Deep temporal models and active inference

Karl J. Friston[a,*], Richard Rosch[a], Thomas Parr[a], Cathy Price[a], Howard Bowman[b,c]

[a] Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, WC1N 3BG, United Kingdom
[b] Centre for Cognitive Neuroscience and Cognitive Systems and the School of Computing, University of Kent at Canterbury, Canterbury, Kent, CT2 7NF, United Kingdom
[c] School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

ABSTRACT

How do we navigate a deeply structured world? Why are you reading this sentence first – and did you actually look at the fifth word? This review offers some answers by appealing to active inference based on deep temporal models. It builds on previous formulations of active inference to simulate behavioural and electrophysiological responses under hierarchical generative models of state transitions. Inverting these models corresponds to sequential inference, such that the state at any hierarchical level entails a sequence of transitions in the level below. The deep temporal aspect of these models means that evidence is accumulated over nested time scales, enabling inferences about narratives (i.e., temporal scenes). We illustrate this behaviour with Bayesian belief updating – and neuronal process theories – to simulate the epistemic foraging seen in reading. These simulations reproduce perisaccadic delay period activity and local field potentials seen empirically. Finally, we exploit the deep structure of these models to simulate responses to local (e.g., font type) and global (e.g., semantic) violations; reproducing mismatch negativity and P300 responses respectively.

1. Introduction

In recent years, we have applied the free energy principle to generative models of worlds that can be described in terms of discrete states in an attempt to understand the embodied Bayesian brain. The resulting active inference scheme (for Markov decision processes) has been applied in a variety of domains (see Table 1). This paper takes active inference to the next level and considers hierarchical models with deep temporal structure (George and Hawkins, 2009; Kiebel et al., 2009; LeCun et al., 2015). This structure follows from generative models that entertain state transitions or sequences over time. The resulting model enables inference about narratives with deep temporal structure (c.f., sequential scene construction) of the sort seen in reading. In short, equipping an agent or simulated subject with deep temporal models allows them to accumulate evidence over different temporal scales to find the best explanation for their sensations.

This paper has two agendas: to introduce hierarchical (deep) generative models for active inference under Markov decision processes (or hidden Markov models) and to show how their belief updating can be understood in terms of neuronal processes. The problem we focus on is how subjects deploy active vision to disambiguate the causes of their sensations. In other words, we ask how people choose where to look

next, when resolving uncertainty about the underlying conceptual, semantic or lexical causes of sensory input. This means that we are not concerned with computational linguistics *per se* but the more general problem of *epistemic foraging*, while using reading as an example.

Epistemics is at the heart of active inference, which is all about reducing surprise or uncertainty, where uncertainty is expected surprise. Technically, this means that one can describe both inference (perception) and behaviour (action) in terms of minimising a free energy functional of probabilistic or Bayesian beliefs. In this setting, variational free energy approximates surprise and expected free energy approximates uncertainty (a.k.a. entropy). This single imperative provides an inclusive account of established (normative) approaches to perception and action; for example, the principle of maximum mutual information, the principle of minimum redundancy, formulations of saliency as Bayesian surprise, risk sensitive or KL control, expected utility theory, and so on (Barlow, 1974; Itti and Baldi, 2009; Kappen et al., 2012; Ortega and Braun, 2013). Our focus here is on how subjects use accumulated beliefs about the hidden states of the world to prescribe active sampling of new information to resolve their uncertainty quickly and efficiently (Ferro et al., 2010).

Our second agenda is to translate these normative (variational) principles into neurobiology by trying to establish the construct validity

Table 1

Applications of active inference for Markov decision processes.

Application	Comment	References
Decision making under uncertainty	Initial formulation of active inference for *Markov decision processes* and *sequential policy optimisation*	(Friston et al., 2012b)
Optimal control (the mountain car problem)	Illustration of *risk sensitive or KL control* in an engineering benchmark	(Friston et al., 2012a)
Evidence accumulation: Urns task	Demonstration of how beliefs states are absorbed into a generative model	(FitzGerald et al., 2015b,c)
Addiction	Application to psychopathology	(Schwartenbeck et al., 2015c)
Dopaminergic responses	Associating dopamine with the encoding of (expected) precision provides a plausible account of dopaminergic discharges	(Friston et al., 2014 ; FitzGerald et al., 2015a)
Computational fMRI	Using Bayes optimal precision to predict activity in dopaminergic areas	(Schwartenbeck et al., 2015a)
Choice preferences and epistemics	Empirical testing of the hypothesis that people prefer to keep options open	(Schwartenbeck et al., 2015b)
Behavioural economics and trust games	Examining the effects of prior beliefs about self and others	(Moutoussis et al., 2014)
Foraging and two step mazes	Formulation of epistemic and pragmatic value in terms of *expected free energy*	(Friston et al., 2015)
Habit learning, reversal learning and devaluation	Learning as minimising variational free energy with respect to model parameters – and action selection as *Bayesian model averaging*	(FitzGerald et al., 2014; Friston et al., 2016)
Saccadic searches and scene construction	*Mean field approximation* for multifactorial hidden states, enabling high dimensional beliefs and outcomes: c.f., functional segregation	(Friston and Buzsaki, 2016; Mirza et al., 2016)
Electrophysiological responses: *place-cell activity, omission related responses, mismatch negativity, P300, phase-procession, theta-gamma coupling*	Simulating neuronal processing with a gradient descent on variational free energy; c.f., dynamic *Bayesian belief propagation* based on marginal free energy	In press
Structure learning, sleep and insight	Inclusion of parameters into expected free energy to enable structure learning via *Bayesian model reduction*	Under review
Narrative construction and reading	Hierarchical generalisation of generative model with *deep temporal structure*	Current paper

of active inference in terms of behaviour and electrophysiological responses. We do this at three levels: first, by highlighting the similarity between the message passing implied by minimising variational free energy and the neurobiology of neuronal circuits. Specifically, we try to associate the dynamics of a gradient descent on variational free energy with neuronal dynamics based upon neural mass models (Lopes da Silva, 1991). Furthermore, the exchange of sufficient statistics implicit in belief propagation is compared with the known characteristics of extrinsic (between cortical area) and intrinsic (within cortical area) neuronal connectivity. Second, we try to reproduce reading-like behaviour – in which epistemically rich information is sampled by sparse, judicious saccadic eye movements. This enables us to associate peri-saccadic updating with empirical phenomena, such as delay period activity and perisaccadic local field potentials (Kojima and Goldman-Rakic, 1982; Purpura et al., 2003; Pastalkova et al., 2008). Finally, in terms of the non-invasive electrophysiology, we try to reproduce the well-known violation responses indexed by phenomena like the mismatch negativity (MMN) and P300 waveforms in event related potential research (Strauss et al., 2015).

This paper comprises four sections. The first (Active inference and free energy) briefly reviews active inference, establishing the normative principles that underlie action and perception. The second section (Belief propagation and neuronal networks) considers action and perception, paying special attention to hierarchical generative models and how the minimisation of free energy could be implemented in the brain. The third section (Simulations of reading) introduces a particular generative model used to simulate reading and provides an illustration of the ensuing behaviour – and simulated electrophysiological responses. The final section (Simulations of classical violation responses) rehearses the reading simulations using different prior beliefs to simulate responses to violations at different hierarchical levels in the model.

2. Active inference and free energy

Active inference rests upon a generative model that is used to infer the most likely causes of observable outcomes in terms of expected states of the world. A generative model is just a probabilistic specification of how consequences (outcomes) follow from causes (states).

These states are called latent or *hidden* because they can only be inferred through observations. Clearly, observations depend upon action (e.g., where you are looking). This requires the generative model to represent outcomes under different actions or policies. Technically, expectations about (future) outcomes and their hidden causes are optimised by minimising variational free energy, which renders them the most likely (posterior) expectations about the (future) states of the world, given (past) observations. This follows because the variational free energy is an upper bound on (negative) log Bayesian model evidence; also known as surprise, surprisal or self-information (Dayan et al., 1995). Crucially, the prior probability of each policy (i.e., action or plan) is the free energy expected under that policy (Friston et al., 2015). This means that policies are more probable if they minimise expected surprise or resolve uncertainty.

Evaluating the expected free energy of plausible policies – and implicitly their posterior probabilities – enables the most likely action to the selected. This action generates a new outcome and the cycle of perception and action starts again. The resulting behaviour represents a principled sampling of sensory cues that has epistemic, uncertainty reducing and pragmatic, surprise reducing aspects. The pragmatic aspect follows from prior beliefs or preferences about future outcomes that makes some outcomes more surprising than others. For example, I would not expect to find myself dismembered or humiliated – and would therefore avoid these surprising state of affairs. On this view, behaviour is dominated by epistemic imperatives until there is no further uncertainty to resolve. At this point pragmatic (prior) preferences predominate, such that explorative behaviour gives way to exploitative behaviour. In this paper, we focus on epistemic behaviour and only use prior preferences to establish a task or instruction set. Namely, to report a categorical decision when sufficiently confident; i.e., under the prior belief one does not make mistakes.

2.1. Hierarchical generative models

We are concerned here with hierarchical generative models in which the outcomes of one level generate the hidden states at a lower level. Fig. 1 provides a schematic of this sort of model. Outcomes depend upon hidden states, while hidden states unfold in a way that
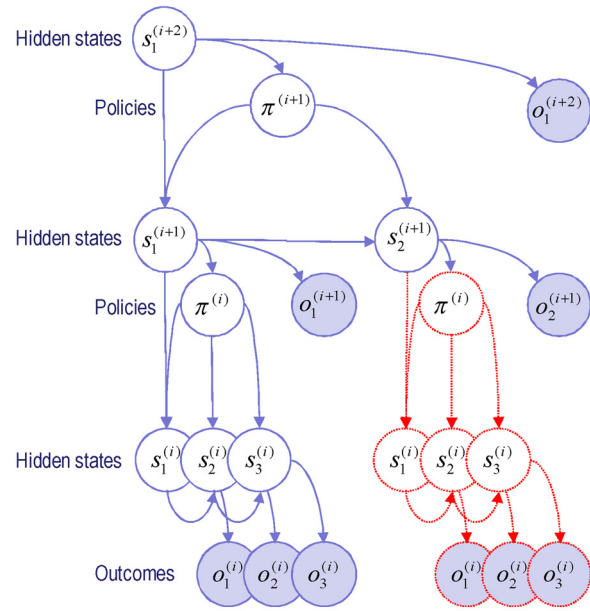
Fig. 1. Generative model and (approximate) posterior. Left panel: these equations specify the generative model. A generative model is the joint probability of outcomes or consequences and their (latent or hidden) causes, see top equation. Usually, the model is expressed in terms of a *likelihood* (the probability of consequences given causes) and *priors* over causes. When a prior depends upon a random variable it is called an *empirical prior*. Here, the likelihood is specified by an array **A** whose elements are the probability of an outcome under every combination of hidden states. The empirical priors pertain to probabilistic transitions (in the **B** arrays) among hidden states that can depend upon action, which is determined probabilistically by policies (sequences of actions encoded by π). The key aspect of this generative model is that policies are more probable *a priori* if they minimise the (path integral of) expected free energy **G,** which depends upon our prior preferences about outcomes encoded by the array **C**. Finally, the **D** arrays specified the initial state, given the state of the level above. This completes the specification of the model in terms of parameter arrays that constitute **A–D**. Bayesian model inversion refers to the inverse mapping from consequences to causes; i.e., estimating the hidden states and other variables that cause outcomes. In variational Bayesian inversion, one has to specify the form of an approximate posterior distribution, which is provided in the lower panel. This particular form uses a mean field approximation, in which posterior beliefs are approximated by the product of marginal distributions over hierarchical levels and points in time. Subscripts index time (or policy), while (bracketed) superscripts index hierarchical level. See the main text and Table 2 for a detailed explanation of the variables (italic variables represent hidden states, while bold variables indicate expectations about those states). Right panel: this Bayesian graph represents the conditional dependencies among hidden states and how they cause outcomes. Open circles are random variables (hidden states and policies) while filled circles denote observable outcomes. The key aspect of this model is its hierarchical structure that represents sequences of hidden states over time or epochs. In this model, hidden states at higher levels generate the initial states for lower levels – that then unfold to generate a sequence of outcomes: c.f., associative chaining (Page and Norris, 1998). Crucially, lower levels cycle over a sequence for each transition of the level above. This is indicated by the variables outlined in red, which are 'reused' as higher levels unfold. It is this scheduling that endows the model with deep temporal structure. Note that hidden states at any level can generate outcomes and hidden states at the lower level. Furthermore, the policies at each level depend upon the hidden states of the level above – and are in play for the sequence of state transitions at the level below. This means that hidden states can influence subordinate states in two ways: by specifying the initial states – or via policy-dependent state transitions. Please see main text and Table 2 for a definition of the variables. For clarity, time subscripts have been omitted from hidden states at level $i + 1$ (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

depends upon a sequence of actions or a *policy*. The generative model is specified by two sets of matrices (or arrays). The first set $\mathbf{A}^{(i,m)}$, maps from hidden states to the *m*-th outcome or modality at the *i*-th level; for example, exteroceptive (e.g., visual) or proprioceptive (e.g., eye position) observations. The second set: $\mathbf{B}^{(i,n)}(u)$, prescribes the transitions among the *n*-th hidden state or factor, at the *i*-th level, under action *u*. Hidden factors correspond to different states of the world, such as the location (i.e., where) and category (i.e., what) of an object. Hierarchical levels are linked by $\mathbf{D}^{(i,n)}$ that play a similar role to $\mathbf{A}^{(i,m)}$. However, instead of mapping from hidden states to outcomes they map from hidden states at the given level to the *initial states* of the *n*-th factor at the level below. A more detailed description of these parameters can be found in Table 2 and the Appendix. For simplicity, Fig. 1 assumes there is a single hidden factor and outcome modality.

