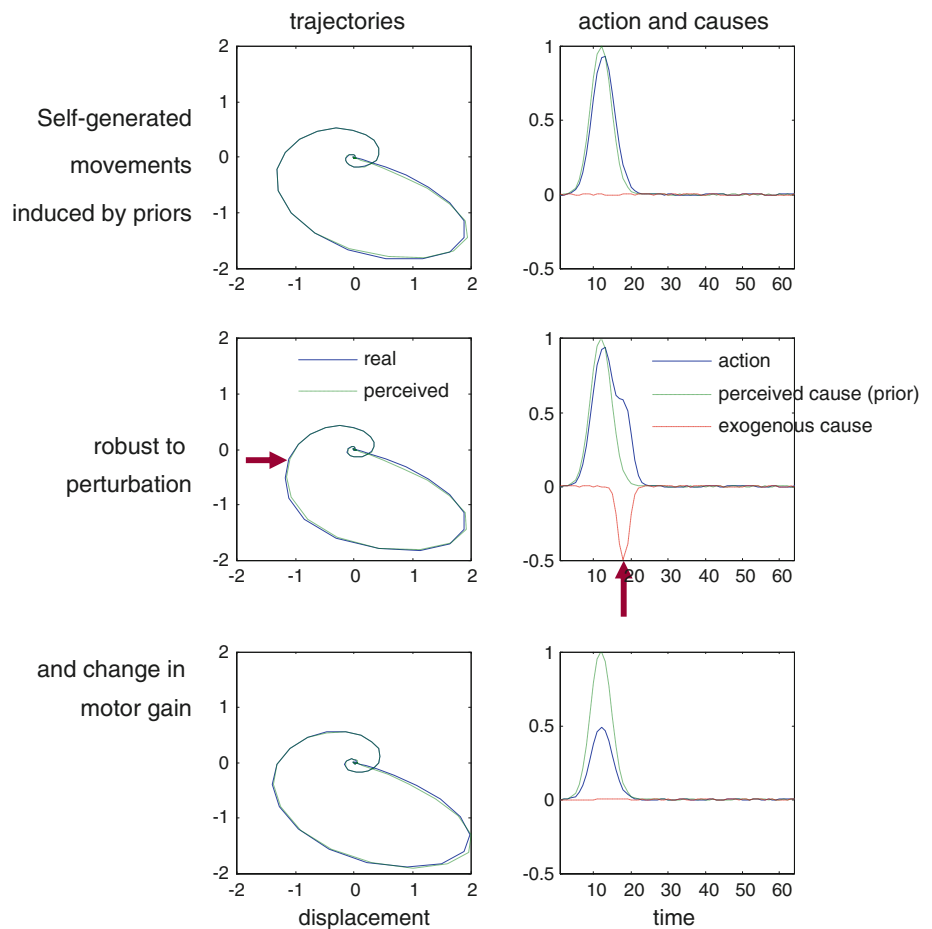
explains away the unexpected perturbation with relatively little impact on the percept (green line). This means that the perturbation was automatically discounted in terms of the movement trajectory. As a final example of the robustness of movement specification under active sampling, we changed the motor plant by changing the sensitivity of movement to action in the generative process; by replacing $a \rightarrow 2a$ in Eq. 11. Again, because action is simply minimizing prediction error, this had no effect on the perceived excursion of the target. As seen in Fig. 6 (lower panels), the only effect was evident in the amplitude of the action elicited. The agent does not need to know about this change in the dynamics of the motor plant because action is effectively exploiting closed-loop feedback (cf. the equilibrium-point hypothesis discussed later).

4.1.3 Summary

In these simulations, we have seen how active sampling can mimic retinal stabilization or visual tracking behavior by implementing a prior expectation that a visual target is stationary. Conversely, we can induce movement through action by invoking prior expectations about movements. Critically, specifying an action in terms of its sensory consequences

renders action robust to variations in the motor plant or unexpected exogenous perturbations. This is because action explains away prediction error, which is prescribed by the prediction and prior expectations. The simulations in this section used unimodal (visual or proprioceptive) models. In the next section, we turn to sensorimotor integration in the context of multimodal (visual and proprioceptive) inference.

4.2 Sensorimotor integration

In this section, we illustrate the Bayesian integration of proprioceptive and visual signals to optimize computational motor control of cued reaching movements (Körding and Wolpert 2004; Disney et al. 2007; Kreisel et al. 2007; Bruyn and Mason 2009; Diedrichsen and Dowling 2009). This method highlights the ability of active inference to explain multimodal integration in perception: through the convergence of bottom-up sensory prediction errors that optimize perceptual representations and multimodal integration in action; through the convergence of top-down sensory prediction error onto motor commands that optimize action. This bilateral role of prediction errors joins sensory and motor processing and provides a nice perspective on sensorimotor integration. Our particular focus here will be the weighting of

visual and proprioceptive signals to optimize action, which emerges from the optimization of precision parameters during recognition.

The model used in this section involves reaching to a target using a two joint arm. The proprioceptive input corresponds to the angle of both joints, (x_1, x_2), which are the hidden states, while the visual input is modeled as the position and brightness of a target and the location of the arm's extremity. We will treat the location (v_1, v_2) and brightness v_3 of the target as autonomous causal states and the location of the arm as a function of hidden states, $J(x)$. We ignore the complexities of extracting positional information from retinotopically mapped visual input and assume that the agent has direct access to locations of the target and arm in visual space:

$$\mathbf{g}(x, v) = g(x, v) = \begin{bmatrix} x \\ v \\ J \end{bmatrix} \quad (12)$$

The dynamics of the arm conform to Newtonian laws, under which action forces the angular position of each joint, with angular inertia, m_i. Both joints have an equilibrium position at 90 degrees; with elasticity $k = [8, 4]$ and viscosity $\kappa = [4, 2]$ giving the following equations of motion

$$\mathbf{f}(\tilde{\mathbf{x}}, \mathbf{v}) = \begin{bmatrix} \dot{\mathbf{x}}_1 \\ \dot{\mathbf{x}}_2 \\ \dot{\mathbf{x}}'_1 \\ \dot{\mathbf{x}}'_2 \end{bmatrix} = \begin{bmatrix} \mathbf{x}'_1 \\ \mathbf{x}'_2 \\ \left(a_1 - k_1\left(\mathbf{x}_1 - \frac{\pi}{2}\right) - \kappa \mathbf{x}'_1\right)/m_1 \\ \left(a_2 - k_2\left(\mathbf{x}_1 - \frac{\pi}{2}\right) - \kappa \mathbf{x}'_2\right)/m_2 \end{bmatrix} \quad (13)$$

However, the agent's empirical priors had a very different form: The agents' generative model ignores elasticity and assumes that the arm's distal extremity accelerates toward the target when, and only, when the target is illuminated brightly. This prior expectation is encoded as an expected force $\varphi = v_3([v_1, v_2]^T - J(x))$ that acts on both joints

$$f(\tilde{x}, v) = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}'_1 \\ \dot{x}'_2 \end{bmatrix} = \begin{bmatrix} x'_1 \\ x'_2 \\ (\varphi^T J_2 J_2^T O J_1 - \kappa_1 x'_1)/m_1 \\ (\phi^T O J_2 - \kappa_2 x'_2)/m_2 \end{bmatrix} \quad (14)$$

Here, $J_i(x)$ corresponds to the relative position of the i-th joint such that $J = J_1 + J_2$ and

$$J_1 = \begin{bmatrix} \cos(x_1) \\ \sin(x_1) \end{bmatrix} \quad J_2 = \begin{bmatrix} -\cos(-x_2 - x_1) \\ \sin(-x_2 - x_1) \end{bmatrix} \quad O = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (15)$$

These priors on the dynamics of the hidden states cause the arm to reach for the target in a reasonably graceful and plausible way, when the target is brightened. Figure 8 shows a typical simulation under fairly noisy proprioception. Here, we modulated the brightness of the target with a Gaussian bump function. When the target becomes bright the agent believes its hand is drawn to the target. This belief induces prediction errors on the generalized angular position (i.e., angular force), which action quenches by moving the joints to comply with prior expectations. Note that sensory predictions are visual and proprioceptive because the generative model renders both modalities a function of hidden states, like forces and torques. In other words, not only does the agent expect to feel a torque but it also expects to see acceleration of its arm. Once the brightness cue has disappeared, the arm remains on target because the agent's generative model does not include elastic forces that would return to the arm to its resting position. This necessitates persistent action to resist elastic forces (see lower middle panel of Fig. 8). This type of control is quite similar to the Passive Motion Paradigm (Mussa Ivaldi et al. 1988) and the ensuing Kinematic Network model. This model expresses kinematic transformations implied by elasticity to regularize the ill-posed problem that arises from redundancy in motor control.

In this simulation, we used a relatively large amount of sensory noise on the proprioceptive input (with a log precision of four), while visual information was relatively high precision (a log precision of eight). Despite this, the movement trajectory was remarkably smooth and accurate. This is because action is driven by both proprioceptive and visual prediction errors (descending black arrows in Fig. 7). Although the proprioceptive errors are noisy, the visual errors are precise and can elaborate accurate predictions of hidden states (see top left panel of Fig. 8) and precise motor commands. This robustness to proprioceptive noise rests on optimizing the expected precisions in both sensory modalities as in Eq. 9. The ensuing optimization boosts visual prediction errors relative to proprioception and ensures that the relative contribution of both modalities is balanced in a Bayes optimal fashion (for both action and perception). This aspect of multimodal integration can be illustrated by examining the expected precisions under factorial manipulations of visual and proprioceptive noise.

Figure 9 shows the expected precisions for proprioceptive μ_γ^{pro} and visual μ_γ^{vis} modalities and the associated movements under all combinations of low (log precision of eight) and high (log precision of four) levels of sensory noise. It can be seen that the estimated precisions are almost exactly correct for all four combinations. In terms of movement, it is clear that vision can substitute for proprioception and vice versa. Only when both modalities are noisy (i.e., imperfect proprioception and obscuring the arm from sight), then reaching is compromised. This is seen on the lower left, when the arm flails around erratically, as might be seen in an excited neonate.

4.2.1 Summary

In this section, we have reiterated the idea that apparently purposeful or intentional movements can be prescribed purely in
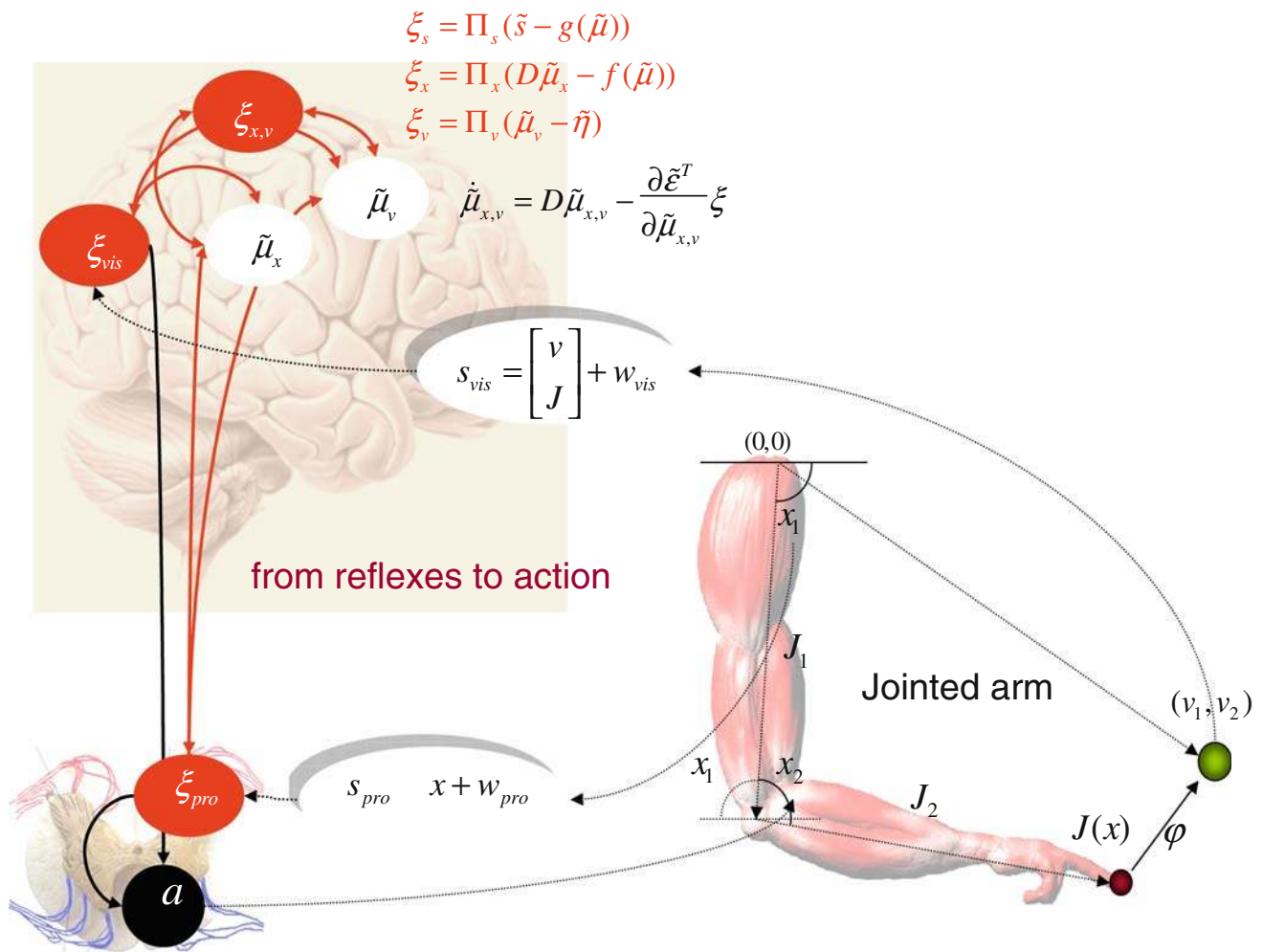
$$\xi_s = \Pi_s(\tilde{s} - g(\tilde{\mu}))$$
$$\xi_x = \Pi_x(D\tilde{\mu}_x - f(\tilde{\mu}))$$
$$\xi_v = \Pi_v(\tilde{\mu}_v - \tilde{\eta})$$

