Diss. ETH No. 20909

HIERARCHICAL GAUSSIAN FILTERING

Construction and variational inversion of a generic Bayesian model of individual learning under uncertainty

Thesis submitted for the degree of

DOCTOR OF SCIENCES
ETH ZURICH
by

Christoph Daniel Mathys

M.Sc., ETH Zurich (Dipl. Natw. ETH)
M.Sc., University of Zurich (lic. phil.)

of Erlenbach ZH, Zürich ZH, and Rohrbachgraben BE

accepted on the recommendation of

Prof. Klaas Prüssmann, Ph.D., examiner, ETH Zurich, Switzerland
Prof. Klaas Enno Stephan, M.D., Ph.D., co-examiner, ETH Zurich, Switzerland
Timothy E.J. Behrens, Ph.D., co-examiner, University of Oxford, England

2012

Abstract

Computational learning models are critical for understanding mechanisms of adaptive behavior. However, the two major current frameworks, reinforcement learning (RL) and Bayesian learning, both have certain limitations. For example, many Bayesian models are agnostic of inter-individual variability and involve complicated integrals, making online learning difficult. Here, I introduce the hierarchical Gaussian Filter (HGF), a generic hierarchical Bayesian framework for individual learning under multiple forms of uncertainty (e.g., environmental volatility and sensory uncertainty). The HGF assumes Gaussian random walks of states, with the step size determined by the next higher level. The coupling between levels is controlled by parameters that shape the influence of uncertainty on learning in a subject-specific fashion. Using variational Bayes under a mean field approximation and a novel approximation to the posterior energy function, I derive trial-by-trial update equations which (i) are analytical and extremely efficient, enabling real-time learning, (ii) have a natural interpretation in terms of RL, and (iii) contain parameters representing processes which play a key role in current theories of learning, e.g., precision-weighting of prediction error. These parameters allow for the expression of individual differences in learning and may relate to specific neuromodulatory mechanisms in the brain. The HGF is very general: it can deal with both discrete and continuous states and equally accounts for deterministic and probabilistic relations between environmental events and perceptual states (i.e., situations with and without sensory uncertainty). These properties are illustrated by simulations and analyses of empirical time series from financial markets, behavioral experiments, and fMRI measurements. Overall, this framework provides a novel foundation for understanding normal and pathological learning that contextualizes RL within a generic Bayesian scheme and thus connects it to principles of optimality from probability theory.

Contents

1 INTRODUCTION

1.1 Inference and Bayesian learning

Learning can be understood as the process of updating an agent's beliefs about the world by integrating new and old information. This enables the agent to exploit past experience and improve predictions about the future; e.g., the consequences of chosen actions. Understanding how biological agents, such as humans or animals, learn requires a specification of both the computational principles and their neurophysiological implementation in the brain. This can be approached in a bottom-up fashion, building a neuronal circuit from neurons and synapses and studying what forms of learning are supported by the ensuing neuronal architecture. Alternatively, one can choose a top-down approach, using generic computational principles to construct generative models of learning and use these to infer on underlying mechanisms. (Daunizeau et al., 2010a, 2010b). The latter approach is the one that I pursue in this thesis.

The laws of inductive inference, prescribing an optimal way to learn from new information, have long been known (Laplace, 1774, 1812). They have a unique mathematical form, i.e. it has been proven that there is no alternative formulation of inductive reasoning that does not violate either consistency or common sense (Cox, 1946). Inductive reasoning is also known as Bayesian learning because the requisite updating of conditional probabilities is described by Bayes' theorem. Since a Bayesian learner processes information optimally, it should have an evolutionary advantage over other types of agents, and one might therefore expect the human brain to have evolved such that it implements an ideal Bayesian learner (Geisler and Diehl, 2002). Indeed, there is substantial evidence from studies on various domains of learning and perception that human behavior is better described by Bayesian models than by other theories (e.g., Behrens et al., 2007; Bresciani et al., 2006; den Ouden et al., 2010; Kording and Wolpert, 2004; Orbán et al., 2008; Xu and Tenenbaum, 2007; Yuille and Kersten, 2006). However, there remain at least three serious difficulties with the hypothesis that humans act as ideal Bayesian learners. The first problem is that in all but the simplest cases, the application of

Bayes' rule involves complicated integrals that require burdensome and time-consuming numerical calculations. This makes online learning a challenging task for Bayesian models, and any evolutionary advantage conferred by optimal learning might be outweighed by these computational costs. A second and associated problem is how ideal Bayesian learning, with its requirement to evaluate high-dimensional integrals, would be implemented neuronally (cf. Beck et al., 2008; Deneve, 2008; Yang and Shadlen, 2007). The third difficulty is that Bayesian learning constitutes a normative framework that prescribes how information *should* be dealt with. In reality, though, even when endowed with equal prior knowledge, not all agents process new information alike. Instead, even under carefully controlled conditions, animals and humans display considerable inter-individual variability in learning (e.g., Daunizeau et al., 2010a; Gluck et al., 2002). Despite previous attempts of Bayesian models to deal with individual variability (e.g., Steyvers et al., 2009; Nassar et al., 2010), the failure of orthodox Bayesian learning theory to account for these individual differences remains a key problem for understanding (mal)adaptive behavior of humans. Formal and mechanistic characterizations of this inter-subject variability are needed to comprehend fundamental aspects of brain function and disease. For example, individual differences in learning may result from inter-individual variability in basic physiological mechanisms, such as the neuromodulatory regulation of synaptic plasticity (Thiel et al., 1998), and such differences may explain the heterogeneous nature of psychiatric diseases (Stephan et al., 2009a).

1.2 Reinforcement learning

The difficulties of Bayesian learning have been avoided by descriptive approaches to learning, which are not grounded in probability theory, notably some forms of reinforcement learning (RL), where agents learn the "value" of different stimuli and actions (Sutton and Barto, 1998; Dayan and Niv, 2008). While RL is a wide field encompassing a variety of schemes, perhaps the most prototypical and widely used model is that by (Rescorla and Wagner, 1972). In this description, predictions of value are updated in relation to the current prediction error, weighted by a learning rate (which may differ across individuals and contexts). The great

advantage of this scheme is its conceptual simplicity and computational efficiency. It has been applied empirically to various learning tasks (Sutton and Barto, 1998) and has played a major role in attempts to explain electrophysiological and functional magnetic resonance imaging (fMRI) measures of brain activity during reward learning (e.g., O'Doherty et al., 2004; Schultz et al., 1997; Daw et al., 2006; Montague et al., 2004). Furthermore, its non-normative descriptive nature allows for modeling aberrant modes of learning, such as in schizophrenia or depression (Dayan and Huys, 2009; Frank, 2008; Murray et al., 2007; Smith et al., 2006). Similarly, it has found widespread use in modeling the effects of neuromodulatory transmitters, such as dopamine, on learning (e.g., Yu and Dayan, 2005; Doya, 2008; Pessiglione et al., 2006).

Despite these advantages, RL also suffers from major limitations. On the theoretical side, it is a heuristic approach that does not follow from the principles of probability theory. In practical terms, it often performs badly in real-world situations where environmental states and the outcomes of actions are not known to the agent, but must also be inferred or learned. These practical limitations have led some authors to argue that Bayesian principles and "structure learning" are essential in improving RL approaches (Gershman and Niv, 2010). In this thesis, I introduce a novel model of Bayesian learning that overcomes the three limitations of ideal Bayesian learning discussed above (i.e., computational complexity, questionable biological implementation, and failure to account for individual differences) and that connects Bayesian learning to RL schemes.

1.3 The proposed hierarchical Bayesian generative model

Any Bayesian learning scheme relies upon the definition of a so-called "generative model", i.e. a set of probabilistic assumptions about how sensory signals are generated. The generative model I propose is inspired by the seminal work of (Behrens et al., 2007) and comprises a hierarchy of states that evolve in time as Gaussian random walks, with each walk's step size determined by the next higher level of the hierarchy. This model can be inverted (fitted) by an agent using a mean field approximation and a novel approximation to the conditional probabilities over unknown

quantities that replaces the conventional Laplace approximation. This enables us to derive closed-form update equations for the posterior expectations of all hidden states governing contingencies in the environment. This results in extremely efficient computations that allow for real-time learning. The form of these update equations is similar to those of Rescorla-Wagner learning, providing a Bayesian analogon to RL theory. Finally, by introducing parameters that determine the nature of the coupling between the levels of the hierarchical model, the optimality of an update is made conditional upon parameter values that may vary from agent to agent. These parameters encode prior beliefs about higher-order structure in the environment and enable the model to account for inter-individual (and inter-temporal intra-individual) differences in learning. In other words, the model is capable of describing behavior that is subjectively optimal (in relation to the agent's prior beliefs) but objectively maladaptive. Importantly, the model parameters that determine the nature of learning may relate to specific physiological processes, such as the neuromodulation of synaptic plasticity. For example, it has been hypothesized that dopamine, which regulates plasticity of glutamatergic synapses (Gu, 2002), may encode the precision of prediction errors (Friston, 2009). In my model, this precision-weighting of prediction errors is determined by the model's parameters (cf. Figure 4). Ultimately, this approach may therefore be useful for model-based inference on subject-specific computational and physiological mechanisms of learning, with potential clinical applications for diagnostic classifications of psychiatric spectrum disorders (Stephan et al., 2009a).

1.4 Overview

This thesis is structured as follows:

- Chapter 2: This chapter lays the groundwork by introducing the main contribution of this work: the hierarchical Gaussian filter (HGF). Along with the general formulation, a simple example model is introduced that will serve as a practical tool in many applications throughout the thesis.

- Chapter 3: Here, we deal with the inversion of the HGF. Exact inversion is discussed, variational inversion is explained in the context of the mean field approximation and the corresponding variational energies are derived.

- Chapter 4: Closed-form update equations are derived and interpreted. After their introduction, the quadratic approximation that leads to them is discussed in detail, and the structure of the equations is analyzed extensively. The updates are then discussed in the context of a very general feature of Bayesian updating: precision-weighting of prediction errors. Finally, their properties are explored in a series of simulations.

- Chapter 5: This chapter deals with sensory input, that is with the interface between the perceptual prediction-generating machine that the HGF is and the hidden states of the world that it is predicting. Sensory uncertainty is discussed both in the case of binary categorical and in the case of continuous hidden states.

- Chapter 6: Response models are introduced, completing the "observing the observer" framework. Two simple examples of response models are derived from first principles (i.e., from a loss function).

- Chapter 7: In this chapter, we deal with parameter estimation. First, we establish theoretically, illustrated by an example, what parameters we can estimate under what circumstances. We then lay out the mechanism of inference and discuss the choice of priors and parameter transformations. To establish that parameter estimation really works, I then present a simulation study that we conducted, comparing four different optimization strategies. We then look at the optimization of parameters, that is at how to find the parameter values that imply least surprise for an agent receiving a given set of inputs. Finally, we discuss Bayesian model comparison and model selection.

- Chapter 8: The HGF toolbox, a collection of Matlab routines implementing the HGF and related models (and their estimation), is introduced.

- Chapter 9: We report a study applying the HGF to saccadic reaction times in a modified Posner paradigm. The HGF here serves as the framework in which to assess competing response models and is itself subjected to comparison with competing models.

- Chapter 10: In this chapter, the HGF is used in the analysis of an fMRI study that introduces a novel paradigm: the observed iterated trust game. This amounts to the observation of an observing observer and provides for the modeling of complex mental inferences. Despite their complexity, these inferred mental states appear to have associated neuronal activations as revealed by fMRI.

2 THE GENERATIVE MODEL

2.1 The Hierarchical Gaussian Filter (HGF)

The goal of the model I introduce here is simple and general: to describe how an agent learns about a continuous quantity x that moves. One simple way of describing this motion is a Gaussian random walk:

$$x^{(k)} \sim \mathcal{N}\left(x^{(k-1)}, \vartheta\right) \tag{1}$$

where k is a time index, and $x^{(k-1)}$ and ϑ are the mean and variance (not standard deviation) of a Gaussian distribution, respectively. In this formulation, the volatility in x is governed by the positive constant ϑ; however, there is in principle no reason to assume that volatility is constant. To allow for changes in volatility, we replace ϑ by a positive function f of a second quantity, x_2, while x becomes x_1:

$$x_1^{(k)} \sim \mathcal{N}\left(x_1^{(k-1)}, f(x_2)\right) \tag{2}$$

We may now further assume that x_2 performs a Gaussian random walk of its own, albeit with a constant variance ϑ (Eq. (1)), taking the former role of x_1. This adding of levels of Gaussian random walks coupled by their variances can now continue up to any number n of levels in the hierarchy, as illustrated in Figure 1. At each level i, the coupling to the next highest level $i+1$ is given by a positive function $f_i(x_{i+1})$ which represents the variance or step size of the random walk:

$$x_i^{(k)} \sim \mathcal{N}\left(x_i^{(k-1)}, f_i(x_{i+1})\right), \quad i = 1, \ldots, n-1. \tag{3}$$

At the top level, instead of f_n, we have ϑ:

$$x_n^{(k)} \sim \mathcal{N}\left(x_n^{(k-1)}, \vartheta\right). \tag{4}$$

The estimation of such a time-series model to yield a prediction on $x_1^{(k)}$ given inputs (i.e., measurements of x_1) $u^{(1)}, u^{(2)}, \ldots, u^{(k-1)}$ is called *filtering* (as opposed to *smoothing*, which refers to estimating $x_1^{(k)}$ on the basis of inputs earlier and also later than $k-1$). Since the model consists of a hierarchy of Gaussian random walks and its purpose is filtering, I call it the *Hierarchical Gaussian Filter (HGF).*

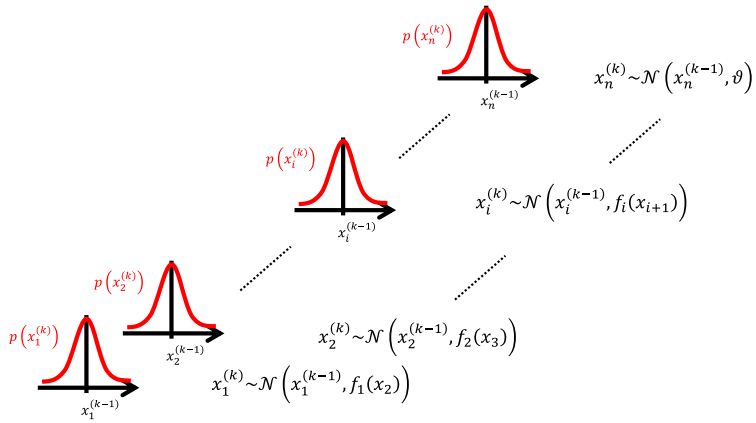

Figure 1. ***Overview of the hierarchical generative model.*** *The model represents a hierarchy of coupled Gaussian random walks.. The levels of the hierarchy relate to each other by determining the step size (volatility or variance) of a random walk. The topmost step size is a constant parameter ϑ.*

2.2 The coupling between levels

To complete the model, we still need to define the f_i in Eq. (3). A flexible and straightforward approach to this is to allow any positive analytic f_i, but to expand it in powers to first order to give it a simple functional form. However, since f_i has to be everywhere positive, we cannot approximate it by expanding it directly. Instead, dropping indices for clarity, we expand its logarithm.

$$f(x) > 0 \; \forall \, x \implies \exists \; g \colon f(x) = \exp\big(g(x)\big) \; \forall \, x \tag{5}$$

$$\begin{aligned} g(x) &= g(a) + g'(a) \cdot (x - a) + \mathrm{O}(2) = \log f(x) = \\ &= \log f(a) + \tfrac{f'(a)}{f(a)} \cdot (x - a) + \mathrm{O}(2) = \\ &= \underbrace{\frac{f'(a)}{f(a)}}_{\stackrel{\text{def}}{=} \kappa} \cdot x + \underbrace{\log f(a) - a \cdot \frac{f'(a)}{f(a)}}_{\stackrel{\text{def}}{=} \omega} + \mathrm{O}(2) = \\ &= \kappa x + \omega + \mathrm{O}(2) \end{aligned} \tag{6}$$

$$\implies f(x) \approx \exp(\kappa x + \omega) \tag{7}$$

This motivates my definition of coupling between levels:

$$f_i(x_{i+1}) \stackrel{\text{def}}{=} \exp(\kappa_i x_{i+1} + \omega_i) \tag{8}$$

Moreover, this form of coupling has the advantage of enabling the derivation of simple one-step update equations for an approximate inversion of the model. Before I derive these update equations in Chapter 3, I will first give an example of this model under minimal assumptions: noiseless binary inputs and a minimal number of levels.

2.3 The generative model under minimal assumptions

A full generative model consists of the joint probability density of all states and parameters. This joint density is the product of the likelihood function defined in Eqs (3), (4), and (8) and a prior density over states and parameters. To make the construction of the joint density clearer and also to introduce a first concrete example of a hierarchical Gaussian filtering model, I first give the joint density of a very simple model before turning to the general case. This simple example will also serve us in further sections and chapters.

An overview of this simple form of the generative model is given by Figure 2. We imagine an agent who receives a sequence of sensory inputs $u^{(1)}, u^{(2)}, \dots, u^{(K)}$. To keep the presentation as simple as possible, we initially deal with an environment where the sensory input $u^{(k)} \in \{0,1\}$ on trial k is of a binary form. In this form, the generative model can be seen as an extension of the model proposed by (Daunizeau et al., 2010b). It is then also similar to the model of (Behrens et al., 2007). However, in its general form, it can deal with states and inputs that are discrete or continuous and uni- or multivariate, and it equally accounts for deterministic and probabilistic relationships between environmental events and perceptual states (i.e., situations with and without sensory uncertainty). In fact, at a later stage we will also deal with stochastic mappings (i.e., sensory uncertainty; cf. Eq. (72)) and continuous (real-valued) inputs and states (cf. Eq. (75)).

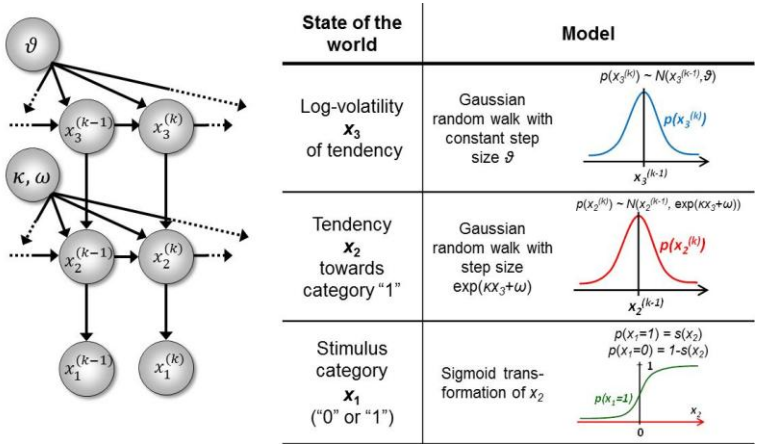

Figure 2. ***The simple example model.*** *There are three levels; the lowest level x_1 is binary and corresponds, in the absence of sensory noise, to sensory input u. Left: schematic representation of the generative model as a Bayesian network. $x_1^{(k)}$, $x_2^{(k)}$, $x_3^{(k)}$ are hidden states of the environment at time point k. They generate $u^{(k)}$, the input at time point k, and depend on their immediately preceding values $x_2^{(k-1)}$, $x_3^{(k-1)}$ and on the on parameters κ, ω, ϑ. Right: model definition.*

To illustrate the situation we are modeling here, imagine a case where the agent is only interested in a single (binary) state of its environment; e.g., whether it is light or dark. In this model, the environmental state x_1 at time k, denoted by $x_1^{(k)} \in \{0,1\}$, causes input $u^{(k)}$. Here, x_1 could represent the on/off state of a light switch and u the sensation of light or darkness. (For simplicity, I will often omit the time index k.) Note that while in the previous sections, all level x_i were continuous, we now have a binary level x_1 at the bottom of the hierarchy. The equations in the previous sections accordingly apply to levels $i > 1$.

Since we assume that the binary state x_1 can be observed directly, without sensory uncertainty, we have

$$x_1 = u. \tag{9}$$

This will later be replaced by a stochastic mapping when dealing with sensory uncertainty (cf. Eq. (72), Section 5.1).

Since x_1 is binary, its probability distribution can be described by a single real number, the state x_2 at the next level of the hierarchy. We map x_2 to the probability of x_1 such that $x_2 = 0$ means that $x_1 = 0$ and $x_1 = 1$ are equally probable. For $x_2 \to \infty$ the probability for $x_1 = 1$ and $x_1 = 0$ should approach 1 and 0, respectively. Conversely, for $x_2 \to -\infty$ the probabilities for $x_1 = 1$ and $x_1 = 0$ should approach 0 and1, respectively. This can be achieved with the following conditional density:

$$p(x_1|x_2) = s(x_2)^{x_1}\big(1 - s(x_2)\big)^{1-x_1} = \text{Bernoulli}\big(x_1; s(x_2)\big), \tag{10}$$

where $s(\cdot)$ is the logistic sigmoid (softmax) function:

$$s(x) \stackrel{\text{def}}{=} \frac{1}{1 + \exp(-x)} \tag{11}$$

Put simply, $x_2 \in \mathbb{R}$ is an unbounded real parameter of the probability that $x_1 = 1$. In the light/dark example, one might interpret x_2 as the tendency of the light to be on.

For the sake of generality, we make no assumptions about the probability of x_2 except that it may change with time as a Gaussian random walk. This means that the value of x_2 at time k will be normally distributed around its value at the previous time point, $x_2^{(k-1)}$:

$$p\left(x_2^{(k)}\middle|x_2^{(k-1)}, x_3^{(k)}\right) = \mathcal{N}\left(x_2^{(k)}; x_2^{(k-1)}, \exp(\kappa x_3^{(k)} + \omega)\right) \tag{12}$$

Importantly, the dispersion of the random walk (i.e., the variance $\exp(\kappa \cdot x_3 + \omega)$ of the conditional probability) is determined by the parameters κ and ω as well as by the state x_3. Here, this state determines the log-volatility of the environment (cf. Behrens et al., 2007, 2008). In other words, the tendency x_2 of the light switch to be on performs a Gaussian random walk with volatility $\exp(\kappa \cdot x_3 + \omega)$. Introducing ω allows for a volatility that scales independently of the state x_3. Everything applying to x_2 now equally applies to x_3, such that we could add as many levels as we please. Here, we stop at the fourth level, and set the volatility of x_3 to ϑ, a constant parameter (which may again differ across agents):

$$p\left(x_3^{(k)}\middle|x_3^{(k-1)}, \vartheta\right) = \mathcal{N}\left(x_3^{(k)}; x_3^{(k-1)}, \vartheta\right) \tag{13}$$

Given full priors on the parameters, i.e. $p(\vartheta, \omega, \kappa)$, we can now write the full generative model

$$\begin{aligned} &p\left(u^{(k)}, x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, x_2^{(k-1)}, x_3^{(k-1)}, \kappa, \omega, \vartheta\right) \\ &= p\left(u^{(k)}\middle|x_1^{(k)}\right) p\left(x_1^{(k)}\middle|x_2^{(k)}\right) p\left(x_2^{(k)}\middle|x_2^{(k-1)}, x_3^{(k)}, \kappa, \omega\right) \\ &\quad \cdot p\left(x_3^{(k)}\middle|x_3^{(k-1)}, \vartheta\right) p\left(x_2^{(k-1)}, x_3^{(k-1)}\right) p(\kappa, \omega, \vartheta) \end{aligned} \tag{14}$$

Given priors on the initial state $p\left(x_2^{(0)}, x_3^{(0)}\right)$, the generative model is defined for all times k by recursion to $k = 1$. To perform the recursion, the density $p\left(x_2^{(k-1)}, x_3^{(k-1)}\right)$ has to be computed by marginalizing over the joint distribution for the previous time point:

$$
\begin{aligned}
& p\left(x_2^{(k-1)}, x_3^{(k-1)}\right) \\
& = \int p\left(u^{(k-1)}, x_1^{(k-1)}, x_2^{(k-1)}, x_3^{(k-1)}, x_2^{(k-2)}, x_3^{(k-2)}, \kappa, \omega, \vartheta\right) \\
& \quad \cdot \mathrm{d}x_2^{(k-2)} \mathrm{d}x_3^{(k-2)} \mathrm{d}\kappa \mathrm{d}\omega \mathrm{d}\vartheta
\end{aligned} \tag{15}
$$

Note that there is no marginalization over u and x_1 since these quantities are known because they are directly observed.

Inverting this model corresponds to optimizing the posterior densities over the unknown (hidden) states $x \stackrel{\text{def}}{=} \{x_1, x_2, x_3\}$ and parameters $\chi \stackrel{\text{def}}{=} \{\kappa, \omega, \vartheta\}$. This corresponds to perceptual inference and learning, respectively. In the next section, I consider the nature of this inversion or optimization.

Turning from this simple example to the general model defined in Eqs (3), (4), and (8), we have $\chi \stackrel{\text{def}}{=} \{\kappa_1, \omega_1, \ldots, \kappa_n, \omega_n, \vartheta\}$ and $x^{(k)} \stackrel{\text{def}}{=} \left\{x_1^{(k)}, \ldots, x_n^{(k)}\right\}$. Combining all levels and introducing priors on the states and parameters yields

$$
\begin{aligned}
& p\left(x^{(k)}, x^{(k-1)}, \chi\right) \\
& = p(\chi) p\left(x^{(k-1)}\right) p\left(x_n^{(k)} \middle| x_n^{(k-1)}, \vartheta\right) \prod_{i=1}^{n-1} p\left(x_i^{(k)} \middle| x_i^{(k-1)}, x_{i+1}^{(k)}, \kappa_i, \omega_i\right) \\
& = p(\chi) p\left(x^{(k-1)}\right) \mathcal{N}\left(x_n^{(k)}; x_n^{(k-1)}, \vartheta\right) \prod_{i=1}^{n-1} \mathcal{N}\left(x_i^{(k)}; x_i^{(k-1)}, \exp\left(\kappa_i x_{i+1}^{(k)} + \omega_i\right)\right).
\end{aligned} \tag{16}
$$

Here again, this is fully defined by recursion to $k = 1$ given a state prior $p\left(x_1^{(0)}, \ldots, x_n^{(0)}\right)$ and marginalization at each time step:

$$p\left(x^{(k-1)}\right) = \int p\left(x^{(k-1)}, x^{(k-2)}, \chi\right) \mathrm{d}x^{(k-2)} \mathrm{d}\chi. \tag{17}$$

3 INVERSION

3.1 Exact inversion

Returning to the simple example of the previous chapter, it is instructive to consider the factorization of the generative density

$$\begin{aligned} & p\left(u^{(k)}, x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, x_2^{(k-1)}, x_3^{(k-1)}, \chi\right) \\ & = p\left(u^{(k)}, x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, \chi \middle| x_2^{(k-1)}, x_3^{(k-1)}\right) p\left(x_2^{(k-1)}, x_3^{(k-1)}\right) \end{aligned} \tag{18}$$

In this form, the Markovian structure of the model becomes apparent: the joint probability of the input and the states at time k depends only on the states at the immediately preceding time $k-1$. It is the probability distribution of these states that contains the information conveyed by previous inputs $u^{(1\ldots k-1)} = \left(u^{(1)}, \ldots, u^{(k-1)}\right)$; i.e.:

$$p\left(x_2^{(k-1)}, x_3^{(k-1)}\right) = p\left(x_2^{(k-1)}, x_3^{(k-1)} \middle| u^{(1\ldots k-1)}\right) \tag{19}$$

By integrating out $x_2^{(k-1)}$ and $x_3^{(k-1)}$ we obtain the following compact form of the generative model at time k:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p\left(u^{(k)}, x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, \chi \middle| x_2^{(k-1)}, x_3^{(k-1)}\right) \cdot p\left(x_2^{(k-1)}, x_3^{(k-1)} \middle| u^{(1\ldots k-1)}\right) \mathrm{d}x_2^{(k-1)} \mathrm{d}x_3^{(k-1)} = p\left(u^{(k)}, x^{(k)}, \chi \middle| u^{(1\ldots k-1)}\right) \tag{20}$$

Once $u^{(k)}$ is observed, we can plug it into this expression and obtain

$$p\left(x^{(k)}, \chi \middle| u^{(1\ldots k)}\right) \tag{21}$$

This is the quantity of interest to us, because it describes the posterior probability of the time-dependent states, $x^{(k)}$ (and time-independent parameters, χ) in the agent's environment. This is what the agent infers (and learns), given the history of previous inputs. Computing this probability is called model inversion: unlike the likelihood $p\left(u^{(k)} \middle| x^{(k)}, \chi, u^{(1\ldots k-1)}\right)$ the posterior does not predict data ($u^{(k)}$) from hidden states and parameters but predicts states and parameters from data.

In the framework I introduce here, I model the individual variability between agents by putting delta-function priors on the parameters:

$$\begin{aligned} p\left(x^{(k)}, \chi \middle| u^{(1\ldots k)}\right) &= p\left(x^{(k)} \middle| \chi, u^{(1\ldots k)}\right) p\left(\chi \middle| u^{(1\ldots k)}\right) \\ p\left(\chi \middle| u^{(1\ldots k)}\right) &= \delta(\chi = \chi_a) \end{aligned} \tag{22}$$

Where χ_a are the fixed parameter values that characterize a particular agent at a particular time (e.g., during the experimental session). This corresponds to the common distinction between states (as variables that

change quickly) and parameters (as variables that change slowly or not at all). In other words, I assume that the timescale at which parameter estimates change is much larger than the one on which state estimates change, and also larger than the periods during which we observe agents in a single experiment. This is not a strong assumption, given that the model has multiple hierarchical levels of states that give the agent the required flexibility to adapt its beliefs to a changing environment. In effect, this approach gives us a family of Bayesian learners whose (slowly changing) individuality is captured by their priors on the parameters. For other examples where subjects' beliefs about the nature of their environment were modeled as priors on parameters, see Daunizeau et al. (2010b) and Steyvers et al. (2009).

In principle, model inversion can proceed in an online fashion: By (numerical) marginalization, we can obtain the (marginal) posteriors $p\left(x_2^{(k)}\middle|u^{(1\ldots k)}\right)$ and $p\left(x_3^{(k)}\middle|u^{(1\ldots k)}\right)$; this is the approach adopted by (Behrens et al., 2007), allowing one to compute $p\left(u^{(k+1)}, x^{(k+1)}, \chi\middle|u^{(1\ldots k)}\right)$ according to Eqs (14) and (15), and subsequently $p\left(x^{(k+1)}, \chi\middle|u^{(1\ldots k+1)}\right)$ once $u^{(k+1)}$ becomes known, and so on. Unfortunately, this (exact) inversion involves integrals that cannot be solved analytically for every new input, rendering exact Bayesian inversion unsuitable for real-time learning in a biological setting, and inefficient in all settings. If the brain uses a Bayesian scheme, it is likely that it relies on some sufficiently accurate, but fast, approximation to Eq. (20). As described in the next section, a generic and efficient approach is to employ a mean field approximation within a variational scheme. This furnishes an efficient solution with biological plausibility and interpretability.

3.2 Variational inversion

Variational Bayesian (VB) inversion determines the posterior distributions $p\left(x^{(k)}, \chi\middle|u^{(1\ldots k)}\right)$ by maximizing the log-model evidence. The log-evidence corresponds to the negative surprise about the data, given a model, and is approximated by a lower bound, the negative free energy. Detailed treatments of the general principles of the VB procedure can be found in numerous papers (e.g., Beal, 2003; Friston and Stephan, 2007); they are

summarized in Appendix A. The approximations inherent in VB enable a computationally efficient inversion scheme with closed-form single-step probability updates from trial to trial. In particular, VB can incorporate the so-called *mean field approximation* which turns the joint posterior distribution into the product of approximate marginal posterior distributions:

$$\begin{aligned} p\left(x^{(k)}, \chi \middle| u^{(1\ldots k)}\right) &= p\left(x^{(k)} \middle| \chi, u^{(1\ldots k)}\right) p\left(\chi \middle| u^{(1\ldots k)}\right) \\ p\left(x^{(k)} \middle| \chi, u^{(1\ldots k)}\right) &\approx \prod_{i=1}^{n} \hat{q}\left(x_i^{(k)}\right) \\ q\left(x_i^{(k)}\right) &\approx \hat{q}\left(x_i^{(k)}\right) \end{aligned} \tag{23}$$

Based on this assumption, the variational maximization of the negative free energy is implemented in a series of variational updates for each level i of the model separately. The second line in Eq. (23) represents the mean field assumption (factorization of the posterior), while the third line reflects the fact that we assume a fixed form $q(\cdot)$ for the approximate marginal $\hat{q}(\cdot)$. We make minimal assumptions about the form of the approximate posteriors by following the maximum entropy principle: given knowledge of, or assumptions about, constraints on a distribution, the least arbitrary choice of distribution is the one that maximizes entropy (Jaynes, 1957). To keep the description of the posteriors simple and biologically plausible, we take them to be characterized only by their first two moments; i.e., by their mean and variance.