The generative model in Fig. 1 generates outcomes in the following way: first, a policy (action or plan) is selected at the highest level using a softmax function of their expected free energies. Sequences of hidden states are then generated using the probability transitions specified by the selected policy (encoded in **B** matrices). These hidden states generate outcomes and initial hidden states in the level below (according to **A** and **D** matrices). In addition, hidden states can influence the expected free energy (through **C** matrices) and therefore influence the policies

that determine transitions among subordinate states. The key aspect of this generative model is that state transitions proceed at different rates at different levels of the hierarchy. In other words, the hidden state at a particular level entails a sequence of hidden states at the level below. This is a necessary consequence of conditioning the initial state at any level on the hidden states in the level above. Heuristically, this hierarchical model generates outcomes over nested timescales; like the second-hand of a clock that completes a cycle for every tick of the minute-hand that, in turn precesses more quickly than the hour hand. It is this particular construction that lends the generative model a deep temporal architecture. In other words, hidden states at higher levels contextualise transitions or trajectories of hidden states at lower levels; generating a deep dynamic narrative.

2.2. Variational free energy and inference

For any given generative model, active inference corresponds to optimising expectations of hidden states and policies with respect to variational free energy. These expectations constitute the sufficient statistics of posterior beliefs, usually denoted by the probability distribution $Q(\tilde{s}, \pi)$, where $(\tilde{s}, \pi)$ are hidden or unknown states and policies. This optimisation can be expressed mathematically as:

Table 2
Glossary of expressions (for the i-th hierarchical level of a generative model).

Expression	Description
$o_\tau^{(i,m)} \in \{0, 1\} \in \mathbb{N}^{D(i,m)}$	Outcomes in M modalities at each time point, taken to be 'one-in-K' vectors of dimension $D(i,m)$
$\tilde{o}^{(i,m)} = (o_1^{(i,m)}, ..., o_t^{(i,m)})$	Sequences of outcomes until the current time point
$s_\tau^{(i,n)} \in \{0, 1\} \in \mathbb{N}^{D(i,n)}$	Hidden states of the n-th factor at each time point and their posterior expectations under each policy
$\mathbf{s}_{\pi,\tau}^{(i,n)} \in [0, 1] \in \mathbb{R}^{D(i,n)}$	
$\tilde{s}^{(i,n)} = (s_1^{(i,n)}, ..., s_T^{(i,n)})$	Sequences of hidden states until the end of the current sequence
$\pi^{(i)} \in \{1, ..., K\} \in \mathbb{N}$	Sequential policies specifying controlled transitions within N hidden factors over time and their posterior expectations
$\boldsymbol{\pi}^{(i)} = (\pi_1^{(i)}, ..., \pi_K^{(i)}) \in [0, 1] \in \mathbb{R}^K$	
$u_{\pi,\tau}^{(i,n)} = \pi^{(i)}(t, n) \in \mathbb{N}$	Action or control variables for the n-th factor of hidden states at a particular time specified by a policy
$\boldsymbol{\nu}_{\pi,\tau}^{(i,n)} = -\ln \mathbf{s}_{\pi,\tau}^{(i,n)}$	Auxiliary (depolarisation) variable corresponding to the surprise of an expected state – a softmax function of depolarisation
$\mathbf{s}_{\pi,\tau}^{(i,n)} = \sigma(-\boldsymbol{\nu}_{\pi,\tau}^{(i,n)})$	
$\mathbf{o}_{\pi,\tau}^{(i,m)} = \mathbf{A}^{(i,m)} \circ (\mathbf{s}_{\pi,\tau}^{(i,1)}, \mathbf{s}_{\pi,\tau}^{(i,2)}...)$	Predictive posterior over future outcomes using a generalised dot product (sum of products) operator
$\mathbf{s}_\tau^{(i,n)} = \sum_\pi \pi_\pi^{(i)} \cdot \mathbf{s}_{\pi,\tau}^{(i,n)}$	Bayesian model average of hidden states over policies
$\mathbf{A}^{(i,m)} \in [0, 1] \in \mathbb{R}^{D(i,m) \times D(i,1) \times D(i,2)...}$	Likelihood tensor mapping from hidden states to the m-th modality
$\mathbf{B}_{\pi,\tau}^{(i,n)} \triangleq \mathbf{B}^{(i,n)}(u_{\pi,\tau}^{(i,n)}) \in [0, 1] \in \mathbb{R}^{D(i,n) \times D(i,n)}$	Transition probability for the n-th hidden state under an action (prescribed by a policy at a particular time)
$\mathbf{C}_\tau^{(i,m)} \in [0, 1] \in \mathbb{R}^{D(i,m) \times D(i+1,n)}$	Prior probability of the m-th outcome at the i-th level conditioned on the n-th (hierarchical) context
$\mathbf{D}^{(i,n)} \triangleq \mathbf{B}_{\pi,0}^{(i,n)} \in [0, 1] \in \mathbb{R}^{D(i,n) \times D(i+1,n)}$	Prior probability of the n-th initial state at the i-th level conditioned on the n-th (hierarchical) context
$\mathbf{F}^{(i)} = (\mathbf{F}_1^{(i)}, ..., \mathbf{F}_K^{(i)}) \in \mathbb{R}^K$	Marginal free energy for each policy
$\mathbf{G}^{(i)} = (\mathbf{G}_1^{(i)}, ..., \mathbf{G}_K^{(i)}) \in \mathbb{R}^K$	Expected free energy for each policy
$\mathbf{H}_{kl...}^{(i,m)} = -\sum_j \mathbf{A}_{jkl...}^{(i,m)} \ln \mathbf{A}_{jkl...}^{(i,m)} \in \mathbb{R}$	Entropy of outcomes under each combination of states in the m-th modality

$$Q(\tilde{s}, \pi) = \arg\min_Q F$$
$$\approx P(\tilde{s}, \pi \mid \tilde{o})$$

$$F = \mathrm{E}_Q[\ln Q(\tilde{s}, \pi) - \ln P(\tilde{o}, \tilde{s}, \pi)]$$

$$= \underbrace{D[Q(\tilde{s}, \pi) \| P(\tilde{s}, \pi \mid \tilde{o})]}_{\text{divergence}} - \underbrace{\ln P(\tilde{o})}_{\text{log evidence}}$$
$$= \underbrace{D[Q(\tilde{s}, \pi) \| P(\tilde{s}, \pi)]}_{\text{complexity}} - \underbrace{\mathrm{E}_Q[\ln P(\tilde{o} \mid \tilde{s})]}_{\text{accuracy}} \qquad (1)$$

where $\tilde{o} = (o_1, ..., o_t)$ denotes observations up until the current time point and $\tilde{s} = (s_1, ..., s_T)$ represents hidden states over all the time points in a sequence. Because the (KL) divergence between a subject's beliefs and the true posterior cannot be less than zero, the penultimate equality means that free energy is minimised when the two are the same. At this point, the free energy becomes the *surprise* or negative log evidence for the generative model (Beal, 2003). In other words, minimising free energy is equivalent to minimising the complexity of accurate explanations for observed outcomes.

In active inference, both beliefs and action minimise free energy. However, beliefs cannot affect outcomes. This means that action affords the only means of minimising surprise, where action minimises expected free energy; i.e. expected surprise or uncertainty. In turn, this rests on equipping subjects with the prior beliefs that their policies will minimise expected free energy (Friston et al., 2015):

$$P(\pi) = \sigma(-G(\pi))$$
$$G(\pi) = \sum_\tau G(\pi, \tau)$$

$$G(\pi, \tau) = \mathrm{E}_{\tilde{Q}}[\ln Q(s_\tau \mid \pi) - \ln P(o_\tau, s_\tau \mid \tilde{o}, \pi)]$$

$$= -\underbrace{E_{\tilde{Q}}[\ln Q(s_\tau \mid o_\tau, \pi) - \ln Q(s_\tau \mid \pi)]}_{\text{epistemic value}} - \underbrace{E_{\tilde{Q}}[\ln P(o_\tau)]}_{\text{expected value}}$$
$$= \underbrace{D[Q(o_\tau \mid \pi) \| P(o_\tau)]}_{\text{risk}} + \underbrace{E_\tau[H[P(o_\tau \mid s_\tau)]]}_{\text{ambiguity}} \qquad (2)$$

Here, $G(\pi, \tau)$ denotes the expected free energy of a particular policy at a particular time, and $\tilde{Q} = Q(o_\tau, s_\tau \mid \pi) = P(o_\tau \mid s_\tau) Q(s_\tau \mid \pi)$ is the

predictive distribution over hidden states and outcomes under that policy. Comparing the expressions for expected free energy (Eq. (2)) with variational free energy (Eq. (1)), we see that the (negative) divergence becomes *epistemic value* and the *log evidence* becomes *expected value* – provided we associate the prior preference over future outcomes with value. In other words, valuable outcomes are those we expect to encounter and costly outcomes are surprising (e.g., being in pain). The last equality provides a complementary interpretation; in which complexity becomes *risk*, and inaccuracy becomes *ambiguity*. Please see the appendices for derivations.

There are several special cases of expected free energy that appeal to (and contextualise) established constructs. For example, maximising epistemic value is equivalent to maximising (expected) Bayesian surprise (Itti and Baldi, 2009), where Bayesian surprise is the divergence between posterior and prior beliefs. This can also be interpreted as the principle of maximum mutual information or minimum redundancy (Barlow, 1961; Linsker, 1990; Olshausen and Field, 1996; Laughlin, 2001). In this context, epistemic value is the expected *mutual information* between future states and their consequences, which is also known as *information gain*. Because epistemic value (i.e., mutual information) cannot be less than zero, it disappears when the (predictive) posterior ceases to be informed by new observations. This means epistemic behaviour will search out observations that resolve uncertainty (e.g., foraging to find a prey or turning on the light in a dark room). However, when the agent is confident about the state of the world, there can be no further information gain and pragmatic (prior) preferences dominate. Crucially, epistemic and expected values have a definitive quantitative relationship, which means there is no need to adjudicate between explorative, epistemic uncertainty reducing and exploitative, pragmatic goal directed behaviour. The switch between behavioural policies emerges naturally from minimising expected free energy. This switch depends on the relative contribution of epistemic and expected value, thereby resolving the exploration-exploitation dilemma. Furthermore, in the absence of any precise preferences, purposeful behaviour is purely epistemic in nature. In what follows, we will see that prior preferences or goals are usually restricted to the highest levels of a hierarchy. This means that active inference at lower levels is purely uncertainty reducing, where action ceases when uncertainty approaches zero (in this paper, a sequence of actions terminates when the

uncertainty about states prescribed by higher levels is $\frac{1}{64}$ nats or less).

2.3. Summary

Minimising expected free energy is essentially the same as avoiding surprises and resolving uncertainty. This resolution of uncertainty is closely related to satisfying artificial curiosity (Schmidhuber, 1991; Still and Precup, 2012) and speaks to the value of information (Howard, 1966). Expected free energy can be expressed in terms of epistemic and expected value – or in terms of risk and ambiguity. The expected complexity or risk is exactly the same quantity minimised in risk sensitive or KL control (Klyubin et al., 2005; van den Broek et al., 2010), and underpins related (free energy) formulations of bounded rationality based on complexity costs (Braun et al., 2011; Ortega and Braun, 2013). In other words, minimising expected complexity renders behaviour risk-sensitive, while maximising expected accuracy induces ambiguity-resolving behaviour. In the next section, we look more closely at how this minimisation is implemented.

3. Belief propagation and neuronal networks

Having defined a generative model, the expectations encoding posterior beliefs (and action) can be optimised by minimising variational free energy. Fig. 2 provides the mathematical expressions for this optimisation or belief updating. Although the updates look a little complicated, they are remarkably plausible in terms of neurobiological process theories (Friston et al., 2014). In brief, minimising variational free energy means that expectations about allowable policies become a softmax function of variational and expected free energy, where the (path integral) of variational free energy scores the evidence that a particular policy is being pursued (Equation 1.c in Fig. 2). Conversely, the expected free energy plays the role of a prior over policies that reflect their ability to resolve uncertainty (Equation 1.d). The resulting policy expectations are used to predict the state at each level in the form of a *Bayesian model average*; in other words, the expected states under each policy are combined in proportion to the expected probability of each policy (Equation 2.d). These Bayesian model averages then provide (top-down) prior constraints on the initial states of the level below. Finally, expectations about policies enable the most likely action to be selected at each level of the hierarchy. Fig. 2 only shows action selection for the lowest (first) level.

Of special interest here, are the updates for expectations of hidden states (for each policy and time). These have been formulated as a gradient descent on variational free energy (see Appendix 1). This furnishes a dynamical process theory that can be tested against empirical measures of neuronal dynamics. Specifically, the Bayesian updating or belief propagation (see Appendix 2) has been expressed so that it can be understood in terms of neurophysiology. Under this interpretation, expected states are a softmax function of log expectations that can be associated with neuronal depolarisation (Equation 2.b). In other words, the softmax function becomes a firing rate function of depolarisation, where changes in postsynaptic potential are caused by currents induced by presynaptic input from prediction error units (Equation 2.a). In this formulation, *state* prediction errors are the difference between the log expected state and its prediction from observed outcomes, the preceding state and subsequent state (Equation 1.a). Similarly, outcome prediction errors are the difference between the log expected outcome and the outcome predicted by hidden states in the level above (Equation 1.b). Physiologically, this means that when state prediction error unit activity is suppressed, there is no further depolarisation of expectation units and their firing attains a variational free energy minimum. This suggests that for every expectation unit there should be a companion error unit, whose activity is the rate of change of depolarisation of the expectation unit; for example, excitatory (expectation) pyramidal cells and fast spiking inhibitory (error) interneurons (Sohal et al., 2009; Cruikshank et al., 2012; Lee et al., 2013).