$$\dot{\tilde{\mu}}_{x,v} = D\tilde{\mu}_{x,v} - \frac{\partial \tilde{\varepsilon}^T}{\partial \tilde{\mu}_{x,v}}\xi$$

$$s_{vis} = \begin{bmatrix} v \\ J \end{bmatrix} + w_{vis}$$

from reflexes to action

$$s_{pro} \qquad x + w_{pro}$$

Jointed arm

Fig. 7 Schematic of the simulated cued-reaching models. *Right*: Motor plant, comprising a two-jointed arm with two degrees of freedom (hidden states) corresponding to the angular position of each joint. The position of the arm's extremity (*red circle*) is the sum of the vectors describing the location of each joint. Here, exogenous causes are the position and brightness of the target (*green circle*). The masses (angular inertia) of the first (shoulder) and second (elbow) joints were four and two respectably, with viscosity coefficients of eight and four. The elasticity of both was one quarter. *Left*: The agent senses the hidden states directly and indirectly through the position of the finger in visual space. In addition, visual input includes the causes representing target location and brightness. Sensory prediction errors are passed to higher brain levels to optimize the conditional expectations of the hidden states (angular position of the joints). 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 red lines denote reciprocal message passing among neuronal populations encoding prediction error and conditional expectations, however, the *black lines* represent descending motor control signals from sensory prediction error units. The equations implement free-energy minimization or recognition, as described in the main text

terms of prior expectations. Here, these priors were encoded by cue-dependent priors on the generalized motion of hidden states (angular position) of an agent's motor plant (a two-joint arm). These priors are combined with multimodal sensory evidence to provide predictions, which drive action. Critically, the evidence from different modalities is integrated in an optimal way by weighting the ensuing prediction errors by their optimized precisions. Both recognition and action rest on prediction errors and their relative weighting. Sensory prediction errors can derive from may modalities and sensory epithelia. In general, these will converge through hierarchical message passing to optimize perception and action in a bilateral fashion. In active inference, perception speaks to action through sensory prediction errors which optimize perception through their bottom-up effects and optimize action by projecting down to motor control units in the oculomotor system and spinal cord.

This section has shown how simple cued movements can be elaborated by priors under the free-energy principle, with a special focus on the role precision in optimizing multisensory integration in motor control. In the next section, we show how autonomous goal-directed movements can be learned by
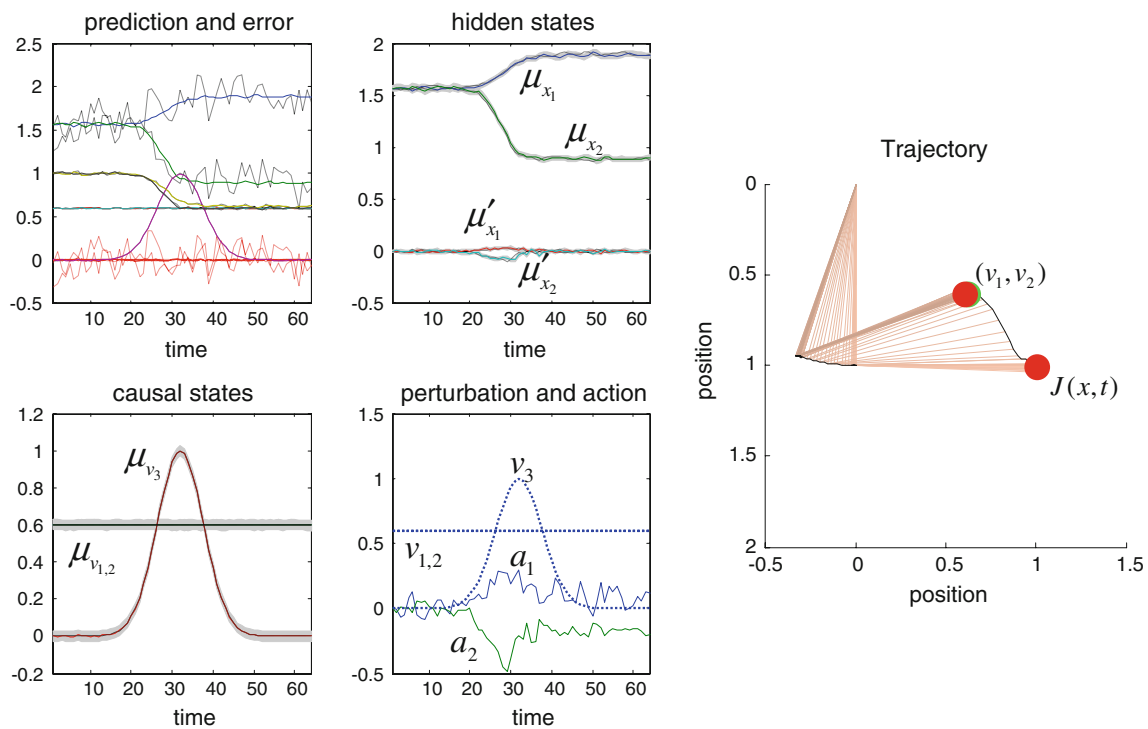
Fig. 8 Simulating cued movements: *Left* conditional expectations of causal and hidden states using the same format as in Figs. 2, 3, and 5. In this example, the proprioceptive input was quite noisy (*gray lines* in the *top left panel*); necessarily producing large proprioceptive prediction errors (*red lines* in the *top left panel*). *Right*: Ensuing movement trajec-tory caused by action (*solid lines* in the *adjacent lower panel*). The *red circles* indicate the initial and final position of the arm, which is seen to reach the target (*green circle*) reasonably quickly and smoothly. The *black lines* show the trajectory of the arm, and the *colored lines* show the position vectors at each time step

optimizing the parameters encoding prior expectations about movement trajectories.

5 Goal-directed behavior

This section deals with the optimization of action or poli-cies under perceptual learning (cf. Grafton et al. 2008). In the previous section, we specified the priors that led to cued-movements. Here, we consider how these priors are learned. In brief, we will show that quite complicated goal-directed behaviors can be instilled through supervised learning by placing an agent in a controlled environment so that it learns what to expect. When replaced in an uncontrolled or native environment, the agent seeks out the sensory states that it has learned to expect. The result is an optimum policy that is robust to perturbations and constrained only by the agent's prior expectations that have been established during training. A controlled environment can be thought of as one in which certain behaviors are enforced. In the context of reaching movements, this control could be exerted by force channels (e.g., Scheidt et al. 2000; Diedrichsen and Dowling 2009).

The model used in this section is a benchmark example in dynamic programming: namely, the mountain-car problem. These simulations have already been reported in Friston et al.

(2009) to illustrate how reinforcement learning can be cast as free-energy optimization. Here, we reprise the simula-tions but interpret them in terms of learning a motor policy that manifests as goal-directed movements. The mountain-car problem can be envisaged as follows: one has to move a heavy weight away from the bottom of the well and keep it there (Fig. 10). However, the weight is too heavy to sim-ply to push it out of the well. This means that the only way to move the weight to its desired location is to push it up one side of the well and use its momentum to get it past the steepest part of the other side. This represents an interesting problem, when considered in the state–space of position and velocity, (x, x'); the agent has to move *away* from the target location $(x = 1, x' = 0)$ to attain its goal and execute a very circuitous movement (cf., avoiding obstacles; Jax and Rosenbaum 2007). This problem can be specified with the sensory mapping and equations of motion

$$\mathbf{g} = \tilde{\mathbf{x}}$$

$$\mathbf{f} = \begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{x}}' \end{bmatrix} = \begin{bmatrix} \mathbf{x}' \\ b(\mathbf{x}) - \frac{1}{4}\mathbf{x}' + \mathbf{v} + \sigma(a + c(\tilde{\mathbf{x}})) \end{bmatrix}$$

$$b(\mathbf{x}) = \begin{cases} -2\mathbf{x} - 1 & : \mathbf{x} \le 0 \\ -\mathbf{x}^2(1 + 5\mathbf{x}^2)^{-3/2} - \mathbf{x}^4/16 & : \mathbf{x} > 0 \end{cases} \quad (16)$$

$$c(\tilde{\mathbf{x}}) = \theta_1 + \theta_2\mathbf{x} + \theta_3\mathbf{x}' + \theta_4\mathbf{x}\mathbf{x} + \theta_5\mathbf{x}\mathbf{x}' + \theta_6\mathbf{x}'\mathbf{x}'$$

Fig. 9 Reaching under different levels of sensory noise. All panels show movement trajectories using the same format as in the previous figure. The *inset* for each panel shows the conditional expectations of the precision (inverse variance or noise level) of proprioceptive and visual input. It can be seen that movement is robust to noise, unless both modalities deliver degraded sensory information. *Lower row*: Noisy proprioceptive (angular joint position) input. *Right column*: Noisy visual information about the location of the arm (e.g., partial occlusion)
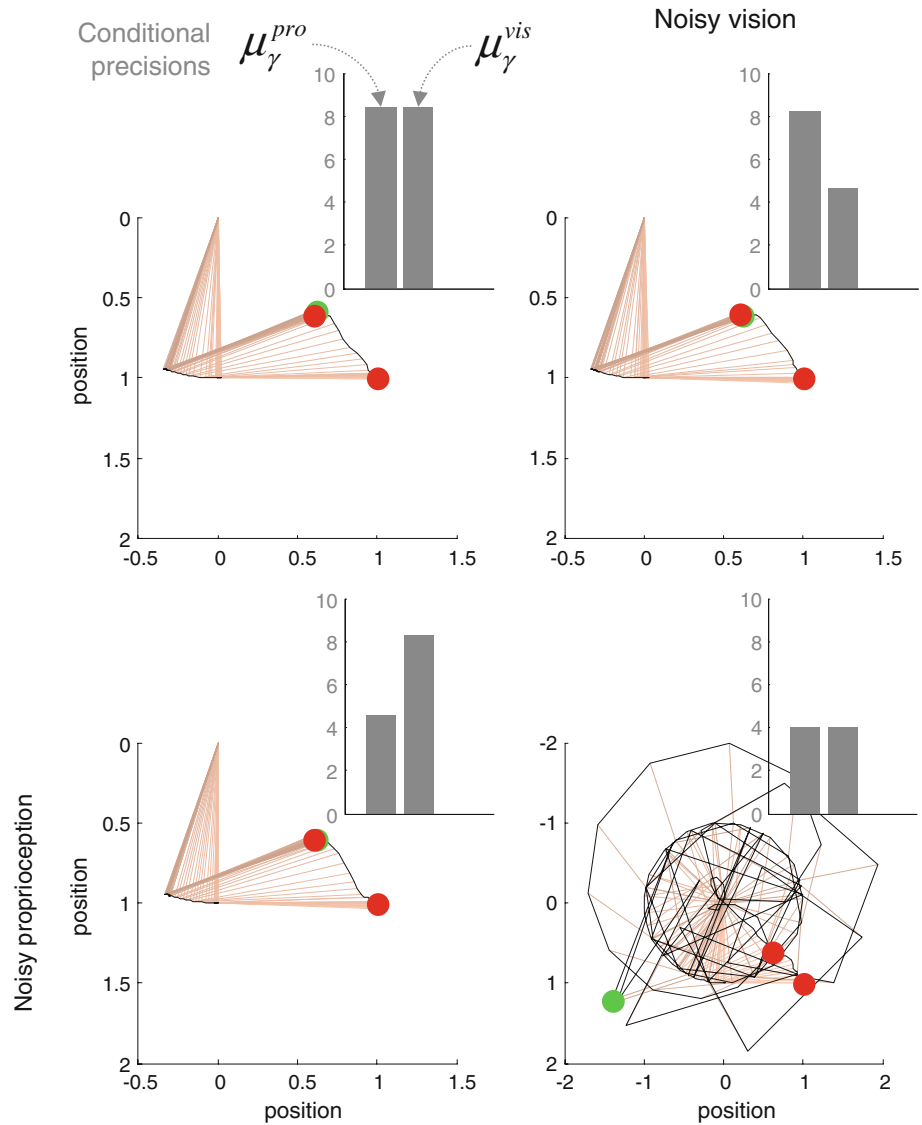