All of the above applies to our simple example as well as to the full model. Before inverting the model variationally in the most general terms, I again first demonstrate the procedure for our simple example. There, at the first level, we have a binary state x_1 with a mean $\mu_1 = p(x_1 = 1)$. Under this constraint, the maximum entropy distribution is the Bernoulli distribution with parameter μ_1 (where the variance $\mu_1(1-\mu_1)$ is a function of the mean):

$$p\left(x_1^{(k)}\middle|u^{(1\ldots k)}\right) \approx q\left(x_1^{(k)}\right) = \text{Bernoulli}\left(x_1^{(k)};\mu_1^{(k)}\right) = \left(\mu_1^{(k)}\right)^{x_1^{(k)}}\left(1-\mu_1^{(k)}\right)^{1-x_1^{(k)}} \tag{24}$$

At the second and third level, the maximum entropy distribution of the unbounded real variables x_2 and x_3, given their means and variances, is Gaussian. Note that the choice of a Gaussian distribution for the approximate posteriors is not due simply to computational expediency (or the law of large numbers) but follows from the fact that, given the assumption that the posterior is encoded by its first two moments, the maximum entropy principle prescribes a Gaussian distribution. Labeling the means μ_2, μ_3 and the variances σ_2, σ_3, we obtain

$$\begin{aligned} p\left(x_2^{(k)}\middle|u^{(1\ldots k)}\right) &\approx q\left(x_2^{(k)}\right) = \mathcal{N}\left(x_2^{(k)};\mu_2^{(k)},\sigma_2^{(k)}\right) \\ p\left(x_3^{(k)}\middle|u^{(1\ldots k)}\right) &\approx q\left(x_3^{(k)}\right) = \mathcal{N}\left(x_3^{(k)};\mu_3^{(k)},\sigma_3^{(k)}\right) \end{aligned} \tag{25}$$

Now that we have approximate posteriors q that we treat as known for all but the ith level, the next step is to determine the variational posterior $\hat{q}(x_i)$ for this level i. Variational calculus shows that given $q(x_j)$ ($j \in \{1,2,3\}$ and $j \neq i$), the approximate posterior $\hat{q}\left(x_i^{(k)}\right) \approx p\left(x_i^{(k)}\middle|u^{(1\ldots k)},\chi\right)$ under the mean field approximation is proportional to the exponential of the *variational energy* $I(x_i)$:

$$\hat{q}\left(x_i^{(k)}\right) = \frac{1}{\mathcal{Z}_i}\exp\left(I\left(x_i^{(k)}\right)\right) \tag{26}$$

$\mathcal{Z}_i$ is a normalization constant that ensures that the integral (or sum, in the discrete case) of $\hat{q}$ over x equals unity. This fundamental relation is

derived in detail in Appendix B (Beal, 2003). Under our generative model, the variational energy is

$$I\left(x_i^{(k)}\right) = \int_{X_{\setminus i}} q\left(x_{\setminus i}^{(k)}\right) \ln p\left(x_i^{(k)}, x_{\setminus i}^{(k)}, \chi \middle| u^{(1\ldots k)}\right) \mathrm{d}x_{\setminus i}^{(k)}$$
$$q(x_{\setminus i}) = \prod_{j \neq i} q(x_j) \tag{27}$$

where $x_{\setminus i}$ denotes all x_j with $j \neq i$ and $\mathrm{q}(\mathrm{x_i}) = \prod_{\mathrm{j}\neq\mathrm{i}} \mathrm{q}(\mathrm{x_j}) X_{\setminus i}$ is the direct product of the ranges (or values in the discrete case) of the x_j contained in $x_{\setminus i}$. The integral over discrete values is again a sum. In what follows, I take this general theory and unpack it using the generative model for sequential learning above. Our special focus here will be on the form of the variational updates that underlie inference and learning.

3.3 The variational energies

To compute $\hat{q}\left(x_i^{(k)}\right)$, we need $q\left(x_{\setminus i}^{(k)}\right)$ and therefore the sufficient statistics $\lambda_{\setminus i} = \{\mu_{\setminus i}, \sigma_{\setminus i}\}$ for the posteriors at all but the i-th level. One could try to extract them from $\hat{q}\left(x_{\setminus i}^{(k)}\right)$ but that would constitute a circular problem. We avoid this by exploiting the hierarchical form of the model: for the first level, we use the sufficient statistics of the higher levels from the previous time point $k-1$, since information about input $u^{(k)}$ cannot yet have reached those levels. From there we proceed upward through the hierarchy of levels, always using the updated parameters $\lambda_{\setminus i}^{(k)}$ for levels lower than the current level and pre-update values $\lambda_{\setminus i}^{(k-1)}$ for higher levels. Extending the approach suggested by Daunizeau et al. (2010b), and using power series approximations where necessary, the variational energy integrals can then be calculated for all x_i (see Appendix C for details), giving, in our example

$$I\left(x_1^{(k)}\right) = x_1^{(k)} \ln s\left(\mu_2^{(k-1)}\right) + \left(1 - x_1^{(k)}\right) \ln\left(1 - s\left(\mu_2^{(k-1)}\right)\right) \tag{28}$$

$$I\left(x_2^{(k)}\right) = \ln s\left(x_2^{(k)}\right) + x_2^{(k)}\left(\mu_1^{(k)} - 1\right) - \frac{1}{2\left(\sigma_2^{(k-1)} + \mathrm{e}^{\kappa\mu_3^{(k-1)}+\omega}\right)}\left(x_2^{(k)} - \mu_2^{(k-1)}\right)^2 \tag{29}$$

$$I\left(x_3^{(k)}\right) = -\frac{1}{2}\ln\left(\sigma_2^{(k-1)} + e^{\kappa x_3^{(k)}+\omega}\right) - \frac{1}{2}\frac{\sigma_2^{(k)} + \left(\mu_2^{(k)} - \mu_2^{(k-1)}\right)^2}{\sigma_2^{(k-1)} + \mathrm{e}^{\kappa x_3^{(k)}+\omega}} - \frac{1}{2\left(\sigma_3^{(k-1)} + \vartheta\right)}\left(x_3^{(k)} - \mu_3^{(k-1)}\right)^2. \tag{30}$$

In the general case, the variational energy is

$$I\left(x_i^{(k)}\right) = -\frac{1}{2}\ln\left(\sigma_{i-1}^{(k-1)} + \exp\left(\kappa_{i-1}x_i^{(k)} + \omega_{i-1}\right)\right) - \frac{1}{2}\frac{\sigma_i^{(k)} + \left(\mu_i^{(k)} - \mu_i^{(k-1)}\right)^2}{\sigma_{i-1}^{(k-1)} + \exp\left(\kappa_{i-1}x_i^{(k)} + \omega_{i-1}\right)} - \frac{1}{2}\frac{1}{\sigma_i^{(k-1)} + \exp\left(\kappa_i\mu_{i+1}^{(k-1)} + \omega_i\right)}\left(x_i^{(k)} - \mu_i^{(k-1)}\right)^2, \tag{31}$$

$$\text{with} \qquad \exp\left(\kappa_n x_{n+1}^{(k)} + \omega_n\right) \equiv \vartheta.$$

Substituting these variational energies into Eq. (26, we obtain the posterior distribution of x under the mean field approximation.

According to Eqs (26) and (28 to (30, $\hat{q}\left(x_i^{(k)}\right)$ depends on $u^{(k)}$, $\lambda_i^{(k-1)}$, $\lambda_{\backslash i}^{(k-1)}$, $\lambda_{\backslash i}^{(k)}$, and χ. In the next section, I show that it is possible to derive simple closed-form Markovian update equations of the form

$$\lambda_i^{(k)} = f_i\left(u^{(k)}, \lambda_i^{(k-1)}, \lambda_{\backslash i}^{(k-1)}, \lambda_{\backslash i}^{(k)}, \chi\right) \tag{32}$$

Update equations of this form allow the agent to update its approximate posteriors over $x_i^{(k)}$ very efficiently and thus optimize its beliefs about the environment in real time. We now consider the detailed form of these equations and how they relate to classical heuristics from reinforcement learning.

4 THE UPDATE EQUATIONS

4.1 Derivation of the update equations

At the first level of the model, it is simple to determine $q(x_1)$ since $\hat{q}(x_1) = 1/Z_1 \cdot \exp(I(x_1))$ is a Bernoulli distribution with parameter

$$\mu_1^{(k)} = \hat{q}\left(x_1^{(k)} = 1\right) = u^{(k)} \tag{33}$$

and therefore already has the form required of $q(x_1)$ by Eq. (23. We can thus take $q(x_1) = \hat{q}(x_1)$ and have in Eq. (33) an update rule of the form of Eq. (32).

At the second level, $\hat{q}(x_2)$ does not have the form required of $q(x_2)$ by Eq. (24) since it is only approximately Gaussian. It is proportional to the exponential of $I(x_2)$ and would only be Gaussian if $I(x_2)$ were quadratic. The problem of finding a Gaussian approximation $q(x_2) \approx \hat{q}(x_2)$ can therefore be reformulated as finding a quadratic approximation to $I(x_2)I$. The obvious way to achieve this is to expand $I(x_2)$ in powers of x_2 up to second order. The choice of expansion point, however, is not trivial. One possible choice is the mode or maximum of $I(x_2)$, resulting in the frequently used *Laplace approximation* (Friston et al., 2007). This has the disadvantage that the maximum of $I(x_2)$ is unknown and has to be found by numerical optimization methods, precluding a single-step analytical update rule of the form of Eq. (32). This is no problem in a continuous time setting, where the mode of the variational energy can be updated continuously (cf. Friston, 2008). However, for the discrete updates we seek, one can use the expectation $\mu_2^{(k-1)}$ as the expansion point for time k, when the agent receives input $u^{(k)}$ and the expectation of $x_2^{(k)}$ is $\mu_2^{(k-1)}$. In terms of computational and mnemonic costs, this is the most economical choice since this value is known. Moreover, it yields analytical update equations which (i) bear structural resemblance to those used by RL

models (see next section) and (ii) can be computed very efficiently in a single step:

$$\sigma_2^{(k)} = \frac{1}{1/\hat{\sigma}_2^{(k)} + \hat{\sigma}_1^{(k)}} \tag{34}$$

$$\mu_2^{(k)} = \mu_2^{(k-1)} + \sigma_2^{(k)} \delta_1^{(k)} \tag{35}$$

where, for clarity, I have used the definitions

$$\hat{\mu}_1^{(k)} \stackrel{\text{def}}{\equiv} s\left(\sigma_2^{(k-1)}\right) \tag{36}$$

$$\delta_1^{(k)} \stackrel{\text{def}}{\equiv} \mu_1^{(k)} - \hat{\mu}_1^{(k)} \tag{37}$$

$$\hat{\sigma}_1^{(k)} \stackrel{\text{def}}{\equiv} \hat{\mu}_1^{(k)}\left(1 - \hat{\mu}_1^{(k)}\right) \tag{38}$$

$$\hat{\sigma}_2^{(k)} \stackrel{\text{def}}{\equiv} \sigma_2^{(k-1)} + \mathrm{e}^{\kappa\mu_3^{(k-1)}+\omega} \tag{39}$$

Formulated in terms of precisions (inverse variances) $\pi_2^{(k)} \stackrel{\text{def}}{\equiv} 1/\sigma_2^{(k)}$, $\hat{\pi}_1^{(k)} \stackrel{\text{def}}{\equiv} 1/\hat{\sigma}_1^{(k)}$, $\hat{\pi}_2^{(k)} \stackrel{\text{def}}{\equiv} 1/\hat{\sigma}_2^{(k)}$, the variance update (Eq. (34)) takes the simple form

$$\pi_2^{(k)} = \hat{\pi}_2^{(k)} + \frac{1}{\hat{\pi}_1^{(k)}}. \tag{40}$$

In the context of the update equations, I use the hat notation to indicate "referring to prediction". While the $\hat{\mu}_i^{(k)}$ are the predictions before seeing the input $u^{(k)}$, the $\hat{\sigma}_i^{(k)}$ and $\hat{\pi}_i^{(k)}$ are the variances (i.e., uncertainties) and precisions of these predictions (see Structural interpretation of the Update Equations).

The approach that produces these update equations is conceptually similar to a Gauss-Newton ascent on the variational energy that would, by iteration, produce the Laplace approximation (cf. Figure 3). The same approach can be taken at the third level, where we also have a Gaussian approximate posterior:

$$\pi_3^{(k)} = \hat{\pi}_3^{(k)} + \frac{\kappa^2}{2} w_2^{(k)} \left(w_2^{(k)} + r_2^{(k)} \delta_2^{(k)} \right) \tag{41}$$

$$\mu_3^{(k)} = \mu_3^{(k)} + \sigma_3^{(k)} \frac{\kappa}{2} w_2^{(k)} \delta_2^{(k)} \tag{42}$$

with $\pi_2^{(k)} \stackrel{\text{def}}{=} 1/\sigma_2^{(k)}$ and

$$\hat{\pi}_3^{(k)} \stackrel{\text{def}}{=} \frac{1}{\sigma_3^{(k-1)} + \vartheta} \tag{43}$$

$$w_2^{(k)} \stackrel{\text{def}}{=} \frac{\mathrm{e}^{\kappa\mu_3^{(k-1)}+\omega}}{\sigma_2^{(k-1)} + \mathrm{e}^{\kappa\mu_3^{(k-1)}+\omega}} \tag{44}$$

$$r_2^{(k)} \stackrel{\text{def}}{=} \frac{\mathrm{e}^{\kappa\mu_3^{(k-1)}+\omega} - \sigma_2^{(k-1)}}{\sigma_2^{(k-1)} + \mathrm{e}^{\kappa\mu_3^{(k-1)}+\omega}} \tag{45}$$

$$\delta_2^{(k)} \stackrel{\text{def}}{=} \frac{\sigma_2^{(k)} + \left(\mu_2^{(k)} - \mu_2^{(k-1)}\right)^2}{\sigma_2^{(k-1)} + \mathrm{e}^{\kappa\mu_3^{(k-1)}+\omega}} \tag{46}$$

Turning to the general case, I choose a slightly different notation for the updates that will later (cf. Section 4.1) provide a complementary interpretation in terms of precision-weighted prediction errors. We here obtain

$$\pi_i^{(k)} = \hat{\pi}_i^{(k)} + \frac{1}{2}\left(\kappa_{i-1}\, v_{i-1}^{(k)} \hat{\pi}_{i-1}^{(k)}\right)^2 \left(1 + \left(1 - \frac{1}{v_{i-1}^{(k)} \pi_{i-1}^{(k-1)}}\right) \delta_{i-1}^{(k)}\right) \tag{47}$$

$$\mu_i^{(k)} = \mu_i^{(k-1)} + \frac{1}{2}\kappa_{i-1}\, v_{i-1}^{(k)} \frac{\hat{\pi}_{i-1}^{(k)}}{\pi_i^{(k)}} \delta_{i-1}^{(k)} \tag{48}$$

with

$$v_i^{(k)} \stackrel{\text{def}}{=} \exp\left(\kappa_i \mu_{i+1}^{(k-1)} + \omega_i\right) \tag{49}$$

$$\hat{\pi}_i^{(k)} \stackrel{\text{def}}{=} \frac{1}{\sigma_i^{(k-1)} + \exp\left(\kappa_i \mu_{i+1}^{(k-1)} + \omega_i\right)} \tag{50}$$

$$\delta_i^{(k)} \stackrel{\text{def}}{=} \frac{\sigma_i^{(k)} + \left(\mu_i^{(k)} - \mu_i^{(k-1)}\right)^2}{\sigma_i^{(k-1)} + \exp\left(\kappa_i \mu_{i+1}^{(k-1)} + \omega_i\right)} - 1 \tag{51}$$

One may easily verify that these updates correspond exactly to those given for the third level in our example model. From that level up to the n-th, the updates retain this same compact and interpretable form, never becoming more complicated.

Applying Eq. (32) recursively and using the updates of Eqs (47) and (48) for its right-hand side, we can now write the inferences implied by the HGF as the following computationally efficient function

$$\lambda^{(k)} = \lambda^{(k)}\left(\chi, \lambda^{(0)}, u\right), \tag{52}$$

where $\lambda^{(k)} \stackrel{\text{def}}{=} \left\{\lambda_1^{(k)}, \ldots, \lambda_n^{(k)}\right\}$ and $u \stackrel{\text{def}}{=} \left\{u^{(1)}, \ldots, u^{(K)}\right\}$.

The derivation of the update equations, based on the quadratic approximation to $I(x_2)$ and $I(x_3)$ briefly introduced above, is described in detail in the next section, and Figure 3 provides a graphical illustration of the ensuing procedure. The meaning of the terms that appear in the update

equations and the overall structure of the updates will be discussed in detail in the Section 4.1.

4.2 Quadratic approximation to the variational energies

While knowing $I(x_i)$ (with $i = 1,2,3$) gives us the unconstrained posterior $\hat{q}(x_i)$ given $q(x_{\setminus i})$, we still need to determine the constrained posterior $q(x_i)$ for all but the first levels, where $q(x_1) = \hat{q}(x_1)$. Schematically, the approximation procedure can be pictured in the following way:

$$p\left(x_i^{(k)}\middle|u^{(1\ldots k)},\chi\right) \xrightarrow{\text{mean field}} \hat{q}\left(x_i^{(k)}\right) \propto \exp I\left(x_i^{(k)}\right) \xrightarrow{\text{Gaussian}} q\left(x_i^{(k)}\right) \propto \exp \tilde{I}\left(x_i^{(k)}\right) \tag{53}$$

I denote by $\tilde{I}$ the quadratic function obtained by expansion of I around $x_i^{(k)} = \mu_i^{(k-1)}$ (see Figure 3 for a graphical summary). Its exponential has the Gaussian form required by Eqs (24) and (25) (where $\tilde{Z}_{\mathrm{i}}$ is a normalization constant):

$$q\left(x_i^{(k)}\right) = \frac{1}{\sqrt{2\pi\sigma_i^{(k)}}} \exp\left(-\frac{\left(x_i^{(k)} - \mu_i^{(k)}\right)^2}{2\sigma_i^{(k)}}\right) = \frac{1}{\tilde{Z}_{\mathrm{i}}} \exp\left(\tilde{I}\left(x_i^{(k)}\right)\right) \tag{54}$$

This equation lets us find $\mu_i^{(k)}$ and $\sigma_i^{(k)}$. Taking the logarithm on both sides and then differentiating twice with respect to $x_i^{(k)}$ gives

$$\sigma_i^{(k)} = -\frac{1}{\partial^2 \tilde{I}\left(x_i^{(k)}\right)} = \text{const.} \tag{55}$$

where ∂^2 denotes the second derivative, which is constant for a quadratic function. Because $\partial^2 I$ and $\partial^2 \tilde{I}$ agree at the expansion point $x_i^{(k)} = \mu_i^{(k-1)}$, we may write

$$\sigma_i^{(k)} = -\frac{1}{\partial^2 I\left(\mu_i^{(k-1)}\right)} \tag{56}$$

A somewhat different line of reasoning leads to $\mu_i^{(k)}$. Since $\mu_i^{(k)}$ is the argument of the maximum (argmax) of q (and exponentiation preserves the argmax), $\mu_i^{(k)}$ has to be the argmax of $\tilde{I}$. Starting at any point $x_i^{(k)}$, the exact argmax of a quadratic function can be found in one step by Newton's method:

$$\mu_i^{(k)} = \text{argmax}\, \tilde{I}\left(x_i^{(k)}\right) = x_i^{(k)} - \frac{\partial \tilde{I}\left(x_i^{(k)}\right)}{\partial^2 \tilde{I}\left(x_i^{(k)}\right)} \tag{57}$$

If we choose $x_i^{(k)}$ to be the expansion point $\mu_i^{(k-1)}$, we have agreement of I with $\tilde{I}$ up to the second derivative at this point and may therefore write

$$\mu_i^{(k)} = \mu_i^{(k-1)} - \frac{\partial I\left(\mu_i^{(k-1)}\right)}{\partial^2 I\left(\mu_i^{(k-1)}\right)} = \mu_i^{(k-1)} + \sigma_i^{(k)} \partial I\left(\mu_i^{(k-1)}\right) \tag{58}$$

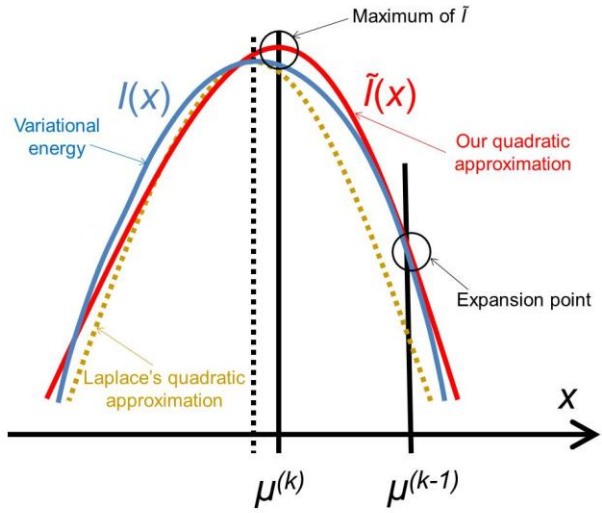

Figure 3. ***Quadratic approximation to the variational energy.*** *Approximating the variational energy $I(x)$ (blue) by a quadratic function leads (by exponentiation) to a Gaussian posterior. To find our approximation $\tilde{I}(x)$ (red), we expand $I(x)$ to second order at the preceding posterior expectation $\mu^{(k-1)}$. The argmax of $\tilde{I}(x)$ is then the new posterior expectation $\mu^{(k)}$. This generally leads to a different result from the Laplace approximation (dashed), but there is a priori no reason to regard either approximation as more exact than the other.*

Plugging $I(x_2)$ from Eq. (29) into Eqs (56) and (58) now yields the parameter update equations (Eqs (34) and (35)) of the form of Eq. (32) for the second level. For the third level, the same procedure gives Eqs (41) and (42), and in the general case of n levels, we obtain Eqs (47) and (48). Note that this method of obtaining closed-form Markovian parameter update equations can readily be applied to multidimensional x_i's with approximately multivariate Gaussian posteriors by reinterpreting ∂I as a gradient and $1/\partial^2 I$ as the inverse of a Hessian in Eqs (56) and (58).

4.3 Structural interpretation of the update equations

As we have seen, variational inversion of our model, using a new quadratic approximation to the variational energies, provides a set of simple trial-by-trial update rules for the sufficient statistics $\lambda_i = \{\mu_i, \sigma_i\}$ of the posterior distributions we seek. These update equations do not increase in complexity with trials, in contrast to exact Bayesian inversion, which requires analytically intractable integrations (cf. Eq. (20)). In our approach, almost all the work is in deriving the update rules, not in doing the actual updates.

Crucially, the update equations for μ_2 and μ_3 have a form that is familiar from RL models such as Rescorla-Wagner learning (cf. also Figure 4). The general structure of RL models can be summarized as:

$$\text{prediction}^{(k)} = \text{prediction}^{(k-1)} + \text{learning rate} \times \text{prediction error}$$

As I explain in detail below, this same structure appears in Eqs (35) and (42) – displayed here in their full forms:

$$\underbrace{\mu_2^{(k)}}_{\text{posterior}^{(k)}} = \underbrace{\mu_2^{(k-1)}}_{\text{posterior}^{(k-1)}} + \underbrace{\sigma_2^{(k)}}_{\text{learning rate}} \underbrace{\left(\mu_1^{(k)} - s\left(\mu_2^{(k-1)}\right)\right)}_{\text{prediction error}}$$

$$\underbrace{\mu_3^{(k)}}_{\text{posterior}^{(k)}} = \underbrace{\mu_3^{(k-1)}}_{\text{posterior}^{(k-1)}} + \underbrace{\sigma_3^{(k)} \frac{\kappa}{2} \frac{e^{\kappa\mu_3^{(k-1)}+\omega}}{\sigma_2^{(k-1)} + e^{\kappa\mu_3^{(k-1)}+\omega}}}_{\text{learning rate}} \underbrace{\left(\frac{\sigma_2^{(k)} + \left(\mu_2^{(k)} - \mu_2^{(k-1)}\right)^2}{\sigma_2^{(k-1)} + e^{\kappa\mu_3^{(k-1)}+\omega}} - 1\right)}_{\text{prediction error}}$$

The term $\delta_1^{(k)} \stackrel{\text{def}}{=} \mu_1^{(k)} - \hat{\mu}_1^{(k)}$ in Eq. (35) corresponds to the prediction error at the first level. This prediction error is the difference between the expectation $\mu_1^{(k)}$ of x_1 having observed input $u^{(k)}$ and the prediction $\hat{\mu}_1^{(k)} = s\left(\mu_2^{(k-1)}\right)$ before receiving $u^{(k)}$; i.e. the softmax transformation of the expectation of x_2 before seeing $u^{(k)}$. Furthermore, σ_2 in Eq. (35) can be interpreted as the equivalent of a (time-varying) learning rate in RL

models (cf. Preuschoff and Bossaerts, 2007). Since σ_2 represents the width of x_2's posterior and thus the degree of our uncertainty about x_2, it makes sense that updates in μ_2 are proportional to this estimate of posterior uncertainty: the less confident the agent is about what it knows, the greater the influence of new information should be.

According to its update equation (Eq. (34)), σ_2 always remains positive since it contains only positive terms. Crucially, σ_2, through $\hat{\sigma}_2$, depends on the log-volatility estimate μ_3 from the third level of the model. For vanishing volatility, i.e. $e^{\kappa\mu_3+\omega} \approx 0$, σ_2 can only decrease from trial to trial. This corresponds to the case in which the agent believes that x_2 is fixed; the information from every trial then has the same weight and new information can only shrink σ_2. On the other hand, even with $e^{\kappa\mu_3+\omega} \approx 0$, σ_2 has a lower bound: when σ_2 approaches zero, the denominator of Eq. (34) approaches $1/\sigma_2$ from above, leading to ever smaller decreases in σ_2. This means that after a long train of inputs, the agent still learns from new input, even when it infers that the environment is stable.

The precision formulation (cf. Eq. (40))

$$\pi_2^{(k)} = \hat{\pi}_2^{(k)} + \frac{1}{\hat{\pi}_1^{(k)}} = \frac{1}{\sigma_2^{(k-1)} + e^{\kappa\mu_3^{(k-1)}+\omega}} + \hat{\sigma}_1^{(k)} \tag{59}$$

illustrates that three forms of uncertainty influence the posterior variance$\sigma_2^{(k)}$: the informational ($\sigma_2^{(k-1)}$) and the environmental ($e^{\kappa\mu_3^{(k-1)}+\omega}$) uncertainty at the second level (see discussion in the next paragraph), and the uncertainty $\hat{\sigma}_1^{(k)}$ of the prediction at the first level. While environmental uncertainty at the second level thus decreases precision $\pi_2^{(k)}$ relative to its previous value $\pi_2^{(k-1)} = 1/\sigma_2^{(k-1)}$, predictive uncertainty at the first level counteracts that decrease, i.e., it keeps the learning rate $\sigma_2^{(k)}$ smaller than it would otherwise be. This makes sense because prediction error should mean less when predictions are more uncertain.

The update rule (Eq. (42)) for μ_3 has a similar structure to that of μ_2 (Eq. (35)) and can also be interpreted in terms of RL. Although perhaps not obvious at first glance, $\delta_2^{(k)}$ (Eq. (46)) represents prediction error. It is positive if the updates at the second level (of μ_2 and σ_2) in response to input $u^{(k)}$indicate that the agent was underestimating x_3. Conversely, it is negative if the agent was overestimating x_3. This can be seen by noting that the uncertainty about x_2 has two sources: *informational*, i.e. the lack of knowledge about x_2 (represented by σ_2), and *environmental*, i.e. the volatility of the environment (represented by $e^{\kappa\mu_3+\omega}$). Before receiving input $u^{(k)}$, the total uncertainty is $\hat{\sigma}_2^{(k)} = \sigma_2^{(k-1)} + e^{\kappa\mu_3^{(k-1)}+\omega}$. After receiving the input, the updated total uncertainty is $\sigma_2^{(k)} + \left(\mu_2^{(k)} - \mu_2^{(k-1)}\right)^2$, where σ_2 has been updated according to Eq. (34) and $e^{\kappa\mu_3^{(k-1)}+\omega}$ has been replaced by the squared update of μ_2. If the total uncertainty is greater after seeing $u^{(k)}$, the fraction in $\delta_2^{(k)}$ is greater than one and μ_3 increases. Conversely, if seeing $u^{(k)}$ reduces total uncertainty, μ_3 decreases. (Since x_3 is on a logarithmic scale with respect to x_2, the ratio and not the difference of quantities referring to x_2 is relevant for the prediction error in x_3.) It is important to note that we did not construct the update equations with any of these properties in mind. It is simply a reflection of Bayes-optimality that emerges on applying our variational update method.

The term corresponding to the learning rate of μ_3 is

$$\sigma_3^{(k)} \cdot \frac{\kappa}{2} \cdot w_2^{(k)} = \sigma_3^{(k)} \cdot \frac{\kappa}{2} \cdot \frac{e^{\kappa\mu_3^{(k-1)}+\omega}}{\sigma_2^{(k-1)} + e^{\kappa\mu_3^{(k-1)}+\omega}} \tag{60}$$

As at the second level, this is proportional to the variance σ_3 of the posterior. But here, the learning rate is also proportional to the parameter κ and a *weighting factor* $w_2^{(k)}$ for prediction error. κ determines the form of the Gaussian random walk at the second level and couples the third level to the second (cf. Eqs (12) (42)); $w_2^{(k)}$ is a measure of the

(environmental) volatility $e^{\kappa\mu_3+\omega}$of x_2 relative to its (informational) conditional uncertainty, σ_2. It is bounded between 0 and 1 and approaches 0 as $e^{\kappa\mu_3+\omega}$ becomes negligibly small relative to σ_2; conversely, it approaches 1 as σ_2 becomes negligibly small relative to $e^{\kappa\mu_3+\omega}$. This means that lack of knowledge about x_2; i.e. conditional uncertaintyσ_2, suppresses updates of μ_3 by reducing the learning rate, reflecting the fact that prediction errors in x_2 are only informative x_3 if the agent is confident about its predictions of x_2. As with the prediction error term, this weighting factor emerged from our variational approximation.

The precision update (Eq. (41)) at third level also has an interpretable form. In addition to $\delta_2^{(k)}$ and $w_2^{(k)}$, we now also have the term $r_2^{(k)}$ (Eq. (45)), which is the *relative difference* of environmental and informational uncertainty (i.e., relative to their sum). Note that it is a simple affine function of the weighting factor $w_2^{(k)}$

$$r_2^{(k)} = 2w_2^{(k)} - 1 \tag{61}$$

As at the second level, the precision update is the sum of the precision $\hat{\pi}_3^{(k)}$ of the prediction, reflecting the informational and environmental uncertainty at the third level, and the term

$$\frac{\kappa^2}{2} w_2^{(k)} \left(w_2^{(k)} + r_2^{(k)} \delta_2^{(k)} \right) \tag{62}$$

Proportionality to κ^2 reflects the fact that stronger coupling between the second and third levels leads to higher posterior precision (i.e., less posterior uncertainty) at the third level, while proportionality to w_2 depresses precision at the third level when informational uncertainty at the second level is high relative to environmental uncertainty; the latter also applies to the first summand in the brackets. The second summand $r_2\delta_2$ means that, when the agent regards environmental uncertainty at the

second level as relatively high ($r_2 > 0$), volatility is held back from rising further if $\delta_2 > 0$ by way of a decrease in the learning rate (which is proportional to the inverse precision), but conversely pushed to fall if $\delta_2 < 0$. If, however, environmental uncertainty is relatively low ($r_2 < 0$), the opposite applies: positive volatility prediction errors increase the learning rate, allowing the environmental uncertainty to rise more easily, while negative prediction errors decrease the learning rate. The term $r_2\delta_2$ thus exerts a stabilizing influence on the estimate of μ_3.

This automatic integration of all the information relevant to a situation is typical of Bayesian methods and brings to mind a remark made by Jaynes (Jaynes, 2003, 517) in a different context: "This is still another example where Bayes' theorem detects a genuinely complicated situation and automatically corrects for it, but in such a slick, efficient way that one is [at first, I would say] unaware of what is happening." In the next section, I use simulations to illustrate the nature of this inference and demonstrate some of its more important properties.

4.4 The updates are precision-weighted prediction errors

A crucial feature of the update equations is emphasized by the notation used in Eq. (48): the updates of the means are precision-weighted prediction errors. For a full understanding of this important property, I will first discuss Bayesian updates in the simplest possible case, where they can be calculated exactly. In this simplest case, there is only one hidden state $x \in \mathbb{R}$ that is the target of our inference, and there is a Gaussian prior on x:

$$p(x) = \mathcal{N}(x; \mu_x, \pi_x), \tag{63}$$

where μ_x is the mean and π_x the precision. The likelihood of x (i.e., the probability of observing the datum $u \in \mathbb{R}$ given x) is also Gaussian, with precision (sc. inverse observation noise) π_u:

$$p(u|x) = \mathcal{N}(u; x, \pi_u). \tag{64}$$

According to Bayes' theorem, the posterior then also turns out to be Gaussian:

$$p(x|u) = \frac{p(u|x)p(x)}{\int p(u|x')p(x')\mathrm{d}x'} = \mathcal{N}\left(x; \mu_{x|u}, \pi_{x|u}\right). \tag{65}$$

The posterior precision $\pi_{x|u}$ and mean $\mu_{x|u}$ can then be written as the following analytical and exact one-step updates:

$$\pi_{x|u} = \pi_x + \pi_u \tag{66}$$

$$\mu_{x|u} = \mu_x + \frac{\pi_u}{\pi_{x|u}}(u - \mu_x) \tag{67}$$

Note that the update in the mean is a precision-weighted prediction error. The prediction error $u - \mu_x$ is weighted *proportionally* to the observation precision π_u, reflecting the fact that the more observation noise there is, the less prediction error can mean since the part of it attributable to noise is growing. On the other hand, prediction error is weighted inverse-proportionally to the posterior precision $\pi_{x|u}$, reflecting the fact that the more precise knowledge of x is, the less impact new information will have.

This same precision-weighting of prediction errors now appears in the update of the means μ_i of the states x_i in the inversion of the general HGF (Eq. (48)):

$$\mu_i^{(k)} - \mu_i^{(k-1)} = \frac{1}{2}\kappa_{i-1}\, v_{i-1}^{(k)} \frac{\hat{\pi}_{i-1}^{(k)}}{\pi_i^{(k)}} \delta_{i-1}^{(k)}, \tag{68}$$

or, in more compact notation,

$$\Delta\mu_i \propto \frac{\hat{\pi}_{i-1}}{\pi_i}\delta_{i-1}. \tag{69}$$

Owing to the hierarchical nature of the HGF, the place of the likelihood precision π_u in Eq. (76) is here taken by the precision of the prediction on the level below, $\hat{\pi}_{i-1}$, while the posterior precision π_i in the HGF corresponds exactly to the posterior precision $\pi_{x|u}$ in Eq. (76).

The prediction error δ_{i-1} is a *volatility prediction error (VOPE)* in the HGF while the prediction errors in the conjugate exponential-family updates above refer to *value prediction errors (VAPEs).* The volatility prediction error is perhaps the most important novelty introduced by the HGF. That the VOPE's role in the updating of inferences on volatilities is the same as that of the VAPE in the updating of inferences on values shows that it is usefully defined and speaks to its value as a concept. These results are summarized in Figure 4.