3.1. Extrinsic and intrinsic connectivity

The graphics in Fig. 2 have assigned various expectations and errors to neuronal populations in specific cortical layers. This (speculative) assignment, allows one to talk about the functional anatomy of extrinsic and intrinsic connectivity in terms of belief propagation. In brief, the mathematical form of Bayesian belief updating tells us which neuronal representations talk to each other. For example, in a hierarchical setting, the only sufficient statistics that are exchanged between levels are the Bayesian model averages of expected states. This means, by definition, that the Bayesian model averages must be encoded by principal cells that send neuronal connections to other areas in the cortical hierarchy. These are the superficial and deep pyramidal cells show in red in Fig. 2. Next, we know that the targets of ascending extrinsic (feedforward) connections from superficial pyramidal cells are the spiny stellate cells in Layer 4 (Felleman and Van Essen, 1991; Bastos et al., 2012; Markov et al., 2013). The only sufficient statistics in receipt of Bayesian model averages from the level below are policy-specific expectations about hidden states. These can be associated with spiny stellate cells (upper cyan layer in Fig. 2). These sufficient statistics are combined to form the Bayesian model average in superficial pyramidal cells, exactly as predicted by quantitative connectivity studies of the canonical cortical microcircuit (Thomson and Bannister, 2003). One can pursue this game and – with some poetic license – reproduce the known quantitative microcircuitry of inter-and intralaminar connections. Fig. 3 illustrates one solution that reproduces not only the major intrinsic connections but also their excitatory and inhibitory nature. This arrangement suggests that inhibitory interneurons play the role of error units (which is consistent with the analysis above), while policy-specific expectations are again encoded by excitatory neurons in Layer 4. Crucially, this requires a modulatory weighting of the intrinsic feedforward connections from expectation units to their Bayesian model averages in the superficial layers. This brings us to extrinsic connections and the neuronal encoding of policies in the cortico-basal ganglia-thalamic loops.

3.2. Extrinsic connectivity and cortico-subcortical loops

According to the belief propagation equations, expected policies rest upon their variational and expected free energy. These free energies comprise (KL) divergences that can always be expressed in terms of an average prediction error. Here, the variational free energy is the expected state prediction error, while the expected free energy is the expected outcome prediction error. These averages are gathered over all policies and time points within a hierarchical level and are passed through a sigmoid (softmax) function to produce policy expectations. If we associate this pooling with cortico-subcortical projections to the basal ganglia – and the subsequent Bayesian model averaging with thalamocortical projections to the cortex – there is a remarkable correspondence between the implicit connectivity (both in terms of its specificity and excitatory versus inhibitory nature) and the connectivity of the cortico-basal ganglia-thalamocortical loops.

The schematic in Fig. 4 is based upon the hierarchical anatomy of cortico-basal ganglia-thalamic loops described in (Jahanshahi et al., 2015). If one subscribes to this functional anatomy, the formal message passing of belief propagation suggests that competing low level (motor executive) policies are evaluated in the putamen; intermediate (associative) policies in the caudate and high level (limbic) policies in the ventral striatum. These representations then send (inhibitory or GABAergic) projections to the globus pallidus interna (GPi) that encodes the expected (selected) policy. These expectations are then communicated via thalamocortical projections to superficial layers encoding Bayesian model averages. From a neurophysiological perspective, the best candidate for the implicit averaging would be matrix thalamocortical circuits that "appear to be specialized for robust transmission over relatively extended periods, consistent with the sort of persistent
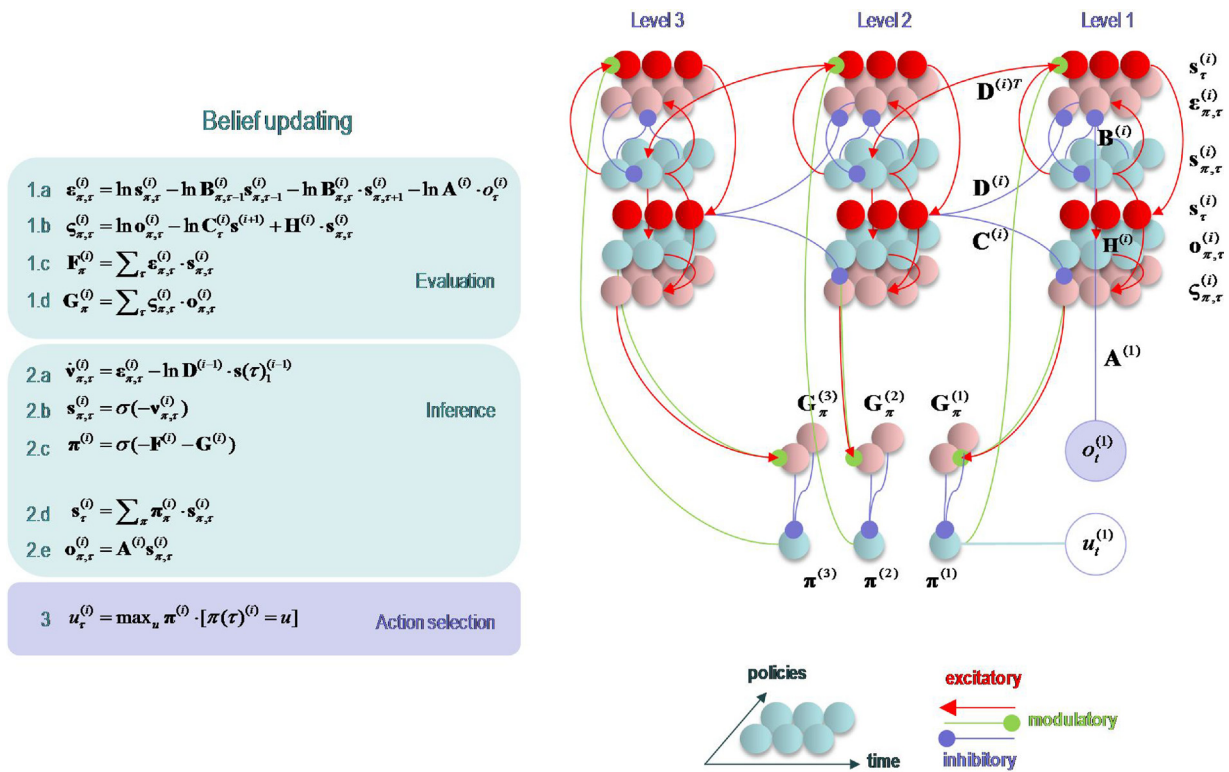
Fig. 2. Schematic overview of belief propagation: left panel: these equalities are the belief updates mediating inference (i.e. state estimation) and action selection. These expressions follow in a fairly straightforward way from a gradient descent on variational free energy. The equations have been expressed in terms of prediction errors that come in two flavours. The first, *state* prediction error scores the difference between the (log) expected states under any policy and time (at each hierarchical level) and the corresponding predictions based upon outcomes and the (preceding and subsequent) hidden states (1.a). These represent likelihood and empirical prior terms respectively. The prediction error drives log-expectations (2.a), where the expectation *per se* is obtained via a softmax operator (2.b). The second, *outcome* prediction error reports the difference between the (log) expected outcome and that predicted under prior preferences set by the level above (plus an ambiguity term – see Appendix) (1.b). This prediction error is weighted by the expected outcomes to evaluate the expected free energy (1.d). Similarly, the free energy *per se* is the expected state prediction error, under current beliefs about hidden states (1.c). These policy-specific free energies are combined to give the policy expectations via a softmax function (2.c). Finally, expectations about hidden states are a Bayesian model average over expected policies (2.d) and expectations about policies specify the action that is most likely to realise the expected outcome (3). The (Iverson) brackets in Equation 3 return one if the condition in square brackets is satisfied and zero otherwise. Right panel: this schematic represents the message passing implicit in the equations on the left. The expectations have been associated with neuronal populations (coloured balls) that are arranged to highlight the correspondence with known intrinsic (within cortical area) and extrinsic (between cortical areas) connections. Red connections are excitatory, blue connections are inhibitory and green connections are modulatory (i.e., involve a multiplication or weighting). This schematic illustrates three hierarchical levels (which are arranged horizontally in this figure, as opposed to vertically in Fig. 1), where each level provides top-down empirical priors for the initial state of the level below, while the lower level supplies evidence for the current state at the level above. The intrinsic connections mediate the empirical priors and Bayesian model averaging. Cyan units correspond to expectations about hidden states and (future) outcomes under each policy, while red states indicate their Bayesian model averages. Pink units correspond to (state and outcome) prediction errors that are averaged to evaluate (variational and expected) free energy and subsequent policy expectations (in the lower part of the network). This (neuronal) network interpretation of belief updating means that connection strengths correspond to the parameters of the generative model in Fig. 1. Please see Table 2 for a definition of the variables. The variational free energy has been omitted from this figure because the policies in this paper differ only in the next action. This means the evidence (i.e. variational free energy) from past outcomes is the same for all policies (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

activation observed during working memory and potentially applicable to state-dependent regulation of excitability" (Cruikshank et al., 2012). This deep temporal hierarchy is apparent in hierarchically structured cortical dynamics – invasive recordings in primates suggest an anteroposterior gradient of spontaneous fluctuation time constants consistent with the architecture in Fig. 4 (Kiebel et al., 2008; Murray et al., 2014). Clearly, there are many anatomical issues that have been ignored here; such as the distinction between direct and indirect pathways (Frank, 2005), the role of dopamine in modulating the precision of beliefs about policies (Friston et al., 2014) and so on. However, the basic architecture suggested by the above treatment speaks to the biological plausibility of belief updating under hierarchical generative models.

3.3. Summary

By assuming a generic (hierarchical Markovian) form for the generative model, it is fairly easy to derive Bayesian updates that clarify the relationship between perception and action selection. In brief, the agent first infers the hidden states under each policy that it entertains. It then evaluates the evidence for each policy based upon observed outcomes and beliefs about future states. The posterior beliefs about each policy are used to form a Bayesian model average of the next outcome, which is realised through action. In hierarchical models, the implicit belief updating (known as belief propagation in machine learning) appears to rest on message passing that bears a remarkable similarity to cortical hierarchies and cortico-basal ganglia-thalamic loops; both in terms of extrinsic connectivity and intrinsic canonical (cortical) microcircuits. In the next section, we use this scheme to simulate reading.
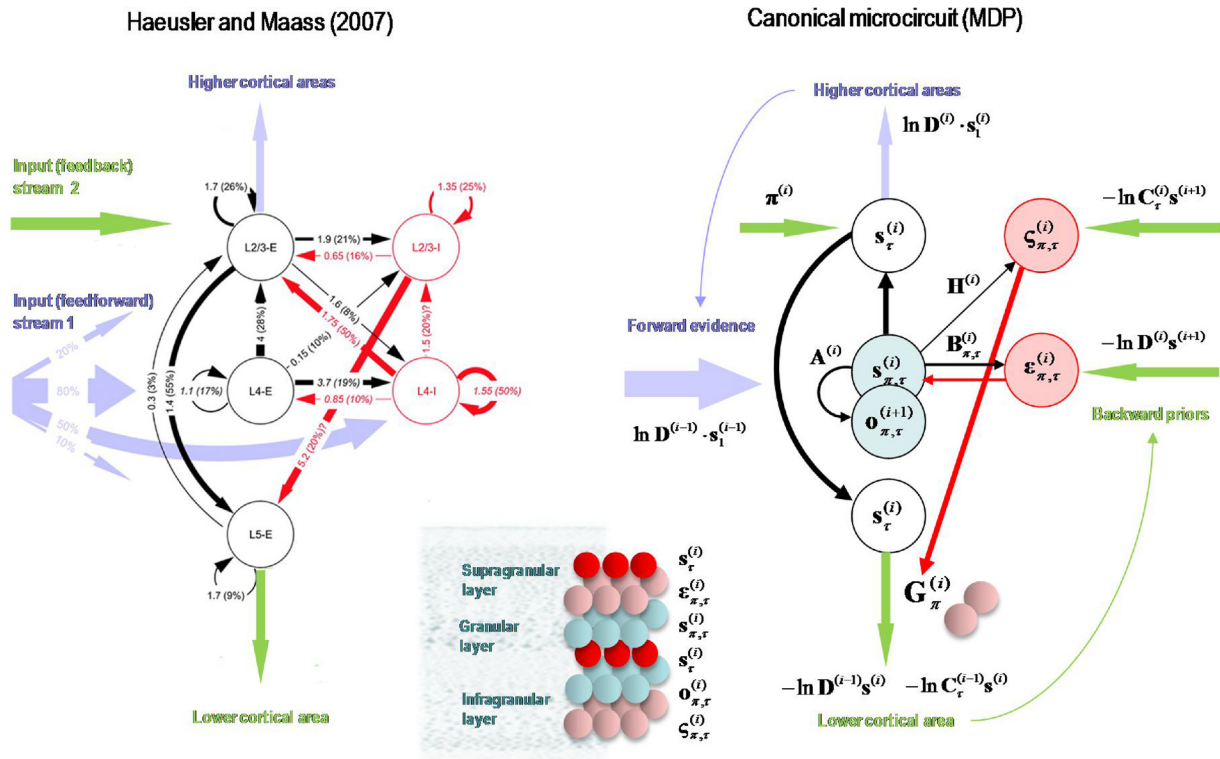
Fig. 3. Belief propagation and intrinsic connectivity. This schematic features the correspondence between known canonical microcircuitry and the belief updates in Fig. 2. Left Panel: a canonical microcircuit based on (Haeusler and Maass, 2007), where inhibitory cells have been omitted from the deep layers – because they have little interlaminar connectivity. The numbers denote connection strengths (mean amplitude of PSPs measured at soma in mV) and connection probabilities (in parentheses) according to (Thomson and Bannister, 2003). Right panel: the equivalent microcircuitry based upon the message passing scheme of the previous figure. Here, we have placed the outcome prediction errors in superficial layers to accommodate the strong descending (inhibitory) connections from superficial to deep layers. This presupposes that descending (interlaminar) projections disinhibit Layer 5 pyramidal cells that project to the medium spiny cells of the striatum (Arikuni and Kubota, 1986). The computational assignments in this figure should be compared with the equivalent scheme for predictive coding in (Bastos et al., 2012). The key difference is that superficial excitatory (e.g., pyramidal) cells encode expectations of hidden states, as opposed to state prediction errors. This is because the prediction error is encoded by their postsynaptic currents, as opposed to their depolarisation or firing rates (see main text). The white circles correspond to the Bayesian model average of state expectations, which are the red balls in the previous figures (and the inset). Black arrows denote excitatory *intrinsic* connections, while red arrows are inhibitory. Blue arrows denote bottom-up of ascending *extrinsic* connections, while green arrows are top-down or descending (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