Fig. 10 Schematic of the mountain car problem: *Left* The landscape or potential energy function that defines forces on the weight. This has a minimum at **x** = −0.5. The weight is shown at its uncontrolled stable position (transparent) and the target position at the top of the well on the right at **x** = 1 (*vertical dotted line*). *Right*: Forces experienced by the weight at different positions due to the slope of the well (*blue*). Critically, at **x** = 0, the force is unity and cannot be overcome by the agent, due to the squashing function applied to action
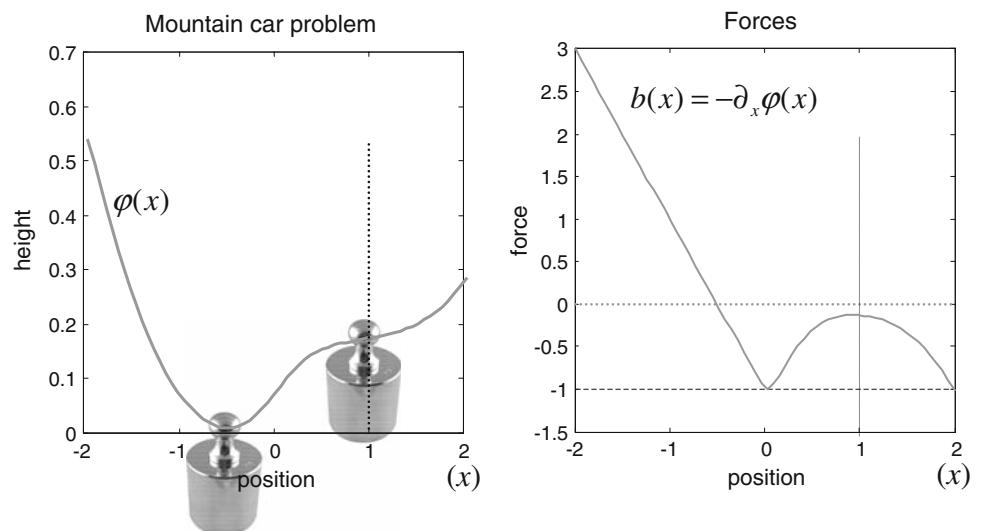
As above, substitution of these functions into Eq. 7 provides the equations used to simulate environmental dynamics and generate sensory data. The first equality means that the agent has a noisy sense of the weight's position and velocity. The second means that the forces on the weight, $\dot{\mathbf{x}}'(t)$ have four components: a gravitational force $b(\mathbf{x}) = -\partial_{\mathbf{x}}\varphi(\mathbf{x})$, which is just the spatial gradient of the potential energy function $\varphi(\mathbf{x})$ in Fig. 10 (left panel). The remaining three forces are friction $\mathbf{x}'/4$, an exogenous force $\mathbf{v}(t)$, and a controlling force that is bounded by a squashing (logistic) function: $-1 \leq \sigma \leq 1$. This force comprises action and state-dependent control, $c(\tilde{\mathbf{x}})$. Control is approximated here with a second-order polynomial expansion of any nonlinear function of the states, whose parameters are $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_6\}$. When $\boldsymbol{\theta} = 0$ the environment is uncontrolled; otherwise, the weight experiences state-dependent forces that enable control.

In order to create a controlled environment, we simply optimize the parameters of its equations of motion to maximize the probability that the goal is attained. This entails adjusting the parameters to maximize the ensemble density, $p(\tilde{\mathbf{x}}|m)$ at the target location, as described in Friston et al. (2009). The upper panels of Fig. 11 show the equilibrium densities without control ($\boldsymbol{\theta} = 0$; top row) and for the controlled environment ($\boldsymbol{\theta} = \theta_Q$; middle row) in which the goal attracts all trajectories. As anticipated, the trajectories in Fig. 11 (middle row) move away from the goal initially and then converge on it. This controlled environment now plays host to a naïve agent, who must learn its dynamics through experience.

5.1 Learning

The agent's generative model of its sensory inputs comprised the functions

$$
\begin{aligned}
g &= \tilde{x} \\
f &= \begin{bmatrix} x' \\ b - \frac{1}{4}x' + v + \sigma(\eta + c) \end{bmatrix}
\end{aligned} \tag{17}
$$

For simplicity, we assumed that $f(\tilde{x}, \tilde{v}, \theta)$ was the same as Eq. 16 but without action. The unknown causes in this model, $\Psi \supseteq \{\tilde{x}, \tilde{v}, \theta, \gamma\}$, comprise the states (position and velocity), exogenous force, parameters controlling state-dependent acceleration, and precisions of the random fluctuations. The agent was exposed to 16 trials of 32 s time-bins. Simulated training involved integrating Eqs. 7 and 9 with $\theta = \theta_Q$ (see Appendix 3). On each trial, the weight was "pushed" with an exogenous force, sampled from a Gaussian density with a standard deviation of eight. This enforced a limited exploration of state–space. The agent was aware of these perturbations, which entered as priors on the forcing term; i.e. $\tilde{\eta} = \tilde{\mathbf{v}}$. During learning, we precluded active inference, such

that the agent sensed the trajectory passively, as it left the target state and returned to it.

Note that the agent does know the true states because we added a small amount of observation error (with a log-precision of eight) to form sensory inputs. Furthermore, the agent's model allows for random fluctuations on both position and velocity. When generating sensory data, we used a small amount of noise on the motion of the velocity (with a log-precision of eight). After 16 trials, the parameters converged roughly to the values that were utilized to construct the control environment. This means that the agent expected the weight to move along state–space trajectories that converge on the target. These optimum dynamics have been learned in terms of empirical priors on the generalized motion of states encoded by μ_θ—the expected parameters of the equations of motion. These expectations are shown in the lower row of Fig. 11 in terms of the trajectories they encode; i.e., $f(\tilde{x}, \tilde{v}, \mu^\theta)$. It can be seen that the nullclines (lower right) based on the parameters after training have a similar topology to the controlled environment (middle right), ensuring that the fixed-points that have been learnt are the same as those required to attain the goal. Therefore, what would happen if the agent was placed in an environment that did not conform to its expectations?

5.2 Active inference

In order to demonstrate that the agent has learnt the optimum policy, we placed it in an uncontrolled environment; i.e., $\boldsymbol{\theta} = 0$, and allowed action to minimize free-energy. We prevented further perceptual learning by fixing the conditional expectations of the parameters to μ_θ and the log-precisions to $\mu_\gamma^x = \mu_\gamma^s = 8$. An example of active inference after learning is presented in Fig. 12. Again, this involved integrating environmental and recognition dynamics (Eqs. 7, 9), where these stochastic differential equations are now coupled through action. The colored lines show the conditional expectations of the states, while the gray areas represent 90% confidence intervals. These are tight because we used low levels of noise. The dotted red line on the upper left corresponds to the prediction error: namely, the discrepancy between the observed and predicted states. The ensuing trajectory is superimposed on the nullclines (middle panels) and shows the agent moving the weight away from its goal initially, to build up the requisite momentum. Once the goal has been attained, action is still required because, in the uncontrolled environment, it is not a fixed-point attractor.

In order to illustrate the robustness of this behavior, we repeated the simulation while applying exogenous forces to the weight (modeled with a random normal variate smoothed with a Gaussian kernel of 8 s). Because the agent did not expect these perturbations, they were explained away by action and not perceived. Goal-directed behavior was
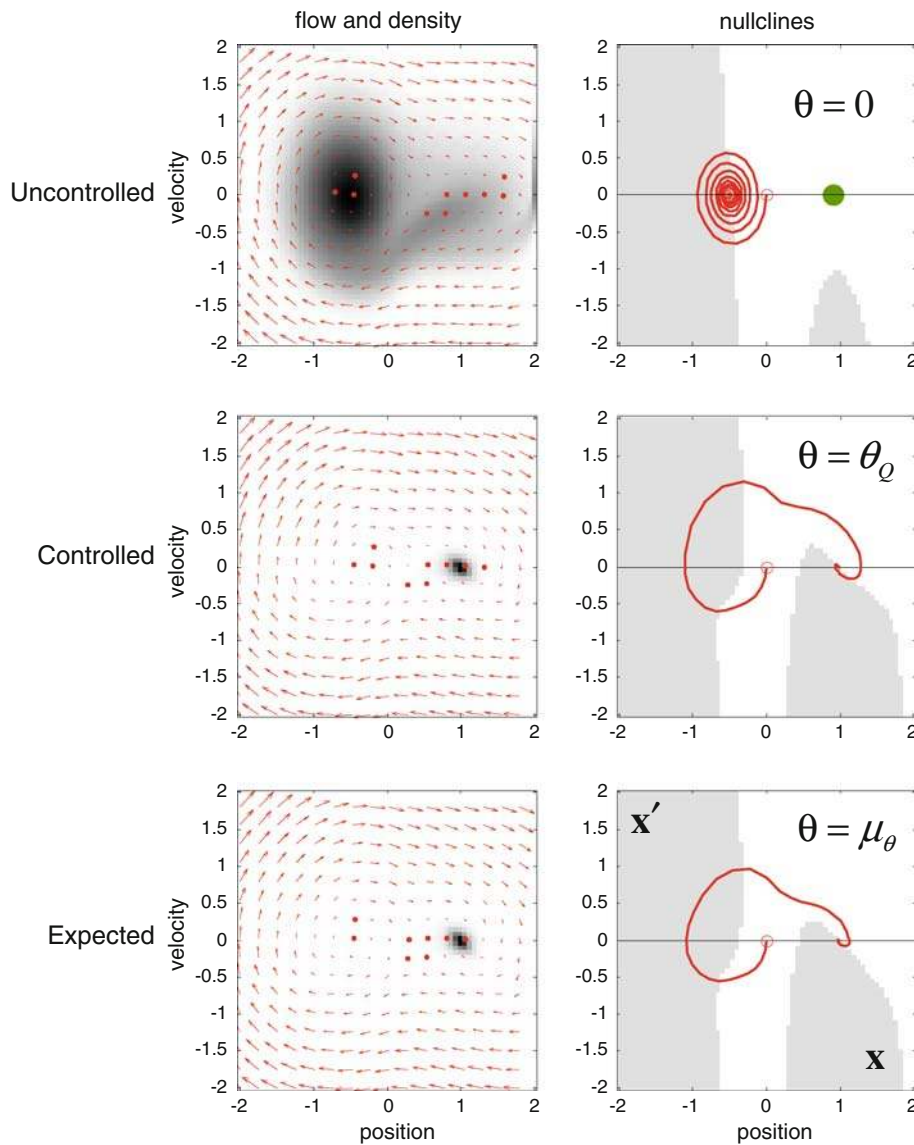
Fig. 11 *Left panels*: Flow-fields and associated equilibrium densities for an uncontrolled environment (*top*), a controlled or optimized environment (*middle*), and under prior expectations after learning (*bottom*). Notice how the flow of states in the controlled environment enforces trajectories that start by moving away from the target location at position **x** = 1 (with a small velocity). The *red arrows* denote the change in state (position and velocity) prescribed by the parameters: this change is $f(\tilde{x}, \tilde{v})$ evaluated at $x_3 = 0.5$. The ensemble density is the principal eigenfunction of the Fokker–Plank operator associated with these parameters. For the controlled and expected environments, these are low entropy equilibria, concentrated on the target location. *Right panels*: These panels show the flow fields in terms of their nullclines. Nullclines correspond to lines in state–space where the rate of change of one variable is zero. Here, the nullcline for position is along the x-axis,

where velocity is zero. The nullcline for velocity is when the change in velocity goes from positive (*gray*) to negative (*white*). Fixed points correspond to the intersection of these nullclines. It can be seen that in an uncontrolled environment (*top*), there is a stable fixed point, where the velocity nullcline intersects the position nullcline with negative slope. Under controlled (*middle*) and expected (*bottom*) dynamics, there are now three fixed points. The rightmost fixed-point is under the maximum of the ensemble density and is stable. The middle fixed-point is halfway up the hill, and the final fixed-point is at the bottom. Both of these are unstable and repel trajectories so that they are ultimately attracted to the target location. The *red lines* depict an exemplar trajectory, under deterministic flow, from $x = x' = 0$. In a controlled environment, this shows the optimum behavior of moving up the opposite side of the well to gain momentum so that the target (*green dot*) can be reached

preserved (lower panels of Fig. 12), with mirror symmetry between action and the displacing force it counters. Note that action is greater because it exerts its effects through a squashing function (see Eq. 16).