Update equation for posterior means

$$\Delta\mu_i^{(k)} = \frac{1}{2}\kappa_{i-1}v_{i-1}^{(k)}\frac{\hat{\pi}_{i-1}^{(k)}}{\pi_i^{(k)}}\delta_{i-1}^{(k)}$$

$$\Longrightarrow \boxed{\Delta\mu_i \propto \frac{\hat{\pi}_{i-1}}{\pi_i}\delta_{i-1}}$$

with

$\hat{\pi}_{i-1}$	Precision of the prediction onto the level below
π_i	Posterior precision at the current level
δ_{i-1}	Prediction error for the volatility of the level below

Figure 4. ***Posterior mean update equation.*** *Updates are precision-weighted prediction errors. This general feature of Bayesian updating is concretized by the HGF for volatility predictions in a hierarchical setting.*

In our example model, we have a transition to binary states x_1 from continuous x_2. This leads to a somewhat different update for μ_2 (cf. Eq. (35)):

$$\mu_2^{(k)} = \mu_2^{(k-1)} + \sigma_2^{(k)}\delta_1^{(k)} \tag{70}$$

At first, this simply looks like an uncertainty-weighted update. However, if we unpack σ_2 according to Eq. (34) and do a Taylor expansion in powers of $\hat{\pi}_1$, we see that it is again proportional to the precision of the prediction on the level below:

$$\sigma_2^{(k)} = \frac{\hat{\pi}_1^{(k)}}{\hat{\pi}_2^{(k)}\hat{\pi}_1^{(k)} + 1} = \hat{\pi}_1^{(k)} - \hat{\pi}_2^{(k)}\left(\hat{\pi}_1^{(k)}\right)^2 + \left(\hat{\pi}_2^{(k)}\right)^2\left(\hat{\pi}_1^{(k)}\right)^3 + O(4). \tag{71}$$

4.5 Simulations

In this section, I present several simulations to illustrate the behavior of the HGF update equations under different values of the parameters ϑ, ω, and κ in our simple example model. Figure 5 depicts a "reference" scenario, which will serve as the basis for subsequent variations. Figures 6, 7, and 8 display the effects of selectively changing one of the parameters ϑ, ω, and κ, leading to distinctly different types of inference. The HGF toolbox, introduced in Chapter 8, makes it easy to perform such simulations.

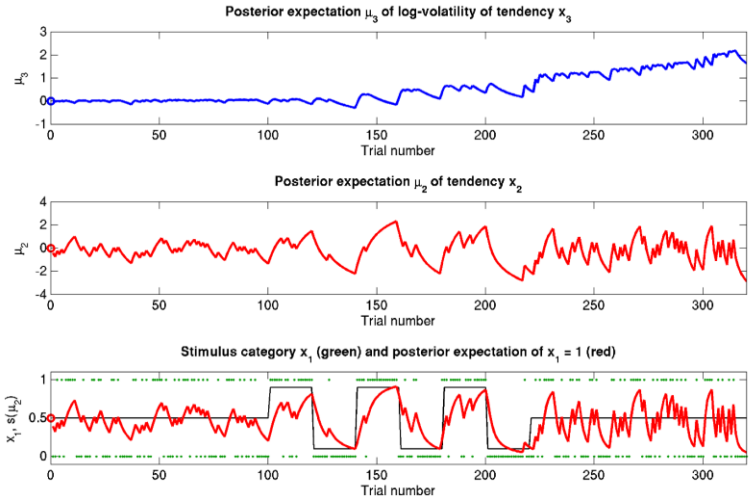

Figure 5. ***Reference scenario:*** $\boldsymbol{\vartheta = 0.5,\ \omega = -2.2,\ \kappa = 1.4.}$ *A simulation of 320 trials. Bottom: the first level. Input u is represented by green dots. In the absence of sensory uncertainty, this corresponds to* x_1*. The fine black line is the true probability (unknown to the agent) that* $x_1 = 1$*. The red line shows* $s(\mu_2)$*; i.e., the agent's posterior expectation that* $x_1 = 1$*. Given the input and update rules, the simulation is uniquely determined by the value of the parameters ϑ, ω, and κ. Middle: the second level with the posterior expectation* μ_2 *of* x_2*. Top: the third level with the posterior expectation* μ_3 *of* x_3*. In all three panels, the initial values of the various μ and σ are indicated by circles at trial* $k = 0$*.*

The reference scenario in Figure 5 (and Figure 9, top row) illustrates some basic features of the model and its update rules. For this reference, I chose the following parameter values: $\vartheta = 0.5$, $\omega = -2.2$, and $\kappa = 1.4$. Overall, the agent is exposed to 320 sensory outcomes (stimuli) that are administered in three stages. In a first stage, it is exposed to 100 trials where the probability that $x_1 = 1$ is 0.5. The posterior expectation of x_1 accordingly fluctuates around 0.5 and that of x_2 around 0; the expected volatility remains relatively stable. There then follows a second period of 120 trials with higher volatility, where the probability that $x_1 = 1$ alternates between 0.9 and 0.1 every 20 trials. After each change, the estimate of x_1 reliably approaches the true value within about 20 trials. In accordance with the changes in probability, the expected outcome tendency x_2 now fluctuates more widely around zero. At the third level, the expected log-volatility x_3 shows a tendency to rise throughout this period, displaying upward jumps whenever the probability of an outcome changes (and thus x_2 experiences sudden updates). As would be anticipated, the expected log-volatility declines during periods of stable outcome probability. In a third and final period, the first 100 trials were repeated in exactly the same order. Note how owing to the higher estimate of volatility (i.e., greater $\exp(\kappa\mu_3 + \omega)$), the learning rate has increased, now causing the same sequence of inputs to have a greater effect on μ_2 than during the first stage of the simulation. As expected, a more volatile environment leads to a higher learning rate.

One may wonder why, in the third stage of the simulation, the expected log-volatility μ_3 continues to rise even after the true x_2 has returned to a stable value of 0 (corresponding to $p(x_1 = 1) = 0.5$; see the fine black line in Figure 5). This is because a series of three $x_1 = 1$ outcomes, followed by three $x_1 = 0$ could just as well reflect a stable $p(x_1 = 1) = 0.5$ or a jump from $p(x_1 = 1) = 1$ to $p(x_1 = 1) = 0$ after the first three trials. Depending on the particular choice of parameters ϑ, ω, and κ, the agent shows qualitatively different updating behavior: Under the parameters in the reference scenario, it has a strong tendency to increase its posterior expectation of volatility in response to unexpected stimuli. For other parameterizations (as in the scenarios described below), this is not the case.

The nature of the simulation in Figure 5 is not only determined by the choice of values for the parameters ϑ, ω, and κ, but also by initial values for μ_2, σ_2, μ_3, and σ_3 (the initial value of μ_1 is $s(\mu_2)$). Any change in the initial value $\mu_3^{(0)}$ of μ_3 can be neutralized by corresponding changes in κ and ω. We may therefore assume $\mu_3^{(0)} = 0$ without loss of generality but remembering that κ and ω are only unique relative to this choice. However, the initial values of μ_2, σ_2, and σ_3 are, in principle, not neutral: They are nevertheless of little consequence in practice since, when chosen reasonably, they make the time series of hidden states $x^{(k)}$ quickly converge to values that do not depend on the choice of initial value in any appreciable extent. In the simulations of Figures 5-8, I used $\mu_2^{(0)} = 0$, $\sigma_2^{(0)} = 1$, and $\sigma_3^{(0)} = 1$.

If we reduce ϑ (the log-volatility of x_3) from 0.5 in the reference scenario to 0.05, we find an agent who is overly confident about its prior estimate of environmental volatility and expects to see little change (Figure 6; Figure 9, second row). This leads to a greatly diminished learning rate for x_3,while learning in x_2 is not directly affected. There is, however, an indirect effect on x_2 in that the learning rate at the second level during the third period is no longer noticeably increased by the preceding period of higher volatility. In other words, this agent shows superficial (low-level) adaptability but has higher-level beliefs that remain impervious to new information.

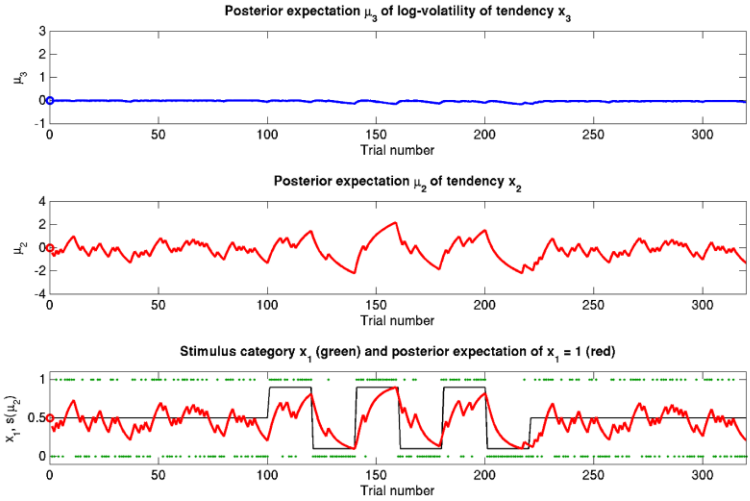

Figure 6. ***Reduced $\vartheta = 0.5$, (unchanged $\omega = -2.2$, $\kappa = 1.4$).*** *Symbols have the same meaning as in Figure 5. Here, the expected x_3 is more stable. The learning rate in x_2 is initially unaffected but owing to more stable x_3 it no longer increases after the period of increased volatility.*

Figure 7 (and Figure 9, third row) illustrate the effect of reducing ω, the absolute (i.e., independent of x_3) component of log-volatility to -4. The multiplicative scaling $\exp(\omega)$ of volatility is thus reduced to a sixth of that in the reference scenario. This leads to a low learning rate for x_2, which in turn leads to little learning in x_3, since the agent can only infer changes in x_3 from changes in x_2 (cf. Eq. (42)). This corresponds to an agent who pays little attention to new information, effectively filtering it out at an early stage.

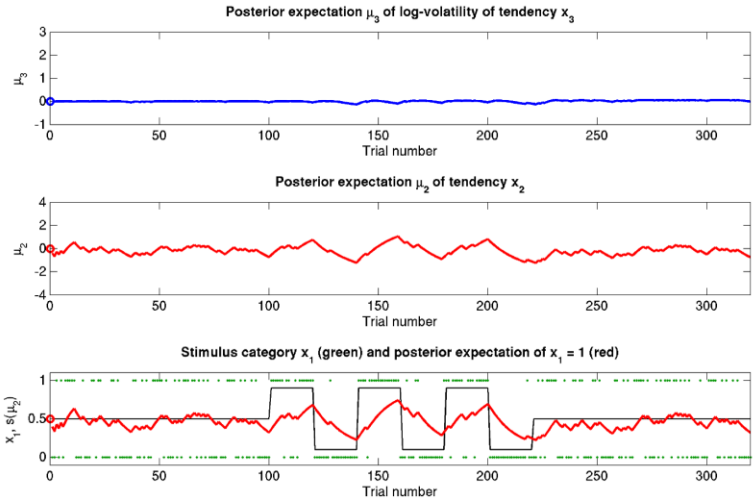

Figure 7. ***Reduced $\omega = -4$, (unchanged $\vartheta = 0.5$, $\kappa = 1.4$).*** *Symbols have the same meaning as in Figure 5. The small learning rate in x_2 leads to an extremely stable expected x_3. Despite prediction errors, the agent makes only small updates to its beliefs about its environment.*

The coupling between x_2 and x_3 can be diminished by reducing the value of κ, the relative (i.e., x_3-dependent) scaling of volatility x_2 ($\kappa = 0.2$ in Figure 8; Figure 9, bottom row). This impedes the flow of information up the hierarchy of levels in such a way that the agent's belief about x_3 is effectively insulated from the effects of prediction error in x_2 (cf. Eq. (42)). This leads to less learning about x_3 and to a much larger posterior variance σ_3 than in any of the above scenarios (see Figure 9, right panel). As with a reduced ϑ (Figure 6), learning about x_2 itself is not directly affected except that the second stage of the simulation, where higher volatility remains without affecting the learning rate of x_2 in the third stage. This time, however, this effect is not caused by overconfidence about x_3 (due to small ϑ) as in the above scenario. Instead, it obtains despite uncertainty about x_3 (large σ_3), which would normally be expected to lead to greater learning because of the dependency of the learning rate on σ_3. This paradoxical effect can be understood by examining Eqs (42) and (41), where smaller κ

exerts opposite direct and indirect effects on the learning rate for μ_3. Indirectly, the learning rate is increased, in that smaller κ increases σ_3. But this is dominated by the dependency of the learning rate on κ, which leads to a decrease in learning for smaller κ. This is quite an important property of the model: it makes it possible to have low learning rates in a highly volatile environment. This scenario describes an agent who is keen to learn but fails because the levels of its model are too weakly coupled for information to be passed efficiently up the hierarchy. In other words, the agent's low-level adaptability is accompanied by uncertainty about higher-level variables (i.e., volatility), leading to inflexibility. In anthropomorphic terms, one might imagine a person who displays rigid behavior because he/she remains uncertain about how volatile the world is (e.g., in anxiety disorders).

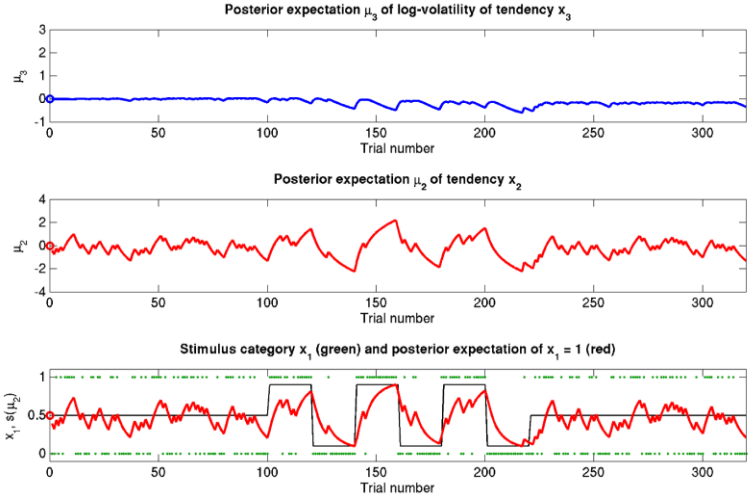

Figure 8. ***Reduced $\kappa = 0.2$, (unchanged $\vartheta = 0.5$, $\omega = -2.2$).*** *Symbols have the same meaning as in Figure 5. x_2 and x_3 are only weakly coupled. Despite uncertainty about x_3, only small updates to μ_3 take place. Sensitivity to changes in volatility is reduced. x_2 is not affected directly, but its learning rate does not increase with volatility.*

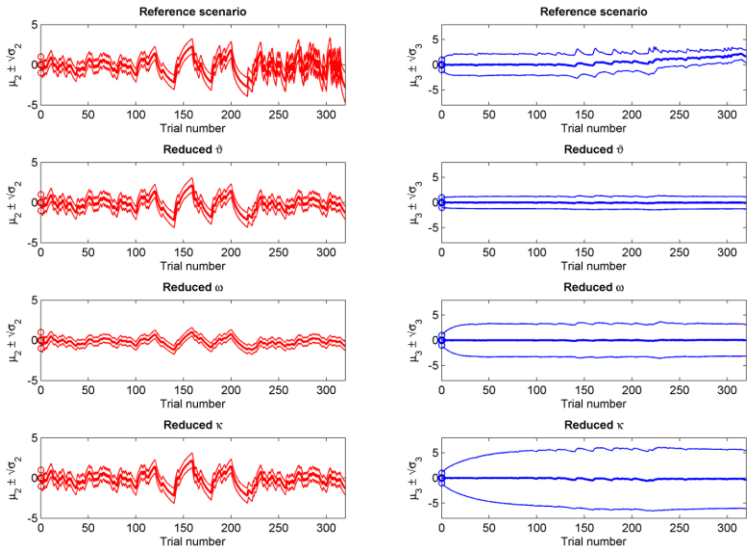

Figure 9. ***Simulations including standard deviations of posterior distributions.*** *Top to bottom: the four scenarios from Figures 5-8. Left:* μ_2 *(bold red line); fine red lines indicate the range of* $\pm\sqrt{\sigma_2}$ *around* μ_2*. Right:* μ_3 *(bold blue line); fine blue lines indicate the range of* $\pm\sqrt{\sigma_3}$ *around* μ_3*. Circles indicate initial values.*

The simulations described above switch between two probability regimes: $p(x_1 = 1) = 0.5$ and $p(x_1 = 1) = 0.9$ or 0.1. The stimulus distributions under these two regimes have different variances (or risk). Figure 10 shows an additional simulation run where risk is constant, i.e. $p(x_1 = 1) = 0.85$ or 0.15, throughout the entire simulation. One recovers the same effects as in the reference scenario. I now consider generalizations of the generative model that relax some of the simplifying assumptions about sensory mappings and outcomes I made during its exposition above.

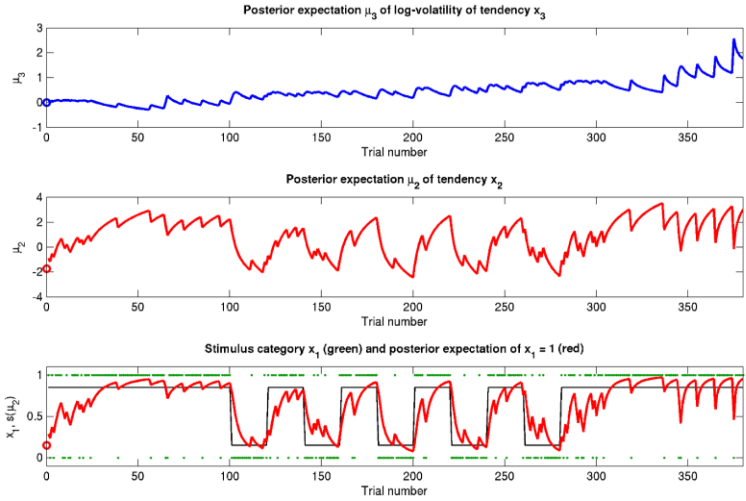

Figure 10. ***A simulation where risk is constant*** ***($\vartheta = 0.2$, $\omega = -2.3$, $\kappa = 1.6$).*** *Symbols have the same meanings as in Figure 5. The same basic phenomena shown in Figure 5 can be observed here.*

5 SENSORY INPUT

5.1 Sensory uncertainty

The HGF can readily accommodate sensory uncertainty at the first level. This pertains to the mapping from stimulus category x_1 to sensory input u. I first deal with the case of binary input.

To allow for sensory uncertainty, for example when the sensory inputs are ambiguous or affected by noise, I replace the deterministic relation in Eq. (9) by a stochastic one (cf. Daunizeau et al., 2010a).

$$p(u|x_1) = \mathcal{N}(u; \eta_a, \alpha)^{x_1} \cdot \mathcal{N}(u; \eta_b, \alpha)^{1-x_1} \tag{72}$$

Here, the input u is no longer binary but a real number whose distribution is a mixture of Gaussians. If $x_1 = 1$, the probability of u is normally distributed with constant variance α around a constant value η_a, corresponding to the most likely sensation if $x_1 = 1$. If however $x_1 = 0$, the most likely sensation is η_b with the probability of u normally distributed with the same variance α. The greater α (relative to the squared distance $(\eta_a - \eta_b)^2$), the greater the sensory uncertainty. The main point here is that with this modification, the model can account for situations where x_1 can no longer be inferred with certainty from u. The variational energy of the first level now is

$$I\left(x_1^{(k)}\right) = \left(-\frac{\left(u^{(k)} - \eta_a\right)^2}{2\alpha} + \ln s\left(\mu_2^{(k)}\right)\right)^{x_1^{(k)}} \cdot \left(-\frac{\left(u^{(k)} - \eta_b\right)^2}{2\alpha} + \ln\left(1 - \mathrm{s}\left(\mu_2^{(\mathrm{k})}\right)\right)\right)^{1-x_1^{(k)}} \tag{73}$$

With Eq. (26) we find the update rule for μ_1:

$$\mu_1^{(k)} = \hat{q}\left(x_1^{(k)} = 1\right) = \frac{\exp\left(-\frac{\left(u^{(k)} - \eta_a\right)^2}{2\alpha}\right) s\left(\mu_2^{(k)}\right)}{\exp\left(-\frac{(u^{(k)} - \eta_a)^2}{2\alpha}\right) s\left(\mu_2^{(k)}\right) + \exp\left(-\frac{(u^{(k)} - \eta_b)^2}{2\alpha}\right)\left(1 - s\left(\mu_2^{(k)}\right)\right)} \tag{74}$$

If $u^{(k)} \approx \eta_a$, one sees that $\mu_1^{(k)} \approx 1$ regardless of $\mu_2^{(k)}$. Likewise, $\mu_1^{(k)} \approx 0$ if $u^{(k)} \approx \eta_b$. This means that if the sensory input is sufficiently discriminable, x_2 has no influence on the agent's belief about x_1. If, however, the sensory input is ambiguous in that $u^{(k)}$ is far from both η_a and η_b, (rendering all exponential terms similarly small) we have $\mu_1^{(k)} \approx s\left(\mu_2^{(k)}\right)$; i.e. the agent has to rely on its belief about x_2 to predict stimulus category. Importantly, the update equations for the higher levels of the model are not affected by this introduction of sensory uncertainty at the first level.

5.2 Inference on continuous-valued states

Here, I turn to the general HGF model where the bottom-level state being inferred is not binary but continuous (i.e., a real number). As a concrete example, I use the exchange rate between the U.S. Dollar (USD) and the Swiss franc (CHF) during the first 180 trading days of the year 2010 (source: http://www.oanda.com/). In this example, the agent represented by our model can be seen as an individual market observer (e.g., a

currency trader), with the model describing how he "perceives" the relative value (and volatility) of USD. The data u are taken to be the closing USD-CHF exchange rate of each trading day with a likelihood model

$$p(u|x_1) = \mathcal{N}(u; x_1, \alpha) \tag{75}$$

where α is the constant variance with which the input u is normally distributed around the true value x_1. This can be regarded as a measure of uncertainty (i.e., how uncertain the trader is about his "perception" of USD value relative to CHF). On top of this input level, we can now add as many coupled random walks as we please. All higher levels will then be described by Eqs. (3), (4), and (8), leading to the general HGF updates of Eqs (47) and (48).

At the first level, we obtain by the method introduced above

$$\begin{aligned} I\left(x_1^{(k)}\right) = &-\frac{1}{2\alpha}\left(x_1^{(k)} - u^{(k)}\right)^2 \\ &-\frac{1}{2\left(\sigma_1^{(k-1)} + \mathrm{e}^{\kappa_1\mu_2^{(k-1)}+\omega_1}\right)}\left(x_1^{(k)} - \mu_1^{(k-1)}\right)^2, \end{aligned} \tag{76}$$

$$\sigma_1^{(k)} = \frac{\alpha\left(\sigma_1^{(k-1)} + \mathrm{e}^{\kappa_1\mu_2^{(k-1)}+\omega_1}\right)}{\alpha + \sigma_1^{(k-1)} + \mathrm{e}^{\kappa_1\mu_2^{(k-1)}+\omega_1}}, \tag{77}$$

$$\begin{aligned}\mu_1^{(k)} &= \mu_1^{(k-1)} + \sigma_1^{(k)}\frac{1}{\alpha}\left(u^{(k)} - \mu_1^{(k-1)}\right) \\ &= \mu_1^{(k-1)} + \frac{\sigma_1^{(k-1)} + \mathrm{e}^{\kappa_1\mu_2^{(k-1)}+\omega_1}}{\alpha + \sigma_1^{(k-1)} + \mathrm{e}^{\kappa_1\mu_2^{(k-1)}+\omega_1}}\left(u^{(k)} - \mu_1^{(k-1)}\right).\end{aligned} \tag{78}$$

For $\alpha = 0$ (no sensory uncertainty), the last two equations reduce, as they should, to $\sigma_1^{(k)} = 0$, $\mu_1^{(k)} = u^{(k)}$. Note also that since $I(x_1)$ is already quadratic here, no further approximation to the mean field approximation is needed, and μ_1 and σ_1 are the exact moments of the posterior under the mean field approximation.

In the equivalent precision formulation, with $\hat{\pi}_u \stackrel{\text{def}}{=} 1/\alpha$, the updates read

$$\pi_1^{(k)} = \hat{\pi}_1^{(k)} + \hat{\pi}_u, \tag{79}$$

$$\mu_1^{(k)} = \mu_1^{(k-1)} + \frac{\hat{\pi}_u}{\pi_1^{(k)}}\left(u^{(k)} - \mu_1^{(k-1)}\right). \tag{80}$$

Here again, we find that the update is a precision-weighted prediction error.

Scenarios with different parameter values for the USD-CHF example are presented in Figures 11-14. To keep the example simple, the number of levels is chosen as $n = 2$. These scenarios can be thought of as corresponding to different individual traders who receive the same market data but process them differently. The reference scenario (Figure 11) is based on the parameter values $\kappa_1 = 1$, $\omega_1 = -12$, $\vartheta = 0.3$, and $\alpha = 2 \cdot 10^{-5}$. This parameterization conveys a small amount of sensory uncertainty that leads to minor but visible deviations of μ_1 from u. The updates to μ_1 are conservative in the sense that they consider prior

information along with new input. Note also that μ_2 rises whenever the prediction error about x_1 is large, that is when the green dots denoting u are outside the range $\mu_1 \pm \sqrt{\sigma_1}$ indicated by the red lines. Conversely, μ_2 falls when predictions of x_1 are more accurate. In the next scenario (Figure 12), the value of α is further reduced to 10^{-6}. This scenario thus shows an agent who is effectively without sensory uncertainty. As prescribed by the update equations above, μ_1 now follows u with great accuracy and μ_2 tracks the amount of change in x_1. In Figure 13, sensory uncertainty is increased by two orders of magnitude ($\alpha = 10^{-4}$). Here, the agent adapts more slowly to changes in the exchange rate since it cannot be sure whether prediction error is due to a change in the true value of x_1 or to misperception. The final scenario in Figure 14 shows an agent with the same sensory uncertainty as in the reference scenario but a prior belief that the environment is not very volatile, i.e., ϑ is reduced from 0.3 to 0.01. Smaller values of ϑ smooth the trajectory of μ_2 in a similar way that sensory uncertainty smoothes the trajectory of μ_1.

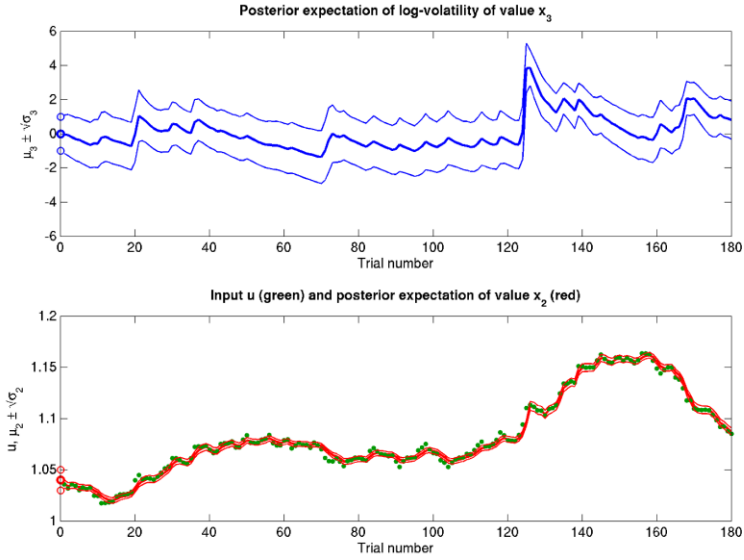

Figure 11. ***Inference on a continuous-valued state*** ***($\vartheta = 0.3$, $\omega = -12$, $\kappa = 1$, $\alpha = 2 \cdot 10^{-5}$).*** *Reference scenario for the model of hierarchical Gaussian random walks applied to a continuous-valued state at the bottom level. The state is the value x_2 of the U.S. Dollar against the Swiss franc during the first 180 trading days of the year 2010. Bottom panel: input u representing closing exchange rates (green dots). The bold red line surrounded by two fine red lines indicates the range $\mu_2 \pm \sqrt{\sigma_2}$. Top panel: The range $\mu_3 \pm \sqrt{\sigma_3}$ of the log-volatility x_3 of x_2.*

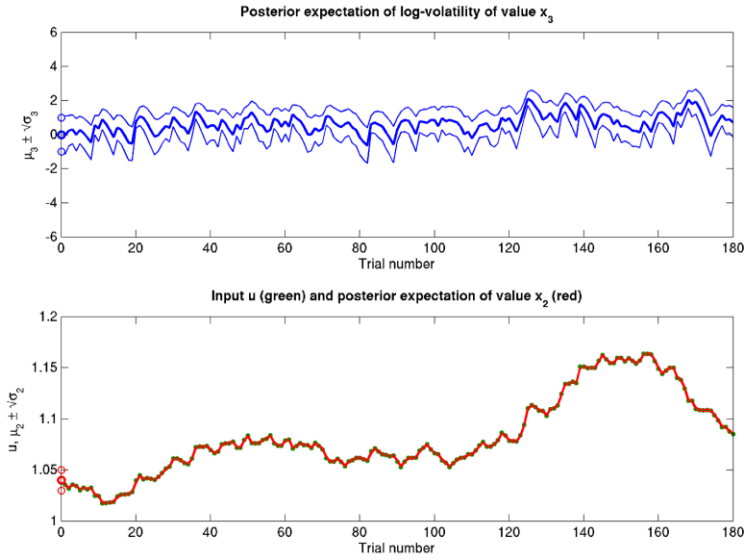

Figure 12. ***Reduced*** $\boldsymbol{\alpha = 10^{-6}}$ $\boldsymbol{(\vartheta = 0.3,\ \omega = -12,\ \kappa = 1)}$. *Reduced sensory uncertainty α with respect to the reference scenario of Figure 11.*

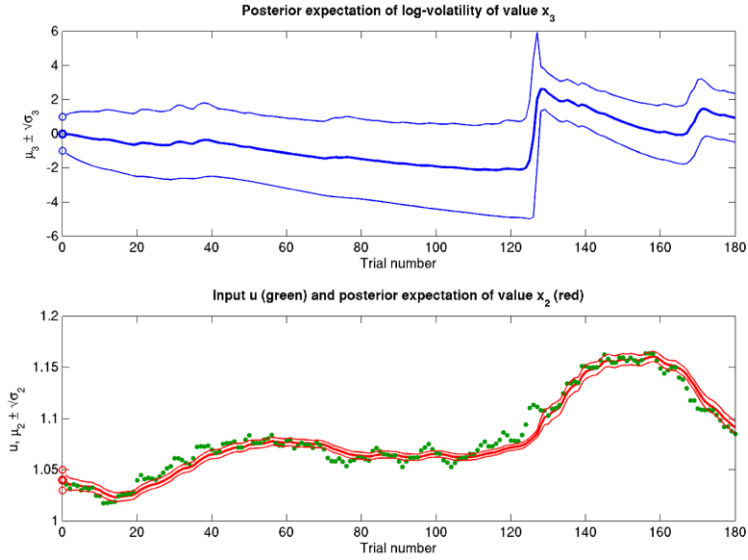

Figure 13. ***Increased*** $\boldsymbol{\alpha = 10^{-4}}$ $\boldsymbol{(\vartheta = 0.3,\ \omega = -12,\ \kappa = 1)}$. *Increased sensory uncertainty α with respect to the reference scenario of Figure 11.*

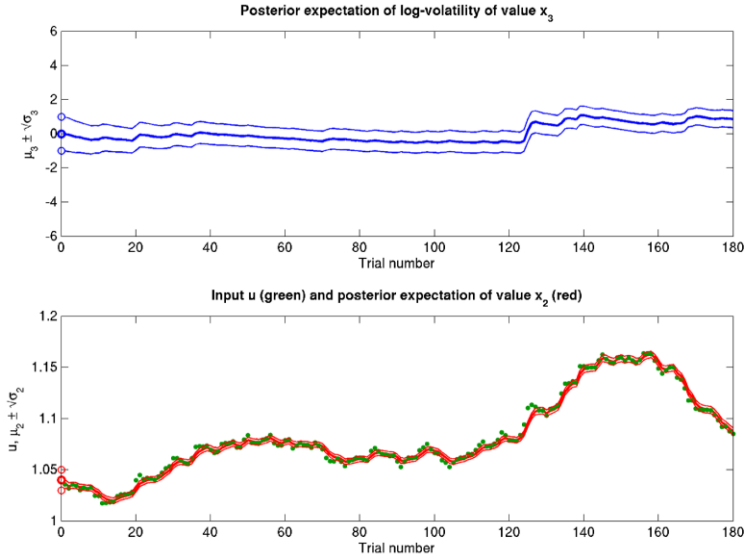

Figure 14. ***Reduced*** $\vartheta = 0.01$ $(\omega = -12,\ \kappa = 1,\ \alpha = 2 \cdot 10^{-5})$. *Reduced* ϑ *with respect to the reference scenario of Figure 11.*

This example again emphasizes the fact that Bayes-optimal behavior can manifest in many diverse forms. The different behaviors emitted by the agents above are all optimal under their implicit prior beliefs encoded by the parameters that control the evolution of high-level hidden states. Clearly, it would be possible to optimize these parameters using the same variational techniques we have considered for the hidden states. This would involve optimizing the free energy bound on the evidence for each agent's model (prior beliefs) integrated over time (i.e., learning). Alternatively, one could optimize performance by selecting those agents with prior beliefs (about the parameters) that had the best free energy (made the most accurate inferences over time). Colloquially, this would be the difference between training an expert to predict financial markets and simply hiring experts whose priors were the closest to the true values.

6 RESPONSE MODELS

6.1 The need for response models

The need for a *response model* (or equivalently, *decision model*) arises whenever we are observing an observer. Modeling such a situation is discussed in detail in Daunizeau et al. (2010a, 2010b). Since two kinds of observations have to be modeled, namely the observations made by the observer and our observations of the observer, and since perceptual models such as the HGF only model the first of these, we need a model of how the decisions the observed agent makes come about given its perception (i.e., its observations and its resulting inferences).

Note that not all possible applications of the HGF fit this description; two cases in particular do not require a response model: first, when simply filtering data with a given parameter setting as in the simulations above, no response is ever made or observed, only an inference is modeled. Second, when optimizing parameters using, for example, Empirical Bayes, the same applies.