4. Simulations of reading

The remainder of this paper considers simulations of reading using a generative model that is a hierarchical extension of a model we have used previously to illustrate scene construction (Mirza et al., 2016). In the original paradigm, (simulated) subjects had to sample four quadrants of a visual scene to classify the arrangement of visual objects (a *bird*, a *cat* and *seeds*) into one of three categories (*flee*, *feed* or *wait*). If the *bird* and *cat* were next to each other (in the upper or lower quadrants) the category was *flee*. If the *bird* was next to the *seeds*, the category was *feed*. Alternatively, if the *bird* and *seeds* occupied diagonal quadrants, the category was *wait*. Here, we treat the visual objects as letters and the scene as a word; enabling us to add a hierarchical level to generate sentences or sequences of words. The subject's task was to categorise sentences of four words into *happy* or *sad* narratives; where *happy* narratives concluded with a *feed* or *wait* in the final two words. Somewhat arbitrarily, we restricted the hypotheses at the second level to 6 sentences (see Fig. 5). By stimulating reading, we hoped to produce realistic sequences of saccadic eye movements, in which the subject interrogated local features (i.e. letters) with sparse and informative foveal sampling; in other words, jumping to key letter features and moving to the next word as soon as the current word could be inferred confidently. Furthermore, because the subject has a deep model, she already has in mind the words and letters that are likely to be sampled in the future; enabling an efficient foraging for information.

To simulate this sort of task, one needs to specify the hidden factors,

allowable policies and prior preferences. Fig. 5 illustrates the factorisation and hierarchical structure of the resulting model. At the highest level there are three hidden factors (only two are shown in the figure for simplicity). These comprise the sentence (with six alternatives), the word the subject is currently examining (one of four words) and the decision (*undecided*, *happy* or *sad*). The word location and current sentence specify the hidden state (word) at the lower hierarchical level. The lower level also includes a letter location state (one of four quadrants) and two spatial transformations (horizontal and vertical *flip*). The current word and letter location specify the outcome (letter or visual object; *cat*, *bird*, *seed* or *nothing*). At both the higher (e.g., sentence) and lower (e.g., word) levels, the hidden locations also specify a proprioceptive outcome in terms of higher (e.g., head) and lower (e.g., eye) movements that sample the word and letter respectively. Finally, the hidden decision state determines (e.g. auditory) feedback with three possibilities; namely, *nothing*, *right* or *wrong*. The decision state and feedback outcomes have been omitted from Fig. 5 for clarity.

This setup defines the state space and mapping from hidden states to outcomes encoded by the **A** parameters. Note that the likelihood mapping involves interactions among hidden states; for example, one has to know both the location being sampled and the word generating outcomes before the letter is specified. These interactions are modelled very simply by placing a one at the appropriate combination of hidden states (and zeros elsewhere) in the row of **A** corresponding to the outcome. Similarly, the **D** parameters specify the outcome in terms of hidden states at the lower level in terms of (combinations of) hidden
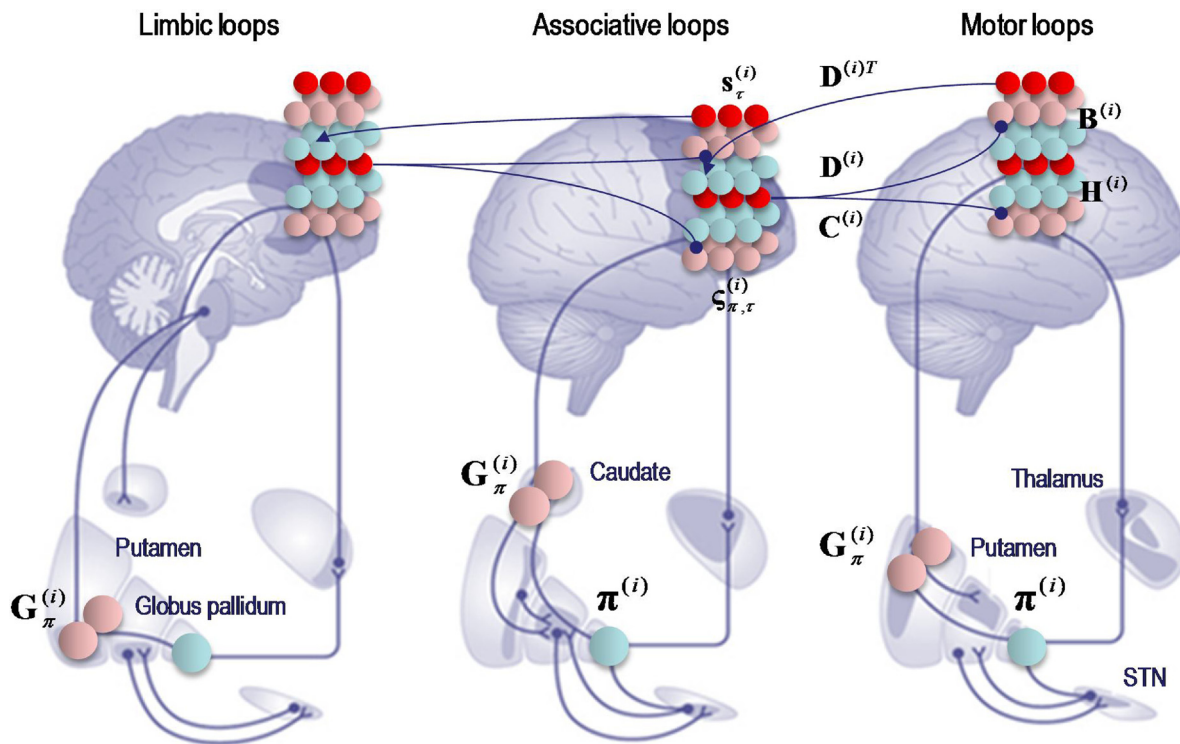
Fig. 4. Belief propagation and extrinsic connectivity. This schematic illustrates a putative mapping between expectations that are updated during belief updating and recurrent interactions within the cortico-basal ganglia-thalamic loops. This figure is based upon the functional neuroanatomy described in (Jahanshahi et al., 2015), which assigns motor updates to motor and premotor cortex projecting to the putamen; associative loops to prefrontal cortical projections to the caudate and limbic loops to projections to the ventral striatum. The correspondence between the message passing implicit in belief propagation and the organisation of these loops is remarkable; even down to the level of the sign (excitatory or inhibitory) of the neuronal connections. The striatum (caudate and putamen) and the subthalamic nucleus (STN) receive inputs from many cortical and subcortical areas. The internal segment of the globus pallidus (GPi) constitutes the main output nucleus from the basal ganglia. The basal ganglia are not only connected to motor areas (motor cortex, supplementary motor cortex, premotor cortex, cingulate motor area and frontal eye fields) but also have connections with associative cortical areas. The basal ganglia nuclei have topologically organized motor, associative and limbic territories; the posterior putamen is engaged in sensorimotor function, while the anterior putamen (or caudate) and the ventral striatum are involved in associative (cognitive) and limbic (motivation and emotion) functions (Jahanshahi et al., 2015).

factors at the higher level.

It is now necessary to specify the contingencies and transitions among hidden states in terms of the **B** parameters. There is a separate **B** matrix for every hidden factor and policy. In this example, these matrices have a very simple form: on any given trial, policies cannot change the sentence or word. This means the corresponding **B** matrices are identity matrices. For hidden locations, the **B** matrices simply encode a transition from the current location to the location specified by the policy. Here the policies were again very simple; namely, where one looks next (in a body and head centred frame of reference at the first and second levels respectively). For simplicity, the preceding actions that constitute each policy were the actions actually selected. In more sophisticated setups, policies can include different sequences of actions; however here, the number of policies and actions were the same. This means we do not have to worry about the evidence the different policies encoded by the variational free energy (as in the right panel of Fig. 2). There were three policies or actions at the second level; proceed to the next word or stop reading and make a categorical decision of *happy* or *sad*; resulting in *right* or *wrong* feedback.

Finally, the prior preferences encoded in the **C** parameters rendered all outcomes equally preferred, with the exception of being wrong, which was set at $\ln \mathbf{C}_\tau^{(i)} = -4$. In other words, the subject thought they were $\exp(4) \approx 54$ times less likely to be wrong than undecided or right. This aversion to making mistakes ensures the subject does not solicit feedback to resolve uncertainty about the category of the sentence. In other words, the subject has to be relatively confident – after epistemic foraging – about the underlying narrative before confirming any inference with feedback. Prior beliefs about first level hidden states, encoded in the **D** parameters, told the subject they would start at the first

quadrant of the first word, with an equal probability of all sentences. Because all six sentences began with either *flee* or *wait*, the prior probability over words was implicitly restricted to *flee* or *wait*, with equal probabilities of horizontal flipping (because these priors do not depend on the higher level). The horizontal flipping corresponds to a spatial transformation, under which the meaning of the word is invariant, much like a palindrome. Conversely, the subject had a strong prior belief that there was no vertical flipping. This (low-level feature) transformation can be regarded as presenting words in upper or lower case. The prior over vertical flipping will become important later, when we switch prior beliefs to make uppercase (vertically flipped) stimuli the prior default to introduce violations of (feature) expectations.

A heuristic motivation for including hidden factors like horizontal flipping appeals to the way that we factorise hidden causes of stimuli; in other words, carve nature at its joints. The fact that we are capable of:

"raeding wrods with jumbled lettres" (Rayner et al., 2006),

suggests that horizontal flipping can be represented in a way that is conditionally independent of grapheme content.

This completes our specification of the generative model. To simulate reading, the equations in Fig. 2 were integrated using 16 iterations for each time point at each level. At the lowest level, an iteration is assumed to take 16 ms, so that each epoch or transition is about 256 ms.[1] This is the approximate frequency of saccadic eye movements