5.3 Optimal behavior and conditional precision

Optimal behavior depends on the expected precision of the hidden state motion encoded by μ_γ^x. In this example, the

Fig. 12 *Top row*: The *left panel* shows the predicted sensory states (position in *blue* and velocity in *green*). The *red lines* correspond to the prediction error based on conditional expectations of the states (*right panel*). The associated conditional covariance is displayed as 90% confidence intervals (*thin gray areas*). *Middle row*: The nullclines and implicit fixed points associated with the parameters learnt by the agent, after exposure to a controlled environment (*left*). The actual trajectory through state–space is shown in *blue*; (the *red line* is the equivalent trajectory under deterministic flow). The action causing this trajectory is shown on the right and shows a polyphasic response until the desired position is reached, after which a small force is required to stop the weight sliding back into the well (see Fig. 10). *Bottom row*: As per the middle row but now in the context of a smoothly varying perturbation (*broken line* in the *right panel*). Note that this exogenous force has very little effect on the navigation of state–space because it is unexpected and countered by action. These simulations used expected log-precisions of eight
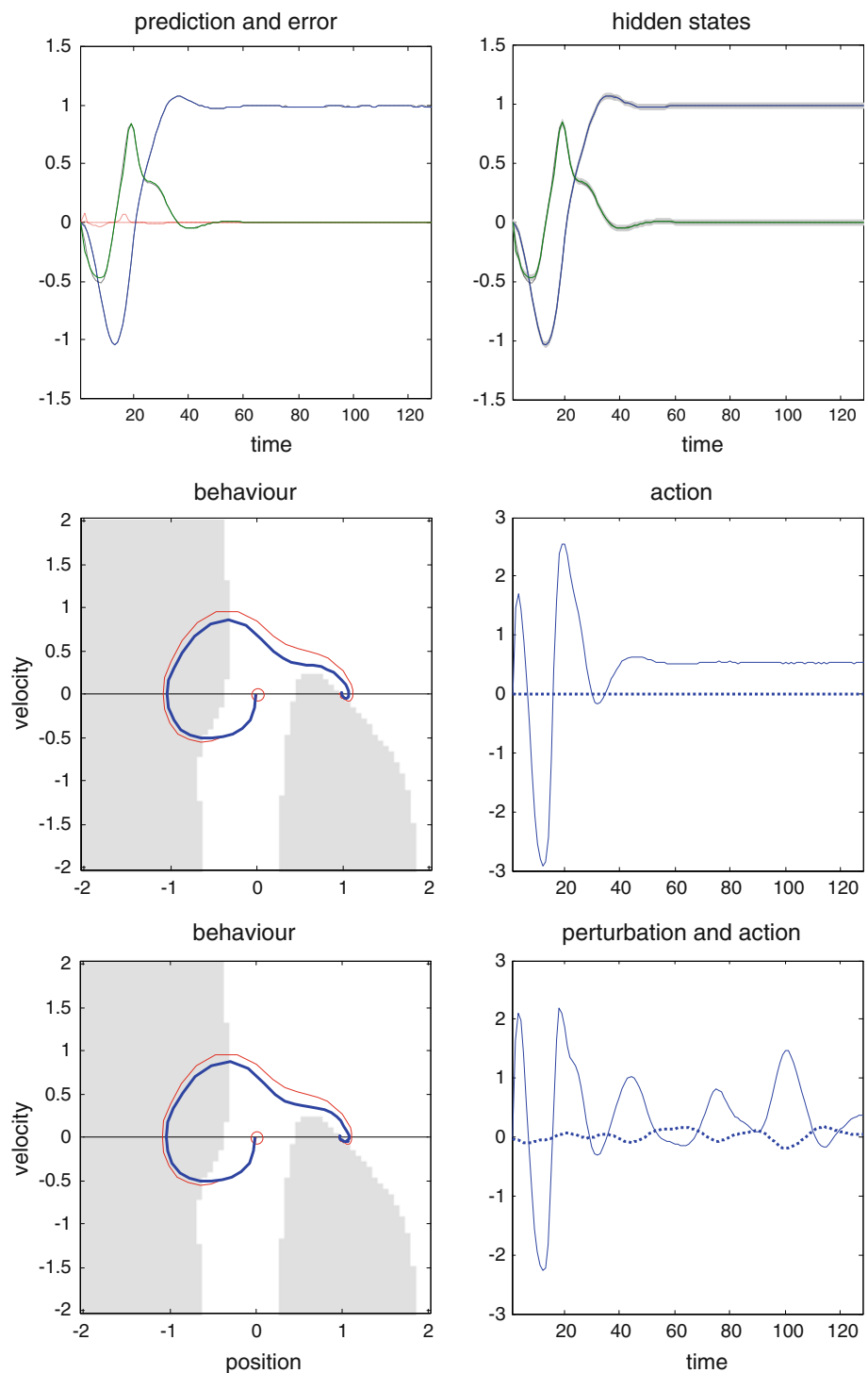
agent was fairly confident about its prior expectations, but did not discount sensory evidence completely (with log-precisions of $\mu_\gamma^v = \mu_\gamma^s = 8$). These conditional precisions are important quantities, which control the relative influence of bottom-up sensory information relative to top-down predictions. In a perceptual setting, they may mediate attentional gain (cf., Abbott et al. 1997; Yu and Dayan 2005; Friston 2008). In active inference, they also control whether an action is emitted or not (i.e., motor intention): Increasing the relative precision of the hidden states causes more confident behavior, whereas reducing it subverts action. This is because prior expectations are discounted by sensory input and are, therefore, not expressed at the level of sensory predictions. In biological formulations of the free-energy principle, modulatory neurotransmitters might encode the precision of prior expectations (Friston 2008). In this context,
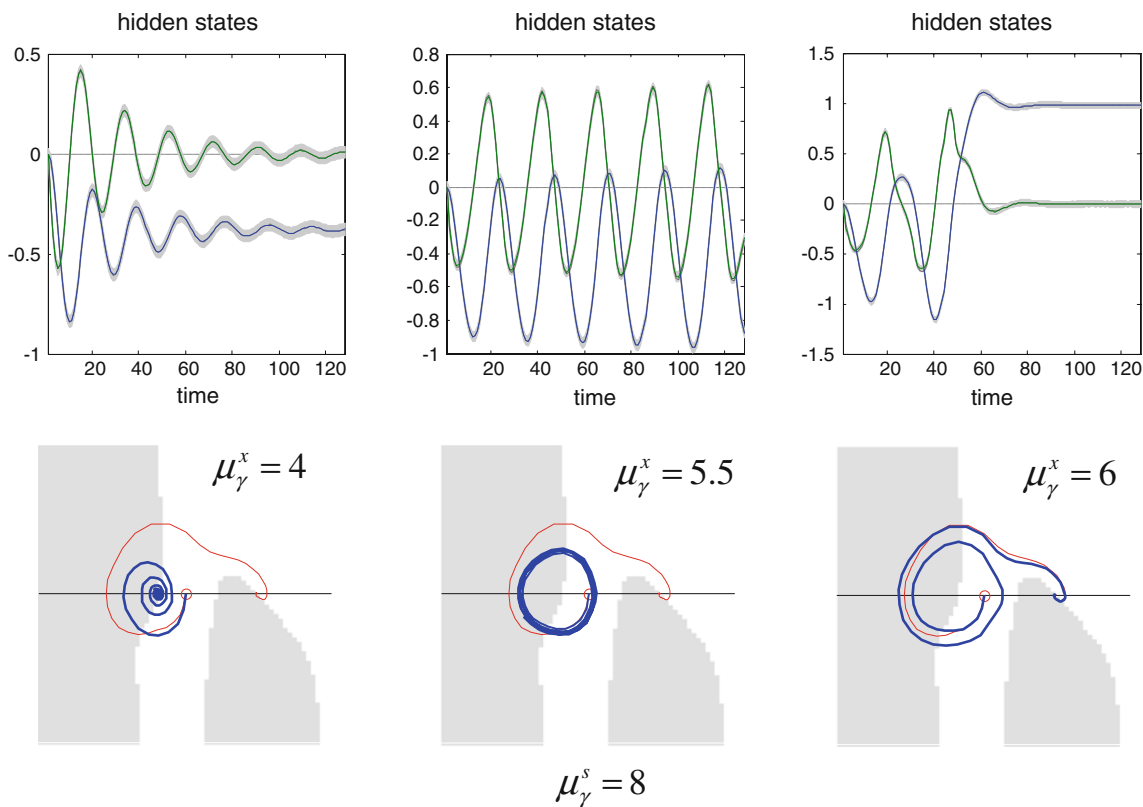
Fig. 13 Inferred states, (*top row*) and trajectories through state–space (*bottom row*) under different levels of conditional uncertainty or expected precision. As in previous figures, the inferred sensory states (position in *blue* and velocity in *green*) are shown with their 90% confidence intervals, and the trajectories are superimposed on nullclines. As the expected precision falls, the inferred dynamics are less accountable to prior expectations, which become less potent in generating prediction errors and action. It is interesting to see that uncertainty about the states (*gray area*) increases, as precision falls and confidence is lost

a deficit in dopaminergic neurotransmission would reduce the operational potency of priors to elicit action and lead to motor poverty; as seen in Parkinson's disease, schizophrenia, and neuroleptic bradykinesia. Through progressively reducing the expected precision of the empirical priors that have been instilled during training, we can simulate this poverty. Figure 13 shows three phases: first a loss of confident behavior, where the agent rocks backward and forward cautiously, until it has more than sufficient momentum to reach its goal. Second, a stereotyped behavior (corresponding to a quasiperiodic attractor), in which the agent prevaricates at the bottom of the well (cf., displacement activity, motor stereotypy, or perseveration). Finally, we get avolitional behavior, where the agent succumbs to gravity (cf., bradykinesia or psychomotor poverty).

5.4 Summary

We have seen that agents can learn causal structure in the environment and sample it in an adaptive and self-supervised fashion. This results in behavioral policies that reproduce those optimized by reinforcement learning and dynamic programming. Critically, we do not need to invoke the notion of reward or value to optimize these behaviors. It is sufficient to induce prior expectations through perceptual learning that prescribe goal-directed behavior in an unsupervised setting. In the previous section, we considered the relative precision among sensory signals in different modalities. Here, we kept the sensory precision fixed and changed the precision of prior expectations. Under active inference, the precision of sensory information and empirical priors determine their relative contribution to action. In this context, it is interesting to note that a "mere expectation to move causes attenuation of sensory signals" (Voss et al. 2008). This is consistent with a relative increase in the precision and consequent amplitude of prediction errors that mediate empirical priors. In this section, we focused on goal-directed movements that had a fixed goal or state. In the final section, we address the learning of behaviors that are specified not in terms of goals but in terms of movement trajectories per se.

6 Learning autonomous behavior

In the previous section, we specified behavior in terms of a target location in state–space. This is fine if we want agents

to occupy specific states: i.e., if we want to induce action with fixed-point attractors. However, real-world behaviors are dynamic, calling for the optimization of state trajectories of an autonomous and itinerant nature. In ethology, these might correspond to a succession of states or sequences of behaviors that are elicited with high-order operant conditioning paradigms. In optimal control, an exemplar problem is bipedal walking (Prinz 2006; Manoonpong et al. 2007). Here, we set ourselves the task of instilling behavioral trajectories in an agent that conform to one of the simplest systems with autonomous dynamics, the Lorenz attractor.

Using the same procedure as above, we used the following generative process and model, whose equations of motion have the form of a Lorenz system

$$\mathbf{g} = \mathbf{x}$$
$$\mathbf{f} = \frac{1}{64} \begin{bmatrix} \boldsymbol{\theta}_1 \mathbf{x}_2 - \boldsymbol{\theta}_1 \mathbf{x}_1 \\ \boldsymbol{\theta}_2 \mathbf{x}_1 - \mathbf{x}_3 \mathbf{x}_1 - \mathbf{x}_2 \\ \mathbf{x}_1 \mathbf{x}_2 - \boldsymbol{\theta}_3 \mathbf{x}_3 \end{bmatrix} + \begin{bmatrix} \mathbf{v} \\ 0 \\ 0 \end{bmatrix} + a$$

$$g = x$$
$$f = \frac{1}{64} \begin{bmatrix} \theta_1 x_2 - \theta_1 x_1 \\ \theta_2 x_1 - x_3 x_1 - x_2 \\ x_1 x_2 - \theta_3 x_3 \end{bmatrix}$$

(18)

We set $\boldsymbol{\theta} = \theta_Q = \left\{ 10, 32, \frac{8}{3} \right\}$ and initialized the conditional expectations of the agent to $\mu_\theta = 0$ (with fixed log-precisions of $\mu_\gamma^s = 8$ and $\mu_\gamma^x = 16$). In these simulations, we took the hidden states to represent the position of an agent (or its arm) in a three-dimensional (3D) Euclidean space. This means that the agent moves under control as if it was on a Lorenz attractor and comes to learn what trajectories to expect. After 16 trials of 512 time-bins (16×8 s of real time), the agent was removed from its controlled environment and examined for evidence of learning. The upper panels of Fig. 14 show the ensemble density and flow of the training environment and a typical 8-s trajectory (upper panels: $\boldsymbol{\theta} = \theta_Q$). The middle panels show the corresponding expectations of the agent before (middle panels: $\boldsymbol{\theta} = 0$) and after (lower panels: $\boldsymbol{\theta} = \mu_\theta$) learning. It can be seen that before learning, the agent expects to move toward a fixed point attractor; however, after learning, it expects to navigate its environment as if it was on a Lorenz attractor.

Figure 15 shows that these expectations are fulfilled by active inference, when placed in an unsupervised environment with $\boldsymbol{\theta} = 0$. Critically, the agent's behavior is impervious to exogenous perturbations applied to the motion of the first state (see Eq. 18); this perturbation was modeled with a random Gaussian time-series smoothed with a Gaussian kernel with a standard deviation of eight time-bins. Despite these perturbations, the agent adheres to its concept of a Lorenz attractor. As above, the precision with which the agent

recapitulates its learned trajectories rests on a high degree of confidence in its dynamical priors, assured by high values of the log-precision, μ_γ^x. Although an almost trivial example, this way of prescribing desired trajectories may have pragmatic applications in engineering and robotics (cf., Prinz 2006; Manoonpong et al. 2007). This is because the trajectories prescribed under active inference are remarkably robust to noise and exogenous perturbations.