In this chapter, I give two examples of simple response models and I demonstrate their principled derivation from a loss function. Other response models will be discussed in Chapters 8 and 10 on applications of the HGF in experimental studies.

6.2 Response model for a simple binary loss function

One of the simplest decision situations for an agent is having to choose between two options, only one of which will be rewarded, but both of which offer the same gain (i.e., negative loss), if rewarded. One may imagine the game sometimes played with children where an adult holds both hands behind his back and asks the child to guess in which of them he is holding a present.

In the HGF, we may code one such binary outcome as $x_1 = 1$ and the other as $x_1 = 0$. This allows us to define a quadratic loss function ℓ where making the wrong choice $y \in \{0,1\}$ leads to a loss of 1 while the right choice leads to a loss of 0:

$$\ell(x_1, y) = (x_1 - y)^2 \tag{81}$$

The expected loss $\mathcal{Q}$ of response y, given the agent's representations λ, is then the expectation of ℓ under the distributions q described by λ:

$$\mathcal{Q}(y;\lambda) = E\lfloor \ell(x_1, y) \rfloor_{q(x_1|\lambda)} = \sum_{x_1 \in \{0,1\}} \ell(x_1, y) \cdot q(x_1|\lambda) \tag{82}$$

To evaluate this, we must remember that the agent has to rely on its beliefs deriving from time $k-1$ to make response $y^{(k)}$ at time k. In the above equation, elements of λ therefore have time index $k-1$, while x_1 and y have time index k. Specifically, the belief on the outcome probability at the first level is $\hat{\mu}_1^{(k)} = s\left(\mu_2^{(k-1)}\right)$. With $q(x_1|\lambda) = (\hat{\mu}_1)^{x_1}(1-\hat{\mu}_1)^{1-x_1}$ ((Mathys et al., 2011), Eq. 12), we then have

$$\mathcal{Q}(y=0;\lambda) = \hat{\mu}_1; \qquad \mathcal{Q}(y=1;\lambda) = 1 - \hat{\mu}_1 \tag{83}$$

The optimal response y^* is the one that minimizes expected loss $\mathcal{Q}$:

$$y^* \stackrel{\text{def}}{=} \arg\min_y Q(y;\lambda) = \left\{ \begin{array}{c} 1 \text{ if } \hat{\mu}_1 > \frac{1}{2} \\ \{0,1\} \text{ if } \hat{\mu}_1 = \frac{1}{2} \\ 0 \text{ if } \hat{\mu}_1 < \frac{1}{2} \end{array} \right\} \tag{84}$$
$$= \lim_{\zeta \to \infty} \frac{\hat{\mu}_1^{\zeta}}{\hat{\mu}_1^{\zeta} + (1 - \hat{\mu}_1)^{\zeta}}$$

This means, quite anticlimactically, that to minimize its losses, the agent should choose the option it believes more likely to be rewarded. It may seem superfluous to go to such lengths to derive such an obvious result, but the purpose of the above is also to give an illustration of the principled way a decision rule can be derived by combining a perceptual model with a loss function in more complicated cases.

It is, however, unreasonable to assume that human agents will always choose the option that minimizes their expected loss in the current trial, for two reasons. First, if there is more than one trial and the probabilities of the different options are independent, there is an exploration/exploitation tradeoff that makes it worth the agent's while (in the long run) sometimes to choose an option that is not expected to minimize loss in the current trial (Daw et al., 2006; Steyvers et al., 2009). Second, biological agents exhibit response noise (Faisal et al., 2008). To allow for exploration and noise, I use a response model that corresponds to the right-hand side of Eq. (84), without taking the limit, instead leaving ζ as a parameter to be estimated from the data:

$$p(y|\lambda,\zeta) = \left(\frac{\hat{\mu}_1^{\zeta}}{\hat{\mu}_1^{\zeta} + (1 - \hat{\mu}_1)^{\zeta}} \right)^{y} \cdot \left(\frac{(1 - \hat{\mu}_1)^{\zeta}}{\hat{\mu}_1^{\zeta} + (1 - \hat{\mu}_1)^{\zeta}} \right)^{1-y} \tag{85}$$

Figure 15 contains a graph of this function for $p(y = 1)$ where ζ plays the role of the noise (or exploration) parameter. This response model was the basis for the simulations described below. I call this response model the

unit square sigmoid model: since both its argument and value are always in the unit interval, its graph is restricted to the unit square, and it has the shape of a sigmoid.

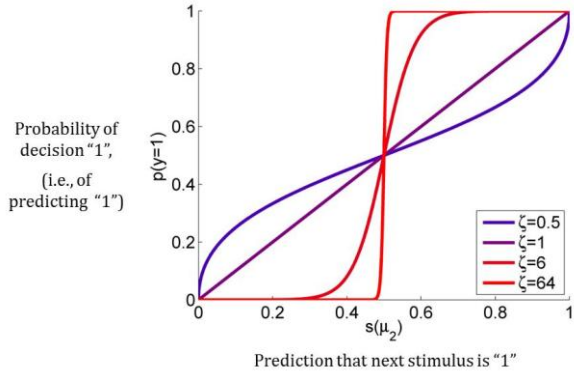

Figure 15. ***The unit square sigmoid*** *(cf. Eqs (84) and (85)). The parameter ζ can be interpreted as inverse response noise because the sigmoid approaches a step function as ζ approaches infinity.*

6.3 Response model for a one-armed bandit

As an additional example, I discuss a more complex binary decision task that we used to collect data from human subjects. In this variant of a one-armed bandit experiment, subjects were asked to play a series of gambles with the goal of maximizing their overall score (Daw et al., 2006; Steyvers et al., 2009). On each trial, subjects chose between two options represented by the same two fractals, which had different and time-varying reward probabilities. At any point in time, these probabilities summed to unity, and exactly one of the two options would be rewarded. Although subjects knew that probabilities varied throughout the course of the experiment, they were not told the schedule that governed these changes. The schedule included both a period of low volatility and a period of high volatility.

In order to encourage subjects to switch options above and beyond normal exploration behavior (Steyvers et al., 2009), the two fractals were associated with varying reward magnitudes. On each trial, magnitudes were drawn from a uniform distribution $\mathcal{U}(1,9)$ (i.e., rewards would take values out of the range {1, 2,…, 9} with equal probability).

Subjects began the experiment with an initial score of 0 points. Once a fractal had been chosen, if that fractal was rewarded, the associated reward would be added to the current score. The final score at the end of the experiment was translated into monetary reimbursement. The experiment consisted of 160 trials.

Calling the two fractals A and B, I parameterize the agent's response by

$$y = \begin{cases} 0 & \text{for choice A} \\ 1 & \text{for choice B} \end{cases} \tag{86}$$

Correspondingly, the state x_1 is

$$x_1 = \begin{cases} 0 & \text{if A rewarded} \\ 1 & \text{if B rewarded} \end{cases} \tag{87}$$

Taking r_A and r_B to be the rewards for A and B, respectively, I introduce the quadratic loss function

$$\begin{aligned} l(x,y) &= -(1-x_1)(y+x_1-1)^2 \cdot r_A - x_1(y+x_1-1)^2 \cdot r_B = \\ &= -(y+x_1-1)^2((r_B-r_A)x_1+r_A) \end{aligned} \tag{88}$$

This corresponds to the following loss table:

$$
\begin{array}{cc|c|c|}
 & & x_1 & \\
 & & 0 & 1 \\
\hline
y & 0 & -r_A & 0 \\
\hline
 & 1 & 0 & -r_B \\
\hline
\end{array}
\tag{89}
$$

Following the same procedure as above, we get:

$$
\begin{aligned}
Q(y;\lambda) &= \sum_{x_1 \in \{0,1\}} \ell(x_1, y) \cdot q(x_1|\lambda) \\
&= -y^2\, r_B\, \hat{\mu}_1 - (1-y)^2\, r_A\, (1-\hat{\mu}_1) \\
&= \begin{cases} -r_B\, \hat{\mu}_1 & \text{for } y = 1, \\ -r_A (1-\hat{\mu}_1) & \text{for } y = 0. \end{cases}
\end{aligned}
\tag{90}
$$

With the expected loss from each option on a continuous scale, a simple but powerful response model is the softmax rule (Daw et al., 2006; Sutton and Barto, 1998)

$$
p(y_i|\lambda, \zeta) = \frac{\exp\big(\zeta \cdot Q(y_i;\lambda)\big)}{\sum_j \exp\Big(\zeta \cdot Q\big(y_j;\lambda\big)\Big)},
\tag{91}
$$

where y_i is one particular option and the sum runs over all options. This means that the response probabilities are Boltzmann-distributed according to their expected rewards (i.e., their expected negative losses) with the parameter ζ serving as the analogon of inverse temperature. In our binary case, this evaluates to

$$
\begin{aligned}
p(y = 1|\lambda, \zeta) &= s\left(-\zeta\left(r_B\,\hat{\mu}_1 - r_A(1 - \hat{\mu}_1)\right)\right), \\
p(y = 0|\lambda, \zeta) &= s\left(+\zeta\left(r_B\,\hat{\mu}_1 - r_A(1 - \hat{\mu}_1)\right)\right).
\end{aligned}
\tag{92}
$$

This is a logistic sigmoid function of the difference $r_B\,\hat{\mu}_1 - r_A(1 - \hat{\mu}_1)$ of expected reward for choice B minus expected reward for choice A. If the expected reward of choice B exceeds that of choice A, the likelihood of choice B is greater than half and vice versa.

In the next chapter, I look at the estimation of the parameters of both the perceptual and response models, either given observed responses or some criterion of optimality.

7 PARAMETER ESTIMATION AND MODEL COMPARISON

7.1 Coordinate choice on higher levels

The definition of response models (or, from the point of view of the experimenter, *observation models*) such as Eqs (85) and (92) completes the "observing-the-observer" framework (Daunizeau et al., 2010b). This permits us to invert the ensemble of perceptual and response models, that is to infer on the response parameters $\zeta \stackrel{\text{def}}{\equiv} \{\zeta_1, \zeta_2, \dots\}$ (just one in the above examples, but in general many) and on the perceptual parameters $\chi \stackrel{\text{def}}{\equiv} \{\kappa_1, \omega_1, \dots, \kappa_n, \omega_n, \vartheta\}$. Before I return to our example perceptual model, where we only have $\kappa_2 \equiv \kappa$ and $\omega_2 \equiv \omega$, and to the response models discussed above, I discuss some of the subtler aspects of the perceptual parameters in more depth and generality.

The question I will deal with in this section is the meaning of the hidden states x_i at higher levels of the HGF. I will again do this by using the by now familiar example of a three-level model with binary x_1. There, the state x_3 at the third level of the model represents the volatility (or, more precisely, the log-volatility) of x_2; this raises the question what units we should measure x_3 in – what effect will moving the origin on x_3 have? – what effect will rescaling x_3 have?

In Section 4.5 I said that any change in the initial value $\mu_3^{(0)}$ of μ_3 can be neutralized by corresponding changes in κ and ω. This means that instead of setting $\mu_3^{(0)}$ to a fixed value, we may just as well set $\kappa = 1$, thereby making it (seemingly) disappear from the model. But we now have a variable $\mu_3^{(0)}$. As an example, let us take data from a behavioral study where parameters have been estimated with $\mu_3^{(0)}$ set to 1. The relevant parameter estimates are:

$$\begin{aligned} \kappa &= 2.49 \\ \sigma_3^{(0)} &= 0.988 \\ \vartheta &= 0.000592 \end{aligned} \tag{93}$$

We can now make κ disappear by setting it to 1. This change should be neutral to the model's predictions of input, which can only be achieved by the following compensatory substitutions:

$$\begin{aligned} \mu_3^{(0)} = 1 &\rightarrow \mu_3^{(0)} = 2.49 \\ \sigma_3^{(0)} = 0.988 &\rightarrow \sigma_3^{(0)} = 0.988 \cdot 2.49^2 \\ \vartheta = 0.000592 &\rightarrow \vartheta = 0.000592 \cdot 2.49^2 \end{aligned} \tag{94}$$

At the first two levels, nothing has changed (the trajectory of μ_2 and therefore the input predictions $\hat{\mu}_1$ are the same); however, at the third level x_3 has been rescaled by the inverse of the factor with which κ has been rescaled (i.e., $1/2.49$). This means that the term

$$v^{(k)} \stackrel{\text{def}}{=} \exp\left(\kappa \mu_3^{(k-1)} + \omega\right) \tag{95}$$

is invariant under the above transformation for all time points k, leading to the same trajectory in μ_2 and σ_2 according to Eqs (34) and (35) as before.

Effectively, because of the transformation to a fixed $\kappa = 1$, our estimate of κ has been reinterpreted as an estimate of $\mu_3^{(0)}$. This is a consequence of our freedom to choose coordinates on x_3. To take an analogy from geometry, the distance between Zurich and London is an objective geometric quantity that does not change whether it is measured in miles or

kilometers; we may even introduce new units where this distance is 1. Likewise, we may always rescale x_3 individually for each agent (and run) such that the coupling κ between the second and third level has value 1. Note, however, that this does not prevent the coupling from existing or from differing between agents.

If the representations of x_3 (i.e., μ_3 and σ_3) were part of the observation model, this would give us a direct handle on x_3 and measures of its representations could immediately be compared between agents, provided we use common coordinates (which we would automatically do without even thinking about it; individually rescaling κ to 1 would then obviously be unwise).

However, as long as we have no measure of x_3, there is a fundamental ambiguity between individual differences in coupling and individual differences in priors for x_3.

Nonetheless, we have to make *some* choice of coordinates. In setting $\mu_3^{(0)} = 1$, we choose to take the belief on environmental volatility that an agent begins inference with as the benchmark. Observed differences in learning are then attributed to, inter alia, differences in coupling. This is one way to obtain comparable measures of belief on x_3 (and consequently, ϑ) between agents, since one may be able to influence their priors while there is usually no way to equalize their coupling levels.

Just as it is possible to set κ to an arbitrary non-zero value while keeping v invariant by compensatory substitutions, one can set ω to an arbitrary value with invariant v using another set of compensatory substitutions. In particular, ω can be set to zero, thereby making it seemingly disappear from the model, just as κ seems to disappear when set to 1.

We have seen that any change of scale in x_3 is expressed in a corresponding change of κ. However, in choosing coordinates on x_3, we have an additional degree of freedom: the choice of origin. Changes in this are expressed in a corresponding change of ω :

$$x_3' \stackrel{\text{def}}{=} x_3 + a \tag{96}$$

then

$$\exp(\kappa x_3 + \omega) = \exp(\kappa(x_3' - a) + \omega) = \exp(\kappa x_3' + \omega') \tag{97}$$

with

$$\omega' \stackrel{\text{def}}{=} -\kappa a + \omega \tag{98}$$

In our example, we have reinterpreted the estimate of κ as one of $\mu_3^{(0)}$. We may now go on to reinterpret it another time, this time as an estimate of ω (up to now fixed to -4) by shifting the origin on x_3 such that $\mu_3^{(0)}$ is again fixed to 1:

$$\begin{gathered} \mu_3^{(0)} = 2.49 \rightarrow \mu_3^{(0)} = 1 \\ \omega = -4 \ \rightarrow \omega = -\kappa(1 - 2.49) + (-4) = -2.51 \end{gathered} \tag{99}$$

Again, the trajectories at the first two levels have not changed. It is now apparent that by rescaling x_3 and shifting its origin, we can choose arbitrary values for two out of the three parameters $\mu_3^{(0)}$, κ, and ω. However, I repeat that this is only possible as long as we do not measure x_3 on any objective scale.

Equivalence classes (with equivalence defined as leading to invariant v) of parameter values are defined by the following *conservation laws:*

$$\begin{aligned} \kappa' \mu_3^{\prime(0)} + \omega' &= \kappa \mu_3^{(0)} + \omega \\ \kappa'^2 \vartheta' &= \kappa^2 \vartheta \\ \kappa'^2 \sigma_3^{\prime(0)} &= \kappa^2 \sigma_3^{(0)} \end{aligned} \tag{100}$$

Using these equations, we can now make κ and ω seemingly disappear by setting them to 1 and 0, respectively. Note that this does not really amount to a simplification of the model: it is only a coordinate choice. In our example this means:

$$\begin{aligned} \omega = -2.51 \ &\rightarrow \omega = 0 \\ \mu_3^{(0)} = 1 &\rightarrow \mu_3^{(0)} = -1.51 \end{aligned} \tag{101}$$

This disappearance of κ and ω (and in the general case, all κ_i and ω_i) from the model may seem convenient. However, a danger here is to confuse coordinates with the underlying reality they describe (in gravitation, it took more than two centuries to see that error). Crucially, it is impossible to discuss the choice of coordinate choice on higher levels in a model that lacks κ_i and ω_i. It is however important to note that this choice does not cease to exist when one stops acknowledging it. The model takes the hidden states at higher levels to be an objective property of the agent's environment. To compare beliefs on it across agents, we need a common scale to measure them on. However, individually adjusting the scale to make things look more convenient (which is what transforming κ_i and ω_i away achieves) moves us away from such a common scale.

This raises the question how a common scale can be found. This is important, for example, for the comparison of estimates such as that of ϑ, which are always relative to the scale of x_n. If we measure inferences on x_n, this would automatically give us a common scale. Likewise, to have a common scale on any of the x_i, inferences on that x_i have to be measured.

The way to measure such inferences is to include them (sc. the representations μ_i) in the response model. In cases where this is not done (it is often either not possible or not desirable), it is important to remember that the choice of coordinates on some higher levels is arbitrary. In order to avoid on overparameterized model, it is then advisable to be explicitly arbitrary by fixing two out of $\mu_i^{(0)}$, κ_i, and ω_i, reflecting a choice of origin and scale on x_i. By contrast, when μ_i is included in the response model, all of $\mu_i^{(0)}$, κ_i, and ω_i can be estimated.

7.2 The MAP estimate

Our goal in this section will be to find an expression for the *maximum-a-posteriori (MAP)* estimate for the parameters $\xi \stackrel{\text{def}}{=} \{\chi, \lambda^{(0)}, \zeta\}$ (with perceptual parameters $\chi \stackrel{\text{def}}{=} \{\kappa_1, \omega_1, \dots, \kappa_n, \omega_n, \vartheta\}$, initial representations $\lambda^{(0)} = \{\mu_1^{(0)}, \sigma_1^{(0)}, \dots, \mu_n^{(0)}, \sigma_n^{(0)}\}$, and response parameters $\zeta \stackrel{\text{def}}{=} \{\zeta_1, \zeta_2, \dots\}$). The MAP estimate ξ^* of ξ is defined as

$$\xi^* \stackrel{\text{def}}{=} \arg\max_{\xi} p(\xi|y, u), \tag{102}$$

where $u \stackrel{\text{def}}{=} \{u^{(1)}, \dots, u^{(K)}\}$ and $y \stackrel{\text{def}}{=} \{y^{(1)}, \dots, y^{(K)}\}$ are the inputs and responses from time points $k = 1$ to $k = K$, respectively. I unpack this to make it tractable:

$$
\begin{aligned}
\xi^* &= \arg\max_{\xi} p(\xi|y,u) = \arg\max_{\xi} \frac{p(\xi,y|u)}{p(y)} \\
&= \arg\max_{\xi} p(\xi,y|u) \; = \arg\max_{\xi} \ln p(\xi,y|u) \\
&= \arg\max_{\xi} \ln\big(p(y|\xi,u)p(\xi)\big) \\
&= \arg\max_{\xi} \left(\ln p(y|\xi,u) + \ln p(\xi)\right) \\
&= \arg\max_{\xi} \left(\sum_k \ln p\big(y^{(k)}\big|\xi,u\big) + \ln p(\xi)\right) \\
&= \arg\max_{\xi} \left(\sum_k \ln p\big(y^{(k)}\big|\lambda^{(k)}\big(\chi,\lambda^{(0)},u\big),\zeta\big) + \ln p(\xi)\right)
\end{aligned}
\tag{103}
$$

The objective function $Z(\xi)$ that needs to be maximized is therefore the log-joint probability density of the parameters ξ and responses y given inputs u:

$$
\begin{aligned}
Z(\xi|u,y) &\stackrel{\text{def}}{=} \ln p(\xi,y|u) \\
&= \sum_k \ln p\big(y^{(k)}\big|\lambda^{(k)}\big(\chi,\lambda^{(0)},u\big),\zeta\big) + \ln p(\xi)
\end{aligned}
\tag{104}
$$

While the response model furnishes $p\big(y^{(k)}\big|\lambda^{(k)},\zeta\big)$ (cf. Eqs (85) and (92)), the perceptual model (i.e., the HGF) provides the representations $\lambda^{(k)}\big(\chi,\lambda^{(0)},u\big)$ (cf. Eq. (52)) by way of its update equations.

The last missing part in Eq. (104) is the prior distribution $p(\xi)$. This will be discussed in the next section.

There are many practical ways to perform the maximization in Eq. (103). We have compared four of them in a simulation study that is discussed in Section 7.4.

7.3 Priors and transformed parameter spaces

A crucial part of Bayesian inference is the specification of a prior distribution, in our case $p(\xi)$. There is in principle no reason why the priors on the different elements of ξ should not be independent; therefore, we may assume

$$p(\xi) = p(\vartheta) \prod_i p(\kappa_i) p(\omega_i) p\left(\mu_i^{(0)}\right) p\left(\sigma_i^{(0)}\right) \prod_j p(\zeta_j). \tag{105}$$

The $p(\zeta_j)$ depend on the response model chosen and will have to be discussed on a case-by-case basis, but the rest are generic and will be discussed in what follows.

The most straightforward case are the priors on the ω_i. Since the ω_i can take values on the whole real line, they can be estimated in their native space with a (possibly wide) Gaussian prior:

$$p(\omega_i) = \mathcal{N}(\omega_i; \mu_{\omega_i}, \sigma_{\omega_i}). \tag{106}$$

The same applies to the $\mu_i^{(0)}$:

$$p\left(\mu_i^{(0)}\right) = \mathcal{N}\left(\mu_i^{(0)}; \mu_{\mu_i^{(0)}}, \sigma_{\mu_i^{(0)}}\right). \tag{107}$$

The $\sigma_i^{(0)}$ have a natural lower bound at zero since they are variances. We can avoid non-positive values by estimating $\sigma_i^{(0)}$ in log-space. In that transformed space, we again choose a Gaussian prior:

$$p\left(\ln \sigma_i^{(0)}\right) = \mathcal{N}\left(\ln \sigma_i^{(0)}; \mu_{\ln \sigma_i^{(0)}}, \sigma_{\ln \sigma_i^{(0)}}\right). \tag{108}$$

Just like $\sigma_i^{(0)}$, ϑ is a variance and has a lower bound at zero. But here, the case is slightly more complicated. In addition to the lower bound, it is desirable to have an upper bound on ϑ. This is because, for a ϑ too large, the assumptions underlying the derivation of the update equations of the HGF no longer hold. It is then possible to get updates that push the precision π_n at the top level below zero, indicating that the agent knows "less than nothing" about x_n. In less extreme cases, a large ϑ may allow μ_n to jump to very high levels, giving rise to improbable inference trajectories. This is due to a violation of the assumption underlying Eqs (56) and (58) that the variational energies $I(x_i)$ are nearly quadratic.

To avoid such violations, it is sensible to place an upper bound on ϑ in addition to the lower bound at zero. This can be achieved by estimating ϑ in "logit-space", a logistic sigmoid transformation of native space with a variable upper bound $a > 0$:

$$\begin{aligned} &\mathrm{logit}_a(x) \stackrel{\text{def}}{=} \ln\left(\frac{x}{a-x}\right); \\ &\Longrightarrow x = \frac{a}{1+\exp\left(-\mathrm{logit}_a(x)\right)} \end{aligned} \tag{109}$$

In that space, the prior on ϑ can then be taken as

$$p\left(\mathrm{logit}_{a_\vartheta}\,\vartheta\right) = \mathcal{N}\left(\mathrm{logit}_{a_\vartheta}\,\vartheta; \mu_{\mathrm{logit}_{a_\vartheta}\,\vartheta}, \sigma_{\mathrm{logit}_{a_\vartheta}\,\vartheta}\right). \tag{110}$$

While the κ_i's can in principle take any real value, flipping the sign of κ_i is equivalent to flipping it on x_{i+1} (cf. Eq. (8)). It is therefore useful to adopt

the convention that all $\kappa_i > 0$. This leads to the more intuitive relation that greater x_{i+1} means greater variability in x_i; in other words, this makes the f_i's in Eq. (8) monotonically increasing functions. A second useful constraint on the κ_i's is that they be bounded above, for the same reason as ϑ. The violations that can occur at the top level for a ϑ too large can occur at lower levels for κ_i's too large. Consequently, we evaluate the κ_i's in logit-space with the following priors:

$$p\left(\text{logit}_{\text{a}_{\kappa_i}}\,\kappa_i\right) = \mathcal{N}\left(\text{logit}_{\text{a}_{\kappa_i}}\,\kappa_i\,;\mu_{\text{logit}_{\text{a}_{\kappa_i}}\,\kappa_i},\sigma_{\text{logit}_{\text{a}_{\kappa_i}}\,\kappa_i}\right). \tag{111}$$

Choosing appropriate upper bounds, means, and variances in a given case is not as tricky (or as arbitrary) as it may seem at first. Since a model is fully defined by the combination of likelihood and prior, choosing a different prior amounts to having a different model. This allows us to justify our choice of prior using model comparison (cf. Section 7.6). As a rule of thumb, one may first try to be generous regarding the widths and upper bounds of priors, only decreasing them as needed to increase model evidence as computed during model comparison. How restrictive one wants to be depends on how general he wants his proposed model to be: while restricting priors further may increase model evidence for the dataset at hand, it may impede generalization to other datasets.

7.4 Simulation study

7.4.1 PURPOSE AND SCOPE

To assess our ability to estimate the parameters $\xi = \left\{\kappa, \omega, \vartheta, \mu_2^{(0)}, \sigma_2^{(0)}, \mu_3^{(0)}, \sigma_3^{(0)}, \zeta\right\}$ in our example model, we conducted a simulation study. Simulations took place in four steps:

1. We chose a particular sequence of 320 binary inputs $u = \left\{u^{(1)}, \ldots, u^{(320)}\right\}$.

2. We chose a particular set of values for the parameters ξ.

3. We generated 320 binary responses $y = \{y^{(1)}, \dots, y^{(320)}\}$ by drawing from the response distribution of Eq. (85).

4. We estimated ξ^* according to Eq. (103).

Step 1 was only performed once, so that u was the same in all simulations. The values of ξ in step 2 were constant for all parameters except κ and ζ. The values of κ and ζ were taken from a two-dimensional grid in which the κ dimension took the values $\{0.5, 1, 1.5, \dots, 3.5\}$ while the ζ dimension took the values $\{0.5, 1, 6, 24\}$. Steps 3 and 4 were then repeated 1'000 times for each value pair on the $\{\kappa, \zeta\}$ grid. The ζ values on the grid were chosen such that they covered the whole range from very little response noise ($\zeta = 24$) to very much ($\zeta = 0.5$, cf. Figure 15). The κ values were chosen to cover the range observed in an empirical behavioral study of 42 human subjects using the same inputs u (Iglesias et al., 2012). The remaining model parameters were held constant (ω=-4, ϑ=0.0025).

This whole procedure was repeated for each of four different methods of performing step 4. The four methods were

a. The Nelder-Mead simplex algorithm (NMSA),

b. Gaussian process-based global optimization (GPGO),

c. Variational Bayes (VB),

d. Markov Chain Monte Carlo estimation (MCMC).

Owing to the computational burden imposed by MCMC, only 100 estimations instead of 1'000 were performed in its case.

κ was chosen as the perceptual parameter to vary because of the interesting effects it has on the nature of the inferential process (cf. Section 4.5). The response parameter ζ was chosen as the second parameter to vary because it represents inverse response noise, with the expectation that the lower ζ was, the more difficult the other parameters would be to estimate.

I will now briefly discuss each of the optimization algorithms employed, before I turn to the results. The discussion of the algorithms will be in terms of minimization, as is customary. Note that every maximization problem can be turned into a minimization problem by taking the negative of the objective function. In our case, maximizing the log-joint $Z(\xi|u,y)$ is equivalent to minimizing the negative log-joint $-Z$.

7.4.2 THE NELDER-MEAD SIMPLEX ALGORITHM

The *Nelder-Mead simplex algorithm* (Nelder and Mead, 1965) is a simple and very popular (15'172 citations of the original paper according to Google Scholar) local optimization algorithm. It is based on the geometric concept of a simplex. A simplex is the convex hull of a set of $d+1$ points in d-dimensional space. It is non-degenerate if and only if its volume is positive. For example, in $\mathbb{R}^2$ a simplex is a triangle; in $\mathbb{R}^3$ it is a tetrahedron, etc.

Given an initial non-degenerate simplex, the NMSA moves downhill by replacing that point of the simplex where the objective function is largest – leading to the next simplex. It does this by reflecting the point to be replaced through the opposing face and choosing the new point on the line defined by the original point and its reflection, according to appropriate criteria, or by contracting the whole simplex toward the point where the objective function is smallest. This is repeated until convergence. We used the NMSA as implemented in the `fminsearch` function in Matlab.

7.4.3 GAUSSIAN-PROCESS BASED GLOBAL OPTIMIZATION

The goal of Gaussian-process based global optimization (GPGO) (Lomakina et al., 2012) is to approximate complicated objective functions by an easy-to-evaluate function modeled as a Gaussian process. A Gaussian process is defined as a process any finite set of observations of which is distributed according to a multivariate Gaussian distribution with a covariance matrix consisting of a pairwise kernel function of corresponding data points. Thus, when modeling a function with a Gaussian process we force it to be smooth. In other words, we force those data points which are similar to each other in terms of a specific kernel function to have similar objective function values. For any as yet unevaluated point, given a set of

observations, we can estimate not only the most likely value of the output but the whole distribution for the value, which is Gaussian. The routine that describes how a Gaussian process approximates an objective function is illustrated in Figure 16.

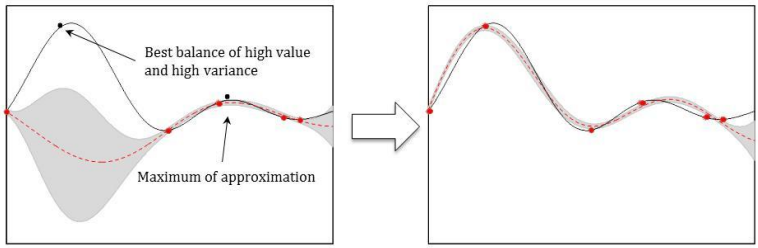

Figure 16. ***Illustration of GPGO****. Left: the optimization (in this illustration, maximization) routine using Gaussian processes is the following: we evaluate the objective function at a set of initial points (red dots) and use these values to approximate the objective function by a Gaussian process (red dashed line). Next, we find the new point which has the best balance, by some criterion, of estimated mean and variance. Since this approach preferentially samples both regions of high value and regions of high uncertainty, it allows for exploration along with exploitation. Right: the objective function is evaluated at the new point, leading to a new approximation. We repeat this procedure until appropriate stopping criteria are satisfied. As a result of this routine we obtain not only the maximum but also a reasonable approximation of the whole objective function.*

Given an initial set of points, this algorithm evaluates the objective function at these points and produces, based on these evaluations, a global estimate of the objective function, consisting of the mean and variance of a Gaussian distribution at each point. It then chooses the next point at which to evaluate the objective function by balancing the competing demands of exploitation (search where the current estimate is minimal) and exploration (search where the current estimate is most uncertain). New evaluation points are added until 110 points have been evaluated. A global estimate of the objective function, including its minimum, is then returned.

7.4.4 *Variational Bayes*

Variational Bayes (VB) is described in detail in Appendices A and B. For our simulation study, we used VB as implemented in Jean Daunizeau's DAVB toolbox, available at http://goo.gl/As8p7 (Daunizeau et al., 2009). In addition to the MAP estimate, this method also furnishes an estimate of the parameter covariance at that point, and thus, under Gaussian assumptions, an estimate of the full posterior distribution.

7.4.5 *Markov Chain Monte Carlo sampling*

Markov chain Monte Carlo (MCMC) sampling leads to a sample of points in parameter space whose distribution converges, in infinite time, to the target distribution (Gelman et al., 2003, 283ff). For this reason, MCMC was the standard against which we compared the other methods: if a method performed as well as MCMC without imposing a comparable computational burden, we would see this as a reason to prefer it and feel confident in its results.

To sample from the posterior over all parameters, we used Gibbs sampling with a one-dimensional Metropolis step for each of the parameters (cf. Gelman et al., 2003, 292). For each of the 100 simulation runs that we ran at each point on our parameter grid, we used one chain with a length of 500'000 samples and a burn-in period of 25'000 samples.

7.4.6 *Results*

All methods could reliably distinguish different values of κ with low or moderate decision noise (Figure 17). Unsurprisingly, this became harder for high levels of noise. The noise level itself could also be determined by all four methods (Figure 18). Figure 19A shows the root mean squared error in κ and $\log(\zeta)$, jointly for all values of κ. Figure 19B displays the accuracy of the confidence with which VB and MCMC make their estimates.