---

[1] This is roughly the amount of time taken per iteration on a personal computer – to less than an order of magnitude.
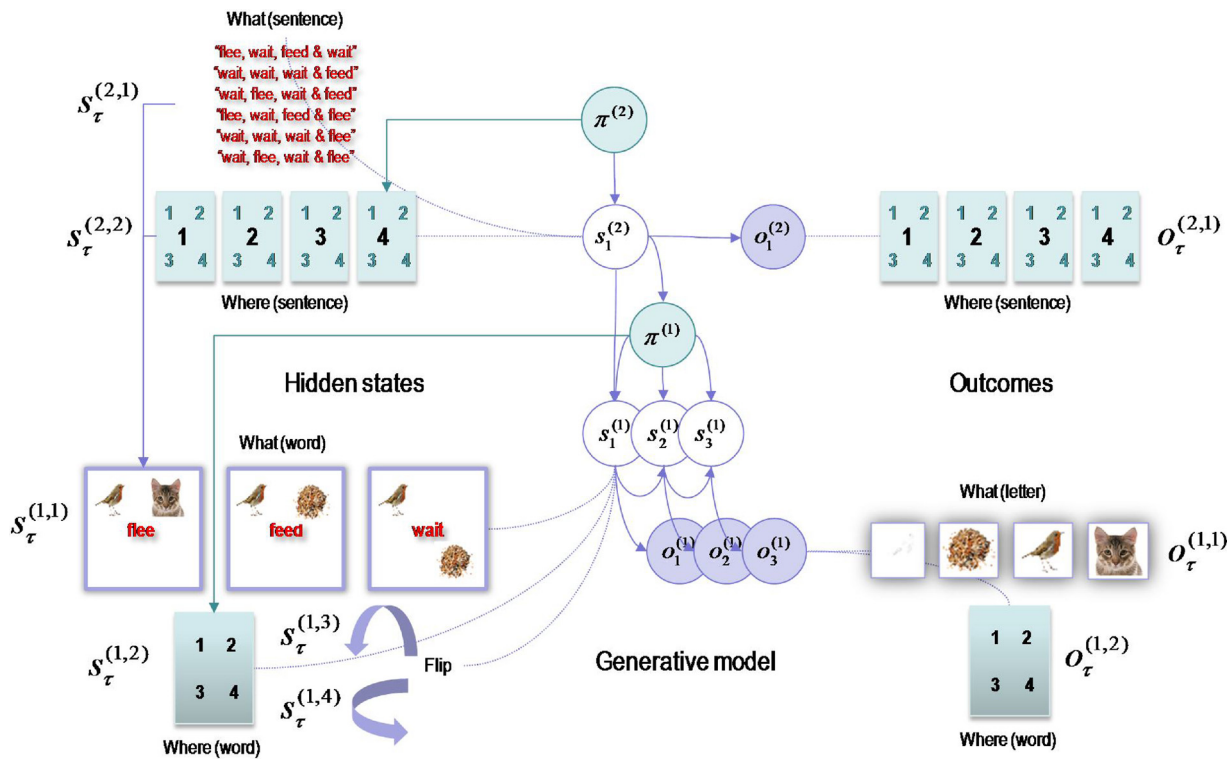
Fig. 5. The generative model used to simulate reading. This graphical model shows the conditional dependencies of the generative model used in subsequent figures, using the same format as Fig. 1. In this model there are two hierarchical levels with three hidden states at the second level and four at the first level (hidden states and outcomes pertaining to categorical decisions and feedback have been omitted for clarity). The hidden states at the higher level correspond to the sentence or narrative – generating sequences of *words* at the first level – and which *word* the agent is currently sampling (with six alternative *sentences* and four *words* respectively). These (higher level) hidden states combine to specify the word generated at the first level (*flee*, *feed* or *wait*). The hidden states at the first level comprise the current *word* and which quadrant the agent is looking at. These hidden states combine to generate outcomes in terms of *letters* or icons that would be sampled if the agent looked at a particular *location* in the current *word*. In addition, two further hidden states provide a local feature context by flipping the locations vertically or horizontally. The vertical flip can be thought of in terms of font substitution (upper case versus lowercase), while the horizontal flip means a word is invariant under changes to the order of the letters (c.f., palindromes that read the same backwards as forwards). In this example, *flee* means that a bird is next to a cat, *feed* means a bird is next to some seeds and *wait* means seeds are above (or below) the bird. Notice that there are outcomes at both levels. At the higher level there is a (proprioceptive) outcome signalling the *word* currently being sampled (e.g., head position), while at the lower level there are two outcome modalities. The first (exteroceptive) outcome corresponds to the observed letter and the second (proprioceptive) outcome reports the letter location (e.g., direction of gaze in a head-centred frame of reference). Similarly, there are policies at both levels. The high-level policy determines which word the agent is currently reading, while the lower level dictates the transitions among the quadrants containing letters.

(and indeed phonemic processing in auditory language processing), meaning that the simulations covered a few seconds of simulated time. The scheduling of updates in hierarchical models presents an interesting issue. In principle, we could implement the belief updating synchronously; enabling second level expectations to be informed by first level expectations as they accumulate evidence. Alternatively, we could wait until the first level convergences and update higher levels asynchronously – so that the high-level waits until the lower level sequence completes before updating and providing (empirical) prior constraints for the initial state at the lower level. We elected to illustrate the latter (asynchronous) updating, noting that alternative (synchronous) schemes could be implemented and compared to empirical neuronal responses. Asynchronous scheduling has the advantage of computational simplicity, because it means each level can be integrated or solved by the same routine (here, **spm_MDP_VB_X.m**). This means that the sequence of posterior expectations following convergence at one level can be passed as (probabilistic) outcomes to the next, while the outcomes from the highest level enter as prior constraints on the initial states of the level below. Furthermore, we will see below that the ensuing updates bear a marked similarity to empirical (neurophysiological) responses.

Fig. 6 shows simulated behavioural responses during reading in terms of eye movements (upper panel) over four transitions at the second level, where each transition entails one or two saccades at the

first. In this exemplar simulation, the stimuli were generated at random using the above generative model. Here, the subject read the first sentence in lower case, apart from the second letter that was in upper case (i.e. with a surprising vertical flipping). In this trial, the subject looks at the first quadrant of the first word and sees a *cat*. She therefore knows immediately that the first word is *flee*. She then turns to the second word and sees nothing. To resolve uncertainty, she samples the fourth quadrant and again finds nothing, which means this word must be *wait* (because the second word of each sentence is either *flee* or *wait* – and the current word cannot be flee because the *cat* cannot be next to the *bird*). The next two saccades, on the subsequent word, confirm the word *feed* (with the *seed* next to the *seed* sampled on the first saccade). Finally, the subject turns to the final word and discovers *seeds* on the second saccade. At this point, residual uncertainty about the sentence is resolved and the subject makes a correct categorisation – a *happy* story. The lower panel shows expected outcomes at the end of sampling each word. The upper row shows the final beliefs about the words under (correct) expectations about the sentence of the second level. This (first) sentence was "*flee, wait, feed* and *wait*".

The key thing to take from these results is that the agent can have precise beliefs about letters without ever seeing them. For example, the subject believes there is a *bird* in the second quadrant of the first word, despite the fact she never looked there. This illustrates the fact that it is not necessary to sample all the constituent letters to identify a word.
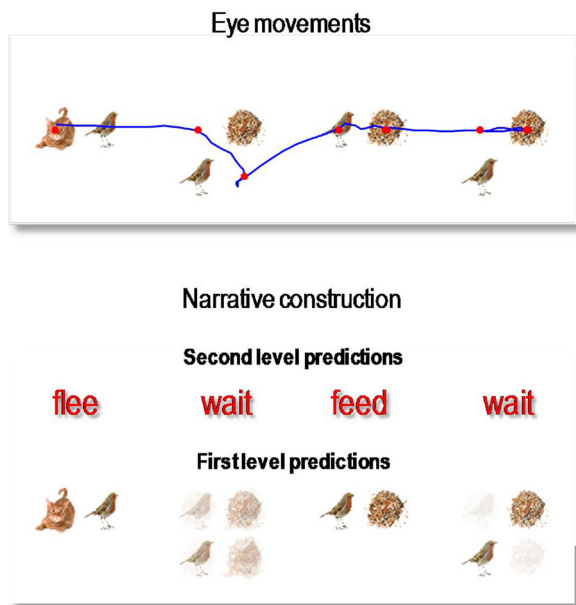
Eye movements

Narrative construction

Second level predictions

flee   wait   feed   wait

First level predictions

Fig. 6. Simulated behavioural responses during reading: upper panel. This shows the trajectory of eye movements over four transitions at the second level that entail one or two saccadic eye movements at the first. In this trial, the subject looks at the first quadrant of the first word and sees a cat. She therefore knows immediately that the first word is *flee*. She then turns to the first quadrant of the second word and sees nothing. To resolve uncertainty she then looks at the fourth quadrant and again finds nothing, which means this word must be *wait* (because the second word of each sentence is either *flee* or *wait* – and the current word cannot be *flee* because the cat cannot be next to the bird). The next two saccades, on the subsequent word confirm the word *feed* (with the *seed* next to the *bird* sampled on the first saccade). Finally, the subject turns to the final word and discloses *seeds* after the second saccade. At this point, uncertainty about the sentence (sentence one versus sentence four) is resolved and the subject makes a correct categorisation – a *happy* story. Lower panel: this panel shows expected outcomes at the end of sampling each word. The upper row shows the final beliefs about the words under (correct) expectations about the sentence at the second level. This (first) sentence was "*flee, wait, feed and wait*". At the first level, expectations about the letters under posterior beliefs about the words are shown in terms of mixtures of icons.

Conversely, there can be uncertainty about particular letters, even though the subject is confident about the word. This is illustrated by the expectations about the letters in the second word. These expectations are consistent with *wait* but reflect a degree of uncertainty about the vertical flip (i.e., lower case or upper case font). This uncertainty – and resulting hesitancy in moving to the next word – reflects the subject's prior belief that letters are usually presented in lower case. However, the actual stimuli were presented in a surprising way (with a vertical flip) that causes the subject to spend an extra saccade on this word, before moving to the next.

Fig. 7 shows the simulated electrophysiological responses associated with the belief updating reported in Fig. 6. Expectations about the hidden state at the higher (upper panel) and lower (middle panel) levels are presented in raster format. The horizontal axis is time over the entire trial, where each iteration corresponds roughly to 16 ms and the trial lasted for three and half seconds. The vertical axis corresponds to the Bayesian model averages or expectations about the six sentences at the higher level and the three words at the lower level. Under the scheduling used in these simulations, higher level expectations wait until lower-level updates have terminated and, reciprocally, lower-level updates are suspended until belief updating in the higher level has been completed. This means the expectations are sustained at the higher level, while the lower level gathers information. The resulting patterns of firing rate over time show a marked resemblance to pre-saccadic delay period activity in the prefrontal cortex. The insert on the upper

right is based upon the empirical results reported in (Funahashi, 2014) and tie in nicely with the putative role of matrix thalamocortical projections during delay period activity (Cruikshank et al., 2012). Note that the expectations are reset at the beginning of each epoch, producing the transients in the lower panel on the left. These fluctuations are the firing rate in the upper panels filtered between 4 Hz and 32 Hz and can be regarded as (band pass filtered) changes in simulated depolarisation. These simulated local field potentials are again remarkably similar to empirical responses. The examples shown in the lower right inset are based on the study of perisaccadic electrophysiological responses in early and inferotemporal cortex during active vision reported in (Purpura et al., 2003).

4.1. Summary

In the previous section, we highlighted the biological plausibility of belief updating based upon deep temporal models. In this section, the biological plausibility is further endorsed in terms of canonical electrophysiological phenomena such as perisaccadic delay period firing activity and local field potentials. Furthermore, these simulations have a high degree of face validity in terms of saccadic eye movements during reading (Rayner, 1978, 2009). In the final section, we focus on the electrophysiological correlates and try to reproduce classical event related potential phenomena such as the mismatch negativity and other responses to violation.

5. Simulations of classical violation responses

Fig. 8 shows simulated electrophysiological correlates of perisaccadic responses after the last saccade prior to the decision epoch, when the subject declared her choice (in this case *happy*). To characterise responses to violations of local and global expectations, we repeated the simulations using exactly the same stimuli and actions but under different prior beliefs. Our hope here was to reproduce the classical mismatch negativity (MMN) response to unexpected (local) stimulus features (Strauss et al., 2015) – and a subsequent P300 (or N400) response to semantic (global) violations (Donchin and Coles, 1988). These distinct violation responses are important correlates of attentional processing and, clinically, conscious level and psychopathology (Morlet and Fischer, 2014; Light et al., 2015).

To simulate local (word or lexical) violations, we reversed the prior expectation of an upper case by switching the priors on the vertical flip for, and only for the last word. This produced greater excursions in the dynamics of belief updating. These can be seen as slight differences between the normal response (dotted lines) and response under local violations (solid lines) in the upper left panel of Fig. 8. The lower-level (lexical) expectations are shown in blue, while high-level (contextual) expectations are shown in red. Belief updating at the lower-level produces a fluctuation at around 100 ms known as an N1 response in ERP research. In contrast to these early (a.k.a. exogenous) responses, later (a.k.a. endogenous) responses appear to be dominated by expectations at the higher level. The difference waveforms (with and without surprising stimulus features) are shown on the upper right panel and look remarkably like a classical mismatch negativity. Note that the mismatch negativity peaks at about 170 ms and slightly postdates the N1. Again, this is exactly what is observed empirically; leading to debates about whether the generators of the N1 and MMM are the same or different. These simulations offer a definitive answer: the generators (neuronal encoding of expectations) are exactly the same; however, evidence accumulation is slightly slower when expectations are violated – leading to a protracted difference waveform.