6.1 Summary

The picture that emerges from these simulations is that an effective way to instantiate desired movements is to induce prior expectations through use-dependent plasticity in the agent's generative model, so that the expected movements are reproduced automatically. The resulting priors might be the basis of hand-path priming effects, in which hand movements around obstacles induce a significant curvature in the trajectory of the movement trajectories that persists after the obstacle has been removed (Jax and Rosenbaum 2007). Indeed, most neuro-rehabilitation therapies rely on motor training and use-dependent plasticity, where the patient is guided in the reproduction or forced to reproduce desired movements with the paretic limb. (e.g., Kreisel et al. 2007). Interestingly, this use-dependent plasticity rests on neuromodulatory transmitter function (Bütefisch et al. 2002) of the sort that may encode precision (Friston 2008).

Optimizing policies under the free-energy principle is not limited to policies that are defined in terms of states with utility. This is because the only things that are required to teach an agent are exemplar trajectories. Heuristically, this corresponds to showing the agent what it is like to be optimal and then letting it fulfill its expectations in an unsupervised setting. In fact, the only things needed to specify behavior are the equations of desired motion. In principle, it may be possible to prescribe desired kinematics, say during walking, and then let the agent reproduce the associated trajectories in an uncontrolled environment with random perturbations. It is easy to imagine the attractor in Fig. 15 being replaced by more complex and possibly itinerant attractors that have desired context-sensitive properties and that could be learned by adaptive agents.

7 Discussion

This article has suggested that active inference is mandated by population dynamics and may provide a plausible explanation for some aspects of motor behavior. In particular, it suggests that motor control can be understood as prior expectations about proprioception, which enslave the peripheral motor system to fulfill them. Furthermore, active inference provides a mechanistic account of how adaptive behavior

Fig. 14 The behavior of an agent that learns to be Lorenz attractor depicted as in Fig. 11, in terms of equilibrium densities (*left*) and exemplar trajectories (*right*). *Top panels*: Dynamics of the supervised environment that offers control of the agent so that it can experience and learn desired behavior. *Middle panels*: Equivalent characterization of behavior before training, when the agent expects to be drawn to a point attractor. *Lower panels*: The same as in the previous row but after learning, when prior expectations about the environment have been transcribed from the controlled environment by learning
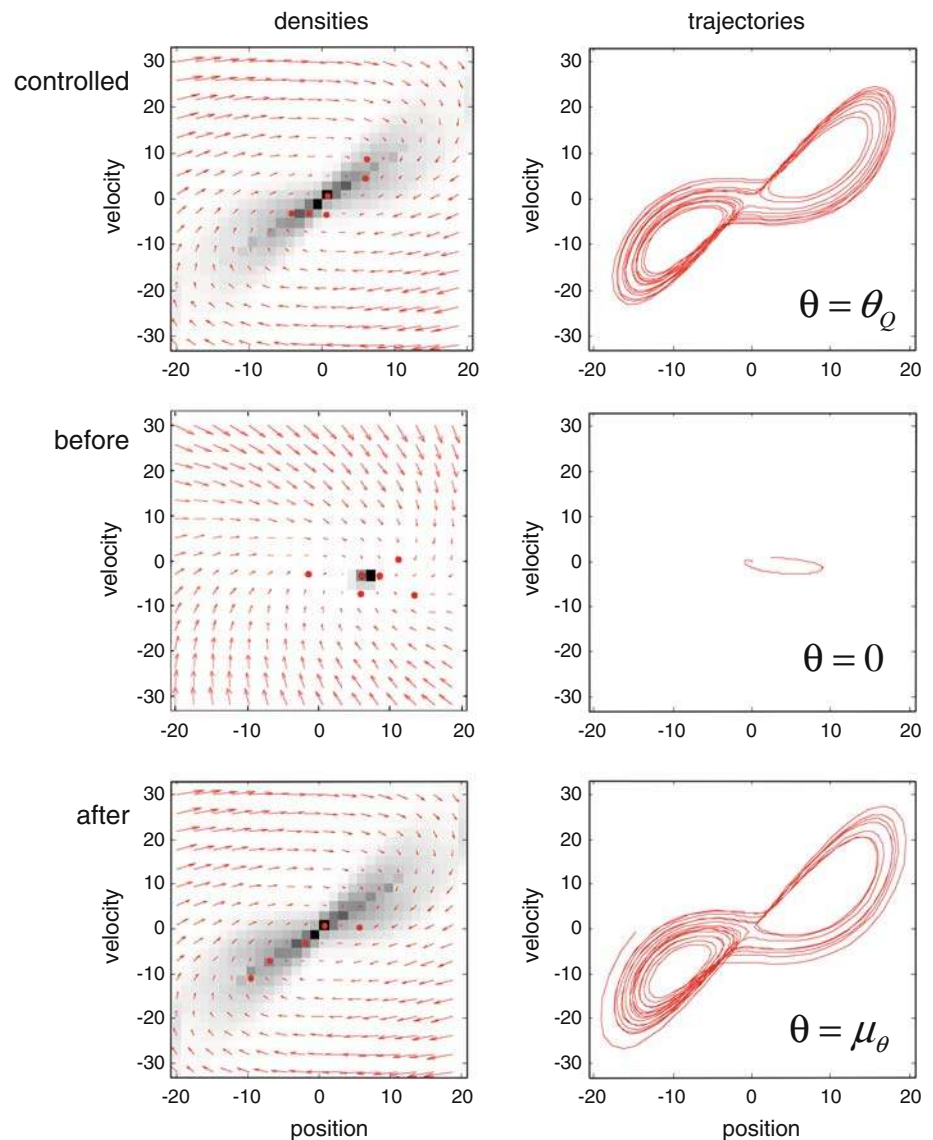
can emerge and suggests a simple alternative to dynamic programming or reinforcement learning, when prescribing desired behaviors. There is a vast literature on motor control and sensorimotor integration, which addresses a host of issues in detail, relating theoretical and empirical perspectives. We have only addressed some of these issues superficially, using simple simulations to impart the intuitions behind active inference. In the next section, we try to highlight the implications of the ideas presented above for some established areas of research.

7.1 Active inference and optimal control

Typically, dynamic programming and related approaches in optimal control theory posit a value-function of every point in state–space. This is the utility or reward expected under the current policy and is the solution to the relevant Bellman equation (Bellman 1952). A policy is then optimized to ensure that states of high value are visited with greater probability. In control theory, value acts as a guiding function by establishing gradients, which the agent can ascend (Bellman 1952; Sutton and Barto 1981; Friston et al. 1994). Similarly, in discrete models, an optimum policy selects states with the highest value (Watkins and Dayan 1992; Todorov 2006). However, under the free-energy principle, there is no value-function or Bellman equation to solve. The only states that agents aspire to are those that they expect to frequent.

In brief, active inference replaces optimal polices with prior expectations about state-transitions (e.g., movement trajectories). These priors may be acquired (empirical priors) through learning or may be innate (full priors) and optimized by natural selection. In the mountain-car example, the prior expectation (optimal policy) was acquired through learning the equations of motion of a controlled environment. This

Fig. 15 The behavior of an agent that thinks it is a Lorenz attractor portrayed using the format in Fig. 2. However, this is no ordinary attractor; its trajectories are supported purely by action (displayed as a function of time in the *lower right panel*). Action tries to suppress prediction errors on motion through its three dimensional state–space (*blue line* in *lower left panel*). These prediction errors (*red line* in *upper left panel*) are the difference between sensed and expected motion based on the agent's generative model and expected hidden states (*upper right panel*). Critically, this autonomous behavior is very resistant to random forces on the agent. This can be seen by noting that the first state has been perturbed with a smooth exogenous force (*broken blue line* on the *lower right*). Note that action counters this perturbation and the ensuing trajectories are essentially unaffected
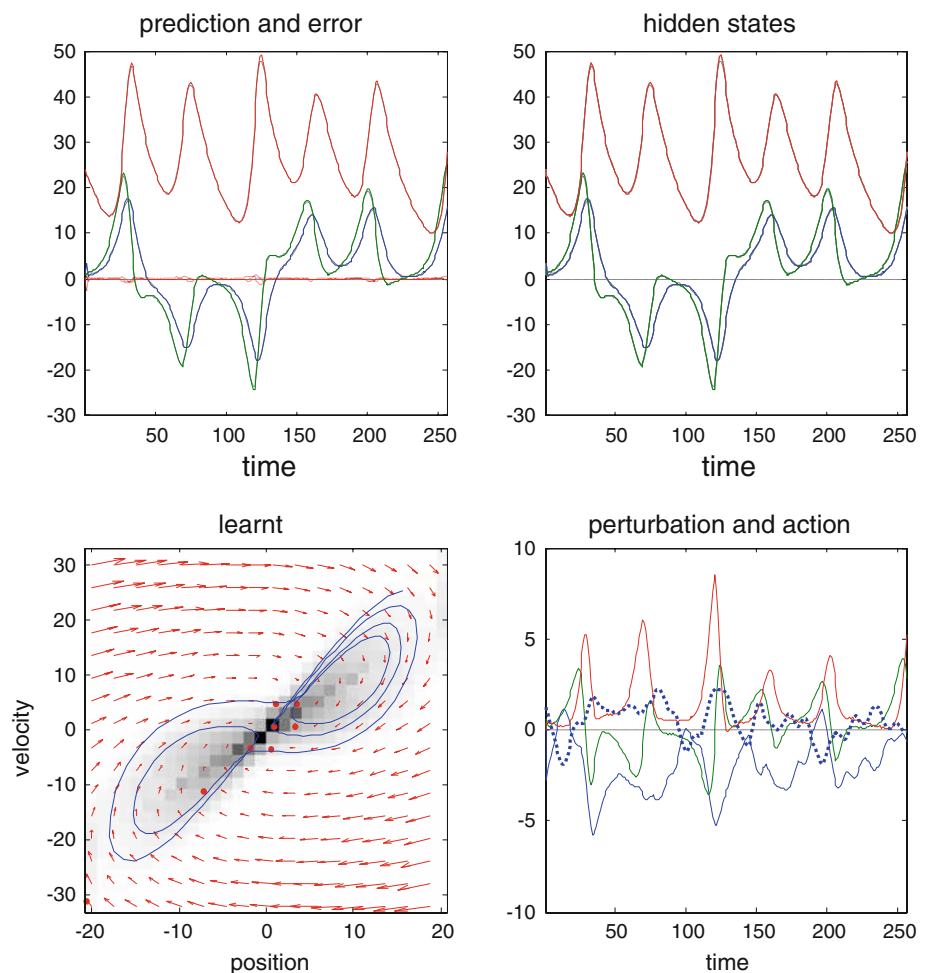
can be regarded as a form of supervised learning, because the environment was supplied by us to instill a particular behavior. In Friston (2010), we show how the same behavior can be elicited in an unsupervised setting, using prior expectations that enforce exploration of state–space until states with high utility are encountered. Similar adaptive behavior emerges under isotropic sequence order learning in a closed-loop setting (Porr and Wörgötter 2003). Both approaches rest on eliminating stationary solutions to sensory-motor dynamics that involve high-cost states. In sequence learning, this can be implemented by eliminating withdrawal reflexes using predictive reflexes that are learned in a self-referenced fashion. Does this mean active inference makes optimal control redundant? Not necessarily: active-inference simply provides a larger context, in which optimal policies become prior expectations. It is possible that the brain solves some Bellman equations to specify optimal expectations (policies) that action could pursue. Interesting developments in machine learning speak this notion; "Recently there is growing interest in using probabilistic inference methods for decision making and planning. Promising about such approaches is that they naturally extend to distributed state representations and effi-

ciently cope with uncertainty"; see Toussaint (2009) for a fuller discussion of probabilistic inference as a model of planned behavior. There are also alternative self-referenced schemes (e.g., Verschure and Voegtlin 1998; Wörgötter and Porr 2005; Tschacher and Haken 2007) that may have greater ethological and neuronal plausibility. This theme will be developed further in a forthcoming article on value-learning and free-energy.