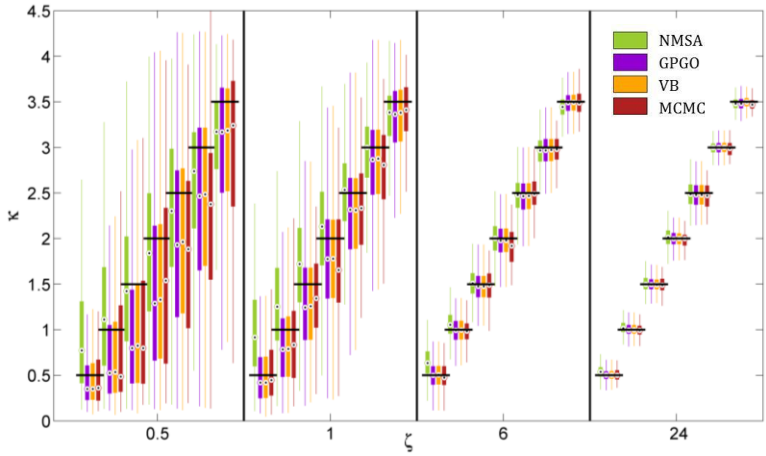

Figure 17. ***Estimation of coupling κ by four methods at different noise levels ζ.*** *A range of κ from 0.5 to 3.5 was chosen based on the range of estimates observed in the analysis of experimental data. Decision noise levels were chosen in a range from very high (0.5) to very low (24). The remaining model parameters were held constant (ω=-4, ϑ=0.0025). For each point of the resulting two-dimensional grid, 1000 task runs with 320 decisions each were simulated. Given the fixed sequence of inputs and simulated sequence of decisions, we then attempted to recover the model parameters, including κ and ζ, by four estimation methods: (1) the function Nelder-Mead simplex algorithm (NMSA), (2) Bayesian global optimization based on Gaussian processes (GPGO), (4) variational Bayes (VB), and Markov chain Monte Carlo sampling (MCMC). The figure shows boxplots of the distributions of the maximum-a-posteriori (MAP) point estimates for the four methods at each grid point. Horizontal shifts within boxes are for readability. Black bars indicate ground truth. The results show that all methods can recover the ground truth reliably when decision noise is moderate to low (ζ=6,24). At higher noise levels, estimates become less reliable. With GP, VB, and MCMC, they then exhibit a tendency to underestimate κ, while FMIN tends to mid-range values. Nonetheless, substantial differences in κ wthin the range observed experimentally will be detected by all four methods even at high levels of noise.*

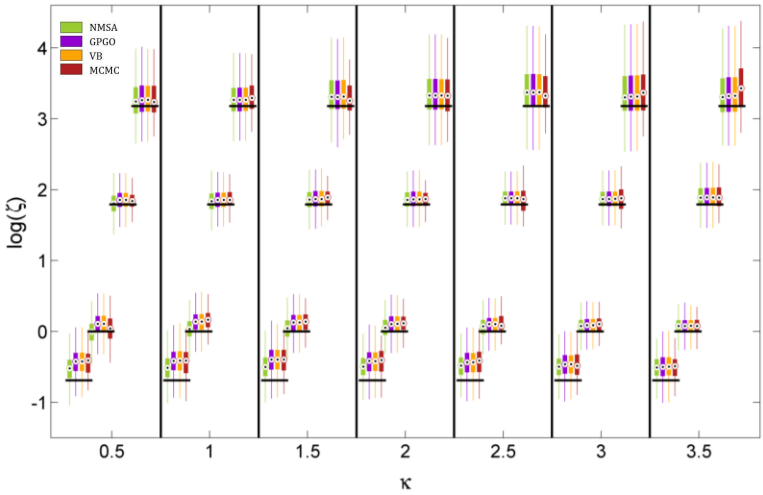

Figure 18. ***Estimation of noise level ζ at different levels of coupling κ.*** *ζ is estimated and displayed here at the logarithmic scale because it has a natural lower bound at 0. See Figure 17 for key to legend. The figure shows boxplots of the distributions of the maximum-a-posteriori (MAP) point estimates for the four methods at each point of the simulation grid. Horizontal shifts within boxes are for readability. Black bars indicate ground truth. The methods do not differ appreciably in their performance. They all tend to underestimate the noise level owing to a mild shrinkage (i.e., large ζ) prior, and errors are smaller for moderate noise levels, increasing for both high and low noise (cf. Figure 19, A2).*

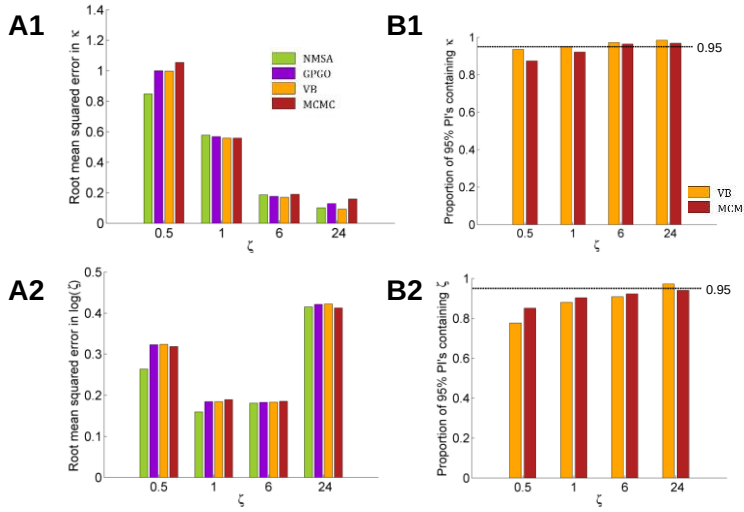

Figure 19. ***Quantitative assessment of parameter estimation.*** *(A) Root mean squared error of MAP estimates by noise level ζ for all four estimation methods (see Figure 17 for key to legends). (A1) Estimates for κ improve with decreasing noise and do not exhibit substantially significant differences between methods although NMSA is somewhat better at very high noise. (A2) As in Figure 18, estimates for ζ were assessed at the logarithmic scale. The results show that the noise level can best be estimated at moderate levels, where in fact most estimates of experimental data are found. Again, the methods perform comparably well, with NMSA best at high noise. (B) VB and MCMC estimate the whole posterior distribution. Parameter estimates can therefore not only be summarized as point estimates, but also as posterior intervals (PI's). If an estimation method were neither over- nor underconfident, 95% of 95% PI's would contain the true parameter value. If the proportion is less than 0.95, this indicates overconfidence; if it is greater than 0.95, underconfidence. (B1) Both methods are realistically confident about their inference on κ across noise levels, with a slight tendency towards overconfidence with higher noise. (B2) This tendency is more pronounced with estimates of ζ.*

7.4.7 DISCUSSION

The results show that all methods can recover the ground truth reliably when decision noise is moderate to low (ζ=6,24). Substantial differences in κ wthin the range observed experimentally will be detected by all four methods even at high levels of noise.

Two of the methods tested have serious drawbacks: the NMSA only comes up with a MAP estimate without giving any indication about the shape of the posterior; MCMC sampling imposes a heavy computational burden. I therefore see a clear advantage for GPGO and VB since they are computationally relatively efficient and furnish an estimate of the full posterior. I have implemented VB in the HGF toolbox (cf. Chapter 8), and the addition of GPGO is planned for the near future.

7.5 Optimization of parameters

It is often desirable to optimize parameters in a broader sense than fitting them, under an appropriate response model, to observed responses. The question such an optimization seeks to answer is: what parameter setting would allow an agent to make optimal predictions? To answer this question, we need to determine the most probable parameter setting given the inputs, that is we have to maximize $p(\xi|u)$.

Adapting Eq. (103), we derive this as

$$\begin{aligned}\xi^* &= \arg\max_{\xi} p(\xi|u) = \arg\max_{\xi} \frac{p(\xi,u)}{p(u)} \\ &= \arg\max_{\xi} p(\xi,u) = \arg\max_{\xi} \ln p(\xi,u) \\ &= \arg\max_{\xi} (\ln p(u|\xi) + \ln p(\xi)) \\ &= \arg\max_{\xi} \left(\sum_k \ln p\left(u^{(k)} \middle| \xi, u^{(1\dots k-1)}\right) + \ln p(\xi) \right) \\ &= \arg\max_{\xi} \left(\sum_k \ln p\left(u^{(k)} \middle| \hat{\mu}_1^{(k)}, \hat{\sigma}_1^{(k)}\right) + \ln p(\xi) \right).\end{aligned} \tag{112}$$

Since $-\ln p\left(u^{(k)} \middle| \hat{\mu}_1^{(k)}, \hat{\sigma}_1^{(k)}\right)$ is the surprise at seeing $u^{(k)}$ given the prediction $\hat{\mu}_1^{(k)}$ and its variance $\hat{\sigma}_1^{(k)}$, and $-\ln p(\xi)$ is the surprise at seeing a particular set of parameter values given the prior, ξ^* can be interpreted as the parameter setting that minimizes the agent's surprise.

7.6 Model comparison

Model comparison calls for a Bayesian approach because classical significance testing does not offer any generic framework for this (Kass and Raftery, 1995). The theory of Bayesian model comparison was developed by Jeffreys (1935, 1961) and involves as its central concept the *Bayes factor;* this is the factor by which the posterior odds of a model over another differ from the prior odds:

$$\frac{p(m_1|y)}{p(m_0|y)} = B \cdot \frac{p(m_1)}{p(m_0)}, \tag{113}$$

where m_1 and m_0 are the models being compared, y are the data observed, and B is the Bayes factor. By application of Bayes' theorem, one sees that

$$B = \frac{p(y|m_1)}{p(y|m_0)}. \tag{114}$$

It is convenient to interpret B in logarithmic space, where, according to the convention introduced by (Kass and Raftery, 1995), $\ln B$ indicates the following weight of evidence in favor of m_1:

$\ln B$	Evidence for m_1
0 to1	Barely mentionable
1 to 3	Positive
3 to 5	Strong
> 5	Very strong

To calculate B, we need to know $p(y|m_1)$ and $p(y|m_0)$. These are the *marginal likelihoods,* or *model evidences,* of m_1 and m_2, respectively. If ξ is the set of parameters of the model m, consisting of the likelihood $p(y|\xi, m)$ and the prior $p(x|m)$, then the model evidence of m is

$$p(y|m) = \int p(y|\xi, m)\, p(\xi|m) \mathrm{d}\xi. \tag{115}$$

This integral is often - when m involves a variant of the HGF, always - intractable. We can, however, derive a lower bound $\mathcal{F}$, the *negative free energy,* of its logarithm:

$$\begin{aligned}\ln p(y|m) \leq \mathcal{F} &\stackrel{\text{def}}{=} \int q(\xi) \underbrace{\ln p(y,\xi|m)}_{\stackrel{\text{def}}{=}\mathcal{L}(\xi)} \mathrm{d}\xi \underbrace{- \int q(\xi) \ln q(\xi) \mathrm{d}\xi}_{\stackrel{\text{def}}{=}S[q]} \\ &= \langle \mathcal{L}(\xi) \rangle_q + S[q]\end{aligned} \tag{116}$$

where $q(\xi)$ is an arbitrary probability density and $\langle \cdot \rangle_q$ is the expectation under q. As its physical analogon, the negative free energy is the sum the expected negative energy ($\mathcal{L}(\xi)$ is the negative Lagrangian) and the entropy S.

As I show in Appendix A, maximizing $\mathcal{F}$ (thereby approximating $\ln p(y|m)$ by it) is equivalent to approximating the posterior $p(\xi|y,m)$ by $q(\xi)$. In Appendix B I show how the q that maximizes $\mathcal{F}$ can be found by variational calculus. If we constrain q to be Gaussian, $S[q]$ can be calculated analytically, but $\langle \mathcal{L}(\xi) \rangle_q$ is still intractable. We can, however, deal with this by introducing the Laplace approximation, that is expanding $\mathcal{L}(\xi)$ around its mode to second order.

Conveniently, we already know the mode: it is ξ^* from Eq. (103), whose expression $\ln p(\xi, y|u)$ is equivalent to $\mathcal{L}(\xi) = \ln p(y, \xi|m)$, apart from the purely notational difference that instead of the model m the inputs u are made explicit. Besides the mode itself, we need to know the Hessian $\partial^2 \mathcal{L}(\xi^*)$ of $\mathcal{L}$ at the mode for our power expansion (the gradient, of course, vanishes at the mode). In Appendix D, I show that the covariance matrix Σ of the parameters ξ at the mode is negative inverse of the Hessian.

$$\Sigma = -\partial^2 \mathcal{L}(\xi^*)^{-1} \tag{117}$$

I further show there that, under the Laplace approximation, the maximum of the negative free energy can be calculated as

$$\mathcal{F} \approx \mathcal{L}(\xi^*) + \frac{1}{2}\ln|\Sigma| + \frac{d}{2}\ln(2\pi), \tag{118}$$

where $|\Sigma|$ is the determinant of Σ and d is the number of parameters in ξ. Since this approximates $\mathcal{F}$, which in turn approximates the model evidence $\ln p(y|m)$, we can now calculate the log-Bayes factor of m_1 over m_0 as

$$\ln B = \ln\frac{p(y|m_1)}{p(y|m_0)} = \ln p(y|m_1) - \ln p(y|m_0) \approx \mathcal{F}_1 - \mathcal{F}_0$$
$$B \approx \exp(\mathcal{F}_1 - \mathcal{F}_0) \tag{119}$$

This Bayesian framework of model comparison can be extended to comparisons of multiple models, and also to group studies (Stephan et al., 2009b). I give an example of this in Chapter 9.

8 THE HGF TOOLBOX

8.1 Overview

The HGF toolbox is a collection of Matlab functions that implement the models and estimation strategies described in the preceding chapters. It is an integral part of this thesis and can be downloaded from http://www.tnu-zurich.com/topics/software/ or http://goo.gl/Wv0zq. In addition to the HGF and the response models discussed above, it contains models such as the Rescorla-Wagner learning model for binary inputs and the softmax decision rule. For a given application, it may be necessary to extend the functions offered by the toolbox to suit the requirements of that case. However, the toolbox's modular nature should make it easy to add new models (see Section 8.5 below).

The HGF toolbox assumes a framework where an agent in the broadest sense (e.g., a human being, an animal, a machine, the stock market, etc.) receives a time series of inputs to which it reacts by emitting a time series of responses. In particular, this process is modeled by the combination of a perceptual (sc. state space) and an observation (sc. decision or response) model. The perceptual model is the time series model on which the agent bases its responses; the observation model describes how the agent makes decisions based on its perceptual inference. Note that what I refer to here as the observation model describes a "second-order" observation in the sense that the perceptual model already contains a ("first-order") observation part that describes how perceptual states relate to inputs. This implements the "observing the observer" framework described Daunizeau et al. (2010b).

8.2 Usage

There are two main ways to use the HGF toolbox:

1. To fit various combinations of perceptual and observation models to observed responses. In cases where optimal (by some

criterion) parameters are estimated using empirical Bayes, responses may simply be omitted.

2. To simulate the trajectories of perceptual states, and responses. In simpler cases (e.g., when simply filtering inputs), only the evolution of the perceptual inference is of interest. Responses and the specification of an observation model may then simply be omitted.

Currently, the following perceptual models are implemented in the HGF toolbox:

Perceptual model	Description
hgf	The generic HGF for continuous inputs, with or without sensory uncertainty
hgf_binary	The 3-level HGF for binary inputs, without sensory uncertainty
rw_binary	The Rescorla-Wagner model for binary inputs

The list of response models is:

Response model	Description	Compatible with
gaussian_obs	Gaussian noise on responses on a continuous scale	Hgf
empirical_bayes	Estimation of parameter values that lead to least surprise at the inputs (on a continuous scale) for the agent	Hgf
unitsq_sgm	The unit square sigmoid binary response model (cf Eq. (85))	hgf_binary rw_binary

softmax_binary	The softmax response model for binary responses	hgf_binary rw_binary
empirical_bayes_binary	Estimation of parameter values that lead to least surprise at the inputs (binary) for the agent	hgf_binary rw_binary

The toolbox uses the BFGS quasi-Newton optimization algorithm due to Broyden (1970), Fletcher (1970), Goldfarb (1970), and Shanno (1970) to implement the optimization in Eqs (103) and (112).

8.3 Installation and main functions

To install the toolbox, simply move the contents of the zip-file hgfToolBox.zip to a location of your choice.

Each of the two usages has its main function. The function

```
fitModel(...)
```

fits models to observed responses, while the function

```
simModel(...)
```

simulates the trajectories of perceptual states, and responses. The documentation to these functions is located at the top of their respcective files fitModel.m and simModel.m.

8.4 Documentation, configuration, and examples

Documentation is given in the toolbox's README file and in extensive comments at the beginning of the main functions and configuration files. To give a few examples of how one can use the toolbox, the remainder of this section has the form of a short tutorial:

Start Matlab, open the files fitModel.m or simModel.m, and read the documentation there. This will point you to the relevant configuration files.

As a simple example, start Matlab and load the example binary inputs provided in the file example_binary_input.txt:

```
>> u = load('example_binary_input.txt');
```

Now use this to generate simulated responses:

```
>> sim = simModel(u, 'hgf_binary', [0 1 1 1 1 -2.5 0.01],
'unitsq_sgm', 5);

>> hgf_binary_plotTraj(sim)
```

Next, try to recover these parameters by fitting the corresponding models to the simulated data:

```
>> est = fitModel(sim.y, sim.u, 'hgf_binary_config',
'unitsq_sgm_config', 'quasinewton_optim_config');
>> fit_plotCorr(est)
>> hgf_binary_plotTraj(est)
```

You can also try to fit the same data using a different perceptual model:

```
>> est2 = fitModel(sim.y, sim.u, 'rw_binary_config',
'unitsq_sgm_config', 'quasinewton_optim_config');
>> fit_plotCorr(est2)
>> rw_binary_plotTraj(est2)
```

The same procedure can be applied to continuous data. The file example_usdchf.txt contains the value of the US dollar in Swiss francs throughout much of 2010 and 2011 (source: http://www.oanda.com).

```
>> usdchf = load('example_usdchf.txt');
>> sim2 = simModel(usdchf, 'hgf', [1.04 0 0.0001 1 1 -12
0.3 0.0003], 'gaussian_obs', 1e-2);
>> hgf_plotTraj(sim2)
>> sim3 = simModel(usdchf, 'hgf', [1.04 0 0 0.0001 1 1 1 1
-12 -2 0.3 0.0001], 'gaussian_obs', 1e-2);
```

```
>> hgf_plotTraj(sim3)
>>    est3    =    fitModel(sim2.y,    sim2.u,    'hgf_config',
'gaussian_obs_config', 'quasinewton_optim_config');
>> fit_plotCorr(est3)
>> hgf_plotTraj(est3)
```

8.5 Adding models or optimization algorithms

The modularity of the toolbox enables you to add perceptual and observation models of your choice. This requires the following functions that fitModel(...) and simModel(...) will expect to find (replace <modelname> by the name of your model):

<modelname>	contains the model machinery
<modelname>_config	contains the configuration settings (only for fitModel(...))
<modelname>_transp	transforms parameters from the space they are estimated in to their native space (only for fitModel(...))
<modelname>_namep	returns a structure of named parameters (only for simModel(...))

Additionally, for observation models, simModel(...) expects to find a function that performs the simulation of responses:

<modelname>_sim

For details, look at the corresponding files of an existing model (e.g. hgf_binary) and use them as templates.

To add a new optimization algorithm, provide the following functions that fitModel(...) will expect to find (replace <algo> by the name of your algorithm):

<algo>	contains the machinery of your algorithm

<algo>_config contains the configuration settings

For details, look at the corresponding files of an existing algorithm (e.g. quasinewton_optim) and use them as templates.

9 APPLICATION TO SACCADIC REACTION TIMES

9.1 Overview

Inferring the environment's statistical structure and adapting behaviour accordingly is a fundamental *modus operandi* of the brain. A simple form of this faculty based on spatial orienting of attention can be studied with Posner's location-cueing paradigm in which a cue indicates the location of a target with a known probability. In a first experimental application of the HGF, we focused on a more complex version of this task, where probabilistic context (the proportion of valid cues) changes unpredictably over time, thereby creating a volatile environment. Saccadic response speeds were recorded in 15 subjects and used to estimate the subject-specific parameters of the HGF (i.e., the simple example model that we used throughout Chapters 1-5) that modeled the subjects' trial-by-trial updates of beliefs. Different response models – specifying how computational states translate into observable behavior – were compared using Bayesian model selection as described in Section 7.6. Saccadic response speed was most plausibly explained as a function of the precision of the belief about the causes of sensory input. This finding is in accordance with current Bayesian theories of brain function, and specifically with the proposal that spatial attention is mediated by a precision-dependent gain modulation of sensory input. Our results therefore provide empirical support for precision-dependent changes in beliefs about saccade target locations and motivate future neuroimaging and neuropharmacological studies of how Bayesian inference may determine spatial attention.

9.2 Introduction

Prior beliefs about the location of a behaviorally relevant stimulus facilitate stimulus detection and speed up reaction times (RTs). One of the first experimental demonstrations of this effect was provided by Posner's

location-cueing paradigm (Posner, 1980). In this task, a spatial cue (e.g., an arrow) indicates the most likely position of a behaviorally relevant target stimulus on a trial-by-trial basis. Average RTs are faster on valid trials – where the target appears at the expected or cued location – than on invalid trials, where target location is unexpected. This reflects covert orienting of attention to the cued location in analogy to an attentional spotlight. Attentional orienting enhances information processing at the cued location at the expense of alternative (uncued) locations.

However, there is accumulating evidence that attentional orienting in response to the spatial cue is not an all-or-none phenomenon, but is critically affected by trial history and by the current probabilistic context. For example, RT costs of invalid cueing are larger after a valid than after an invalid trial (Jongen and Smulders, 2007) – and RTs to invalid targets increase with the number of preceding valid trials (Vossel et al., 2011). Moreover, the RT difference between invalid and valid trials increases, the higher the proportion of validly cued trials (percentage of cue validity, %CV; Eriksen and Yeh, 1985; Giessing et al., 2006; Jonides, 1980; Risko and Stolz, 2010). These results imply that subjects infer and predict the current probabilistic context and adjust their behavior accordingly.

The behavioral effects observed in Posner's location-cueing paradigm can be interpreted within recent theoretical frameworks of perception and attention based on Bayesian principles (Chikkerur et al., 2010; Friston, 2009, 2010; Feldman and Friston, 2010; Itti and Baldi, 2009; Rao, 2005). Here, the brain is considered as a Bayesian inference machine (e.g., Dayan et al., 1995; Friston, 2009) which maintains and updates a generative model of its sensory inputs. In other words, perception can be framed as an 'inverse problem': under a specific generative model, the current state of the world has to be inferred from the noisy signals conveyed by the sensorium. Notably, even when stimuli are presented with very high signal-to-noise, there are many aspects about the state of the world (i.e., the cause of sensory inputs) that are non-trivial to infer, such as its probabilistic structure (the "laws" that relate causes of stimuli to each other) or non-linear interactions among causes (e.g., visual occlusion). The overall goal of this architecture is to minimize surprise about sensory inputs and thus underwrite homeostasis – either by updating model-based

predictions or by eliciting actions to sample the world according to prior expectations. Notably, because surprise about sensory inputs cannot be evaluated directly, it has been proposed that perception and action optimize a free-energy bound on surprise (Friston et al., 2006; Friston, 2009, 2010). Based on this free-energy principle, simulations have demonstrated how spatially selective attention can be understood as a function of precision (confidence or inverse uncertainty) during perceptual inference: attentional selection serves to increase the precision of sensory channels, enabling faster responses to attended stimuli (Feldman and Friston, 2010). Physiologically, this attentional effect may be mediated by an increase in the synaptic gain of neuronal populations encoding prediction error. These populations are assumed to project to higher-level units in the visual hierarchy where faster changes in neuronal activity are engendered in the context of higher precision (for details, see Feldman and Friston, 2010).

An important aspect of Posner's location-cueing task relates to the trial-by-trial uncertainty about the predictive value of the spatial cue (i.e., the probability that the target appears at the cued location in a given trial) (cf. Yu and Dayan, 2005). This becomes particularly important in volatile environments, where the cue predicts the target location with varying probabilities over the course of the experiment – in other words, situations in which probabilistic context changes unpredictably over time. Here, the estimate (representation) of this probability – which we will operationalize in terms of percentage of cue validity, %CV – depends on the integration of information over past events.

A simple description of trial-by-trial learning of cue-target contingencies is provided by reinforcement learning models such as Rescorla-Wagner (Rescorla and Wagner, 1972). There, the update of the probability estimate (in our case, the probability that the target will appear in the cued hemifield) after completing a trial equals the product of a fixed learning rate and the prediction error (i.e., the difference between observed and predicted outcome). The learning rate determines the impact of the prediction error on the belief update and, at the same time, determines to which extent the current belief is affected by past events, since it

determines the slope of the exponential decay function of the influence of previous trials (cf. Rushworth and Behrens, 2008).

While the Rescorla-Wagner rule describes a variety of human and animal behaviors, it is a heuristic approach that does not follow from principles of probability theory. Moreover, it suffers from some practical limitations which might be overcome by the application of Bayesian principles (Gershman and Niv, 2010). For associative learning paradigms, hierarchical Bayesian learning models provide a principled prescription of how beliefs are updated optimally in the presence of new data. These models may provide a more plausible account of behavior than the Rescorla-Wagner rule, particularly in volatile environments where a fixed learning rate is suboptimal (Behrens et al., 2007; den Ouden et al., 2010).

The HGF now grandfathers and extends existing normative models (Mathys et al., 2011). To repeat its main features (cf. Chapters 1-5), the HGF results in analytical update equations that (i) minimize free-energy, (ii) are extremely fast to evaluate, (iii) contain parameters allowing for individual differences in learning, and (iv) directly express the crucial role of prediction errors (and their weighting by uncertainty) that play such a prominent role in predictive coding schemes based on the free-energy principle described above. Crucially, the HGF can be applied to empirical behavioral data, allowing one to compare different models of subject responses and quantify their trial-by-trial estimates of states of the environment that lead to sensory predictions, including the precision of these estimates (cf. Chapters 6 and 7). This enables formal tests of free-energy based accounts of attention using empirically observed behavior that complements simulation work (e.g., Feldman and Friston, 2010). In particular, it can be tested which quantities from the Bayesian learning model are most influential in determining response speed (RS). While one might hypothesize a relationship between precision and RS in the present attentional cueing task (or even more generally; see e.g., Whiteley and Sahani, 2008), other studies (employing different experimental paradigms) have shown that reaction times can be related to the (log) probability estimate per se (Anderson and Carpenter, 2006; Carpenter and Williams, 1995; Brodersen et al., 2008; den Ouden et al., 2010), or to the amount of surprise that is associated with a particular stimulus (Bestmann

et al., 2008). Here, we try to explain the observed RS under these different assumptions (in contrast to RTs, RS tends to have a Gaussian distribution (Brodersen et al., 2008; Carpenter and Williams, 1995) and are thus a preferred behavioral quantity for modeling). To this end, we formulate competing models that embody the different notions above and formally compare their evidence, using Bayesian model selection.

In particular, we here apply this hierarchal Bayesian learning model to saccadic RS data from a variant of Posner's location-cueing paradigm with changes of probabilistic context (%CV) that are unknown to the subject. Saccadic eye movements and covert spatial attention are closely linked and share a common functional neuroanatomy (Beauchamp et al., 2001; Corbetta et al., 1998; de Haan et al., 2008; Nobre et al., 2000; Perry and Zeki, 2000). There is very strong evidence that eye movements to a given location are inevitably preceded by covert attention shifts to this location enhancing local perceptual processing (e.g., Deubel and Schneider, 1996; Deubel, 2008; Doré-Mazars et al., 2004; Godijn and Theeuwes, 2003). The 'premotor theory of attention' (Rizzolatti et al., 1987) states that attentional orienting may be functionally equivalent to saccade planning and initiation, and that therefore programming a saccade *causes* a shift of spatial attention. In a related theory, the 'Visual Attention Model' (Schneider, 1995), a single visual attention mechanism is proposed which controls both the selection for perception and the selection for action. Here, attention shifts are not caused by - but are a precondition for - saccade preparation (Deubel, 2008). The obligatory coupling between spatial attention and saccade programming is also evident in a recent computational model of evidence accumulation in the visuomotor cascade: visually responsive neurons which can be found in the frontal eye fields (FEF), the lateral intraparietal area (LIP) and superior colliculi (SC) provide the source of drive for motor neurons in FEF and SC to elicit a saccade (Schall et al., 2011).

Saccadic RS have been shown to be critically affected by the probability of the saccade target location (Carpenter and Williams, 1995; Chiau et al., 2011; Farrell et al., 2010) and there is initial evidence that trial-by-trial changes in saccadic RS reflect learning of probabilistic context according to Bayesian principles (Anderson and Carpenter, 2006; Brodersen et al.,

2008). Anderson and Carpenter (2006) presented two subjects with multiple trial blocks in which targets initially appeared to the left and right side of fixation with equal probability. After 70-120 trials in each block this probability could abruptly change, so that saccades were more likely to be made to one of the targets. Estimating an exponential decay parameter of the trial-by-trial probability of the target location, the authors showed that saccadic RS is related to the learned prior probability of target appearance. Brodersen et al. (2008) presented three subjects with different experimental blocks of left and right targets with different stochastic properties: the targets were either presented with different fixed probabilities, or the probability of the target location was conditional on the target location in the previous trial (1st order Markov sequence). They used two different learning models to examine each case whether the subjects learned and utilized the marginal probabilites of the targets' locations or their conditional probabilities (and thus a transition matrix).

While both studies (Anderson and Carpenter, 2006; Brodersen et al., 2008) started to address the question of *inter*-trial variability in probabilistic beliefs, they did not deal with the effects of the uncertainty (precision) of these beliefs which has been suggested as a prominent mechanism in spatial attention (Feldman and Friston, 2010). Moreover, both studies employed models that are agnostic about environmental volatility, thereby precluding the possibility that the subjects can adapt their learning rates based on their current belief about the stability of the environment.

Here we extend the previous findings in two ways. First, we show that trial-by-trial saccadic RS in the location-cueing paradigm can be explained as a function of the precision of trial-wise beliefs, as inferred using the HGF. Secondly, the HGF accommodates individual learning processes by introducing subject-specific parameters for coupling the different hierarchical levels of learning and thus provides a novel quantification of and explanation for individual learning differences, expressed by saccadic RS. In what follows, we will refer to the HGF simply as the *perceptual model*, since it provides a mapping from hidden states (or environmental causes) to sensory inputs. Furthermore, we will introduce and compare different *response models* (cf. Chapter 6 and Daunizeau et al. (2010b)) that describe the mapping from the subject's probabilistic representations

(beliefs) – as provided by the perceptual model – to the observed responses (i.e., response speed, RS).

9.3 Methods

9.3.1 SUBJECTS

Sixteen healthy subjects gave written informed consent to participate in the current study. One subject had to be excluded from further analysis due to lack of fixation during the cue-target interval. Therefore, data from fifteen subjects were analyzed (9 males, 6 females; age range from 23-35 years; mean age 27.4 years). All subjects were right-handed and had normal or corrected to normal vision. The study had been approved by the local ethics committee (University College London).

9.3.2 STIMULI AND EXPERIMENTAL PARADIGM

We used a location-cueing paradigm with central predictive cueing (Posner, 1980). Stimuli were presented on a 19-inch monitor (spatial resolution 1024 x 768 pixels, refresh rate 75Hz) with a viewing distance of 60cm. On each trial, two peripherally located boxes were shown (1.9° wide and 8° eccentric in each visual field, see Figure 20) that could contain target stimuli. A central diamond (0.65° eccentric in each visual field) was placed between them, serving as a fixation point. Cues comprised a 200ms increasing brightness of one side of the diamond – creating an arrowhead pointing to one of the peripheral boxes. After a 1200ms stimulus onset asynchrony, a target appeared for 100ms in one of the boxes. The targets were vertical and horizontal circular sinusoidal gratings (1.3° visual angle). Vertical and horizontal grating were presented with equal probability.

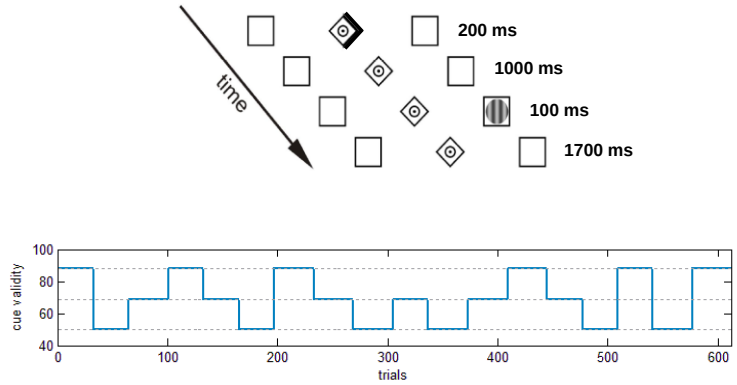

Figure 20. ***Task.*** *Illustration of the experimental task and the manipulation of %CV over the 612 trials.*

Subjects were instructed to maintain central fixation during the cue period and to make a saccade to the target stimulus as fast as possible. They were encouraged to blink and re-fixate the central fixation dot after the saccade. After a short practice session of 64 trials – with constant 88 %CV – the experiment comprised 612 trials with block-wise changes in %CV that were unknown to the subjects. Each block with constant %CV contained an equal number of left and right targets, counterbalanced across valid and invalid trials. %CV changed after either 32 or 36 trials switching unpredictably to levels of 88%, 69% or 50% (see Figure 20). Subjects were told in advance that there would be changes in %CV over the course of the experiment, but were not informed about the levels of these probabilities or when they would change. Each subject was presented with the same sequence of trials. This is a standard procedure in computational studies of learning processes that require inference on conditional probabilities in time series (cf. Behrens et al., 2008; Daunizeau et al., 2010b). In this common case, the parameters of the learning process depend on the exact sequence of trials used. Although this dependency will diminish asymptotically with increasing numbers of trials, for the relatively short

sequences (of a few hundred trials at best) that are feasible within a standard experiment, introducing a different sequence for each participant could increase the variability of parameter estimates, over and beyond the intrinsic inter-individual trait-differences per se. We therefore decided to keep this factor constant to ensure that differences in the model parameters can be attributed to subject-specific rather than task-specific factors. After half of the trials, the subjects had a short rest of 1 min.

9.3.3 *EYE MOVEMENT DATA RECORDING AND ANALYSIS*

Participants sat in a dimly lit sound-proof cabin with their head stabilized by a chinrest. Eye movements were recorded from the right eye with an EyeLink 1000 desktop mounted eye-tracker (SR Research Ltd.) with a sampling rate of 250Hz. A 9-point eye-tracker calibration and validation was performed at the start of the experiment and after the pause in the middle of the experiment. The validation error was less than 1° of visual angle.

Eye movement data were analyzed with MATLAB (Mathworks) and ILAB (Gitelman, 2002). Blinks were filtered out and pupil coordinates within a time window of 20 ms around the blink were removed. Trials with more than 20% missing data were discarded from the analyses. To ensure central fixation after presentation of the spatial cue, the time period between cue and target was analyzed for gaze deviations from the center. After target appearance, only the first saccade was analyzed. Saccades were identified when the eye velocity exceeded 30°/s (Stampe, 1993; Fischer et al., 1993). Once this threshold was reached, the start of the saccade was defined as the time when the velocity exceeded 15% of the trial-specific maximum velocity (Fischer et al., 1993). Likewise, the end of the saccade was defined by the time when the velocity fell below 15% of the trial-specific maximum velocity. Moreover, the saccade amplitude needed to subtend at least 2\3 of the distance between fixation point and the actual target location. Saccadic RT was defined as the latency between target and saccade onset. Saccades in which the starting position was not within a region of 1° from the fixation point and saccades with a latency <90ms were discarded from the analyses. Our analyses focused on inverse RTs (i.e., response speed, RS) since, in contrast to RTs, RS are normally distributed (cf. Brodersen et al., 2008; Carpenter and Williams, 1995).