To emulate global violations, we decreased the prior probability of the inferred (first) sentence by a factor of eight. This global (semantic) violation rendered the sampled word relatively surprising, producing a difference waveform with more protracted dynamics. Again, this is remarkably similar to empirical P300 responses seen with contextual
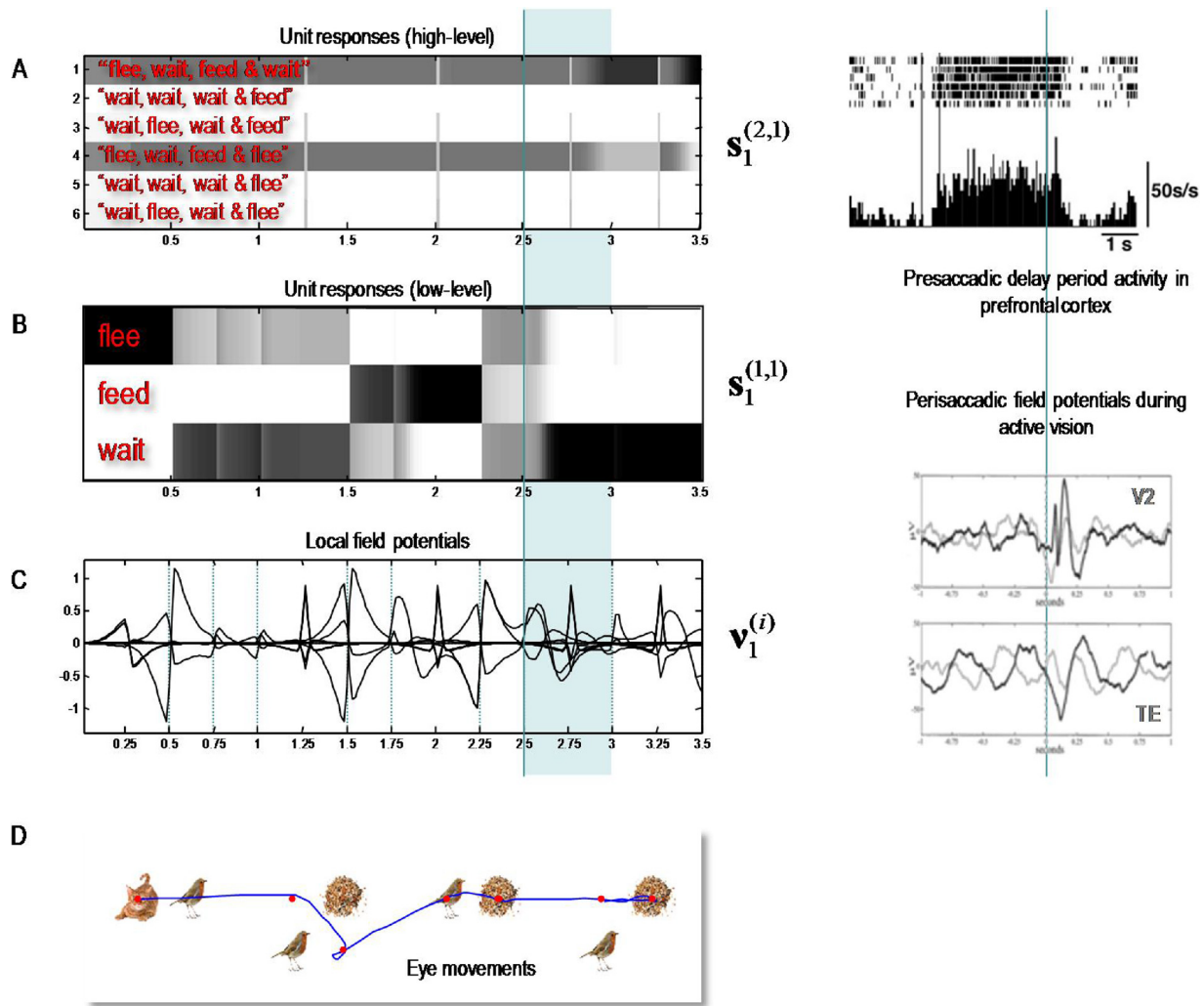
Fig. 7. Simulated electrophysiological responses during reading: these panels show the Bayesian belief updating that underlies the behaviour and expectations reported in the previous figure. Expectations about the initial hidden state (at the first time step) at the higher (upper panel A) and lower (middle panel B) hierarchical levels are presented in raster format, where an expectation of one corresponds to black (i.e., the firing rate activity corresponds to image intensity). The horizontal axis is time over the entire trial, where each iteration corresponds roughly to 16 ms. The vertical axis corresponds to the six sentences at the higher level and the three words at the lower level. The resulting patterns of firing rate over time show a marked resemblance to delay period activity in the prefrontal cortex prior to saccades. Saccade onsets are shown with the vertical (cyan) lines. The inset on the upper right is based upon the empirical results reported in (Funahashi, 2014). The transients in the lower panel (C) are the firing rates in the upper panels filtered between 4 Hz and 32 Hz – and can be regarded as (band pass filtered) fluctuations in depolarisation. These simulated local field potentials are again remarkably similar to empirical responses. The examples shown in the inset are based on the study of perisaccadic electrophysiological responses during activation reported in (Purpura et al., 2003). The upper traces come from early visual cortex (V2), while the lower traces come from inferotemporal cortex (TE). These can be thought of as first and second level empirical responses respectively. The lower panel (D) reproduces the eye movement trajectories of the previous figure. The simulated electrophysiological responses highlighted in cyan are characterised in more detail in the next figure.
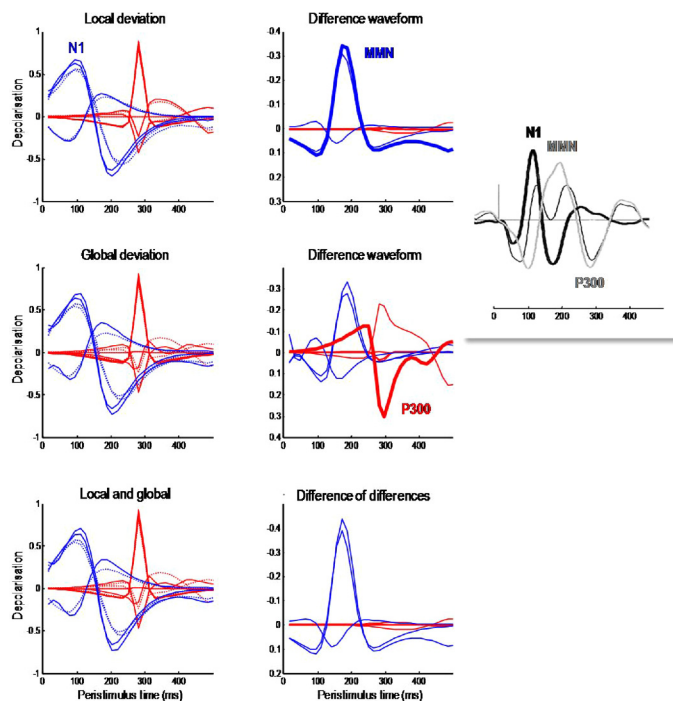
violations. It is well known that the amplitude of the P300 component is inversely related to the probability of stimuli (Donchin and Coles, 1988). The anterior P3a is generally evoked by stimuli that deviate from expectations. Indeed, novel stimuli generate a higher-amplitude P3a component than deviant but repeated stimuli. The P3b is a late positive component with a parietal (posterior) distribution seen in oddball paradigms and is thought to represent a context-updating operation (Donchin and Coles, 1988; Morlet and Fischer, 2014). Here, this context is operationalised in terms of changes in (sentence) context, under which lexical features are accumulated.

Finally, we combined the local and global priors to examine the interaction between local and global violations in terms of the difference of difference waveforms. The results are shown on the lower right and suggest that the effect of a global violation on the effects of a local violation (and *vice versa*) look similar to the mismatch negativity. This means, that the effect of a local violation on a global violation is

manifest as an increase in the amplitude of mismatch negativity and the positive P300 differences. Interestingly, this interaction appears to be restricted to low (lexical) representations. This suggests that, empirically, a late peak P300 like response to global violations may appear to be generated by sources normally associated with mismatch negativity (e.g., a shift to more anterior sources of the sort that define the P3a).

5.1. Summary

The opportunity to simulate these classical waveforms rests upon having a computationally and neurophysiologically plausible process theory that accommodates notions of violations and expectations. Happily, this is exactly the sort of framework offered by active inference. The MMN and P300 are particularly interesting from the point of view of clinical research and computational psychiatry (Montague et al., 2012). Indeed, their use in schizophrenia research (Umbricht and

Fig. 8. Simulated electrophysiological responses to violations: these simulated electrophysiological correlates focus on the perisaccadic responses around the last saccade prior to the final epoch, i.e. the cyan region highlighted in Fig. 7. The blue lines report the (filtered) expectations over peristimulus time at the first level (with one line for each of the three words), while the red lines show the evolution of expectations at the second level (with one line for each of the six sentences). Here, we repeated the simulations using the same stimuli and actions but under different prior beliefs. First, we reversed the prior expectation of a lower case by switching the informative priors on the vertical flip for, and only for, the last word. This means that the stimuli violated local expectations, producing slightly greater excursions in the dynamics of belief updating. These can be seen in slight differences between the normal standard response (dotted line) and the response to the surprising letter (solid lines) in the upper left panel. The ensuing difference waveform is shown on the upper right panel and looks remarkably like the classical mismatch negativity. Second, we made the inferred sentence surprising by decreasing its prior probability by a factor of eight. This global violation rendered the sampled word relatively surprising, producing a difference waveform with more protracted dynamics. Again, this is remarkably similar to empirical P300 responses seen with global violations. Finally, we combined the local and global priors to examine the interaction between local and global violations in terms of the difference in difference waveforms (these are not the differences in the difference waveforms above, which are referenced to the same normal response). The results are shown on the lower right and suggest that the effect of a global violation on the effects of a local violation (and *vice versa*) is largely restricted to early responses. The insert illustrates empirical event related potentials to unexpected stimuli elicited in patients with altered levels of consciousness to illustrate the form of empirical MMN and P300 responses (Fischer et al., 2000) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Krljes, 2005; Wang and Krystal, 2014; Light et al., 2015) was a partial motivation for the work reported in this paper.

6. Discussion

This paper has introduced the form and variational inversion of deep (hierarchical) temporal models for discrete (Markovian) hidden states and outcomes. This form of modelling has been important in machine learning; e.g. (Nefian et al., 2002; George and Hawkins, 2009), with a special focus on hierarchical or deep architectures (Salakhutdinov et al., 2013; Zorzi et al., 2013; Testolin and Zorzi, 2016). The technical contribution of this work is a formal treatment of discrete time in a hierarchical (nested or deep) setting and a simple set of belief update rules that follow from minimising variational free energy. Furthermore, this minimisation is contextualised within active inference to generate purposeful (epistemic and pragmatic) behaviour based on planning as inference (Botvinick and Toussaint, 2012).

The inference scheme presented here takes a potentially important step in explaining hierarchical temporal behaviour and how it may be orchestrated by the brain. There are a number of directions in which the scope of hierarchical schemes of the sort could be expanded. Firstly, to fully capture the dynamic character of language comprehension and production, means one has to handle systems of compositional recursive rules (Fodor and Pylyshyn, 1988; Pinker and Ullman, 2002), that underlie language grammars (Chomsky, 2006). This is likely to require deeper generative models that may entail some structure learning or nonparametric Bayesian methods (MacKay and Peto, 1995; Goldwater, 2006; Gershman and Niv, 2010; Collins and Frank, 2013). Secondly, there are subtle aspects to processing of serial order that have been identified and modelled (Page and Norris, 1998; Burgess and Hitch, 1999; Brown et al., 2000; Botvinick and Plaut, 2006). For example, without additional mechanisms, associative chaining models – in which chains are constructed with one-to-one dependencies between items – have difficulty modelling repetition (Lashley, 1951). This is because dependencies from an item typically change when it is repeated, requiring context dependent mechanisms to be added. The types-tokens framework (Kanwisher, 1987, Bowman and Wyble, 2007) may be a particularly useful way to handle repetition. In addition to

repetitions, error patterns in serial order recall seem inconsistent with a vanilla associative chaining model (Henson, 1998; Page and Norris, 1998). In further work on the deep temporal model presented here, we will explore extensions that enable language grammars and classic serial order recall data to be simulated.

From a neurobiological perspective, the belief updating appears to be sufficiently simple to be biologically plausible; resting on simple operators such as softmax functions, logarithmic transforms and linear algebra (that can be implemented using firing rate functions, nonlinear postsynaptic responses and neuronal connectivity respectively). Furthermore, the intrinsic and extrinsic connectivity required by the belief propagation appears to map gracefully to intrinsic and extrinsic connectivity within canonical microcircuits – and in the cortical-basal ganglia-thalamic loops responsible for action selection in the brain. The computational architecture that emerges from applying standard (variational) Bayesian belief updating to hierarchical models relates observable neuronal dynamics to underlying computational processes; an approach that might be applicable to temporally structured neurophysiological responses across different measurements and cognitive domains (Dehaene-Lambertz et al., 2006; Hasson et al., 2008; Cocchi et al., 2016). Finally, the biological plausibility of the resulting scheme acquires a predictive validity; in the sense that it reproduces some specific violation responses studied in state-of-the-art cognitive neuroscience (Strauss et al., 2015; Uhrig et al., 2016).

Software note

Although the generative model – specified by the (**A**, **B**, C, **D**) matrices – changes from application to application, the belief updates in Fig. 2 are generic and can be implemented using standard routines (here **spm_MDP_VB_X.m**). These routines are available as Matlab code in the SPM academic software: http://www.fil.ion.ucl.ac.uk/spm/. The simulations in this paper can be reproduced (and customised) via a graphical user interface: by typing **DEM** and selecting the **reading** demo.

Disclosure statement

The authors have no disclosures or conflict of interest.

Acknowledgement

Appendices

These appendices are provided for readers who are interested in technical details. In particular, they provide the derivations behind the (Matlab) implementation of active inference for Markov decision processes used to illustrate reading in this paper. A more detailed description of the update schemes can be found in the annot ated (Matlab) routines that are freely available as academic software.