The perspective on optimum policies as prior expectations may call for a reappraisal of reinforcement learning as an explanation for behavior and, in particular, the role of dopamine in coding unpredicted reward (Montague et al. 1995; Schultz et al. 1997; Doya 2002): If dopamine encodes precision through its classical neuromodulatory effects (cf., Yu and Dayan 2005) how can this be reconciled with the view that it encodes prediction error on reward? From a neurobiological perspective, it may be that dopamine does not encode the *prediction error of value* but the *value of prediction error*: i.e., the precision of prediction errors that drive perception and action. If this idea is right, then it speaks to a symmetry between the role of dopamine in optimizing precision in anterior (e.g., mesocortical and mesolimbic) systems
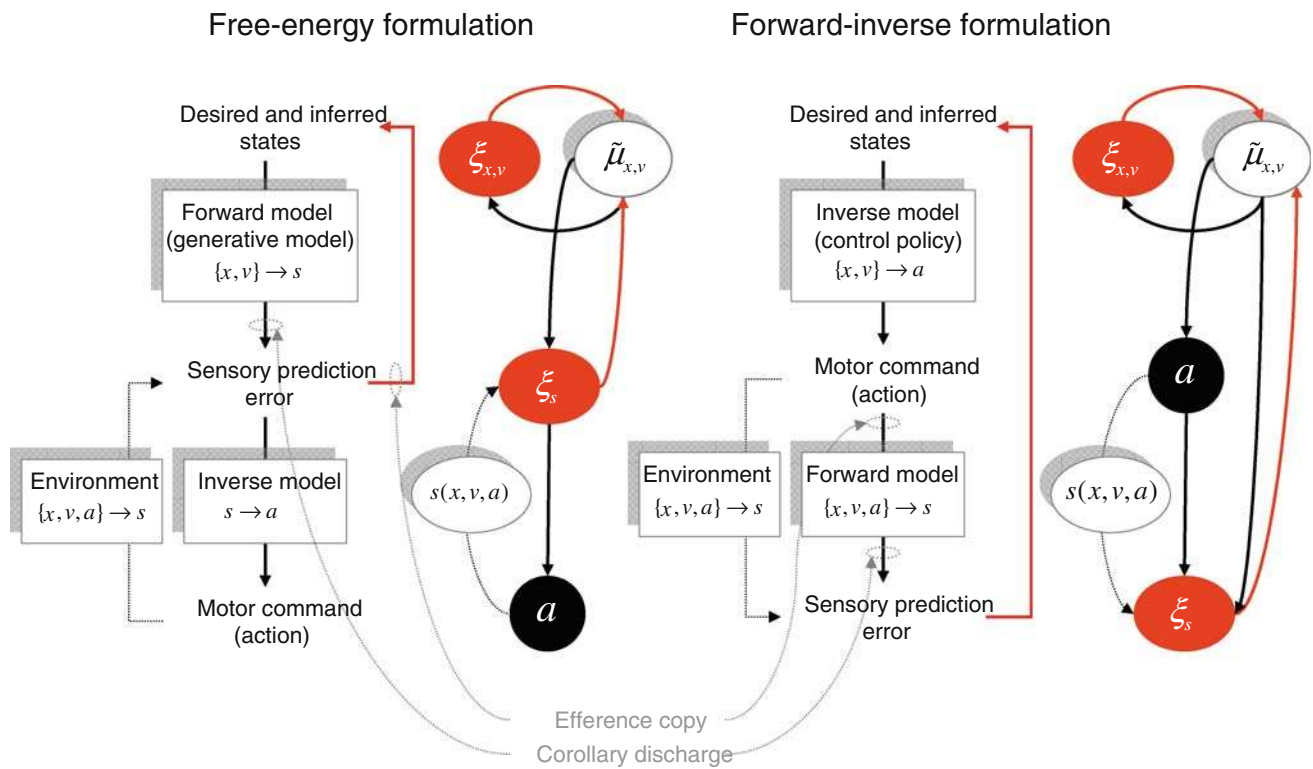
Free-energy formulation

Forward-inverse formulation

Fig. 16 Schematic highlighting the differences in functional architecture implied by the free-energy formulation and conventional motor control theories. Both schemes rest on forward and inverse models (i.e., internal models) but their nature is fundamentally different. In the free-energy formulation, action or motor commands are generated through an inverse model by sensory prediction errors, whereas in conventional formulations, motor commands are produced by an inverse model that is driven by inferred and desired states. The *red arrows* denote ascending or bottom-up effects and *black arrows* mean descending or top-down message passing. The equations in the *text boxes* indicate the mapping implicit in the corresponding model

trying to predict proprioceptive and interoceptive sensations (i.e., value-learning) and the role of acetylcholine in optimizing hierarchical inference on exteroceptive input in posterior (e.g., paralimbic and parietal) systems (i.e., attention; Disney et al. 2007). Furthermore, this perspective on dopaminergic function fits comfortably with a gating role for dopamine (O'Reilly et al. 2002) in selecting the percepts that guide action (Redgrave et al. 1999).

Another interesting aspect of precision is that it covers generalized motion. This means that action is trying to suppress a mixture of [squared] sensory prediction error on position, velocity, acceleration, and jerk etc., where the relative contributions, encoded by μ_γ^s control the smoothness of the trajectory. This mixture is formally similar to cost-functions in the motor-control literature that try to account for smooth motor trajectories that are observed empirically (Todorov and Jordan 1998; Nakano et al. 1999; Berret et al. 2008).

7.2 Inverse models in motor control

In conventional motor control theory (Wolpert and Miall 1996; Todorov and Jordan 2002; Shadmehr and Krakauer 2008), there are two models, an *inverse model* and a *forward model*. The inverse model maps from desired sensory states to action or motor commands, while the forward model maps from action to sensory consequences. In order to learn the inverse model, one needs to evaluate the consequences of action, which may be sensed with delays and noise. Forward models are used to finesse this learning, by providing surrogate prediction errors (i.e., the difference between desired sensory consequences and those predicted by the forward model). However, these forward models are not generative models of sensory dynamics that are learned in active inference. There is a fundamental difference between the functional architecture implied by active inference and conventional models of motor control (see Fig. 16). In conventional models, action is not driven by sensory prediction error but is specified by a *control policy* or *inverse* model. This inverse model computes the optimum control for achieving some desired states, which entails inverting a generative model of how action changes hidden states. In active inference, the corresponding inverse model determines how sensory prediction errors influence action. This implicit inverse model corresponds to the partial derivatives $\tilde{\varepsilon}_a$ in

Eq. 9, which describes the change in prediction error with respect to action; i.e., the change in sensory states with action. In the examples above, we used the true generative process to compute this mapping (see Eq. A3.3 in Appendix 3). However, real agents do not have access to the generative process: Does this mean they have to learn an inverse model as in conventional motor control theories? This is certainly possible; however, in active inference, the inverse model is much simpler than in conventional formulations, because it maps from sensations to action, not from hidden states to action (see Fig. 16). This is important because simple inverse mappings can be hard-wired by evolution. For example, the mapping from proprioception to action may be part of classical motor reflex arcs.

In Fig. 7, the inverse mapping from proprioceptive (e.g., stretch receptor) signals to action is simple because specific alpha and gamma motor neurons stretch specific muscle spindles. In contrast, the mapping between action and visual input is highly nonlinear and one-to-many. This means that the inverse mapping from visual prediction error to action is unlikely to be implemented in the spinal cord, as implied by the descending black connection in Fig. 7. Would removing this connection preclude visual input from contributing to sensorimotor integration? No; the proprioceptive predictions driving action depend on conditional expectations about hidden states of the body, and these expectations depend on vision (they explain away visual prediction error). This means that sensory prediction errors in all modalities can affect proprioceptive predictions through conditional expectations about the deployment and trajectory of the motor plant. This key interplay between action and perception will be the subject of a future article.

In short, the free-energy formulation replaces an inverse model, mapping from inferred states of the world to action, with the inversion of a much simpler mapping between action and sensory consequences. This may be one way of understanding the prevalence of retinotopic maps (such as those in the superficial and deep layers of the superior colliculus), which enable changes in sensory input with small eye movements to be inverted quickly and simply. The mapping between action and proprioception is even simpler: Motor neuron discharges in the ventral horn cause muscle spindles to contract and change sensory prediction errors in dorsal horn cells through classical motor reflex arcs (see Fig. 1). The argument here is that action changes proprioceptive prediction errors in a stereotyped and simple way. This may enable the inverse mapping to be specified epigenetically in peripheral or subcortical systems and conserved over context and phylogeny. The advantages of feedback control at the peripheral (e.g., spinal) level include robustness to noise or incomplete specification of the inverse model (Wolpert and Miall 1996). Furthermore, generating command signals in the periphery finesses many problems induced by temporal delays.

7.3 Forward models in motor control

In the free-energy formulation, recognition proceeds by inversion of a generative or forward model mapping from hidden states to sensations. In conventional motor control, the forward model is used to generate predicted sensory consequences of action. These ensuing sensory prediction errors are used in state-estimation and as teaching signals to optimize the inverse model or control policy. The mechanisms of this estimation and learning vary but are formally related to perceptual inference and learning (e.g., Kalman filtering and associative plasticity); however, conventional schemes only represent states, as opposed to the trajectories implicit in generalized coordinates, which renders them sensitive to temporal delays.

The key difference is that forward models in conventional motor control include the effects of action whereas, forward models in active inference do not. Action is not part of inference (i.e., inversion of a generative model) because it is a known quantity; action is just there to explain away unexpected sensory prediction errors. This does not mean that agents cannot learn about their motor plant. Indeed, the motor plant is probably one of the most important aspects of the environment for predicting sensory input (see Grafton and Hamilton 2007). This may be reflected in the preoccupation of infants with moving their limbs (and the role of rattles in promoting multimodal learning). However, generative models of the motor plant are not conventional forward models of motor control, because they map from causes to consequences (not action to consequences). In other words, generative models allow exogenous causes of movement to be perceived but not action per se (imagine a percept of action in the absence of sensations).

7.4 Efference copy and corollary discharge

In modern treatments, *efference copy* refers to a copy of the signals from the control policy (inverse) model that are passed to the forward model to create *corollary discharge* signals (predicted sensory consequences of action; see Fig. 16 and Wolpert and Miall 1996). Corollary discharge is then used to explain sensory input that can be attributed to action (*reafference*). This enables the consequences of exogenous influences (*exafference*) to be isolated. An example in lower vertebrates is seen in electric fish (Bell and Grant 1989), which send top-down corollary discharges to inhibit ascending sensory pathways. This allows the fish to disambiguate between self-generated electric organ discharges and those from other fish. However, in the free-energy formulation, there is no need to remove the sensory consequences of action because action is only there to explain away sensory prediction errors. In the free-energy scheme, efference copy and corollary discharge are simply the bottom-up and top-down

signals that mediate perception. In other words, corollary discharge can be understood as top-down predictions that inhibit ascending sensory pathways by explaining away sensory prediction error. Conversely, sensory prediction error units send top-down signals to motor control units and ascending prediction errors (efference copy) to optimize predictions (see Fig. 16). If this interpretation of efference copy and corollary discharge holds, then it highlights again the intimate relationship between action and perception. A relationship that is highlighted by formulations of the equilibrium-point hypothesis, which suggest "action and perception are accomplished in a common spatial frame of reference" (Feldman 2009).

7.5 The equilibrium-point hypothesis

The equilibrium-point hypothesis (Feldman and Levin 1995) suggests that movement is not controlled directly by the central nervous system but by top-down signals that supply the final point of a movement (in terms of the lengths of tensor and extensor muscles). In this model, muscles and the peripheral nervous system accommodate un-modeled forces and viscosity automatically. This is closely related to the current formulation, in which movements are prescribed by top-down predictions and peripheral systems respond by suppressing prediction error. Under the equilibrium-point hypothesis, these predictions correspond to *sensory input at the equilibrium point*. Under active inference, the predictions are of the *sensory trajectories that lead to the equilibrium point*. This difference is crucial, because it might account for some difficulties with the equilibrium-point hypothesis: for example, in explaining fast movements and overshoots.

Clearly, we have not touched on the detailed physiology or mechanistic issues that are addressed by conventional efference copy or equilibrium-point models. However, the formulation of sensory prediction error in terms of trajectories (i.e., in generalized coordinates of motion) may be useful for movement specification in terms of fixed-point attractors. Generalized coordinates mean that movements are specified as predicted trajectories, which include speed, acceleration, jerk etc. This resolves some of the difficulties encountered when considering neuronal propagation delays or unpredicted changes in force or viscosity. In particular, it addresses the challenge posed to the equilibrium-point hypothesis: "A theory of movement should offer insight into why we make movements the way we do and why we activate muscles in particular patterns" (Gottlieb 1998). It is interesting to note that recent equilibrium-point formulations focus on the (coordinate) frames of reference that enable anticipation and prediction: "Experimental data also imply that once a frame of reference is chosen, its attributes are modified in a feedforward way, thus enabling the brain to act in an anticipatory and predictive manner" (Feldman 2009).

7.6 Dynamic systems and optimal control

Generative models in the free-energy formulation are perceptual models that preclude action as a potential cause of sensory perturbations. Action per se is used to suppress unexplained prediction errors. This means that, from the point of view of the agent, there is no difference between navigating in a controlled and uncontrolled environment; the agent does not know that the expected sensory trajectories are being caused by its own action; it experiences the world as if it were delivering predicable streams of sensory input. This perspective could reconcile some differences between dynamic systems and optimal control treatments of computational motor control: "The dynamic system approach emphasizes motor control as a process of self-organization between an animal and its environment. … In contrast, optimal control approaches view motor control as the evolutionary or development result of a nervous system that tries to optimize rather general organizational principles" (see Schaal et al. 2007). In active inference, recognition dynamics optimize free-energy: i.e., self-organizing dynamical exchanges with the environment that optimizes neural processes and connectivity under a general free-energy principle. Crucially, free-energy and surprise rest on prior expectations, which can only be optimized by evolutionary or developmental processes.

8 Conclusion

In summary, we have shown how the free-energy principle can be motivated from the need for agents to maintain their exchange with the environment in equilibrium. We have considered behavior that would emerge under this principle and have shown how it can be harnessed to optimize policies usually addressed with optimum control theory. Underpinning this study is a unifying approach to action and perception: perceptual learning and inference is necessary to induce prior expectations about the sensorium and action is engaged to resample the world to fulfill these expectations. This places perception and action in intimate relation and accounts for the both with the same principle. Furthermore, this principle can be implemented in a simple and biologically plausible fashion. The same scheme used in this article has been used to simulate a range of biological processes: ranging from perceptual categorization of bird-song (Kiebel et al. 2008) to perceptual learning during the mismatch negativity paradigm (Friston et al. 2006). Furthermore, all the simulations in those cited articles and in this article use just one Matlab routine (see Appendix 3). If these ideas are valid, then they speak to a fundamental role for perception in action (Wolpert et al. 1995; Shadmehr and Krakauer 2008; Bays and Wolpert 2007; Tseng et al. 2007; Wei and Körding 2009).