To assess the effect of probabilistic context (true %CV), mean RS for each subject and for each %CV condition were entered into a 2 (*cue*: valid, invalid) x 3 (*%CV*: 50, 69, 88%) within-subjects ANOVA. In this analysis, evidence for an impact of probabilistic context would be reflected in a significant *cue* x *%CV* interaction effect – with increasing differences between valid and invalid RS with higher %CV. Results from this analysis are reported in the Results section at a significance level of $p<0.05$ after Greenhouse-Geisser correction. Condition-specific mean RS was also calculated separately for the two halves of the experiment and analyzed with a 2 (*cue*: valid, invalid) × 3 (*%CV*: 50, 69, 88%) × 2 (*time*: 1st half, 2nd half) within-subjects ANOVA (note that each %CV condition was presented 3 times in each half, cf. Figure 20).

Having established the significance of the experimental effects we then sought to model them in terms of hierarchical Bayesian updating.

9.3.4 PERCEPTUAL MODEL

In what follows, we briefly outline the particular variant of the HGF used in this study. The perceptual model (dark grey panel in Figure 21) here comprises a hierarchy of 3 hidden states (denoted by x), with states 2 and 3 evolving in time as Gaussian random walks.

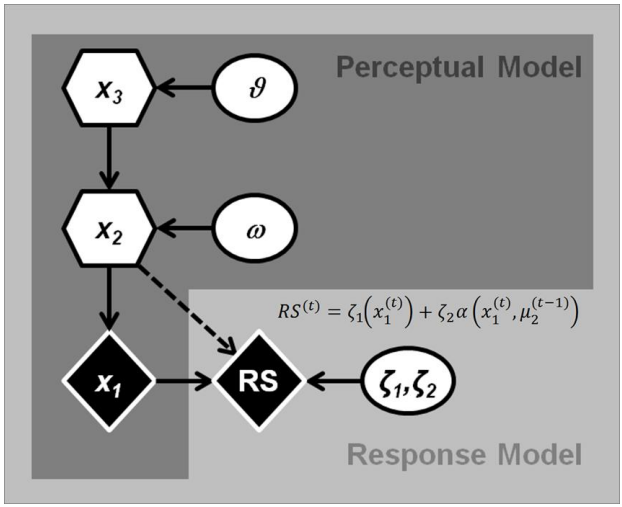

Figure 21. ***Models.*** *Graphical illustration of the perceptual (generative) model with states x_1, x_2 and x_3. The model parameters ω and ϑ impact on the time course of subjects' inferred belief about the states x and are estimated from the individual subject RS data. Circles represent constants, while diamonds represent quantities that change in time (i.e., that carry a time (or trial) index). Hexagons, like diamonds, represent quantities that change in time but that additionally depend on their previous state in time in a Markovian fashion.*

The probability of a target appearing at the cued location in a given trial (k) (represented by the state $x_1^{(k)}$, with $x_1 = 1$ for valid and $x_1 = 0$ for invalid targets) is governed by a state x_2 at the next level of the hierarchy. Hence, in the current location-cueing paradigm, x_2 determines the trial-specific estimate for %CV. Note that in this particular experiment, the target stimulus was visible without any ambiguity (very high signal-to-noise ratio); this means there is a simple deterministic mapping between the (mean of) x_1 and input u of the general model, which allows for situations with perceptual ambiguity (e.g., visual noise).

Since the response models used in this study did not involve the representations μ_3 and σ_3, we chose a scale on x_3 by fixing κ to 1 (cf. the

discussion of coordinate choice on higher levels in Section 7.1). Since κ is a multiplicative constant, this effectively removes it from the model.

As described in detail in Chapter 4, the precision-weighting of the updates at the second level can be understood as a time-varying learning rate, which varies with the state-dependent component μ_3 of the log-volatility. An alternative – but equally useful – perspective on the generic precision-weighted prediction error update scheme of Chapter 4 (particularly Eq. (69)) is in terms of Bayesian filtering, for example Kalman filtering. The Kalman filter can be regarded as an extension of the Rescorla-Wagner rule. It formalizes the predictive relationship between events, but also comprises expectations about how this relationship is expected to change over time and takes into account the uncertainty about this prediction (Dayan et al., 2000). Note, however, that the model underlying the Kalman filter is linear and therefore more limited than the generic non-linear model we use here. From this perspective, the precision–dependent weighting of prediction errors in our scheme correspond to the Kalman gain. These two perspectives (reinforcement learning rates and Kalman gain) illustrate the close relationship between reinforcement learning and Bayesian inference disclosed by the general scheme used here.

In addition to the full three-level HGF, we employed two reduced versions of the perceptual model. This was done to evaluate whether the relatively complex hierarchical model was truly needed to explain subjects' behavior. Specifically, the full hierarchical model assumes that (i) subjects are indeed capable of learning the hierarchical structure of the probabilities in this experiment, and (ii) exploit this knowledge to dynamically adapt the speed at which they update beliefs (i.e., learning rate) by using precision-weighted prediction errors. Although these assumptions are theoretically well-founded, it needs to be shown that equivalent explanations of the data could not be afforded by simpler, non-hierarchical learning models. Therefore, we specified two alternative perceptual Bayesian models which eschewed assumptions about hierarchically structured learning, but in different ways. The first alternative model assumed that subjects ignored the instructions that the environment was volatile, expecting negligible changes in log-volatility (third level): ϑ was thus fixed to zero, and only ω was estimated. The

second perceptual model did not exploit the estimates of environmental volatility for adapting learning optimally by precision-weighted prediction errors. In this model, the influence of x_3 on the variance of x_2 was therefore fixed to zero, so that levels 2 and 3 of the model became decoupled and rendered the values at the third level of the hierarchical model irrelevant (an equivalent effect is obtained by fixing $x_3^{(k)}$to zero).

9.3.5 *RESPONSE MODELS*

To map from the subject's posterior beliefs to observed responses, three different response models were compared. A detailed analysis and motivation of their functional forms can be found in Appendix E. All response models predict inverse RT (response speed, RS), since the distribution of RS is typically normal, in contrast to RTs themselves (Carpenter and Williams, 1995). Furthermore, all response models describe trial-wise RS as a linear function of an attentional factor α, based on the posterior beliefs of the perceptual model. This factor can be regarded as the proportion of attentional resources allocated to the cued location (i.e., α is normalized to the unit interval):

$$RS = \begin{cases} \zeta_{1\_valid} + \zeta_2 \alpha & \text{for } x_1{=}1 \text{ (i.e., valid trial)} \\ \zeta_{1\_invalid} + \zeta_2 (1-\alpha) & \text{for } x_1{=}0 \text{ (i.e., invalid trial)} \end{cases} \quad (120)$$

Note that in all cases, RS is the same function of attentional resources allocated to the outcome location: on valid trials, this is the amount of attentional resources α allocated to the cued location, while – on invalid trials – it is the amount of attentional resources $1-\alpha$ allocated to the uncued location (cf. Figure 22). Here, ζ_{1_valid}, $\zeta_{1\_invalid}$ and ζ_2 are subject-specific parameters that are estimated from the data. Minimal and maximal RS for valid and invalid trials are then defined by $\zeta_{1_{valid/invalid}}$ and $\zeta_{1valid/invalid} + \zeta_2$, respectively.

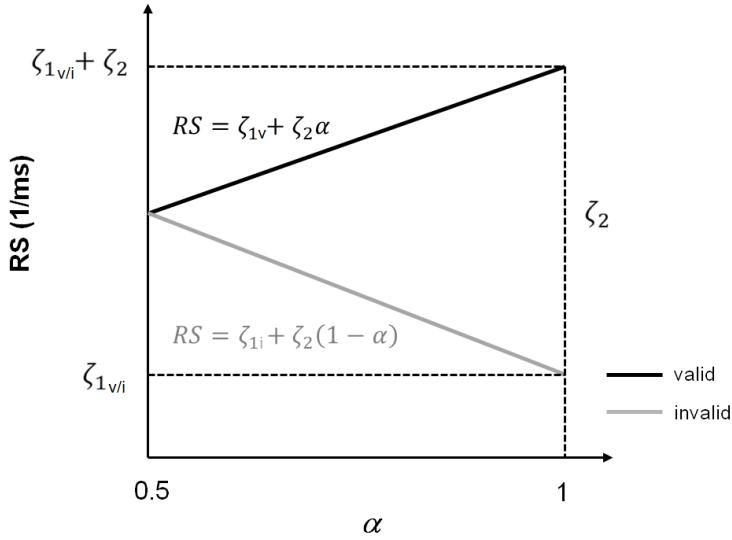

Figure 22. ***Reaction speed as a function of attention.*** *Illustration of the relationship between RS (inverse RT) and the quantity α, representing the amount of attentional resources allocated to the cued location. For each response model, RS were assumed to be linearly related to α (which differs between the three models, see Appendix). Note the opposite behaviour of RS for increasing α on valid (black line and equation) and invalid (grey line and equation) trials (cf. Eq. (120)).*

Crucially, the three competing response models differ in how they specify the dependence of α on computational quantities from the perceptual model: these are precision, belief, and surprise about the sensory signal, respectively. All three models respected the same boundary conditions, i.e., α remained confined to the unit interval with $\alpha = 0.5$ when $\hat{\mu}_1 = 0.5$ (cf. Appendix and Figure 9).

The first response model focused on the precision estimate at the first level of the perceptual model – following the recent proposal by Feldman and Friston (2010) concerning the role of precision for spatially selective

attention in the location-cueing paradigm. Here, we assumed that on a given trial k the attentional factor $\alpha^{(k)}$ was determined by a sigmoid transformation (s) of $\hat{\pi}_1^{(k)}$, the precision of the prediction at the first level, relative to its minimal value (i.e., 4 when $\hat{\mu}_1 = 0.5$):

$$\alpha^{(t)} = s\left(\hat{\pi}_1^{(t)} - 4\right) \tag{121}$$

In the second response model, the 'belief' model, the attentional factor α depended on the strength of the prediction about cue validity:

$$\alpha^{(t)} = \hat{\mu}_1^{(t)} = s\left(\mu_2^{(t-1)}\right) \tag{122}$$

The third response model ('surprise') was based upon the (Shannon) surprise associated with the target stimulus. The Shannon surprise (Shannon, 1948) is the negative logarithm of a probability estimate (here $\hat{\mu}_1^{(k)}$). This response model was inspired by a previous study on cueing of motor responses in which RTs were examined in relation to trial-wise surprise (Bestmann et al., 2008). Here, we defined α as a nonlinear function of Shannon surprise:

$$\alpha^{(k)} = \frac{1}{\left(1 + surprise\left(\hat{\mu}_1^{(k)}\right)\right)} \tag{123}$$

with $surprise\left(\hat{\mu}_1^{(k)}\right) = -\log_2 p\left(x_1^{(k)} = 1 \middle| \hat{\mu}_1^{(t)}\right) = -\log_2\left(\hat{\mu}_1^{(k)}\right)$

In summary, we specified three alternative perceptual models and three alternative response models. This resulted in a 3×3 factorial model space.

We compared the relative plausibility of these models using a random effects Bayesian model selection (BMS) procedure at the group level, both for individual models and model families (Penny et al., 2010; Stephan et al., 2009b). In addition, we compared these models to a standard Rescorla-Wagner learning model as well as to a model assuming that the true underlying (categorical) probabilities were known to subjects – in other words, they did not have to be inferred on the basis of experience. In the latter two models, trial-wise RS was supposed to be linearly related to the estimated or true %CV, respectively.

9.3.6 ESTIMATION OF THE MODEL PARAMETERS

The perceptual model parameters ω and ϑ, as well as the response model parameters ζ_{1_valid}, $\zeta_{1\_invalid}$ and ζ_2 were estimated from the trial-wise RS measures using variational Bayes as described in Chapter 7 and implemented in Jean Daunizeau's DAVB toolbox, available at http://goo.gl/As8p7 (Daunizeau et al., 2009). This approach is analogous to the Bayesian inversion of Dynamic Causal Models (DCM's) for functional imaging or electrophysiological data (Friston et al., 2003; Daunizeau et al., 2011).

As any Bayesian approach, variational Bayesian inversion requires the definition of priors on the parameters. Importantly, the prior (co)variance influences the estimability of parameters, e.g., their degree of independence; also by choosing a very small prior variance (very high prior precision) one can effectively fix the value of a parameter. Table 1 provides the priors used for inverting the full hierarchical model. In the perceptual model, initial values for μ and σ of states 2 and 3 were fixed and an upper bound of 1 was defined for the parameter ϑ. In the response model, the prior variance for ζ_2, which parameterizes the relationship between the attentional factor α and RS (Figure 22), was set to a fairly small value (10^{-3}). In other words, we assumed that the relation between RS and α (see equation 5) did not differ greatly across subjects. In contrast, to account for individual baseline differences in response speed (i.e., the intercept of the linear slope); the response model parameters ζ_{1_valid} and $\zeta_{1\_invalid}$ were given a larger prior variance, allowing for substantial individual differences between subjects.

While trials with missing responses did not contribute to parameter estimation, they did contribute to estimating the evolution of the states *x*, since they still provided the subject with an observation about the cue-target contingency. In other words, we used what the subject saw to estimate the Bayes optimal estimate of hidden states over the experiment – under a particular set of parameters and used subject responses to optimize the parameters of the perceptual and response models.

Table 1. ***Priors.*** *Prior mean and variance for the parameters of the perceptual and response models, and the noise parameter. ϑ is estimated in logit-space, while ζ_{1_valid}, $\zeta_{1\_invalid}$ and ζ_2 are estimated in log-space.*

Parameter	Prior mean	Prior variance
Perceptual Model		
ω	-6	100
ϑ	0.1	100
Response model		
ζ_{1_valid}	0.0052	0.1
$\zeta_{1\_invalid}$	0.0052	0.1
ζ_2	0.0006	0.001
Noise parameters		
ζ_3	0.001	1000

9.3.7 BAYESIAN MODEL SELECTION (BMS)

BMS is described in detail in Section 7.6. It evaluates the relative log-evidence (or log-marginal likelihood) of alternative models. The log-evidence of a model is the negative surprise about the data, given a model, and represents a generic trade-off between the accuracy and complexity of a model that can be derived from first principles of probability theory. Over the past decade, BMS has become a standard approach to assess the relative plausibility of competing models that describe how neurophysiological or behavioral responses are generated (cf. Daunizeau et al., 2010a, 2010b; Stephan et al., 2009b). Here, we use it to disambiguate different hypotheses about how learning (as described by the perceptual

models) and decision making (as described by the response models) evolve across and within trials.

Above, we introduced three perceptual models and three response models ('precision', 'belief' and 'surprise'). Combining these alternatives provides 9 models in a 3×3 factorial model space, plus the additional two control models (standard Rescorla-Wagner model and a model assuming that the true probabilities were known to the subjects). To assess the relative plausibility of our models at the group level, we used random effects BMS (Stephan et al., 2009b) and report both posterior probabilities and the exceedance probabilities of the competing models. Importantly, random effects BMS treats the model itself as being probabilistically distributed in the population (i.e., as a random variable following a Dirichlet distribution) using a full hierarchical model for multi-subject data. In brief, this enables group-level inference while taking into account inter-individual differences (e.g., the optimal model can vary across subjects). Critically, random effects BMS not only assesses the relative goodness of competing models but also quantifies (via the Dirichlet parameter estimates) the degree of heterogeneity in the sample studied (Stephan et al., 2009b).

The exceedance probability of a model is the probability that it is more likely than any other model considered, given the data. For example, an exceedance probability of 95% for a particular model means that one has 95% confidence that this model has a greater posterior probability than any other model tested (Stephan et al., 2009b). Both posterior probabilities and exceedance probabilities sum to unity over all models tested.

9.3.8 REPRODUCIBILITY OF RESULTS

To examine the replicability and hence generalizability of our findings, we performed an additional analysis, using an independent set of subjects (n=16, 8 males, 8 females: age range from 19-30 years; mean age 23.4 years). Again, all subjects were right-handed and had normal or corrected to normal vision. The subjects were tested as part of a separate psychopharmacological study employing a within-subject cross-over design. The data presented here were taken from the placebo session only,

during which the subjects received a multivitamin tablet. This study had been approved by the NHS Research Ethics Committee.

The subjects were presented with exactly the same trial sequence as the subjects in the original dataset. The within-trial structure was also almost identical, with slight modifications to the timing of the task: the cue-target SOA was reduced to 800ms and the target was presented for 200ms. Moreover, the trials were interspersed with 106 'null-trials' where only the baseline display (the fixation point and peripheral boxes) was shown. The task had a duration of 35 minutes and comprised 4 short rest periods. Finally, the subjects received a slightly longer training than the original group (one session with 100 trials with constant 80 %CV and one session with 121 trials with changes in %CV).

The same procedures and analyses as outlined above were applied to the eye movement data, except that the data here were recorded with a sampling rate of 1000Hz. Using trial-wise RS, we again fitted the parameters of the perceptual and response models as outlined above.

9.4 Results

9.4.1 FIXATION DURING THE CUE-TARGET INTERVAL AND MISSING TRIAL DATA

Between the appearance of the cue and the target, the subjects fixated the centre of the display in 87.7± 2.3% (mean ± SEM) of the trials – within a region of interest of 1° – and in 95.4 ± 1.2% of the trials within in a region of 2° from the fixation point. The proportion of trials with missing eye data or missing or incorrect saccades amounted to 20.0 ± 3%, so that on average 80% of the trials (487 of 612 trials) were analyzed. Trials needed to be excluded due to anticipated responses (3 ± 1%), incorrect or absent saccades (5 ± 1%), saccades not starting from the fixation zone (8 ± 1%), or missing data points, e.g., due to blinks (4 ± 1%). There was no significant difference in the percentage of correct trials between the first and second half of the experiment (paired t-test, p=0.895).

9.4.2 *CLASSICAL INFERENCE ABOUT THE EFFECTS OF PROBABILITY ON RS*

The 2 (*cue*: valid, invalid) × 3 (*%CV*: 50, 69, 88%) ANOVA on RS data revealed a significant main effect of cue ($F_{(1,14)}$=8.8, p=0.01) reflecting faster responses (higher RS) on valid than on invalid trials. The main effect of *%CV* was not significant – in other words, averaging over valid and invalid trials removed any effect of probability. Crucially, we observed a significant *cue* × *%CV* interaction effect ($F_{(1.9,26.6)}$=9.5, p=0.001) reflecting a differential impact of %CV on valid and invalid trials (Figure 23). A separate analysis also considered general trends in the data over time, e.g. due to fatigue, by including *time* (first vs. second half of the experiment) as additional factor. This resulted in a three-factorial *cue* (valid, invalid) × *%CV* (50, 69, 88%) × *time* (1st, 2nd half) ANOVA. Again, this analysis revealed a main effect of *cue* ($F_{(1,14)}$=8.2, p=0.013) and a significant *cue* × *%CV* interaction ($F_{(1.6, 22.5)}$=10.5, p=0.001). The main effect of *%CV* was not significant. Importantly, there was neither a significant main effect of *time* nor interaction effects of the factor *time* with any of the other factors (all p-values >0.4).

The *cue* × *%CV* interaction effect indicates a significant influence of probabilistic context on the subjects' responses, with stronger attentional orienting to the cue (and higher RT costs after invalid cueing) with higher %CV. However, Figure 23 does not show a strictly monotonic relationship between RS and true %CV for valid cues. This probably results from the fact that the underlying probabilistic structure (i.e., %CV) was unknown to the subjects and was changing in time fairly rapidly. It therefore had to be inferred by the subjects online, and these subject-specific and dynamic estimates should be the relevant predictors of observed RS, not %CV. In other words, the ANOVAs above (and the visualization in Figure 23) average across trials which are heterogeneous in terms of subjective probability estimates, and a model predicting the subjective estimates should be superior in explaining behavior (cf. Figure 27). In what follows, we test this, evaluating whether the empirically observed RS might reflect trial-by-trial updating of the subjects' expectancies according to our Bayesian perceptual model. Additionally, we compare a systematic set of models that combine different putative learning processes (perceptual

models) with different ways in which the learned quantities drive behavior (response models).

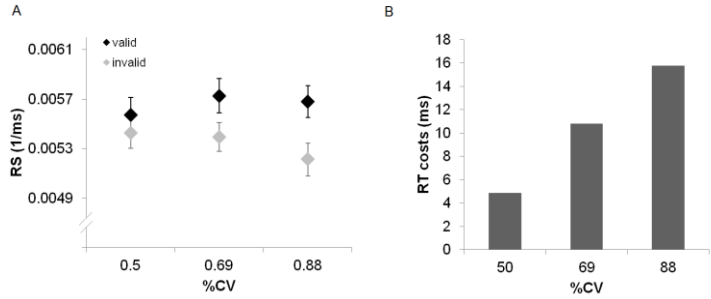

Figure 23. ***Reaction speed/time and %CV.*** *(A) Average RS in valid and invalid trials for the three (true) %CV levels. Error bars depict standard errors of the mean (SEM). (B) Illustration of how the observed RS costs after invalid cueing translate into RT differences (in ms).*

9.4.3 BAYESIAN MODEL SELECTION

Random effects BMS among the 3 perceptual model families (i.e., the full models and the two reduced model versions for each of the 3 response models) revealed that the full hierarchical Bayesian model had substantially higher model evidence than the two reduced (null) versions (Table 2).

Table 2. ***Results of the Bayesian model selection (BMS).*** *PP is posterior probability and XP is exceedance probability.*

	Main dataset (n=15)		**Replication dataset (n=16)**	
Model	**PP**	**XP**	**PP**	**XP**
Model family comparison – Perceptual Models				
Full hierarchical Bayesian family	.873	.999	.777	.997
Reduced model family (ϑ =0)	.064	<.001	.105	.001
Reduced model family ($x_3^{(k)}$=0)	.063	<.001	.118	.002
Model family comparison – Response Models				
'Precision' family	.756	.991	.642	.930
'Belief' family	.076	.001	.251	.066
'Surprise' family	.168	.008	.107	.004
Model comparison of all 11 models				
Full hierarchical Bayesian model 'Precision'	.499	.995	.381	.914
Reduced model (ϑ =0) 'Precision'	.006	<.001	.182	.074
Reduced model ($x_3^{(k)}$=0) 'Precision'	.119	.004	.047	<.001
Full hierarchical Bayesian model 'Belief'	.040	<.001	.041	<.001
Reduced model (ϑ =0) 'Belief'	.040	<.001	.042	<.001
Reduced model ($x_3^{(k)}$=0) 'Belief'	.040	<.001	.074	.004
Full hierarchical Bayesian model 'Surprise'	.040	<.001	.079	.004
Reduced model (ϑ =0) 'Surprise'	.040	<.001	.039	<.001
Reduced model ($x_3^{(k)}$=0) 'Surprise'	.040	<.001	.039	<.001
Rescorla-Wagner model	.040	<.001	.038	<.001
True categorical probability model	.040	<.001	.038	<.001

Comparing the 3 response model families (i.e., the precision, belief and surprise models for each of the 3 versions of the perceptual model) showed that the response model based upon precision was clearly superior to the 'belief' and the 'surprise' model (Table 2). Finally, comparison of all 11 individual models revealed that the full hierarchical Bayesian model combined with the precision response model was clearly

superior to all other models we considered (Table 2, supplementary Figure S1).

9.4.4 PARAMETERS OF THE WINNING MODEL

The subject-specific values for log-volatility ω and meta-volatility ϑ derived from the full hierarchical perceptual model – based upon precision – are depicted in Figure 24A. Figure 24B shows the minimal and maximal RS for each subject as derived from the response model parameters ζ_{1_valid}, $\zeta_{1\_invalid}$ and ζ_2 in relation to the subject's overall (mean) RS. The graph shows that there were considerable differences in the absolute speed of responding across subjects, as parameterized by averaged values for ζ_{1_valid} and $\zeta_{1\_invalid}$ which were estimated from the individual datasets.

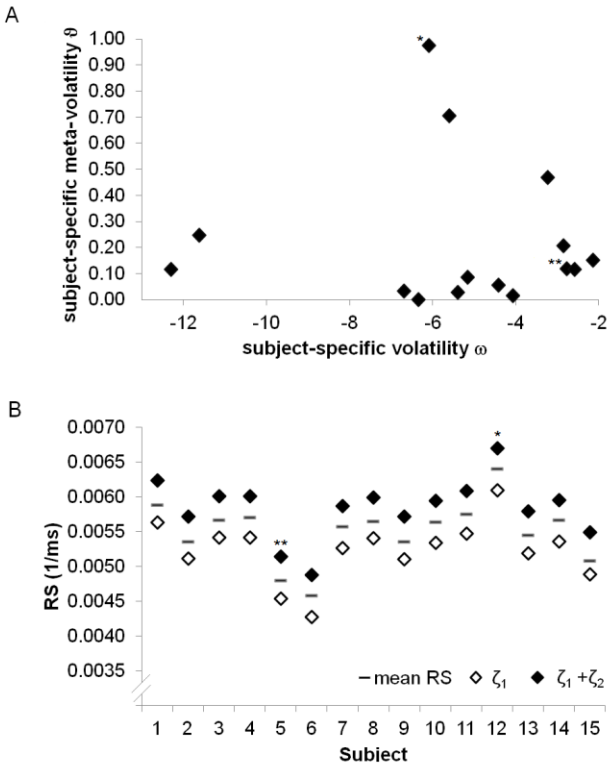

Figure 24. ***Subject-specific parameters.*** *(A) Illustration of the subject-specific patterns for the values of the volatility estimate ω and the meta-volatility estimate ϑ. (B) Illustration of minimal and maximal RS (as derived from the response model parameters ζ_1 (averaged for valid and invalid trials) and ζ_2) in relation to overall (mean) RS. The symbols * and ** denote the data from subjects A and B depicted in Figure 26, respectively.*

In our hierarchical Bayesian scheme, the precision-weighting $\hat{\pi}_1^{(t)}/\left(\hat{\pi}_2^{(t)}\hat{\pi}_1^{(t)}+1\right)$ (cf. Eq. (98)) at the second level plays the role of a

(time-varying) learning rate that depends on the log-volatility, determined by ω and μ_3. This dependence – on higher-order knowledge about change points in the environment – enables more adaptive learning in volatile environments, such as our paradigm. This is also reflected by the BMS results described above, where the hierarchical Bayesian model clearly outperformed a standard Rescorla-Wagner model with a fixed learning rate. However, given the formal similarity of the two models, one may expect to find a correlation between the fixed learning rate of the Rescorla-Wagner model and the parameters determining the learning rate of our hierarchical Bayesian model. Figure 25 depicts this relationship between the perceptual parameters ω and ϑ, and the learning rate ε derived from the Rescorla-Wagner model. While there was a significantly positive correlation between the subject-specific volatility estimate ω and learning rate ε (r=0.69; p=0.004), no relationship was observed between ε and the meta-volatility ϑ (p>0.25) (Figure 25).

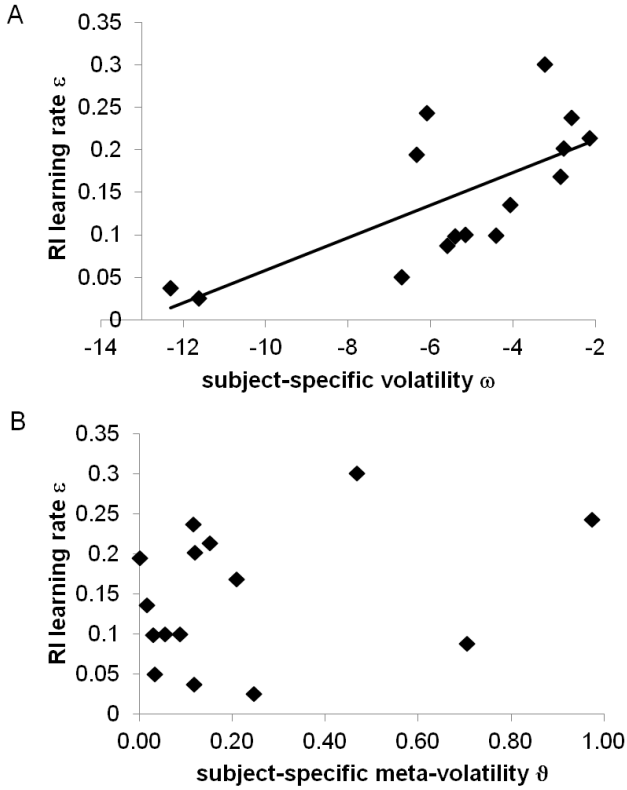

Figure 25. ***HGF – Rescorla-Wagner parameter comparison.*** *Relationship between the perceptual parameters ω and ϑ, and the Rescorla-Wagner learning rate ε.*

To illustrate different individual learning styles, Figure 26 shows the exemplary time courses of the third and first levels of the Bayesian model for two subjects with distinct updating behavior. The two subjects show differences in the volatility estimate ω as well as the meta-volatility estimate ϑ (cf. Figure 24 where these subjects are indicated by stars). Despite the meta-volatility estimate ϑ is higher in subject A than in subject B, subject B shows faster updating of beliefs due to a higher value of the

volatility estimate ω. In other words, our model indicates that the first subject perceives the environment as substantially less volatile than the second subject. Since the updates of $\mu_2^{(k)}$, the estimated cue validity, are coupled to the estimated log-volatility $\mu_3^{(k-1)}$, this translates into a higher learning rate and visibly more rapid updating behaviour in the second subject when the true underlying %CV changes.

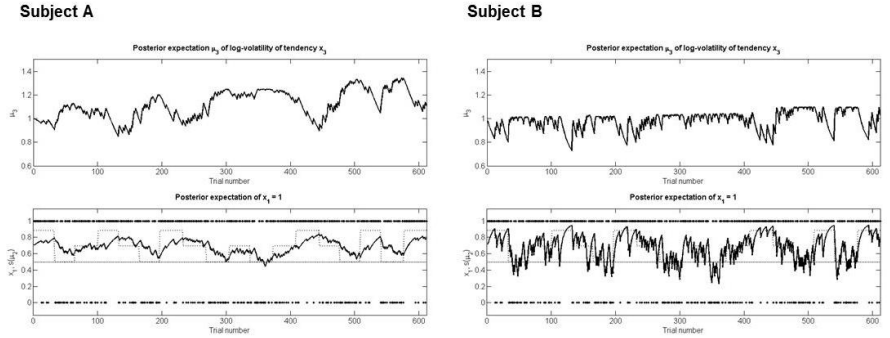

Figure 26. ***Illustration of the time course of μ_3 (upper panels) and $s(\mu_3)$ (lower panels)*** *during observation of x_1 (diamonds) for two exemplary subjects with different parameters for ω and ϑ. The true %CV is depicted as a dotted line. It can be seen that subject A (ω = -6.09; ϑ = 0.97) shows slower updating of the probability estimate that the target will appear at the cued location than subject B (ω = -2.78; ϑ = 0.12). This can be attributed to subject A's lower value of ω (reflecting the subject's belief in a less volatile environment).*

To illustrate how reaction times are related to the precision-based attentional factor α, we pooled response speeds over different bins of the attentional factor (using bins of 0.1, separately for valid and invalid trials) using estimates of trial-specific α based on the group average values for ω and ϑ. Figure 27 depicts the binned RS over subjects as a function of α. A 2 (*cue*: valid, invalid) x 6 (precision-based quantity α: 0.5, 0.6, 0.7, 0.8, 0.9, 1.0) ANOVA revealed a significant main effect of *cue* ($F_{(1,14)}$=11.8, p=.004 and a significant *cue* x α interaction effect ($F_{(3.44,48.09)}$=10.5, p<.001). We

compared these empirical RS values to the RS predicted by the model. For this, we computed the expected RS as a function of α on the basis of the group average values for ζ_{1_valid} and $\zeta_{1\_invalid}$ and ζ_2 (see Figure 27). It can be seen that the observed RS shows a similar pattern as the predicted RS. As expected, as precision (confidence in the validity of the cue) increases, there is a reaction time benefit for valid trials and an equivalent cost for invalid trials. This illustrates that one can explain attention formally in terms of optimizing or learning the relative precision of competing sensory channels.

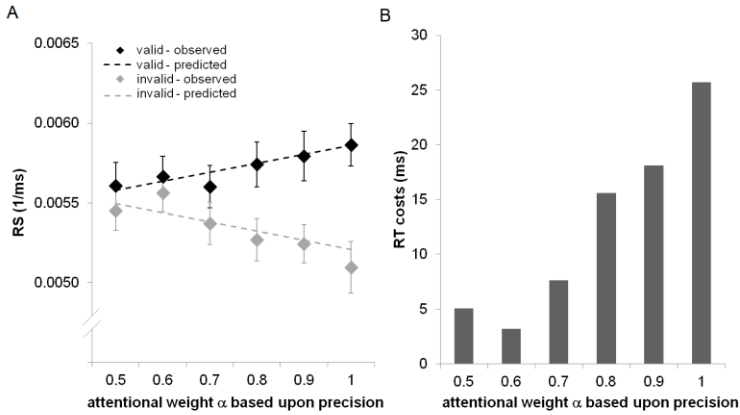

Figure 27. ***Observed and predicted reaction speeds****. (A) Observed and predicted average RS in valid and invalid trials as a function of the precision-dependent attentional weight parameter α (attention to cued location; calculated for the group average values). Error bars depict standard errors of the mean (SEM). The lines correspond to the predictions using the average response model parameters, over subjects. (B) Illustration of how the observed RS costs after invalid cueing translate into RT differences (in ms).*

9.4.5 REPLICABILITY OF RESULTS

In the independent replication dataset, the proportion of trials with missing eye data or missing or incorrect saccades amounted to 7.6 ± 2%, so that on average 92.4% of the trials (566 of 612 trials) were analyzed. Trials needed to be excluded due to anticipated responses (0.6 ± 0.2%), incorrect or absent saccades (0.6 ± 0.2%), saccades not starting from the fixation zone (3.3 ± 1%), or missing data points (e.g. due to blinks) (3.1 ± 1%). Note that due to the extended training and the increased number of resting periods, the amount of usable trials was higher than in the original dataset.