Appendix 1

Belief updating: Bayesian inference is equivalent to maximising the marginal likelihood – or minimising variational free energy – with respect to beliefs (here, expectations) about hidden or latent variables. Free energy is essentially the difference between complexity and accuracy. The free energy for the generative model in Fig. 1 can be expressed as follows (ignoring constants):

$$
\begin{aligned}
F &= \underbrace{D[Q(\tilde{s},\pi)\,\|\,P(\tilde{s},\pi\,|\,\tilde{o})]}_{divergence} - \underbrace{\ln P(\tilde{o})}_{log\ evidence} \\
&= \underbrace{D[Q(\tilde{s},\pi)\,\|\,P(\tilde{s},\pi)]}_{complexity} - \underbrace{E_Q[\ln P(\tilde{o}\,|\,\tilde{s})]}_{accuracy} \\
&= \underbrace{D[Q(\tilde{s}\,|\,\pi)\,\|\,P(\tilde{s}\,|\,\pi)]}_{(state)\ complexity} + \underbrace{D[Q(\pi)\,\|\,P(\pi)]}_{(policy)\ complexity} - \underbrace{E_Q[\ln P(\tilde{o}\,|\,\tilde{s})]}_{accuracy} \\
&= \sum_i \boldsymbol{\pi}^{(i)} \cdot (\ln \boldsymbol{\pi}^{(i)} + \mathbf{F}^{(i)} + \mathbf{G}^{(i)})
\end{aligned}
$$

$$
\begin{aligned}
Q(\tilde{s},\pi) &= \prod_i Q(\tilde{s}^{(i)}\,|\,\pi^{(i)})Q(\pi^{(i)}) \\
Q(\tilde{s}^{(i)}\,|\,\pi^{(i)}) &= \prod_\tau Q(s_\tau^{(i)}\,|\,\pi^{(i)}) \\
Q(s_\tau^{(i)}\,|\,\pi^{(i)}) &= \prod_n Q(s_\tau^{(i,n)}\,|\,\pi^{(i)}) \\
Q(s_\tau^{(i,n)}\,|\,\pi^{(i)}) &= Cat(\mathbf{s}_{\pi,\tau}^{(i,n)}) = Q_{\pi,\tau}^{(i,n)} \\
Q(\pi^{(i)}) &= Cat(\boldsymbol{\pi}^{(i)})
\end{aligned}
$$

(A.1)

The (marginal) free energy of N hidden factors at the i-th level, under the π-th policy, is given by its path integral (see Appendix 2):

$$
\mathbf{F}_\pi^{(i)} = \tfrac{1}{N}\sum_{\tau,n}\mathbf{F}_{\pi,\tau}^{(i,n)}
$$

$$
\mathbf{F}_{\pi,\tau}^{(i,n)} = \underbrace{D[Q_{\pi,\tau}^{(i,n)}\,\|\,E_{Q_{\pi,\tau-1}^{(i,n)}}[P(s_\tau^{(i,n)}\,|\,s_{\tau-1}^{(i,n)},\pi^{(i)})]]}_{(forward)\ complexity} + \underbrace{D[Q_{\pi,\tau}^{(i,n)}\,\|\,E_{Q_{\pi,\tau+1}^{(i,n)}}[P(s_\tau^{(i,n)}\,|\,s_{\tau+1}^{(i,n)},\pi^{(i)})]]}_{(backward)\ complexity}
$$

$$
\underbrace{-\ln E_{Q_{\pi,\tau}^{(i,\backslash n)}}[P(o_\tau^{(i)}\,|\,s_\tau^{(i)})]}_{accuracy}
$$

$$
\mathbf{F}_{\pi,\tau}^{(i,n)} = \mathbf{s}_{\pi,\tau}^{(i,n)} \cdot \boldsymbol{\varepsilon}_{\pi,\tau}^{(i,n)}
$$

$$
\boldsymbol{\varepsilon}_{\pi,\tau}^{(i,n)} = 2\ln \mathbf{s}_{\pi,\tau}^{(i,n)} - \ln \mathbf{B}_{\pi,\tau-1}^{(i,n)}\mathbf{s}_{\pi,\tau-1}^{(i,n)} - \ln \mathbf{B}_{\pi,\tau}^{(i,n)} \cdot \mathbf{S}_{\pi,\tau+1}^{(i,n)} - \ln \mathbf{A}^{(i)} \circ \mathbf{s}_{\pi,\tau}^{(i,\backslash n)} \cdot o_\tau^{(i)}
$$

(A.2)

Eq. (A.2) expresses free energy as an expected *prediction error*, where the prediction error is the log expectation of the hidden state minus the prediction from its Markov blanket; namely, the previous state, the next state and other factors at the current time (i.e., parents of the current outcome). For the first hidden state, the prediction from the previous state is replaced by the prediction from the hierarchical level above; $\mathbf{B}_{\pi,0}^{(i,n)}\mathbf{s}_{\pi,0}^{(i,n)} \equiv \mathbf{D}^{(i,n)}\mathbf{s}^{(i+1)}$. For simplicity, we have omitted the time subscript for expected states at the higher level and have assumed a single outcome modality and a single higher-level factor.

The expected free energy of a policy (see Appendix 3) has the same form as (A.2), now allowing for M multiple outcome modalities:

$$
\mathbf{G}_\pi^{(i)} = \tfrac{1}{M}\sum_{\tau,m}\mathbf{G}_{\pi,\tau}^{(i,m)}
$$

$$
\mathbf{G}_{\pi,\tau}^{(i,m)} = \underbrace{D[Q(o_\tau^{(i,m)}\,|\,\pi^{(i)})\,\|\,P(o_\tau^{(i,m)}\,|\,s^{(i+1)})]}_{expected\ complexity\ (risk)} - \underbrace{E_{P(o_\tau^{(i,m)}|s_\tau^{(i)})Q(s_\tau^{(i)}|\pi^{(i)})}[\ln P(o_\tau^{(i,m)}\,|\,s_\tau^{(i)})]}_{expected\ accuracy\ (unambiguity)}
$$

$$
= (\mathbf{A}^{(i,m)} \circ \mathbf{s}_{\pi,\tau}^{(i)}) \cdot (\ln \mathbf{A}^{(i,m)} \circ \mathbf{s}_{\pi,\tau}^{(i)} - \ln \mathbf{C}_\tau^{(i,m)} \circ \mathbf{s}^{(i+1)}) + \mathbf{H}^{(i,m)} \circ \mathbf{s}_{\tau,\pi}^{(i)}
$$

$$
\mathbf{s}^{(i)} = \sum_\pi \boldsymbol{\pi}^{(i)} \cdot \mathbf{s}_\pi^{(i)}
$$

(A.3)

The updates for expectations of hidden states can now be formulated as a gradient descent on free energy, while policy expectations minimises free energy explicitly (ignoring constants):

$$
\dot{\mathbf{v}}_{\pi,\tau}^{(i,n)} = \partial_s F \approx \boldsymbol{\varepsilon}_{\pi,\tau}^{(i,n)} - \ln \mathbf{D}^{(i-1,n)} \circ \mathbf{s}^{(i/n)} \cdot \mathbf{s}(\tau)_1^{(i-1)}
$$

$$
\mathbf{s}_{\pi,\tau}^{(i,n)} = \sigma(-\mathbf{v}_{\pi,\tau}^{(i,n)}) \Rightarrow \mathbf{v}_{\pi,\tau}^{(i,n)} = -\ln \mathbf{s}_{\pi,\tau}^{(i,n)}
$$

$$
\boldsymbol{\pi}^{(i)} = \sigma(-\mathbf{F}^{(i)} - \mathbf{G}^{(i)}) \Rightarrow \partial_\pi F = 0
$$

(A.4)

Here, $\mathbf{s}(\tau)_1^{(i-1)}$ denotes the expected initial state in the lower level during epoch τ at the higher level. Note that the prediction error is simply the free energy gradient with respect to each hidden state. This means that optimising free energy through a gradient descent corresponds to destroying free

energy gradients (Tschacher and Haken, 2007).

Appendix 2

Marginal free energy lemma: the variational free energy in (A.2) has been formulated a slightly unusual way to finesse the overconfidence problem that attends mean field approximations. Minimising this marginal free energy, with respect to the approximate posterior, yields the true posterior by the following lemma:

Lemma (marginal free energy): under a mean field approximation $Q(s) = \prod_i Q^{(i)}$, the marginal free energy $\mathbf{F}^{(i)}$ is minimised when the approximate posteriors becomes the true marginal posteriors, at which point (negative) free energy becomes log evidence:

$$P\left(s^{(i)} \mid o\right) = \arg\min_{Q^{(i)}} \mathbf{F}^{(i)}$$

$$Q^{(i)} \triangleq Q(s^{(i)}) = P(s^{(i)} \mid o) : \forall i \Rightarrow \mathbf{F}^{(i)} = \ln P(o)$$

$$\mathbf{F}^{(i)} = \underbrace{D[Q^{(i)} \| \mathrm{E}_{Q^{(\backslash i)}}[P(s^{(i)} \mid s^{(\backslash i)}, o)]]}_{divergence} - \underbrace{\ln P(o)}_{log\ evidence}$$

$$= \underbrace{D[Q^{(i)} \| \mathrm{E}_{Q^{(\backslash i)}}[P(s^{(i)} \mid s^{(\backslash i)})]]}_{complexity} - \underbrace{\mathrm{E}_{Q^{(i)}}[\ln \mathrm{E}_{Q^{(\backslash i)}}[P(o \mid s^{(i)}, s^{(\backslash i)})]]}_{accuracy}$$

(A.5)

Proof: when the approximate posterior is equal to the true posterior, marginalisation over the Markov blanket of any hidden state returns its true posterior and the divergence attains its minimum of zero

$$Q^{(i)} = P(s^{(i)} \mid o) \Rightarrow$$

$$\mathrm{E}_{Q^{(\backslash i)}}\left[P(s^{(i)} \mid s^{(\backslash i)}, o)\right] = P(s^{(i)} \mid o) \Rightarrow \mathbf{F}^{(i)} = \ln P(o)$$

(A.6)

This means the free energy reduces to the log evidence □.

Remarks: to finesse the problem that we do not have an explicit joint posterior over the Markov blanket, one can replace the complexity based upon the joint posterior with the sum of complexities based upon the marginal posteriors – based upon the assumption that they share the same minimum.

$$\mathbf{F}^{(i)} = \underbrace{\sum_{j \neq i} D[Q^{(i)} \| \mathrm{E}_{Q^{(j)}}[P(s^{(i)} \mid s^{(j)})]]}_{complexity} - \underbrace{\mathrm{E}_{Q^{(i)}}[\ln \mathrm{E}_{Q^{(\backslash i)}}[P(o \mid s^{(i)}, s^{(\backslash i)})]]}_{accuracy}$$

$$Q^{(\backslash i)} = \prod_{j \neq i} Q^{(j)}$$

(A.7)

In practice, this appears to offer robust convergence. This leads to the marginal free energy in (A.3) in which the complexity can be separated into prior constraints from the parents and children of states at any point in time (that can be construed of in terms of a forward and backwards complexity).

Appendix 3

Expected free energy: variational free energy is a functional of a distribution over states, given observed outcomes. We can express this as a function of the sufficient statistics of the posterior:

$$F(\mathbf{s}) = E_{Q(s|\mathbf{s})}[\ln Q(s \mid \mathbf{s}) - \ln P(o, s)]$$

(A.8)

In contrast, the expected free energy is the average over (unobserved) outcomes, given some policy that determines the distribution over states. This can be expressed as a function of the policy:

$$F(\pi) = E_{P(o,s|\pi)}[\ln P(s \mid \pi) - \ln P(o, s \mid \pi)]$$

$$= E_{P(o,s|\pi)}[\ln P(s \mid \pi) - \ln P(o \mid s) - \ln P(s \mid \pi)]$$

$$= E_{P(s|\pi)}[H[P(o \mid s)]]$$

(A.9)

The expected free energy is therefore just the expected entropy or uncertainty about outcomes under a particular policy. Things get more interesting if we express the generative model-terms of a prior over outcomes that does not depend upon the policy

$$F(\pi) = E_{P(o,s|\pi)}[\ln P(s \mid \pi) - \ln P(s \mid o, \pi) - \ln P(o)]$$

$$= E_{P(o,s|\pi)}[\ln P(s \mid \pi) - \ln P(s \mid o, \pi)] - E_{P(o|\pi)}[\ln P(o)]$$

$$= E_{P(o,s|\pi)}[\ln P(o \mid \pi) - \ln P(o \mid s, \pi)] - E_{P(o|\pi)}[\ln P(o)]$$

$$= E_{P(o|\pi)}[\ln P(o \mid \pi) - \ln P(o)] - E_{P(o,s|\pi)}[\ln P(o \mid s, \pi)]$$

$$= \underbrace{D[P(o \mid \pi) \| P(o)]}_{risk} + \underbrace{E_{P(s|\pi)}[H[P(o \mid s)]]}_{ambiguity}$$

(A.10)

In this setting, the expected uncertainty or ambiguity is supplemented with a risk that scores the difference between the prior (preferred) outcomes and outcomes under a particular policy. This is the form used in active inference, where all the probabilities in (A.10) are conditioned upon past observations. This enables one to replace the posterior in (A.10) with the approximate posterior that minimises variational free energy based on (observed) outcomes in the past.

References

Arikuni, T., Kubota, K., 1986. The organization of prefrontocaudate projections and their laminar origin in the macaque monkey: a retrograde study using HRP-gel. J. Comp. Neurol. 244, 492–510.

Barlow, H., 1961. Possible principles underlying the transformations of sensory messages. In: Rosenblith, W. (Ed.), Sensory Communication. MIT Press, Cambridge, MA, pp. 217–234.

Barlow, H.B., 1974. Inductive inference, coding, perception, and language. Perception 3, 123–134.

Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., Friston, K.J., 2012. Canonical microcircuits for predictive coding. Neuron 76, 695–711.

Beal, M.J., 2003. Variational Algorithms for Approximate Bayesian Inference. PhD Thesis. University College London.

Botvinick, M., Toussaint, M., 2012. Planning as inference. Trends Cogn. Sci. 16, 485–488.

Botvinick, M.M., Plaut, D.C., 2006. Short-term memory for serial order: a recurrent neural network model. Psychol. Rev. 113, 201–233.

Bowman, H., Wyble, B., 2007. The simultaneous type, serial token model of temporal attention and working memory. Psychol. Rev. 114, 38–70.