Appendices

Appendix 1

Here, we show why the entropy of hidden states is bounded by the entropy of sensory states and a sensory mapping term. We then discuss why the minimizing sensory entropy is sufficient to reduce the entropy of hidden states. The sensory states $\tilde{s} \in S$ are an additive mixture of a function of the hidden states $\tilde{x} \in X$ plus some generalized random fluctuations

$$\tilde{s} = g(\tilde{x}, \theta) + \tilde{z} \tag{A1.1}$$

Because $\tilde{x}$ and $\tilde{z} \in Z$ are statistically independent, we have (Eq. 6.4.6 in Jones 1979, p. 149)

$$
\begin{aligned}
I(S, Z) &= H(S|m) - H(S|Z, m) \\
&= H(S|m) - H(X|m) - \int p(\tilde{x}|m) \ln |g_{\tilde{x}}| \mathrm{d}\tilde{x}
\end{aligned}
\tag{A1.2}
$$

Here and below a subscript denotes differentiation: i.e., $g_x := \partial_x g$. In (A1.1) $I(S, Z) = D(p(\tilde{s}, \tilde{z}|m)||p(\tilde{s}|m)p(\tilde{z})) \geq 0$ is the mutual information between the sensory states and noise. This non-negative cross-entropy or Kullback–Leibler divergence (see Theorem 6.5; Jones 1979, p. 151) means the entropy of the sensory states is always greater than the entropy of the sensory map of hidden states

$$H(S|m) \geq H(X|m) + \int p(\tilde{x}|m) \ln |g_{\tilde{x}}| \mathrm{d}\tilde{x} \tag{A1.3}$$

The gradient $g_{\tilde{x}}$ is of the sensory mapping with respect to the hidden states. The integral n (A1.3) reflects the fact that entropy is not invariant to a change of variables and rests on the assumption that the sensory mapping $g : X \to S$ is diffeomorphic. This assumption could be considered as a prediction that sensory mappings must be diffeomorphic (i.e., bijective and smooth). A diffeomorphism requires the dimensionality of the hidden and sensory state–spaces to be equal. This is assured by the fact we are dealing with generalized states that can be truncated at any arbitrarily high order. For example, if we had n hidden states in m generalized coordinates motion, then we would consider m sensory states in n generalized coordinates, such that $\dim(\tilde{x}) = \dim(\tilde{s}) = n \times m$. Finally, rearranging (A1.3) gives Eq. 1 in the main text.

Minimizing sensory entropy

Because entropy is not invariant under a change of variables, the sensory mapping is a key determinant of sensory entropy. Note that agents cannot change the mapping per se; they can only change the hidden states that are mapped to sensory input. However, the astute reader will note that nonlinearities in the sensory mapping mean that changing some hidden states will affect the sensory mapping of others (e.g., closing ones eyes). Does this mean that sensory entropy can be suppressed by simply avoiding sensory input? No—because entropy is conditioned on the agent (it is the average surprise given a particular agent). Agents who are not surprised by the absence of sensations are unlikely to exist because they are unable to navigate the environment and maintain an equilibrium density. This argument is at the heart of the free-energy formulation and appeals to the same tautology as adaptive fitness in natural selection. In the present context, it suggests that sensory channels that are critical for survival should be privileged and cannot be occluded (e.g., nociceptive pain signals and interoception signaling hunger, thirst etc.). Second, it suggests that agents should find the absence of sensory information surprising. We develop this theme in more detail in the context of value-learning (Friston et al., in preparation). In brief, prior expectations (cf. optimal policies) about the trajectories of hidden states induce exploration and sampling of the environment. These priors are part of the model on which entropy is conditioned, which reduce the entropy of hidden states. This reduces sensory entropy, which is a lower bound on free-energy. Put simply, the free-energy formulation assumes that agents have the right priors. This assumption cannot be violated, because agents who do not have the right priors cannot exist.

Appendix 2

Here, we derive the various formations of free-energy and show how they relate to each other. We start with the quantity we want to bound: namely, surprise or log-evidence associated with some sensory states $\tilde{s}$ that have been caused by some unknown quantities Ψ (dropping conditional dependency on model m for clarity)

$$-\ln p(\tilde{s}) = -\ln \int p(\tilde{s}, \Psi) \mathrm{d}\Psi \tag{A2.1}$$

We now simply add a non-negative cross-entropy or divergence between some arbitrary (recognition) density $q(\Psi) := q(\Psi|\mu)$ and the posterior density $p(\Psi|\tilde{s})$ to create a free-energy bound on surprise

$$
\begin{aligned}
F &= -\ln p(\tilde{s}) + \int q(\Psi) \ln \frac{q(\Psi)}{p(\Psi|\tilde{s})} \mathrm{d}\Psi \\
&= -\ln p(\tilde{s}) + D(q(\Psi)||p(\Psi|\tilde{s}))
\end{aligned}
\tag{A2.2}
$$

Fig. 17 Schematic summary of the relationships among the various energies that participate in the free-energy formulation. The *right hand panel* shows that free-energy upper bounds sensory surprise, which in turn bounds implicit surprise about hidden states of the world (under some simplifying assumptions). These bounding relationships are assured by the divergences that separate them, which are (by construction) non-negative. The associated time or path-integrals of the energies are shown on the *left*. These are called 'actions' and are formally equivalent to entropy (i.e., average surprise). See Appendix 2 for a more detailed discussion
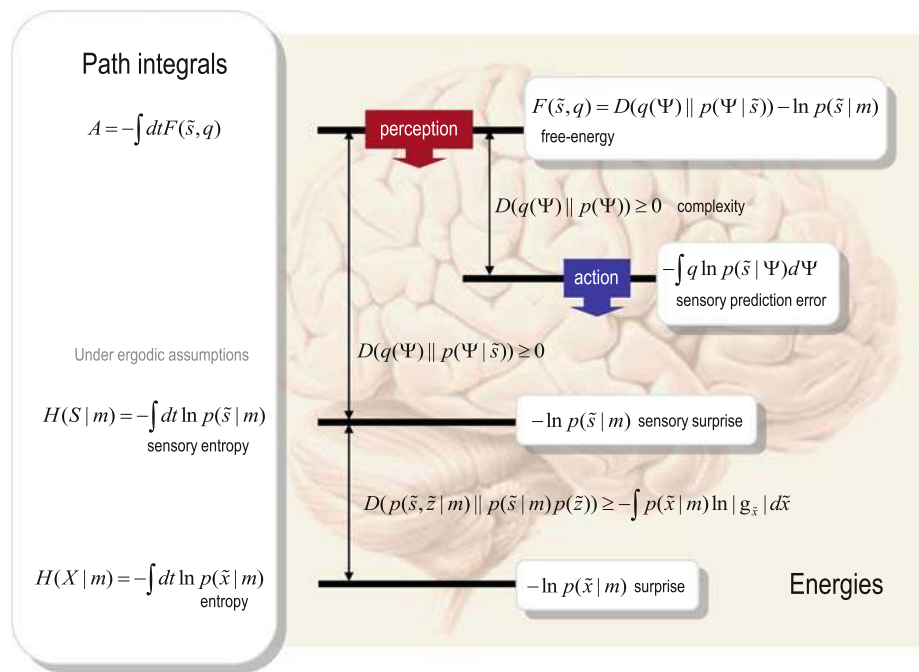
The cross entropy term is non-negative by Gibb's inequality. Because surprise depends only on sensory states, we can bring it inside the integral and use $p(\Psi, \tilde{s}) = p(\Psi|\tilde{s})p(\tilde{s})$ to show free-energy is the Gibb's energy $-\ln p(\Psi, \tilde{s})$ expected under $q(\Psi)$, minus its entropy

$$
\begin{aligned}
F &= \int q(\Psi) \ln \frac{q(\Psi)}{p(\Psi|\tilde{s})p(\tilde{s})} d\Psi \\
&= \int q(\Psi) \ln \frac{q(\Psi)}{p(\Psi, \tilde{s})} d\Psi \\
&= -\int q(\Psi) \ln p(\Psi, \tilde{s}) d\Psi + \int q(\Psi) \ln q(\Psi) d\Psi \\
&= -\langle \ln p(\Psi, \tilde{s}) \rangle_q + \langle \ln q(\Psi) \rangle_q
\end{aligned}
\tag{A2.3}
$$

A final rearrangement, using $p(\Psi, \tilde{s}) = p(\tilde{s}|\Psi)p(\Psi)$, shows that free-energy is also complexity minus accuracy, where complexity is the divergence between the recognition density $q(\Psi)$ and the prior density $p(\Psi)$

$$
\begin{aligned}
F &= \int q(\Psi) \ln \frac{q(\Psi)}{p(\tilde{s}|\Psi)p(\Psi)} d\Psi \\
&= -\int q(\Psi) \ln p(\tilde{s}|\Psi) d\Psi + \int q(\Psi) \ln \frac{q(\Psi)}{p(\Psi)} d\Psi \\
&= -\langle \ln p(\tilde{s}|\Psi) \rangle_q + D(q(\Psi)\|p(\Psi))
\end{aligned}
\tag{A2.4}
$$

Equations (A2.2), (A2.3) and (A2.4) are the three formulations used in the main text. Figure 17 provides a graphical summary of these relationships.

Appendix 3

The simulations in this article involve integrating time-varying states in both the environment and the agent as a single system, which can be modeled with the following ordinary differential equation

$$
\dot{u} = \begin{bmatrix} \dot{\tilde{s}} \\ \dot{\tilde{x}} \\ \dot{\tilde{v}} \\ \dot{\tilde{z}} \\ \dot{\tilde{w}} \\ \dot{\tilde{\mu}}_x \\ \dot{\tilde{\mu}}_v \\ \dot{\tilde{\eta}} \\ \dot{a} \end{bmatrix} = \begin{bmatrix} D\mathbf{g} + D\tilde{\mathbf{z}} \\ \mathbf{f} + \tilde{\mathbf{w}} \\ D\tilde{v} \\ D\tilde{z} \\ D\tilde{w} \\ D\tilde{\mu}^x - F_{\tilde{\mu}_x} \\ D\tilde{\mu}^v - F_{\tilde{\mu}_v} \\ D\tilde{\eta} \\ -\partial_a F \end{bmatrix}
\tag{A3.1}
$$

In order to update these states, we use a local linearization: $\Delta u = (\exp(\Delta t \Im) - I)\Im(t)^{-1}\dot{u}$ over time steps of Δt, where

$$
\Im = \frac{\partial \dot{u}}{\partial u} =
$$

$$
\begin{bmatrix}
0 & D\mathbf{g}_{\tilde{x}} & D\mathbf{g}_{\tilde{v}} & D & 0 & 0 & \cdots & & 0 \\
& \mathbf{f}_{\tilde{x}} & \mathbf{f}_{\tilde{v}} & & I & & & & \mathbf{f}_a \\
\vdots & & D & & \vdots & \vdots & & & 0 \\
& & & D & 0 & & & & \vdots \\
0 & & \cdots & 0 & D & 0 & & \cdots & \\
-F_{\tilde{\mu}_x\tilde{y}} & & \cdots & & 0 & D - F_{\tilde{\mu}_x\tilde{\mu}_x} & -F_{\tilde{\mu}_x\tilde{\mu}_v} & -F_{\tilde{\mu}_x\tilde{\eta}} & -F_{\tilde{\mu}_x a} \\
-F_{\tilde{\mu}_v\tilde{y}} & & & & & -F_{\tilde{\mu}_v\tilde{\mu}_x} & D - F_{\tilde{\mu}_v\tilde{\mu}_v} & -F_{\tilde{\mu}_v\tilde{\eta}} & -F_{\tilde{\mu}_v a} \\
& & & & & 0 & 0 & D & 0 \\
-F_{a\tilde{y}} & & \cdots & & 0 & -F_{a\tilde{\mu}_x} & -F_{a\tilde{\mu}_v} & -F_{a\tilde{\eta}} & -F_{aa}
\end{bmatrix}
\tag{A3.2}
$$

The conditional expectations of the parameters and precisions are updated after every simulated trail as described in Friston (2008). Equation A3.2 may look complicated but can be evaluated automatically using numerical derivatives. All the simulations in this article used just one routine—spm_ADEM.m. Demonstrations of this scheme are available as part of the SPM software (http://www.fil.ion.ion.ucl.ac.uk/spm; DEM_demo.m), and reproduce the examples in the main text.

In order to include action in the free-energy scheme, we simply add action to the states of the generative process producing sensory data and specify its dynamics as a gradient descent on free-energy. Because action can only affect the free-energy through the sensory data that are sampled, it can only affect sensory prediction error. Therefore, action dynamics are prescribed by

$$
\begin{aligned}
\dot{a} &= -F_a = -\tilde{\varepsilon}_a^T \xi \\
\tilde{\varepsilon}_a &= \mathbf{g}_{\tilde{\mathbf{x}}} \sum_i D^{-i} \mathbf{f}_{\tilde{\mathbf{x}}}^{i-1} \mathbf{f}_a
\end{aligned}
\tag{A3.3}
$$

The partial derivative of the error with respect to action is the partial derivative of the sensory data with respect to action and is specified by the generative process. In biologically plausible instances of this scheme, this partial derivative would have to be computed on the basis of a mapping from action to sensory consequences (see Sect. 7). One might assume that this mapping was sufficiently simple to be hardwired or, when dependent upon sensory states, was based on some interpolation or Gaussian process model. This will be demonstrated in future simulations.