The 2 (*cue*: valid, invalid) × 3 (*%CV*: 50, 69, 88%) ANOVA on RS data gave the same results as for the original dataset. Specifically, it revealed a significant main effect of *cue* ($F_{(1,15)}$=17.6, p=0.001) reflecting faster responses (higher RS) on valid than on invalid trials. As before, the main effect of *%CV* was not significant but we observed a significant *cue* × *%CV* interaction effect ($F_{(1.99,29.88)}$=4.7, p=0.017). Since the data were derived from a within-subject cross-over design (where half of the subjects received the placebo tablet in the first session while the placebo session for the other half of subjects was the second experimental session), we additionally tested for an effect of session order by adding this variable as a between-subject factor to the ANOVA. No main effect of session order (p=0.15) or interaction of session order with any of the other factors (all p>0.28) was observed.

The results of the Bayesian model comparison are depicted in Table 2. Again, the full Bayesian model based upon precision showed the highest exceedance probability (0.914) when compared to alternative models. For the winning model, we again observed a significant positive correlation between the Rescorla-Wagner learning rate ε and ω (r=0.59, p=0.017), while no such relationship was observed between ε and ϑ (p=0.97). In summary, this second dataset provided a full replication of our original results.

9.5 Discussion

The present study analyzed saccadic RTs in a location-cueing paradigm with a volatile probabilistic context, probing Bayesian theories of perceptual inference. Extending previous theoretical work (Feldman and Friston, 2010), we were able to provide empirical evidence for the free-energy formulation of attention in the context of a Posner paradigm – where cue validity changed unpredictably in time, thus requiring the subject to learn about environmental volatility. Specifically, using the HGF, we compared 3 alternative models of how subjects might update estimates of cue validity across trials (perceptual models) and crossed these with 3 alternative hypotheses about how posterior beliefs (precision, belief, surprise) might inform decision making within trials (response models). The resulting 9 models – and two control models – were optimized using empirical measures of saccadic response speed and their relative plausibility was evaluated using Bayesian model selection (BMS). The results of this model comparison provided strong evidence in favor of the hierarchical Bayesian model combined with the precision response model (Table 2) and this finding was replicated in an independent dataset. This supports the notion that attention can be formulated as optimizing the confidence in (or precision of) the inference on sensory input (Friston, 2009). In the following, we examine our results in more detail, discuss them in the context of previous work, and outline future extensions.

Our experimental paradigm differed from a conventional Posner task, in that the spatial cues predicted the target location with different probabilities at different times during the experiment, thus requiring the subject to infer cue validity while accounting for environmental volatility. Indeed, a conventional ANOVA showed that the subjects' response speed varied as a function of the (unknown) true probabilities, reflecting adaptation to the changing environmental statistics. In other words, probabilistic context significantly influenced saccadic latencies, although the probabilistic structure of the task was changing in a way that was unknown to subjects.

This relates to previous work in so far as it has been shown that (inverse) saccadic RTs are sensitive to the probability of the saccade target location

when abrupt changes in location probability occur within an experimental block (Anderson and Carpenter, 2006), or when different experimental blocks employ saccade targets with different probabilities and/or different stochastic properties (Brodersen et al., 2008). In contrast to our task, both studies presented targets without preceding cues, and the latter study also examined learning of sequential (conditional) dependencies between successive stimuli according to a 1st order Markov sequence. The present task used explicit cues to elicit spatial attention shifts and investigated how the impact of these cues depended on the subject's current belief (and its precision) about the cue-target contingency. Moreover, instead of presenting different experimental blocks with different probabilistic contexts, here we introduced a volatile environment with frequent but hidden changes of probabilistic context within one continuous trial sequence. A natural modeling framework for explaining the ensuing saccadic reactions is a hierarchical Bayesian learning model – where the subject's belief about the environment's volatility affects the updating of beliefs about the most likely saccade target location. Indeed, comparison of competing perceptual models showed that a full hierarchical perceptual model had higher evidence than reduced models; assuming either that subjects ignored prior knowledge about the volatile nature of the environment or that they did not use them for updating beliefs about current cue validity. Moreover, the optimal full hierarchical Bayesian learning model showed higher model evidence than a Rescorla-Wagner learning model or a model which assumes that the subjects knew the true underlying probabilities. Interestingly, however, the subject-specific volatility parameter ω significantly correlated with the learning rate ε of the Rescorla-Wagner model, while no such relationship was observed for the meta-volatility parameter ϑ. The effects of the Bayesian model selection as well as the relationship to the learning parameter of a Rescorla-Wagner model could be replicated in an independent dataset.

Hierarchical Bayesian models have been used previously to successfully explain various aspects of human behavior under uncertainty, such as binary choices (Behrens et al., 2007) or manual RTs (den Ouden et al., 2010). These previous studies, however, assumed an ideal Bayesian observer with no inter-individual variation in the learning process *per se*. In contrast, we followed the meta-Bayesian approach of Daunizeau et al.

(2010a, 2010b) and inferred subject-specific parameters of the HGF from empirical RS. Our results showed that there is considerable inter-individual variability, even within our group of young healthy subjects (cf. Figures 5 and 7). An obvious and important extension of the present work is to relate this variability to demographic or neurobiological factors. In fact, the present work is a prelude to future psychopharmacological and patient studies in which we will examine the putative relation of individual differences in learning and attention (as encoded by our model parameters) to individual differences in neuromodulatory processes (as induced by medication, aging or disease).

Moreover, we introduced and tested different response models; i.e., mappings from posterior beliefs provided by the perceptual model to observable behavior. These response models account for the individual variability in the overall speed of responding (Figure 24B), but differ in assuming whether precision of predictions, strength of the prediction about cue validity or surprise, respectively, determine saccadic response speed. Our results showed that model evidence was highest for the response model in which RS was determined by the precision of the prediction.

In one sense, our findings from the Bayesian model comparison – that precision was the most plausible account for reaction time benefits – should not be surprising. This is because precision plays the role of a rate-constant in evidence accumulation schemes based upon predictive coding (Feldman and Friston, 2010). In other words, precision modulates the gain of prediction error in driving changes in conditional representations or expectations. This means that sensory channels that enjoy greater precision will engender faster changes in high-level representations and lead to more rapid perceptual convergence. Behaviorally, this should be manifest in speeded up reaction times. Exactly the same theme is seen at higher levels of the hierarchy – that concern slower timescales – such as inference about the probabilistic (trial-to-trial) contingencies we manipulated in our volatility paradigm. Here, the rate-constant corresponds to a learning rate in conventional (reinforcement learning) formulations. In short, sensory evidence and empirical priors that are afforded greater precision have preferential access to higher levels in

hierarchical inference. This is expressed as more efficient and faster convergence in those processing streams – and provides a nice metaphor for attention.

In other words, attention corresponds to optimizing estimates of precision in sensory hierarchies and is implemented by changing the post-synaptic gain of neuronal prediction error units. Hence, attention determines which part of the sensorium is treated as furnishing precise information. In this respect, this approach is perfectly congruent with spotlight or zoomlens theories of attention (Posner, 1980; Eriksen and James, 1986) as well as with the biased competition model (Desimone and Duncan, 1995): the limitation of processing capacities demands a selection of stimulus locations or features, so that only the most relevant receive full attention. Neurobiologically, this is likely reflected in increased synaptic gain and neuronal synchronization, manifesting itself in enhanced firing rates (e.g., Luck et al., 1997) or BOLD responses (e.g., Brefczynski and DeYoe, 1999; Kastner et al., 1999) in visual cortex, when attention is directed to a particular spatial location. It may also be noteworthy that, neurochemically, precision-dependent synaptic gain (e.g., at superficial pyramidal cells) may be controlled by classical neuromodulators such as dopamine or acetylcholine (Friston, 2009). Increased gain may engender faster changes in neuronal activity in higher-level areas (such as the intraparietal sulcus, IPS, or the FEF), so that evidence can accumulate more rapidly and saccades are elicited more quickly. This notion resonates with findings from several recent studies. For example, Saproo and Serences (2010) showed that spatial attention increases the mutual information of population response profiles in early visual cortex and suggested that this should enable higher visual areas to read out this information more quickly and efficiently. This is similar to the proposals by Feldman and Friston (2010) and in this thesis, where higher precision at lower levels induces more rapid changes in the activity of higher-level areas. Others have suggested that attention produces behavioral improvements by efficiently selecting the "relevant" sensory signals (Pestilli et al., 2011); the suggested mechanism (focusing on the magnitudes of signals and employing pooling operations) differs from ours, and it would be interesting to see whether the results obtained by Pestilli et al. on behavioral contrast-discrimination performance could be

replicated when trials are grouped according to precision estimates. Finally, it has been shown that electrical stimulation of direction-selective neurons in MT elicits faster perceptual decisions due to faster evidence accumulation (Ditterich et al., 2003).

According to hierarchical Bayesian schemes as employed here, precision controls the gain of prediction error associated with bottom-up signals. Physiologically, precision may be encoded by the gain of superficial pyramidal cells (Brown and Friston, 2012). Accordingly, our computational model would predict that during spatial attention, activity in hierarchically related visual areas should exhibit precision-dependent modulatory effects that result from the enhanced gain of superficial pyramidal cells. This hypothesis - as well as the question where in the spatial attention/saccade network precision exerts this effect - could be tested with models that infer on mechanisms of measured EEG or MEG data (Bastos et al., 2011; Brown and Friston, 2012). Interestingly, a recent fMRI study, using a simpler and less fine-grained DCM for fMRI, has highlighted the importance of the modulation of inhibitory self-connections in visual areas by attention and prediction. This type of modulation corresponds (at a coarser level) to a simple gain control mechanism that may reflect the precision-dependent modulation of pyramidal cells described above.

Given the involvement of common areas (FEF, IPS) in both covert attentional orienting of attention and overt eye movements (Beauchamp et al., 2001; Corbetta et al., 1998; de Haan et al., 2008; Nobre et al., 2000; Perry and Zeki, 2000), the psychophysical evidence for an inherent link between attention shifts and saccade programming (Deubel and Schneider, 1996; Deubel, 2008; Doré-Mazars et al., 2004; Godijn and Theeuwes, 2003), and the existence of both visual and motor neurons in key structures such as the FEF (e.g., Bruce and Goldberg, 1985; Schall and Hanes, 1993), it seems plausible that precision should affect both sensory-perceptual as well as motor preparatory processes (cf. the model proposed by Schall et al., 2011). Hence, one could also frame the processes studied here in the broader context of visual-saccadic decision making (see Glimcher, 2001, 2003 for comprehensive reviews).

The focus of the present study was on explaining observed trial-wise saccadic RS by a generative (hierarchical Bayesian) model and use model selection to disambiguate potential computations that govern the updating of beliefs about upcoming saccade target locations in a volatile environment. While our analyses suggest a precision-based mechanism for spatial attention, it remains to be investigated where these precision estimates are computed within the hierarchical visual attention/saccade network. The present behavioral modeling results are a foundation for future imaging studies that will exploit the across-trial and between-subject variation in model states and parameters to identify the network of regions in which precision plays a role for belief updating in spatial attention. Concretely, future neuroimaging studies could use the time-series of the states of our perceptual model as predictor variables to identify brain responses that covary with these computational processes (cf. Behrens et al., 2007; den Ouden et al., 2010). Furthermore, as mentioned above, subject-specific estimates of the parameters encoding individual learning style can be used in group analyses to reveal the neuronal substrates of inter-individual differences.

In conclusion, we have used the HGF for characterizing Bayes-optimal trial-by-trial updating of probabilistic beliefs under uncertainty for explaining attentional mechanisms. Specifically, we characterized saccadic response speed during an extended Posner paradigm with variable cue validity. Comparing 11 alternative models, we found that empirical responses are most plausibly explained as a function of precision (of the beliefs about the causes of sensory input). This finding is consistent with attention theories derived from Bayesian theories of brain function (the free-energy principle) that equate spatial attention to a precision-dependent gain modulation of sensory input. Future neuroimaging work could use the modeling approach introduced in this thesis to identify the neural and neurochemical basis of attentional selection and saccadic eye movements, in relation to probabilistic expectancies.

10 APPLICATION TO THE OBSERVATION OF AN ITERATED TRUST GAME

10.1 Overview

Trust games are widespread in behavioral economic research and have recently been shown to have relevant applications in psychopathological research. In that context, the dynamics of the participants' unfolding interaction is crucial, as opposed to more static attitudes that may be revealed. While the nature of such interaction dynamics has been quantified previously in a model-free way, in the study I report here, we used the HGF to model the perception, by the observer of a trust game, of the interaction dynamics. 44 healthy human subjects observed the interactions in a trust game that had previously taken place between two additional subjects. While lying in an fMRI scanner, they had to predict the next decision by one of the participants in the trust game (the trustee) before that decision was revealed to them. The game went over 70 rounds, and prediction accuracy was rewarded. Several questionnaire- and behavior-based psychological measures were assessed. Performance in the trust game prediction task was significantly correlated with the "openness to experience" scale of the NEO-FFI, and, in women, with the "reward responsiveness" scale of the BIS/BAS. The fMRI analysis revealed strong responses in various brain regions to very abstract inferred measures: cooperativity of the trustee and its volatility (i.e., measures at the first and second hidden level of the HGF). This study thus provides proof of concept for the applicability of the HGF to inference on complex learning and decision making tasks. It also takes the "observing the observer" framework one step further and extends it to the observation of an observing observer. Furthermore, it introduces a task that is quick and simple, and which can crucially be performed by the subject alone while potentially retaining the anticipated advantages of assessing multiplayer interactions for clinical diagnosis and prognosis.

10.2 Introduction

Trust games (Berg et al., 1995) have become an essential empirical tool in behavioral economics (Fehr, 2009). The basic version of a trust game involves two players, an investor and a trustee. Both receive the same initial endowment to prevent effects such as inequality aversion from distorting subsequent behavior. The investor may now transfer some or all of his endowment to the trustee. On the way to the trustee, the transferred amount is multiplied by a factor greater than one – usually, three. The trustee may then repay some or all of what he has (i.e., his initial endowment plus the multiplied transferred amount) to the investor. Note that if the transferred amount was tripled, both players will end up with the same amount if the trustee repays twice what the investor sent him (i.e., two thirds of what he received in addition to his endowment); however, the interaction is profitable to the investor as soon as the trustee sends back more than one third of what he received.

While it is often important to exclude reputation effects in behavioral economics, leading to a preference for one-shot interactions, the dynamics of repeated interactions are of keen interest in psychology and psychopathology. For example, King-Casas et al. (2008) have shown that in a trust game that runs over 10 rounds, the behavior individuals with a diagnosis of borderline personality disorder (BPD) differs from that of healthy controls in important ways. When cooperation is breaking down (sc. investment levels are low) healthy controls are twice as likely to engage in coaxing, that is to repay a large amount nonetheless. This is called coaxing because it is assumed to take place with a view to raising the amounts invested by showing oneself trustworthy and generous. In the study by King-Casas et al. (2008), coaxing did indeed pay off: investment levels subsequently rose.

While letting subjects engage in interactions may prove to be a valuable tool for clinical assessments, it poses logistical and conceptual problems. Concerning the logistics, it is impractical always to need two subjects for a task; and conceptually, it is unclear how to compare the behavior of two subjects who have interacted with two possibly very different partners.

For these reasons, we developed a task that can be performed by one subject alone and that is exactly the same for all subjects.

Iterative (i.e., multi-round) trust games are usually played over no more than 15 rounds (e.g., Buskens et al., 2010). While this may in many cases be enough to lead to interesting behavior, we were concerned that it would only allow us to capture the initial phase of a series of interactions that might have taken further turns. This had us lengthen the game considerably to 70 rounds.

10.3 Methods

10.3.1 TASK

Subjects were asked to observe an iterated trust game that went over 70 rounds. They were informed that the game had actually taken place at our lab in precisely the way they would see it. They were not given any information about the two players they were observing beyond the numbers announcing the investor's investment and subsequently the trustee's repayment, sequentially for each round. Subjects knew that both players had an endowment of 100 points on each round, convertible into Swiss francs at the end of the game at a certain fixed, but unmentioned rate. They were asked to do a number of example calculations before beginning the task to make sure they had understood it.

While the subjects were lying in the MRI scanner, they first saw the amount invested by the investor displayed on a screen for 2 s. After a delay jittered around 3 s, they then had 6 s to announce by a button press that they had reached a decision on their prediction of the amount the trustee would repay in that round. They were informed that they would not be rewarded on that trial if they failed to press a button within the allotted 6 s. After this decision period of 6 s (the screen simply went blank after the button press for the remainder of the period) a horizontal scale from 0 to 400 appeared on the screen, with two vertical lines across. The red line on the right-hand side indicated the maximal possible repayment by the trustee on that trial, given the investors investment. This could range from 100 for an investment of 0 (all the trustee then has is his endowment) to

400 for an investment of 100. The second, green, vertical line was placed at a random initial point between zero and the red line. This indicated the subject's prediction and could be moved around quasi-continuously (steps of 1) by button presses. Subjects had 10 s to move the green bar to the location of their prediction and finalize their choice by a button press. If they failed to finalize a prediction, a missed trial was recorded and subjects were not rewarded for that trial. After a delay jittered around 6 s and beginning immediately with the finalization of their prediction, they saw the amount the trustee had actually repaid on the screen for 2 s. After an intertribal interval jittered around 4 s, the next trial began.

On each trial, subjects could win 400 points. They received this maximum for exactly accurate predictions, but reward fell with the squared difference between prediction and outcome, so that no reward was given for prediction errors of 20 or more points. At the end of the experiment, points were converted into Swiss francs at a rate of 0.002, resulting in a maximal total reward of 56 francs (about 59 U.S. dollars).

Note that at no point in this experiment were subjects deceived.

10.3.2 PRE-STUDY

In order to acquire the trust game data to be used in the main study, we conducted a pre-study where 24 subjects were randomly matched into pairs to play a trust game over 70 rounds. Subjects each sat in their own cubicles in front of a computer screen. They knew that one of the other people in the room was their partner in the trust game but they did not know who. They were also informed that the identity of their partner would not be revealed after the game, nor would their identity be revealed to their partner. The trust game was implemented in the software z-Tree (Fischbacher, 2007). At the end of the experiment, points were converted into Swiss francs at a rate of 0.002.

10.3.3 SUBJECTS

For the main study, 44 healthy adult human subjects (22 men, 22 women) were recruited via a subject pool administered by the Department of Economics at the University of Zurich. They gave written informed consent

to participate in the study which was approved by the Zurich cantonal ethics committee (KEK-ZH 2010-0327). None of the subjects had a history of mental health problems.

For the pre-study, 24 subjects with the same characteristics were recruited.

10.3.4 MRI ACQUISITION

MRI images were acquired on a Philips 3 Tesla Achieva whole body MR scanner (Philips Healthcare, Best, The Netherlands) equipped with an eight-channel Philips SENSE head-coil. Structural images were acquired using a T_1-weighted sequence. For functional imaging, 40 transverse slices were measured in ascending order with a slice thickness of 2.6 mm and a gap of 0.7mm for a voxel size of $2.0 \times 2.0 \times 2.6$ mm and a field of view of $192 \times 192 \times 131.3$ mm. A T_2^*-weighted single-shot echo-planar imaging sequence with a TR of 2500 ms, a TE of 25 ms, and a flip angle of 80° was used.

10.3.5 QUESTIONNAIRES AND ADDITIONAL BEHAVIORAL MEASURES

In addition to the trust game prediction task during fMRI acquisition, subjects were asked to fill in a number of questionnaires and carry out a behavioral task.

The questionnaires were

- the NEO-FFI (McCrae and Costa Jr., 2004), a five-factor personality inventory;
- the Barratt impulsiveness scale (BIS11) (Patton et al., 1995);
- the BIS/BAS scales (Carver and White, 1994), which measure behavioral approach and behavioral inhibition;
- the MACH-IV test of "Machiavellianism" (Christie and Geis, 1970);

- a measure of social value orientation (SVO) (Murphy et al., 2011);
- the Davis empathy scales (Davis, 1980);
- the Yamagishi trust scale (Yamagishi and Yamagishi, 1994);
- and a measure of betrayal aversion (Bohnet and Zeckhauser, 2004).

The behavioral measure taken was the balloon analogue risk task (BART) (Lejuez et al., 2002).

10.3.6 DATA ANALYSIS

Questionnaire and BART data were analyzed using the statistical software R (R Core Team, 2012); data from the trust game prediction task were analyzed using the HGF toolbox introduced in Chapter 8.

The fMRI data were analyzed using Statistical Parametric Mapping (SPM8, Wellcome Trust Center for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm). All functional images and the structural image of each subject were preprocessed prior to the statistical analyses. We applied motion correction to the functional time-series and a 'new segmentation' procedure, including the computation of bias-corrected images and forward deformation fields, of the resliced mean functional and the structural image. For the segmentation of the functional image, tissue probability maps resulting from the segmentation of the subject-specific structural image were used. After coregistration of the bias-corrected functional images to the subject-specific bias-corrected structural image, the images were normalized to MNI space using the forward deformation field of the structural image. Finally, the functional images were smoothed applying a 6 mm full-width at half maximum Gaussian kernel and resampled to $1.5 \times 1.5 \times 1.5$ mm resolution.

The HGF was used to infer on subjects' beliefs regarding the current level of cooperation to be expected from the trustee. This was based on a framework that assumes us, the experimenters, to be the observers of an observing observer. We therefore needed three models:

1. A model for the subject's perception of the trustee's level of cooperation,

2. A decision model for the trustee, and

3. A decision model for the subject.

Before specifying the details of the first (i.e., perceptual) model we made the assumption that the trustee's level of cooperation, which I will call his *cooperativity* x_1, could be quantified on an unbounded continuous scale (cf. Eq. (128) below) and tracked by the HGF. This permitted us to construct the second and third models from very simple assumptions: the trustee's decisions u are characterized by constant Gaussian noise α around his cooperativity:

$$p(u|x_1) = \mathcal{N}(u; x_1, \alpha) \tag{124}$$

This is exactly Eq. (75), meaning that we can model the subject to use the HGF directly with the trustee's decisions to infer on his cooperativity. The HGF then is the first model in the list above.

Note, however, that before we can use u as inputs for the HGF, they have to be transformed from the bounded space in which they occur to an unbounded space. We did this by applying a logit transformation to the observed number of points repaid by the trustee. If $m^{(k)}$ is the maximal possible number of points that can be repaid on trial k, and the observed repayment is $u_{obs}^{(k)}$, then $u^{(k)}$ is defined as

$$u^{(k)} \stackrel{\text{def}}{\equiv} \ln\left(\frac{u_{obs}^{(k)}}{m^{(k)} - u_{obs}^{(k)}}\right), \tag{125}$$

and the space of $u^{(k)}$ is *cooperativity space*. In the event that $u_{obs}^{(k)}$ is either zero or $m^{(k)}$, we move it slightly away from these extremes to $0.005 \cdot m^{(k)}$

or $0.995 \cdot m^{(k)}$, respectively, to keep $u^{(k)}$ within bounds even though it now lives in an unbounded space where it can be modeled by Eq. (124). An illustration of such a transformation to cooperativity space is given in Figure 28 and Figure 29.

If we now apply the transformation of Eq. (128) also to the subject's predictions $y_{obs}^{(k)}$, a simple choice of model (third in the list above) for the transformed predictions $y^{(k)}$ is

$$p(y|\hat{\mu}_1) = \mathcal{N}(y; \hat{\mu}_1, \zeta), \tag{126}$$

where $\hat{\mu}_1$ is the predicted cooperativity as modeled by the HGF and ζ is a constant decision noise parameter. Note that the parameters governing the HGF updates will differ from subject to subject, resulting in subject-specific predictions $\hat{\mu}_1$. It is these subject-specific predictions that the inversion of the HGF detailed in the preceding chapters allows us to model efficiently.

Using the model defined by Eqs (124)-(126), trajectories for μ_1 and μ_2 (i.e., the volatility of cooperativity) were computed for each subject. Weakly informative (i.e., wide) priors were chosen with means suggested by the empirical Bayes estimation of optimal parameters implemented in the HGF toolbox. The μ_1 and μ_2 trajectories were then used as individual parametric modulators in the GLM regression analysis of the fMRI data.

10.4 Results

10.4.1 Pre-study

Out of the twelve subject pairs that played an iterated trust game, about half (seven) showed relatively uninteresting and simple-to-predict patterns of consistently high cooperation. The other five exhibited a more varied series of interactions with cooperation intermittently breaking down before picking up again. The game shown in Figure 28 was chosen for use in this study.

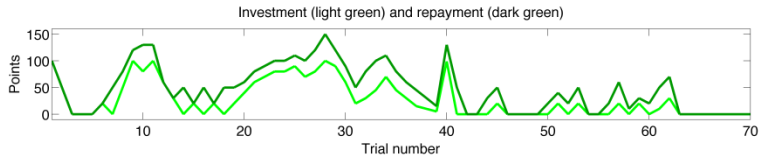

Figure 28. ***Chosen trust game trajectory.*** *Coaxing by the trustee is evident in many places throughout the game, for example around trials 15 and 60.*

The same trajectory transformed into cooperativity space is shown in Figure 29.

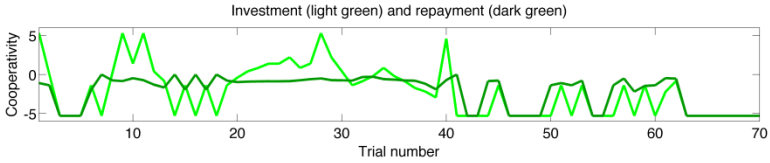

Figure 29. ***Chosen trust game trajectory in cooperativity space.*** *In this presentation, it seems apparent that the trustee's recurring unwillingness to share profits equitably (which would be indicated by a cooperativity of 0) ultimately led to a breakdown of cooperation despite repeated coaxing.*

10.4.2 Questionnaires and behavioral measures

Subjects achieved a median score of 29.64 francs in the prediction task, with an interquartile range of 4.79 francs. This score was regressed onto the various questionnaire scales. The only significant results were a significant positive association of score with the "openness to experience" scale of the NEO-FFI (effect size $d = 0.34$; $p = 0.03$) and a significant positive association with the "reward responsiveness" scale of the BIS/BAS in women (effect size $d = 0.56$; $p = 0.02$), but not in men (effect size $d = 0.13$; $p = 0.54$).

10.4.3 FMRI

An example of an individual HGF estimation of cooperativity and its volatility is given in Figure 30.

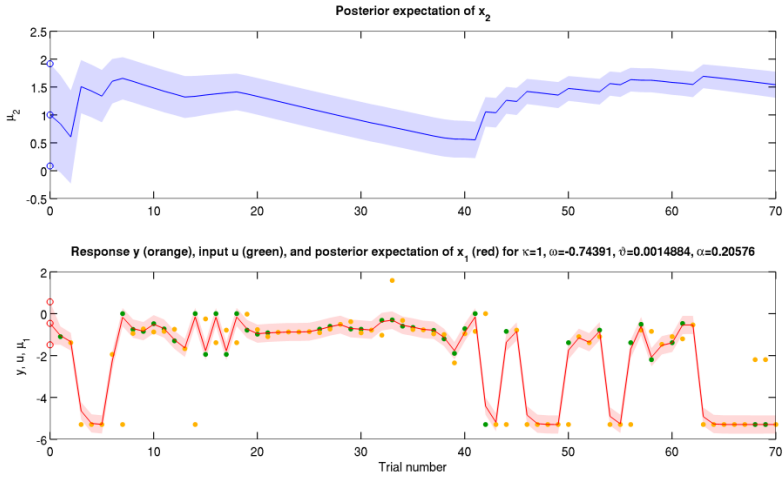

Figure 30. ***HGF estimation of cooperativity and its volatility.*** *Trajectories for one example subject.*

When applied as a parametric modulator to the outcome phase of trials, μ_1 was associated with strong activations ($p < 0.05$ whole brain family-wise error (FWE) corrected) in regions previously associated with reward and learning signals (cf. Figure 31 and Table 3). Most notably, the cooperativity regressor correlates positively with large clusters in the dorsal anterior cingulate cortex (dACC), the anterior insula (AI), and with two clusters in the midbrain and brainstem. The larger one of these (79 voxels) is situated in the superior colliculi while the smaller one (22 voxels) represents the ventral tegmental area (VTA). These regions are more active the higher the cooperativity of the trustee is seen to be by the subject.

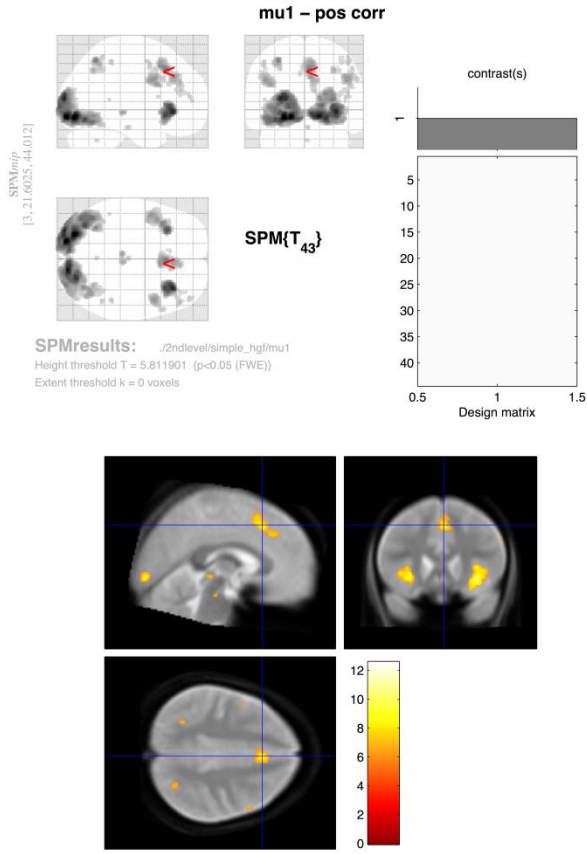

Figure 31. ***Positive correlation with μ_1.*** *Activation clusters are found in the dACC, the anterior insula, and in the brainstem, where one cluster of 22 voxels occupies the VTA while the superior colliculi are represented with 79 voxels.*

Table 3. ***Positive correlation with μ_1: significant clusters at $p < 0.05$ FWE.*** *p_{FWE} is the family-wise error-corrected p-value, p_{FDR} is the false discovery rate-corrected p-value, k_E is the number of voxels in the cluster, and T is the t-statistic at the peak.*

Cluster			Peak			Coordinates (mm)		
p_{FWE}	p_{FDR}	k_E	p_{FWE}	p_{FDR}	T	x	y	z
0	0	9864	0	0	12.53	-18	-91	-11
			0	0	12.47	-24	-82	-11
			0	0	11.84	14	-90	-9
0	0	1055	0	0	11.13	34	30	-5
			0	0	10.17	33	21	-8
			0	0.001	8.6	40	22	1
0	0	491	0	0	9.86	-28	23	-5
			0	0.001	8.42	-38	18	-3
			0.031	0.648	5.98	-28	17	-14
0	0	949	0	0.001	8.81	3	23	42
			0	0.003	8.09	6	15	46
			0	0.006	7.82	6	30	34
0	0	209	0	0.001	8.67	-32	-58	52
			0.005	0.139	6.63	-30	-54	43
0	0	566	0	0.005	7.96	-46	6	31
0	0	365	0	0.005	7.88	46	9	33
			0.006	0.149	6.58	57	18	33
			0.007	0.189	6.49	52	11	43
0	0	79	0	0.012	7.57	-3	-28	-5
			0.001	0.025	7.25	6	-27	-5
0	0	93	0.001	0.025	7.26	32	-61	42
			0.016	0.379	6.21	38	-51	40
0	0.004	39	0.002	0.048	7.01	50	38	27
0.001	0.019	22	0.005	0.132	6.65	2	-21	-21
0	0.007	32	0.005	0.14	6.63	-48	3	49
0.001	0.019	22	0.005	0.14	6.6	48	-40	51
0.002	0.049	15	0.006	0.159	6.55	48	-51	56
0	0.015	26	0.008	0.208	6.45	45	48	13
0.005	0.138	8	0.012	0.293	6.32	-6	27	37
0.005	0.138	8	0.013	0.311	6.3	2	-58	-36
0.002	0.054	14	0.015	0.355	6.25	-39	24	27
0.006	0.161	7	0.016	0.377	6.22	38	38	13
0.009	0.224	5	0.017	0.385	6.2	-46	-43	9
0	0.015	25	0.019	0.437	6.15	40	-54	51
0	0.016	24	0.019	0.437	6.15	44	33	21
0.009	0.224	5	0.02	0.438	6.14	50	12	12
0.028	0.561	1	0.031	0.648	5.98	-8	23	40
0.028	0.561	1	0.031	0.648	5.98	39	44	18
0.02	0.466	2	0.033	0.68	5.96	39	42	15
0.028	0.561	1	0.035	0.714	5.94	38	45	16
0.028	0.561	1	0.039	0.769	5.9	36	-48	45

At the second level (volatility of cooperativity), it is the negative correlation with μ_2 that produces interesting results: as shown in Table 4, there is a cluster of 12 voxels in the ventral striatum that retains significance under whole brain FWE correction at the $p < 0.05$ level.

Table 4. ***Negative correlation with μ_2: significant clusters at $p < 0.05$ FWE.*** *p_{FWE} is the family-wise error-corrected p-value, p_{FDR} is the false discovery rate-corrected p-value, k_E is the number of voxels in the cluster, and T is the t-statistic at the peak.*

Cluster			Peak			Coordinates (mm)		
p_{FWE}	p_{FDR}	k_E	p_{FWE}	p_{FDR}	T	x	y	z
0.002	0.239	12	0.009	0.855	6.45	15	9	-8
0.027	0.541	1	0.034	0.954	5.98	20	16	-6
0.027	0.541	1	0.034	0.954	5.97	34	47	30
0.019	0.541	2	0.035	0.954	5.97	-14	18	-5
0.027	0.541	1	0.045	0.954	5.87	-32	50	27
0.027	0.541	1	0.048	0.954	5.86	16	21	10

If we look at the same contrast at the $p < 0.001$ uncorrected level, we find large clusters of activation (significant at the cluster level) in the ventral striatum (VS), the dorsolateral prefrontal cortex (dlPFC), and the OFC (cf. Figure 32 and Figure 33, and Table 5) These regions are more active the lower the volatility of the cooperativity of the trustee is seen to be by the subjects.