Braun, D.A., Ortega, P.A., Theodorou, E., Schaal, S., 2011. Path integral control and bounded rationality. In: 2011 IEEE Symposium on Adaptive Dynamic Programming And Reinforcement Learning (ADPRL). Paris: IEEE. pp. 202–209.

Brown, G.D., Preece, T., Hulme, C., 2000. Oscillator-based memory for serial order. Psychol. Rev. 107, 127–181.

Burgess, N., Hitch, G.J., 1999. Memory for serial order: a network model of the phonological loop and its timing. Psychol. Rev. 106, 551.

Chomsky, N., 2006. Language and Mind. Cambridge University Press.

Cocchi, L., Sale, M.V., LG, L., Bell, P.T., Nguyen, V.T., Zalesky, A., Breakspear, M., Mattingley, J.B., 2016. A hierarchy of timescales explains distinct effects of local inhibition of primary visual cortex and frontal eye fields. Elife 5, e15252. http://dx. doi.org/10.7554/eLife.15252.

Collins, A.G.E., Frank, M.J., 2013. Cognitive control over learning: creating, clustering and generalizing task-set structure. Psychol. Rev. 120, 190–229.

Cruikshank, S.J., Ahmed, O.J., Stevens, T.R., Patrick, S.L., Gonzalez, A.N., Elmaleh, M., Connors, B.W., 2012. Thalamic control of layer 1 circuits in prefrontal cortex. J. Neurosci. 32, 17813–17823.

Dayan, P., Hinton, G.E., Neal, R., 1995. The Helmholtz machine. Neural Comput. 7, 889–904.

Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Meriaux, S., Roche, A., Sigman, M., Dehaene, S., 2006. Functional organization of perisylvian activation during presentation of sentences in preverbal infants. Proc. Natl. Acad. Sci. U. S. A. 103, 14240–14245.

Donchin, E., Coles, M.G.H., 1988. Is the P300 component a manifestation of context updating? Behav. Brain Sci. 11, 357.

Felleman, D., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47.

Ferro, M., Ognibene, D., Pezzulo, G., Pirrelli, V., 2010. Reading as active sensing: a computational model of gaze planning during word recognition. Front. Neurorobot. 4, 1.

Fischer, C., Morlet, D., Giard, M., 2000. Mismatch negativity and N100 in comatose patients. Audiol. Neuro-Otol. 5, 192–197.

FitzGerald, T., Dolan, R., Friston, K., 2014. Model averaging, optimal inference, and habit formation. Front. Hum. Neurosci. 26 (8), 457. http://dx.doi.org/10.3389/fnhum. 2014.00457.

FitzGerald, T.H., Dolan, R.J., Friston, K., 2015a. Dopamine, reward learning, and active inference. Front. Comput. Neurosci. 9, 136.

FitzGerald, T.H., Moran, R.J., Friston, K.J., Dolan, R.J., 2015b. Precision and neuronal dynamics in the human posterior parietal cortex during evidence accumulation. Neuroimage 107, 219–228.

FitzGerald, T.H., Schwartenbeck, P., Moutoussis, M., Dolan, R.J., Friston, K., 2015c. Active inference, evidence accumulation, and the urn task. Neural Comput. 27, 306–328.

Fodor, J.A., Pylyshyn, Z.W., 1988. Connectionism and cognitive architecture: a critical analysis. Cognition 28, 3–71.

Frank, M.J., 2005. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J. Cogn. Neurosci. 1, 51–72.

Friston, K., Adams, R., Montague, R., 2012a. What is value—accumulated reward or evidence? Front. Neurorobot. 6, 11.

Friston, K., Samothrakis, S., Montague, R., 2012b. Active inference and agency: optimal control without cost functions. Biol. Cybernet. 106 (8–9), 523–541 [Epub ahead of print].

Friston, K., Buzsaki, G., 2016. The functional anatomy of time: what and when in the brain. Trends Cogn. Sci. 20 (7), 500–511.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., O'Doherty, J., Pezzulo, G., 2016. Active inference and learning. Neurosci. Biobehav. Rev. 68, 862–879.

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., Pezzulo, G., 2015. Active inference and epistemic value. Cogn. Neurosci. 1–28.

Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., Dolan, R.J., 2014. The anatomy of choice: dopamine and decision-making. Philos. Trans. R. Soc. Lond. B Biol. Sci. 369.

Funahashi, S., 2014. Saccade-related activity in the prefrontal cortex: its role in eye movement control and cognitive functions. Front. Integr. Neurosci. 8, 54.

George, D., Hawkins, J., 2009. Towards a mathematical theory of cortical micro-circuits.

PLoS Comput. Biol. 5, e1000532.

Gershman, S.J., Niv, Y., 2010. Learning latent structure: carving nature at its joints. Curr. Opin. Neurobiol. 20, 251–256.

Goldwater, S., 2006. Nonparametric Bayesian Models of Lexical Acquisition. Brown University.

Haeusler, S., Maass, W., 2007. A statistical analysis of information-processing properties of lamina-specific cortical microcircuit models. Cereb. Cortex 17, 149–162.

Hasson, U., Yang, E., Vallines, I., Heeger, D.J., Rubin, N., 2008. A hierarchy of temporal receptive windows in human cortex. J. Neurosci. 28, 2539–2550.

Henson, R.N., 1998. Short-term memory for serial order: the start-end model. Cognit. Psychol. 36, 73–137.

Howard, R., 1966. Information value theory. IEEE Trans. Syst. Sci. Cybernet. SSC 2, 22–26.

Itti, L., Baldi, P., 2009. Bayesian surprise attracts human attention. Vis. Res. 49, 1295–1306.

Jahanshahi, M., Obeso, I., Rothwell, J.C., Obeso, J.A., 2015. A fronto-striato-subthalamic-pallidal network for goal-directed and habitual inhibition. Nat. Rev. Neurosci. 16, 719–732.

Kanwisher, N.G., 1987. Repetition blindness: type recognition without token individuation. Cognition 27, 117–143.

Kappen, H.J., Gomez, Y., Opper, M., 2012. Optimal control as a graphical model inference problem. Mach. Learn. 87, 159–182.

Kiebel, S.J., Daunizeau, J., Friston, K., 2008. A hierarchy of time-scales and the brain. PLoS Comput. Biol. 4, e1000209.

Kiebel, S.J., Daunizeau, J., Friston, K.J., 2009. Perception and hierarchical dynamics. Front. Neuroinform. 3, 20.

Klyubin, A.S., Polani, D., Nehaniv, C.L., 2005. Empowerment: a universal agent-centric measure of control. Proc CEC 2005 IEEE 1, 128–135.

Kojima, S., Goldman-Rakic, P.S., 1982. Delay-related activity of prefrontal neurons in rhesus monkeys performing delayed response. Brain Res. 248, 43–49.

Lashley, K.S., 1951. The Problem of Serial Order in Behavior. Bobbs-Merrill.

Laughlin, S.B., 2001. Efficiency and complexity in neural coding. Novartis Found. Symp. 239, 177–187.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444.

Lee, J., Whittington, M., Kopell, N., 2013. Top-down beta rhythms support selective attention via interlaminar interaction: a model. PLoS Comput. Biol. 9, e1003164.

Light, G.A., Swerdlow, N.R., Thomas, M.L., Calkins, M.E., Green, M.F., Greenwood, T.A., Gur, R.E., Gur, R.C., Lazzeroni, L.C., Nuechterlein, K.H., Pela, M., Radant, A.D., Seidman, L.J., Sharp, R.F., Siever, L.J., Silverman, J.M., Sprock, J., Stone, W.S., Sugar, C.A., Tsuang, D.W., Tsuang, M.T., Braff, D.L., Turetsky, B.I., 2015. Validation of mismatch negativity and P3a for use in multi-site studies of schizophrenia: characterization of demographic, clinical, cognitive, and functional correlates in COGS-2. Schizophr. Res. 163, 63–72.

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

Lopes da Silva, F., 1991. Neural mechanisms underlying brain waves: from neural membranes to networks. Electroencephalogr. Clin. Neurophysiol. 79, 81–93.

MacKay, D.J.C., Peto, L.C.B., 1995. A hierarchical Dirichlet language model. Nat. Lang. Eng. 1, 289–308.

Markov, N., Ercsey-Ravasz, M., Van Essen, D., Knoblauch, K., Toroczkai, Z., Kennedy, H., 2013. Cortical high-density counterstream architectures. Science 342, 1238406.

Mirza, M.B., Adams, R.A., Mathys, C.D., Friston, K.J., 2016. Scene construction, visual foraging, and active inference. Front. Comput. Neurosci. 10, 56.

Montague, P.R., Dolan, R.J., Friston, K.J., Dayan, P., 2012. Computational psychiatry. Trends Cogn. Sci. 16, 72–80.

Morlet, D., Fischer, C., 2014. MMN and novelty P3 in coma and other altered states of consciousness: a review. Brain Topogr. 27, 467–479.

Moutoussis, M., Trujillo-Barreto, N.J., El-Deredy, W., Dolan, R.J., Friston, K.J., 2014. A formal model of interpersonal inference. Front. Hum. Neurosci. 8 (160).

Murray, J.D., Bernacchia, A., Freedman, D.J., Romo, R., Wallis, J.D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., Wang, X.J., 2014. A hierarchy of intrinsic timescales across primate cortex. Nat. Neurosci. 17, 1661–1663.

Nefian, A.V., Liang, L.H., Pi, X.B., Liu, X.X., Murphy, K., 2002. Dynamic Bayesian networks for audio-visual speech recognition. EURASIP J. Appl. Signal Process. 2002, 1274–1288.

Olshausen, B.A., Field, D.J., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381, 607–609.

Ortega, P.A., Braun, D.A., 2013. Thermodynamics as a theory of decision-making with information-processing costs. Proc. R. Soc. A 469, 2153.

Page, M.P., Norris, D., 1998. The primacy model: a new model of immediate serial recall. Psychol. Rev. 105, 761–781.

Pastalkova, E., Itskov, V., Amarasingham, A., Buzsaki, G., 2008. Internally generated cell assembly sequences in the rat hippocampus. Science 321, 1322–1327.

Pinker, S., Ullman, M.T., 2002. The past and future of the past tense. Trends Cogn. Sci. 6, 456–463.

Purpura, K.P., Kalik, S.F., Schiff, N.D., 2003. Analysis of perisaccadic field potentials in the occipitotemporal pathway during active vision. J. Neurophysiol. 90, 3455–3478.

Rayner, K., 1978. Eye movements in reading and information processing. Psychol. Bull. 85, 618–660.

Rayner, K., 2009. Eye movements in reading: models and data. J. Eye Mov. Res. 2, 1–10.

Rayner, K., White, S.J., Johnson, R.L., Liversedge, S.P., 2006. Raeding wrods with jumbled lettres: there is a cost. Psychol. Sci. 17, 192–193.

Salakhutdinov, R., Tenenbaum, J.B., Torralba, A., 2013. Learning with hierarchical-deep models. IEEE Trans. Pattern Anal. Mach. Intell. 35, 1958–1971.

Schmidhuber, J., 1991. Curious model-building control systems. In: Singapore IEEE. Proc

International Joint Conference on Neural Networks, vol. 2. pp. 1458–1463.

Schwartenbeck, P., FitzGerald, T.H., Mathys, C., Dolan, R., Friston, K., 2015a. The dopaminergic midbrain encodes the expected certainty about desired outcomes. Cereb. Cortex 25, 3434–3445.

Schwartenbeck, P., FitzGerald, T.H., Mathys, C., Dolan, R., Kronbichler, M., Friston, K., 2015b. Evidence for surprise minimization over value maximization in choice behavior. Sci. Rep. 5, 16575.

Schwartenbeck, P., FitzGerald, T.H., Mathys, C., Dolan, R., Wurst, F., Kronbichler, M., Friston, K., 2015c. Optimal inference with suboptimal models: addiction and active Bayesian inference. Med. Hypotheses 84, 109–117.

Sohal, V.S., Zhang, F., Yizhar, O., Deisseroth, K., 2009. Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. Nature 459, 698–702.

Still, S., Precup, D., 2012. An information-theoretic approach to curiosity-driven reinforcement learning. Theory Biosci. = Theorie in den Biowissenschaften 131, 139–148.

Strauss, M., Sitt, J.D., King, J.R., Elbaz, M., Azizi, L., Buiatti, M., Naccache, L., van Wassenhove, V., Dehaene, S., 2015. Disruption of hierarchical predictive coding

during sleep. Proc. Natl. Acad. Sci. U. S. A. 112, E1353–E1362.

Testolin, A., Zorzi, M., 2016. Probabilistic models and generative neural networks: towards an unified framework for modeling normal and impaired neurocognitive functions. Front. Comput. Neurosci. 10, 73.

Thomson, A.M., Bannister, A.P., 2003. Interlaminar connections in the neocortex. Cereb. Cortex 13, 5–14.

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

Uhrig, L., Janssen, D., Dehaene, S., Jarraya, B., 2016. Cerebral responses to local and global auditory novelty under general anesthesia. Neuroimage 141, 326–340.

Umbricht, D., Krljes, S., 2005. Mismatch negativity in schizophrenia: a meta-analysis. Schizophr. Res. 76, 1–23.

van den Broek, J.L., Wiegerinck, W.A.J.J., Kappen, H.J., 2010. Risk-sensitive Path Integral Control, vol. 6. UAI, pp. 1–8.

Wang, X.J., Krystal, J.H., 2014. Computational psychiatry. Neuron 84, 638–654.

Zorzi, M., Testolin, A., Stoianov, I.P., 2013. Modeling language and cognition with deep unsupervised learning: a tutorial overview. Front. Psychol. 4, 515.