Note that (A3.1) is formulated in generalized coordinates of motion. Although there is no need to express action in generalized coordinates (because it is real-world state), it can change the high-order motion of other environmental states. For example, in the mountain-car system, action changes the second-order motion of position (i.e., acceleration) and plays the role of a Newtonian force.

References

Abbott LF, Varela JA, Sen K, Nelson SB (1997) Synaptic depression and cortical gain control. Science 275(5297):220–224

Andersen RA (1989) Visual and eye movement factions of the posterior parietal cortex. Annu Rev Neurosci 12: 377–405

Anosov DV (2001) Ergodic theory. In Hazewinkel M (ed) Encyclopaedia of mathematics. Kluwer Academic Publishers. ISBN 978-1556080104

Ballard DH, Hinton GE, Sejnowski TJ (1983) Parallel visual computation. Nature306:21–26

Barlow HB (1969) Pattern recognition and the responses of sensory neurons. Ann N Y Acad Sci156:872–881

Bays PM, Wolpert DM (2007) Computational principles of sensorimotor control that minimize uncertainty and variability. J Physiol 578(Pt 2):387–396

Bell CC, Grant K (1989) Corollary discharge inhibition and preservation of temporal information in a sensory nucleus of mormyrid electric fish. J Neurosci 9(3):1029–1044

Bellman R (1952) On the theory of dynamic programming. Proc Natl Acad Sci USA 38:716–719

Bernard C (1974) Lectures on the phenomena common to animals and plants (trans: Hoff HE, Guillemin R, Guillemin L). Charles C Thomas, Springfield. ISBN 978-0398028572

Berret B, Darlot C, Jean F, Pozzo T, Papaxanthis C, Gauthier JP (2008) The inactivation principle: mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements. PLoS Comput Biol 4(10): e1000194

Bruyn JL, Mason AH (2009) Temporal coordination during bimanual reach-to-grasp movements: the role of vision. Q J Exp Psychol (Colchester) 5:1–15

Bütefisch CM, Davis BC, Sawaki L, Waldvogel D, Classen J, Kopylev L, Cohen LG (2002) Modulation of use-dependent plasticity by d-amphetamine. Ann Neurol 51(1):59–68

Camerer CF (2003) Behavioural studies of strategic thinking in games. Trends Cogn Sci 7(5): 225–231

Crauel H, Flandoli F (1994) Attractor for random dynamical systems. Probab Theory Relat Fields 100:365–393

Daw ND, Doya K (2006) The computational neurobiology of learning and reward. Curr Opin Neurobiol 16(2): 199–204

Dayan P, Hinton GE, Neal RM (1995) The Helmholtz machine. Neural Comput 7:889–904

Deneve S (2008) Bayesian spiking neurons I: inference. Neural Comput 20(1):91–117

Diedrichsen J, Dowling N (2009) Bimanual coordination as task-dependent linear control policies. Hum Mov Sci [Epub ahead of print]

Diedrichsen J, Verstynen T, Hon A, Zhang Y, Ivry RB (2007) Illusions of force perception: the role of sensori-motor predictions, visual information, and motor errors. J Neurophysiol 97(5):3305–3313

Disney AA, Aoki C, Hawken MJ (2007) Gain modulation by nicotine in macaque v1. Neuron 56(4):701–713

Doya K (2002) Metalearning and neuromodulation. Neural Netw 15(4–6):495–506

Evans DJ (2003) A non-equilibrium free-energy theorem for deterministic systems. Mol Phys 101:1551–1554

Feldman AG (2009) New insights into action-perception coupling. Exp Brain Res 194(1):39–58

Feldman AG, Levin MF (1995) The origin and use of positional frames of reference in motor control. Behav Brain Sci 18:723–806

Feynman RP (1972) Statistical mechanics. Benjamin, Reading

Fourneret P, Jeannerod M (1998) Limited conscious monitoring of motor performance in normal subjects. Neuropsychologia 36(11):1133–1140

Friston K (2005) A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci 360(1456):815–836

Friston K (2008) Hierarchical models in the brain. PLoS Comput Biol 4(11):e1000211. PMID: 18989391

Friston K (2010) The free-energy principle: a unified brain theory? Nat Rev Neurosci 11(2):127–138

Friston K, Stephan KE (2007) Free-energy and the brain. Synthese 159:417–458

Friston KJ, Tononi G, Reeke GNJr, Sporns O, Edelman GM (1994) Value-dependent selection in the brain: simulation in a synthetic neural model. Neuroscience 59(2):229–243

Friston K, Kilner J, Harrison L (2006) A free-energy principle for the brain. J Physiol (Paris) 100(1–3):70–87

Friston KJ, Trujillo-Barreto N, Daunizeau J (2008) DEM: a variational treatment of dynamic systems. NeuroImage 41(3): 849–885

Friston KJ, Daunizeau J, Kiebel SJ (2009) Reinforcement learning or active inference? PLoS One 4(7):e6421

Gontar V (2000) Entropy principle of extremality as a driving force in the discrete dynamics of complex and living systems. Chaos Solitons Fractals 11:231–236

Gottlieb GL (1998) Rejecting the equilibrium-point hypothesis. Motor Control 2(1):10–12

Grafton ST, Hamilton AF (2007) Evidence for a distributed hierarchy of action representation in the brain. Hum Mov Sci 26(4):590–616

Grafton ST, Schmitt P, Van Horn J, Diedrichsen J (2008) Neural substrates of visuomotor learning based on improved feedback control and prediction. Neuroimage 39(3):1383–1395

Helmholtz H (1860/1962). Handbuch der physiologischen optik, vol 3 (English trans: Southall JPC, ed). Dover, New York

Hinton GE, von Camp D (1993) Keeping neural networks simple by minimizing the description length of weights. In: Proceedings of COLT-93, pp 5–13

Huffman KJ, Krubitzer L (2001) Area 3a: topographic organization and cortical connections in marmoset monkeys. Cereb Cortex 11:849–867

Jax SA, Rosenbaum DA (2007) Hand path priming in manual obstacle avoidance: evidence that the dorsal stream does not only control visually guided actions in real time. J Exp Psychol Hum Percept Perform 33(2):425–441

Jones DS (1979) Elementary information theory. Clarendon Press, New York

Kawato M, Hayakawa H, Inui T (1993) A forward-inverse optics model of reciprocal connections between visual cortical areas. Network 4:415–422

Kersten D, Mamassian P, Yuille A (2004) Object perception as Bayesian inference. Annu Rev Psychol 55:271–304

Kiebel SJ, Daunizeau J, Friston KJ (2008) A hierarchy of time-scales and the brain. PLoS Comput Biol 4(11):e1000209. PMID: 19008936

Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 27(12):712–719

Körding KP, Wolpert DM (2004) Bayesian integration in sensorimotor learning. Nature 427(6971):244–247

Körding KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L (2007) Causal inference in multisensory perception. PLoS ONE 2(9):e943

Kreisel SH, Hennerici MG, Bäzner H (2007) Pathophysiology of stroke rehabilitation: the natural course of clinical recovery, use-dependent plasticity and rehabilitative outcome. Cerebrovasc Dis 23(4):243–255

Kulvicius T, Porr B, Wörgötter F (2007) Development of receptive fields in a closed-loop behavioral system. Neurocomputing 70:2046–2049

Lee TS, Mumford D (2003) Hierarchical Bayesian inference in the visual cortex. J Opt Soc Am A Opt Image Sci Vis 20:1434–1448

Linsker R (1990) Perceptual neural organization: some approaches based on network models and information theory. Annu Rev Neurosci 13:257–281

Liu D, Todorov E (2007) Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. J Neurosci 27(35):9354–9368

Mackay DJC (1992) Information-based objective functions for active data selection. Neural Comput 4:590–604

MacKay DJC (1995) Free-energy minimization algorithm for decoding and cryptoanalysis. Electron Lett 31:445–447

Manoonpong P, Geng T, Kulvicius T, Porr B, Wörgötter F (2007) Adaptive, fast walking in a biped robot under neuronal control and learning. PLoS Comput Biol 3(7):e134

Maturana HR, Varela F (1972) De máquinas y seres vivos. Editorial Universitaria, Santiago. English version: "Autopoiesis: the orga-

nization of the living," in Maturana HR, Varela FG (1980) Auto-poiesis and cognition. Reidel, Dordrecht

Montague PR, Dayan P, Person C, Sejnowski TJ (1995) Bee foraging in uncertain environments using predictive Hebbian learning. Nature 377(6551):725–728

Mumford D (1992) On the computational architecture of the neocortex. II. The role of cortico-cortical loops. Biol Cybern 66:241–251

Mussa Ivaldi FA, Morasso P, Zaccaria R (1988) A distributed model for representing and regularizing motor redundancy. Biol Cybern 60(1):1–16

Nakano E, Imamizu H, Osu R, Uno Y, Gomi H, Yoshioka T, Kawato M (1999) Quantitative examinations of internal representations for arm trajectory planning: minimum commanded torque change model. J Neurophysiol 81(5):2140–2155

Neal RM, Hinton GE (1998) A view of the EM algorithm that justifies incremental, sparse, and other variants. In: Jordan MI (ed) Learning in graphical models. Kluwer Academic Publishers, Dordrecht pp 355–368

Olshausen BA, Field DJ (1996) Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381:607–609

O'Reilly RC, Noelle DC, Braver TS, Cohen JD (2002) Prefrontal cortex and dynamic categorization tasks: representational organization and neuromodulatory control. Cereb Cortex 12(3):246–257

Paulignan Y, MacKenzie C, Marteniuk R, Jeannerod M (1991) Selective perturbation of visual input during prehension movements. 1. The effects of changing object position. Exp Brain Res 83(3):502–512

Paulin MG (2005) Evolution of the cerebellum as a neuronal machine for Bayesian state estimation. J Neural Eng 2(3):S219–S234

Porr B, Wörgötter F (2003) Isotropic sequence order learning. Neural Comput 15(4):831–864

Prinz AA (2006) Insights from models of rhythmic motor systems. Curr Opin Neurobiol 16(6):615–620

Rao RP, Ballard DH (1998) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive field effects. Nat Neurosci 2:79–87

Redgrave P, Prescott TJ, Gurney K (1999) The basal ganglia: a vertebrate solution to the selection problem?. Neuroscience 89(4):1009–1023

Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforce-ment. In: Black AH, Prokasy WF (eds) Classical conditioning II: current research and theory. Appleton Century Crofts, New York pp 64–99

Schaal S, Mohajerian P, Ijspeert A (2007) Dynamics systems vs. optimal control—a unifying view. Prog Brain Res 165:425–445

Scheidt RA, Reinkensmeyer DJ, Conditt MA, Rymer WZ, Mussa-Ivaldi FA (2000) Persistence of motor adaptation during constrained, multi-joint, arm movements. J Neurophysiol 84:853–862

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593–1599

Schweitzer F (2003) Brownian agents and active particles: collective dynamics in the natural and social sciences. Springer Series in Synergetics, 1st ed 2003, 2nd printing 2007. ISBN: 978-3-540-73844-2

Shadmehr R, Krakauer JW (2008) A computational neuroanatomy for motor control. Exp Brain Res 185(3):359–381

Sutton RS (1996) Generalization in reinforcement learning: successful examples using sparse coarse coding. Adv Neural Inf Process Syst 8:1038–1044

Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88(2):135–170

Tatler BW, Wade NJ (2003) On nystagmus, saccades, and fixations. Perception 32(2):167–184

Todorov E (2006) Linearly-solvable Markov decision problems. In: Scholkopf B et al (eds) Advances in neural information processing systems, vol 19. MIT Press, Cambridge, pp 1369–1376

Todorov E, Jordan MI (1998) Smoothness maximization along a pre-defined path accurately predicts the speed profiles of complex arm movements. J Neurophysiol 80(2):696–714

Todorov E, Jordan MI (2002) Optimal feedback control as a theory of motor coordination. Nat Neurosci 5(11):1226–1235

Toussaint M (2009) Probabilistic inference as a model of planned behavior. Künstliche Intelligenz Ger Artif Intell J (in press)

Tschacher W, Haken H (2007) Intentionality in non-equilibrium systems? The functional aspects of self-organized pattern formation. New Ideas Psychol 25:1–15

Tseng YW, Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ (2007) Sensory prediction errors drive cerebellum-dependent adaptation of reaching. J Neurophysiol 98(1):54–62

Verschure PF, Voegtlin T (1998) A bottom up approach towards the acquisition and expression of sequential representations applied to a behaving real-world device: Distributed Adaptive Control III. Neural Netw 11(7–8):1531–1549

Verschure PF, Voegtlin T, Douglas RJ (2003) Environmentally mediated synergy between perception and behavior in mobile robots. Nature 425:620–624

Voss M, Ingram JN, Wolpert DM, Haggard P (2008) Mere expectation to move causes attenuation of sensory signals. PLoS ONE 3(8):e2866

Watkins CJCH, Dayan P (1992) Q-learning. Mach Learn 8:279–292

Wei K, Körding KP (2009) Relevance of error: what drives motor adaptation?. J Neurophysiol 101(2):655–664

Wolpert DM, Miall RC (1996) Forward models for physiological motor control. Neural Netw 9(8):1265–1279

Wolpert DM, Ghahramani Z, Jordan MI (1995) An internal model for sensorimotor integration. Science 269(5232):1880–1882

Wörgötter F, Porr B (2005) Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. Neural Comput 17(2):245–319

Yu AJ, Dayan P (2005) Uncertainty, neuromodulation and attention. Neuron 46:681–692