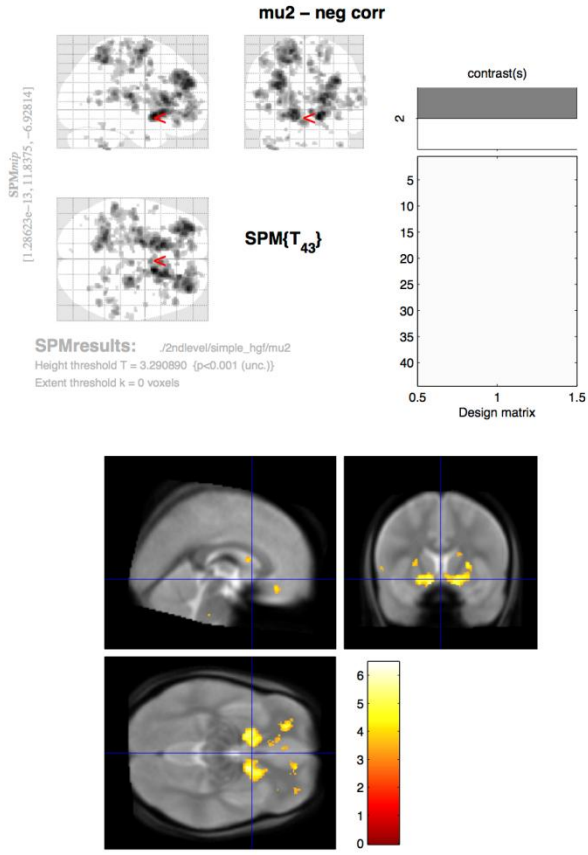

Figure 32. ***Negative correlation with μ_2.*** *Activations are displayed at a threshold of* $p < 0.001$ *uncorrected. Both large clusters in the ventral striatum are significant at the cluster level.*

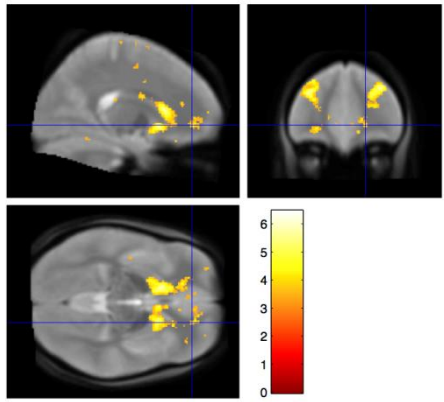

Figure 33. ***Negative correlation with μ_2 (cont'd).*** *Further clusters (significant at the cluster level) in OFC and dlPFC displayed at a threshold of $p < 0.001$ uncorrected.*

‘

Table 5. ***Negative correlation with μ_2: selection of significant clusters at $p < 0.001$ uncorrected.*** *p_{FWE} is the family-wise error-corrected p-value, p_{FDR} is the false discovery rate-corrected p-value, k_E is the number of voxels in the cluster, and T is the t-statistic at the peak.*

Cluster			Peak			Coordinates (mm)		
p_{FWE}	p_{FDR}	k_E	p_{FWE}	p_{FWE}	T	x	y	z
0	0	2804	0.009	0.077	6.45	15	9	-8
			0.034	0.077	5.98	20	16	-6
			0.034	0.077	5.97	34	47	30
0	0	1337	0.035	0.077	5.97	-14	18	-5
			0.076	0.077	5.69	-15	6	-9
			0.288	0.14	5.16	-26	44	-8
0	0	1551	0.045	0.077	5.87	-32	50	27
			0.077	0.077	5.68	-30	36	36
			0.139	0.103	5.46	-26	41	27
0	0	891	0.059	0.077	5.78	-16	-6	63
			0.412	0.183	5	-28	-9	55
			0.987	0.479	4.18	-20	5	72
0.003	0.002	328	0.084	0.077	5.65	24	-19	13
			1	0.716	3.74	20	-25	19
0	0	1278	0.103	0.084	5.57	-21	-48	51
			0.242	0.129	5.24	-12	-54	57
			0.42	0.183	4.99	-24	-48	63
0.009	0.004	277	0.442	0.187	4.96	-30	-15	7
			0.977	0.468	4.24	-24	9	10
			0.995	0.479	4.1	-24	2	7
0.028	0.013	218	0.454	0.188	4.95	-56	-36	15
			0.996	0.492	4.07	-45	-34	13

0.001	0.001	391	0.627	0.254	4.76	20	46	-5
			0.854	0.345	4.51	18	54	1
			0.927	0.397	4.38	14	52	-8
0.264	0.073	117	0.722	0.276	4.66	15	-3	60
			0.998	0.508	4.04	14	-10	63
			0.999	0.58	3.96	16	-3	69
0.089	0.032	165	0.739	0.276	4.65	-18	23	10
0.427	0.106	95	0.766	0.289	4.62	-44	-13	1
0.82	0.257	58	0.817	0.323	4.55	4	5	49
0.401	0.104	98	0.853	0.345	4.51	6	-18	42
			1	0.825	3.54	9	-31	43
0.061	0.023	182	0.868	0.354	4.49	-14	51	3
			1	0.637	3.87	-16	51	12
0.436	0.106	94	0.903	0.381	4.43	-48	-15	51
			1	0.716	3.74	-46	-19	60
0.886	0.285	51	0.909	0.387	4.42	16	2	37
1	0.825	14	0.92	0.394	4.4	33	-10	22
0.157	0.051	140	0.922	0.394	4.39	60	-36	10
			0.989	0.479	4.17	57	-33	19
			1	0.854	3.5	50	-37	21
0.911	0.306	48	0.923	0.394	4.39	50	33	3
0.83	0.257	57	0.934	0.407	4.37	-15	-40	72
0.337	0.088	106	0.944	0.412	4.34	22	-48	54
			1	0.61	3.92	26	-48	63
0.252	0.073	119	0.945	0.412	4.34	-36	-45	-45
			0.951	0.42	4.33	-42	-51	-42
			1	0.706	3.77	-48	-52	-35
0.965	0.401	39	0.956	0.427	4.31	34	52	-11
0.16	0.051	139	0.976	0.468	4.24	-20	-21	67
			0.991	0.479	4.14	-33	-28	70
			1	0.918	3.39	-20	-18	76
1	0.825	13	0.984	0.479	4.2	-21	0	-33
0.83	0.257	57	0.985	0.479	4.2	50	-55	6
0.999	0.737	20	0.987	0.479	4.18	39	-46	-41
0.886	0.285	51	0.988	0.479	4.17	18	0	22
0.315	0.085	109	0.992	0.479	4.13	24	-9	58
			1	0.706	3.77	21	-7	51
0.04	0.017	201	0.992	0.479	4.13	-46	-31	51
			1	0.788	3.62	-42	-28	45
			1	0.791	3.62	-50	-24	46
0.86	0.274	54	0.993	0.479	4.12	48	-43	10
0.82	0.257	58	0.993	0.479	4.12	-48	-18	12
0.999	0.737	21	0.994	0.479	4.11	34	-66	7
0.236	0.073	122	0.995	0.479	4.1	9	8	16
			0.998	0.508	4.04	0	9	12
			0.999	0.562	3.98	-8	6	16
1	0.825	3	0.996	0.492	4.07	-18	38	10
1	0.799	16	0.997	0.492	4.07	28	20	30
0.999	0.724	22	0.998	0.508	4.04	33	-55	16
1	0.825	11	0.998	0.52	4.02	-16	-30	63
0.599	0.161	78	0.999	0.562	3.98	-16	-28	40
0.983	0.472	34	1	0.605	3.94	-54	11	3
0.999	0.737	20	1	0.61	3.92	12	18	34
0.745	0.231	65	1	0.611	3.9	44	22	3
			1	0.832	3.53	38	15	1
			1	0.854	3.48	34	22	10
1	0.825	10	1	0.627	3.88	18	-43	19
1	0.825	10	1	0.631	3.88	36	40	-6
etc.								

10.5 Discussion

This is to my knowledge the first study that has applied complex individual inferential modeling to a mentalizing task and found whole-brain corrected activations associated with individually inferred cognitive states of the subjects. Moreover, activations are found in the midbrain and brainstem, which are notoriously difficult to image functionally.

In previous work, Behrens et al. (2007, 2008) found that learning in a volatile environment entails activations in the dACC, the ventral striatum, and the VTA – a collection of regions notable for their dopaminergic connectivity. Interestingly, they found similar patterns of activation for learning in the social domain as in the non-social. In the non-social domain, this agreed with earlier findings that dopaminergic neurons in the VTA encode reward prediction error (Schultz et al., 1997), but it suggested that social learning depended on much the same mechanisms. Our results confirm these findings and extend them in several ways: first, we take social inference one step further: instead of inferring on a one-way relationship where one partner is always the giver and the other the taker (of advice, in Behrens et al. (2008)), we study inference on a situation that is truly interactive – an iterated trust game. Second, we model the subject's beliefs individually, allowing for different parameter values (underpinned by different notions of optimality) that govern the Bayesian updates in the HGF. Third, the value that the subject is tracking – the cooperativity of the trustee – does not relate to the subject himself, but to the investor. Note that VTA activity correlates positively with cooperativity of the trustee despite the fact that this is not directly rewarding to the subject – the sequences of the experiment where it is easiest for subjects to make money are in fact those where cooperation has broken down and the trustee predictably repays nothing. This raises the possibility that rewarding interactions among others – and be they complete strangers – can elicit strong reward prediction error signals in those observing them, perhaps on condition that these interactions happen to be relevant to the observer. Strong activations in the anterior insula and the dorsal ACC point in the same direction since they have been found to be associated with

empathic reactions (Singer et al., 2004). Activations in the ventral striatum in response to low volatility may indicate yet another way in which reward can be experienced by the subject: the less volatility, the more predictability and therefore the more reward from predicting the repayment.

A second strand of inquiry this study builds upon is that of King-Casas et al. (2008). In that regard, the innovation of this study is to set subjects the task of observing a trust game and having them make predictions. The strong activations found in regions previously associated with reward, learning, and decision making are an indication that subjects were heavily engaged in the task; many spontaneously reported that during the task they had developed a feeling of familiarity with the players they had been observing. This is remarkable since the players were complete strangers about whom the subjects knew nothing except the numbers they had entered into a computer in the course of a trust game at some earlier time. The human brain's capacity to make far-reaching and often quite accurate inferences based on very little information was on vivid display here. The very astoundingness of this capacity could be an indication of its fragility, suggesting that disruptions of it could lead to subtle deficits in social interactions that may underlie psychopathological symptoms. If that is the case, one may hope to find such abnormalities reflected in the individual parameter estimates that inversion of the HGF in combination with decision models entails.

In the pre-study, letting a trust game go over 70 rounds has proved a success. In more than a third of the observed cases, interesting long-term interaction patterns were visible, raising the possibility that the study of such interaction patterns might yield insights into the mechanisms of social interactions as they manifest themselves under uncertainty and in the face of competing interests.

The results obtained here could be strengthened in various ways. Returning to the three interrelated models defining the data analysis outlined in the previous section, the second and third of them could be improved considerably by taking into account the behavior of the investor along with that of the trustee. Specifically, it would seem obvious that the trustee's decision to repay a certain amount will not only be informed by

his current inclination to cooperate (cooperativity), but also quite directly by his perception of the amount invested by his partner in the game. That is, manifest cooperation may not reveal hidden cooperativity in any straightforward way. This could be accommodated by more sophisticated decision models for the trustee that also involve the cooperativity of the investor. Note that this is not necessarily restricted to the cooperation by the investor on the current trial, but also on his history of cooperation. For example, coaxing becomes important in situations where current cooperation has fallen below levels previously seen, opening the possibility to restore these levels.

Similar concerns apply to the subject's decision model (third in the list above). The subject may not base his prediction of the trustee's repayment solely on the cooperativity he imputes to him. A more sophisticated decision model could capture interaction effects between the perceived cooperativities of both investor and trustee.

11 CONCLUSIONS

11.1 The hierarachical Gaussian filter (HGF)

In this thesis, we have introduced a generic hierarchical Bayesian framework that describes inference under uncertainty; for example, due to environmental volatility or sensory uncertainty. The model assumes that the states evolve as Gaussian random walks at all but the first level, where their volatility (i.e., conditional variance of the state given the previous state) is determined by the next highest level. This coupling across levels is controlled by parameters, whose values may differ across subjects. In contrast to "ideal" Bayesian learning models, which prescribe a fixed process for any agent, this allows for the representation of inter-individual differences in behavior and how it is influenced by uncertainty. This variation is cast in terms of prior beliefs about the parameters coupling hierarchical levels in the generative model.

A major goal of our work was to eschew the complicated integrals in exact Bayesian inference and instead derive analytical update equations with algorithmic efficiency and biological plausibility. For this purpose, we used an approximate (variational) Bayesian approach, under a mean field assumption and a novel approximation to the posterior energy function. The resulting single-step, trial-by-trial update equations have several important properties:

i. They have an analytical form and are extremely efficient, allowing for real-time inference.

ii. They are biologically plausible in that the mathematical operations required for calculating the updates are fairly basic and could be performed by single neurons (Herz et al., 2006; London and Hausser, 2005).

iii. Their structure is remarkably similar to update equations from standard RL models; this enables an interpretation that places

RL heuristics, such as learning rate or prediction error, in a principled (Bayesian) framework.

iv. The model parameters determine processes, such as precision-weighting of prediction errors, which play a key role in current theories of normal and pathological learning and may relate to specific neuromodulatory mechanisms in the brain (see below).

v. They can accommodate states of either discrete or continuous nature and can deal with deterministic and probabilistic mappings between environmental causes and perceptual consequences (i.e., situations with and without sensory uncertainty).

Crucially, the closed-form update equations do not depend on the details of the model but only on its hierarchical structure and the assumptions on which the mean field and quadratic approximation to the posteriors rest. Our method of deriving the update equations may thus be adopted for the inversion of a large class of models. Above, we demonstrated this anecdotally by providing update equations for two extensions of the original model, which accounted for sensory states of a continuous (rather than discrete) nature and sensory uncertainty, respectively.

11.2 Alternatives to the HGF

As alternatives to our variational scheme, one could deal with the complicated integrals of Bayesian inference by sampling methods or eschew them altogether and use simpler reinforcement learning schemes. We did not pursue these options because we wanted to take a principled approach to individual learning under uncertainty; i.e., one that rests on the inversion of a full Bayesian generative model. Furthermore, we wanted to avoid sampling approximations because of the computational burden they impose. Although it is conceivable that neuronal populations could implement sampling methods, it is not clear how exactly they would do that and at what temporal and energetic cost (Beck et al., 2008; Deneve, 2008; Yang and Shadlen, 2007).

We would like to emphasize that the examples of update equations derived here can serve as the building blocks for those of more complicated models. For example, if we have more than two categories at the first level, this can be accommodated by additional random walks at the second and subsequent levels; at those levels, Equations (38) and (40) have a straightforward interpretation in n dimensions. Inference using our update scheme is thus possible on n-categorical discrete states, n-dimensional unbounded continuous states, and (by logarithmic or logistic transformation of variables) n-dimensional bounded continuous states.

One specific problem that has been addressed with Bayesian methods in the recent past concerns online inference of "changepoints", i.e. sudden changes in the statistical structure of the environment (Corrado et al., 2005; Fearnhead and Liu, 2007; Krugel et al., 2009; Wilson et al., 2010; Steyvers et al., 2009). Our generative model based on hierarchically coupled random walks describes belief updating without representing such discrete changepoints explicitly. As illustrated by the simulations in Figures 5-8, this does not diminish its ability to deal with volatile environments. See also previous studies where similar models were applied to data generated by changepoint models (Behrens et al., 2007, 2008) or where RL-type update models were equipped with an adjustable learning rate in order to deal with sudden changes in the environment (Krugel et al., 2009; Steyvers et al., 2009). One such sudden change in the environment that is nicely picked up in the application of our model to the empirical exchange rate data is the outbreak of the Greek financial crisis in spring 2010. The sudden realization of the financial markets that Greece was insolvent led to a flight into the U.S. Dollar which is reflected by a sharply increasing value of the Dollar against the Swiss franc visible in the lower panel of Figure 11. Because this sudden rise of the Dollar is unexpected, it immediately leads to a jump in the agent's belief about the volatility of its environment, as is clearly visible in the upper panel of Figures 11 or 13. In this manner, a sudden event akin to a changepoint is detected without representations of changepoints being an explicit component of the model.

Clearly, our approach is not the first that has tried to derive tractable update equations from the full Bayesian formulation of learning. Although

not described in this way in the original work, even the famous Kalman filter can be interpreted as a Bayesian scheme with RL-like update properties but is restricted to relatively simple (non-hierarchical) learning processes in stable environments (Kalman, 1960). Notably, none of the previous Bayesian learning models we know leads to analytical one-step update equations without resorting to additional assumptions that are specifically tailored for the update equations (Steyvers et al., 2009) or the learning rate (Krugel et al., 2009). In contrast, in our scheme, the update equations and their critical components, such as learning rate or prediction error, emerge naturally by inverting a full Bayesian generative model of arbitrary hierarchical depth under a generic mean field reduction. That is, once we have specified the nature of our approximate inversion, the update equations are fully defined and do not require any further assumptions. This distinguishes our framework conceptually and mathematically from any previously suggested approach to Bayesian learning we are aware of.

The distinction between hidden states, which vary in time and are the dynamic components of the agent's model of the world, and parameters, which are time-invariant and encode stable subject-specific learning styles, is a key component of the HGF. One might compare this to classical reinforcement learning models where value estimates (states) are updated dynamically while the learning rate is an invariant parameter. In our case, however, the (implicit) learning rate is dynamic and results from an interaction between states and parameters: the latter determine how higher-level states influence lower-level ones. This effect of the static parameters on dynamic cross-level coupling can be seen directly from the update equations above (e.g., Eqs (35) and (39)) and is illustrated in Figures 5 and 10 where the learning rate visibly changes while the parameters are fixed. In other words, subject-specific learning mechanisms, represented by cross-level coupling in the HGF, have both dynamic (higher-level states) and static (parameters) components in the HGF.

One could, of course, consider alternative formulations of our model in which individual learning mechanisms, determined by the coupling across levels, are encoded entirely by states. There is a simple reason why we did

not pursue this alternative. Clearly, modeling dynamic aspects of learning, such as rapid updating of learning rates, does require a representation involving states (see above). On the other hand, within an individual agent's brain, the physiological mechanism underlying this coupling must obey some general principles that have been shaped both by (life-long) experience and genetic background (cf. reinforcement learning depends on individual genotype (Frank et al., 2007, 2009; Krugel et al., 2009). Such stable subject-specific learning mechanisms could be represented in two ways. One could choose a relatively deep hierarchical model with high-level states that change very slowly. Alternatively, these mechanisms can be represented by time-invariant parameters. We chose the latter option simply because it provides a more concise and interpretable summary of subject-specific learning mechanisms. For example, when quantifying individual differences in computational learning mechanisms as a function of individual differences in physiology (e.g., pharmacological treatment) or genetics, it is not only statistically easier to deal with time-invariant parameters (i.e., a single number per subject) rather than temporal trajectories of states, but the results are also more readily interpretable.

It should be emphasized that the idea of fitting learning models to subject-specific data and using the ensuing individual parameter estimates for assessing inter-individual variability is not new and has been pursued by many previous studies (e.g., Steyvers and Brown, 2006; 2007, 2009; Krugel et al., 2009). This, however, is less straightforward with those "ideal" Bayesian models that have no free parameters; in this case, parameters can only be taken from adjunct models in which ideal Bayesian models are often embedded (e.g., observation or decision models; (cf. Brodersen et al., 2008)). The novelty of our approach is that we transform, by variational approximation, an ideal Bayesian learner into a near-optimal scheme in which parameters represent individual learning traits as an integral part of Bayesian learning. These parameters shape the ensuing update equations which are analytical and have an RL-like structure. By combining the principled nature of Bayesian approaches and the practical ease of RL models, we hope that our approach will facilitate future empirical investigations of individual variability.

11.3 Limitations of the HGF

Our update scheme also has its limitations. The most important of these is that it depends on the variational energies being approximately quadratic. If they are not, the approximate posterior implied by our update equations might bear little resemblance (e.g., in terms of Kullback-Leibler divergence) to the true posterior. Specifically, the update fails if the curvature of the variational energy at the expansion point is negative (which implies that the conditional variance is negative; see Eq. (56)). According to the update Eq. (34), σ_2 can never become negative; $\sigma_3 = 1/\pi_3$ however could become negative according to Eq. (41). However, the simulations and application to empirical data in this thesis suggest that this is not a problem in practice.

11.4 Validation of the HGF

While the theoretical soundness of its derivation and the intuitive appeal of its update equations provide strong a priori reasons to have confidence in it, this is certainly not enough.

I have therefore striven to provide empirical validation for the HGF from multiple sources. The simplest validation has come in the form of the simulations in Section 4.5 that demonstrate model behavior under different parameterizations (priors). These simulations confirmed the computational efficiency of the HGF: the simulations in Figures 5-8 (with 320 trials) each take about 5 ms on a standard laptop computer. Furthermore, they demonstrate that changes in any of the parameters lead to plausible changes in the evolution of the states (i.e., as predicted from the structure of the model) and that each parameter produces distinctly different behavior.

A second and more important validation has come from the simulation study of parameter estimation in Section 7.4. We could show there that at least four different optimization strategies are able to recover parameter values with tolerable accuracy as long as response noise is not extreme.

A third validation comes from the application to saccadic reaction times in the study reported in Chapter 9, where we deal with experimental data for

the first time. One crucial aspect of this is the validation of the number of levels in the simple example of HGF that has been used throughout this thesis: human subjects seem to exhibit behavior that can in fact better be described by a three-level model than by a less complex one.

Finally, the observed iterated trust game study provides proof of concept that the HGF can be used for inference on complex mental states and may be a valuable tool for the construction of regressor in fMRI studies that capture trajectories of quantities relevant to reward processing, learning, and decision making under uncertainty.

11.5 The HGF as a model of adaptive and maladaptive inference

Maladaptive behavior, due to inappropriate learning and decision making, is at the heart of most psychiatric diseases, and our framework may be particularly useful for modeling the underlying mechanisms. There are two complementary approaches one might consider: phenomenological and neurophysiological. To illustrate a phenomenological approach, I will consider extreme settings of the parameters in terms of psychopathology. In the example of Figures 5-8, variations in the parameters can explain a spectrum of different types of inference, some of which may be interpreted as aberrant or even pathological. Given the scenario in Figure 8, we could adopt the anthropomorphic interpretation of the agent and interpret underconfidence about estimates of environmental volatility (i.e., high σ_3) as the possible cause of anxiety. In other words, knowing that the world is changing quickly is frightening enough, but being uncertain about the extent of this change may be even more upsetting. Anxiety of this sort is often observed prior to (or in the early phase of) psychotic episodes (Häfner et al., 1998). One way to reduce anxiety (that is, to reduce the effects of high σ_3 due to abnormally low κ), would be to reduce ϑ, leading to a scenario akin to that in Figure 6. This, however, would induce a rigid high-level belief that is impervious to prediction error from the lower level. Rigid high-level priors of this sort that provide inappropriate predictions for lower levels may provide a metaphor for delusions and hallucinations that constitute the positive symptoms of schizophrenia. In contrast, negative symptoms could be related to the scenario in Figure 7:

here, a reduction in ω renders the agent completely passive, such that new information is barely taken in and only weakly processed. Notably, in these simple and anecdotal simulations, I chose some parameter settings that lead to superficially similar behavior (e.g., Figs 6 and 8). While this indicates some degree of interdependence among the parameters, this does not mean that the parameters are non-identifiable. Informally, one can intuit this by noting the obvious differences expressed in the evolution of higher-level states of the model; these will be expressed in different behavioral predictions, given a suitably chosen sequence of stimuli. When fitting the model to empirical data, one can test for parameter identifiability more formally using a sensitivity analysis or, equivalently but more conveniently, their posterior covariance.

From a neurophysiological perspective, it has been proposed that dopamine might not encode the prediction error on value (Schultz et al., 1997) but instead the value of prediction error, i.e. the precision-weighting of prediction errors (Friston, 2009). In our model, this process is represented, at the second level, by the parameters κ and ω. It is apparent from the definition Eq. (39) and the update Eq. (34) that these parameters influence the precision of the prediction on the next trial, the precision of the posterior belief, and the learning rate. If dopaminergic midbrain activity encodes the conditional (posterior) precision of beliefs, this dopaminergic activity should be reflected by estimates of κ and ω, obtained from behavioral, fMRI or electrophysiological data. This hypothesis can be tested using neuropharmacological experiments. In short, by harvesting subject-specific parameter estimates for group analyses of physiological measurements, hierarchical generative models (of the sort considered in this work) could be used to test hypotheses about the relations between computational and physiological processes. We are currently pursuing this approach in ongoing research. Alternatively, one can also use our model for analyses at the subject level: the sequence of inferred hidden states, as represented by their sufficient statistics μ and σ, can be used as predictor variables in analyses of fMRI, EEG or behavioral data to shed light on the neurophysiological correlates of inference and learning (cf. den Ouden et al., 2010).

12 SUMMARY

This thesis has introduced the hierarchical Gaussian filter, a novel and generic framework for approximate Bayesian inference with computationally efficient and interpretable closed-form update equations. Simulations show that this approach is applicable to a range of situations beyond classical RL, including inductive inference on discrete and continuous states and situations with perceptual ambiguity. Crucially, my approach accommodates inter-individual differences, in terms of prior beliefs about key model parameters, and quantifies their computational effects: Some of these parameters may map to neurophysiological (neuromodulatory) mechanisms that have been implicated in the neurobiology of learning and psychopathology. As such, it may be a useful framework for modeling individual differences in behavior and to formally characterize behavioral stereotypes and pathophysiologically distinct subgroups in psychiatric spectrum diseases (Stephan et al., 2009a).

We have further validated the HGF in several applications. The first has been a simulation study where the combination of the HGF with a decision model and the subsequent recovery of parameters underlying simulations was proven to be possible, breaking down only at very high levels of response noise. The second validation has been an application to saccadic reaction time data from human subjects. In this application, the HGF has proven able to serve as the underlying perceptual model for various decision models, enabling us to adjudicate among them by Bayesian model selection. Furthermore, in this application, it has been shown that the three-level HGF outperforms simpler versions of itself, suggesting that subjects used hierarchical inference even in this relatively simple task. Finally, I have shown that the HGF can serve as an inferential mechanism describing complex cognitive states as they occur in the observation of an iterated trust game. States inferred by the HGF correlate strongly with neuronal activation as measured by fMRI, suggesting that the HGF captures quantities that are represented by the human brain during reward processing, learning, and decision making.

APPENDICES

A. Variational Bayes

The term *variational Bayes* refers to the approximation of posterior probability densities by means of a variational optimization of a free-energy bound on the log-model evidence. The free energy bound derives its name from a mathematically analogous bound in statistical mechanics (Feynman, 1972, 47).

Given a model $p(u, x|\chi)$ with observed data u, hidden states x, and fixed parameters χ, any arbitrary probability density $q(x)$ of the hidden states gives rise to a lower bound on the logarithm of the model evidence $p(u|\chi)$:

$$\begin{aligned} \ln p(u|\chi) &= \ln \int p(u, x|\chi) \mathrm{d}x = \ln \int q(x) \frac{p(u, x|\chi)}{q(x)} \mathrm{d}x \\ &\geq \int q(x) \ln \left(\frac{p(u, x|\chi)}{q(x)} \right) \mathrm{d}x \stackrel{\text{def}}{=} \mathcal{F} \end{aligned} \tag{127}$$

where we have used Jensen's inequality (which results from the concavity of the logarithm). The right-hand side of the inequality can be rearranged in two ways:

$$\begin{aligned} \mathcal{F} &= \int q(x) \ln \left(\frac{p(u, x|\chi)}{q(x)} \right) \mathrm{d}x \\ &= \int q(x) \ln p(u, x|\chi) \mathrm{d}x - \int q(x) \ln q(x) \mathrm{d}x \\ &\stackrel{\text{def}}{=} \langle \ln p(u, x|\chi) \rangle_q + S[q] \end{aligned} \tag{128}$$

and

$$\begin{aligned}\mathcal{F} &= \int q(x) \ln\left(\frac{p(u,x|\chi)}{q(x)}\right) \mathrm{d}x \\ &= \int q(x) \ln\left(\frac{p(x|u,\chi)p(u|\chi)}{q(x)}\right) \mathrm{d}x \\ &= \int q(x) \ln\left(\frac{p(x|u,\chi)}{q(x)}\right) \mathrm{d}x \\ &\quad + \int q(x) \ln p(u|\chi) \mathrm{d}x \\ &= -D[q(x), p(x|u,\chi)] + \ln p(u|\chi)\end{aligned} \tag{129}$$

The first rearrangement, which in physical terms represents expected negative energy plus entropy (i.e., negative free energy $\mathcal{F}$), can be calculated because the model $p(u,x|\chi)$ and the density $q(x)$ are known. The second rearrangement shows that if we choose $q(x)$ such that $\mathcal{F}$ is maximized, the expression $D[q(x), p(x|u,\chi)]$, which is always positive, is minimized. This expression is the Kullback-Leibler divergence of $p(x|u,\chi)$ from $q(x)$. The better $q(x)$ approximates the posterior $p(x|u,\chi)$, the smaller the divergence; vanishing when the two are equal. Maximizing the (negative) free energy bound on the log-model evidence by varying $q(x)$ is therefore the same as approximating the true posterior $p(x|u,\chi)$ by $q(x)$.

In other words, we find the posterior by introducing $\mathcal{F}$ and q and maximizing $\mathcal{F}$ as a functional $\mathcal{F}[q|u,\chi]$ of q. This procedure is called *variational Bayes* because finding the maximum of $\mathcal{F}$ under variations of q involves variational calculus and leads to an approximation to the posterior distribution in Bayesian inference. The precise mechanics of maximization of $\mathcal{F}$ are explained in Appendix B.

Another way to optimize $q(x)$ is to introduce parametric constraints on q, such that $q(x) = q(x|\lambda)$ with parameters λ. For example, q could be assumed to be Gaussian with sufficient statistics λ. $\mathcal{F}$ can then be seen as a function of λ (cf. Eq. (128)):

$$\mathcal{F}(\lambda|u,\chi) = \int q(x|\lambda)\ln p(u,x|\chi)\mathrm{d}x - \int q(x|\lambda)\ln q(x|\lambda)\mathrm{d}x \qquad (130)$$

According to Eq. (129), the parameter values λ^* that maximize $\mathcal{F}$ with respect to λ now give us the posterior $q(x|\lambda^*)$ that approximates the true posterior $p(x|u,\chi)$ most closely in the sense that it minimizes their Kullback-Leibler divergence under the parametric constraints on q:

$$\begin{aligned} &\lambda^* = \arg\max_{\lambda} \mathcal{F}(\lambda|u,\chi) \\ &\Longrightarrow \quad q(x|\lambda^*) \approx p(x|u,\chi). \end{aligned} \qquad (131)$$

B. Variational optimization of the free energy under the mean field approximation

Given a model $p(u, x|\chi)$, our goal is to optimize an arbitrary density $q(x)$ so that it maximizes negative free energy $\mathcal{F}$ and therefore approximates the posterior density $p(x|u, \chi)$ most closely (cf. Appendix A).

Under the *mean field approximation,* we assume that $q(x)$ can be factored into the densities of pairwise disjoint index sets j:

$$q(x) = \prod_{j \in J} q(x_j) \tag{132}$$

where J is a set of index sets. To simplify our notation, we set

$$x_{\setminus i} \stackrel{\text{def}}{\equiv} \{x_j\}_{j \neq i}, \quad q_i \stackrel{\text{def}}{\equiv} q(x_i), \quad q_{\setminus i} \stackrel{\text{def}}{\equiv} \prod_{j \neq i} q(x_j), \qquad i, j \in J \tag{133}$$

Since $\mathcal{F}$ is a functional of q, the mean field approximation allows us to write $\mathcal{F}[q] = \mathcal{F}[q_i, q_{\setminus i}]$ and thus to optimize each of the q_i separately. When optimizing q, we have to observe the constraints that

$$\int q(x_j) \mathrm{d}x_j = 1 \qquad \forall j \in J \tag{134}$$

We account for these constraints by introducing Lagrange multipliers λ_j and defining

$$
\begin{aligned}
\tilde{\mathcal{F}}[q] &\stackrel{\text{def}}{=} \mathcal{F}[q] + \sum_j \lambda_j \left(\int q(x_j) \mathrm{d}x_j - 1 \right) \\
&= \int q(x) \ln p(u, x|\chi) \, \mathrm{d}x - \int q(x) \ln q(x) \, \mathrm{d}x \\
&\quad + \sum_j \lambda_j \left(\int q(x_j) \mathrm{d}x_j - 1 \right) \\
&= \int q(x_i) \int q(x_{\setminus i}) \ln p(u, x|\chi) \, \mathrm{d}x_{\setminus i} \, \mathrm{d}x_i \\
&\quad - \int q(x_i) \int q(x_{\setminus i}) \ln \left(q(x_i) q(x_{\setminus i}) \right) \mathrm{d}x_{\setminus i} \, \mathrm{d}x_i \\
&\quad + \lambda_i \left(\int q(x_i) \mathrm{d}x_i - 1 \right) + \sum_{\setminus i} \lambda_{\setminus i} \left(\int q(x_{\setminus i}) \mathrm{d}x_{\setminus i} - 1 \right) \\
&= \tilde{\mathcal{F}}[q_i, q_{\setminus i}]
\end{aligned}
\tag{135}
$$

To optimize $\tilde{\mathcal{F}}[q_i, q_{\setminus i}]$ with respect to q_i, we make use of the *fundamental lemma of variational calculus,* which states that

$$
\left. \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \right|_{\varepsilon=0} \tilde{\mathcal{F}}[q_i + \varepsilon \varphi_i, q_{\setminus i}] \stackrel{\text{def}}{=} \langle \frac{\delta \tilde{\mathcal{F}}}{\delta q_i} \rangle_{\varphi_i} = 0 \quad \forall \varphi_i \quad \Leftrightarrow \frac{\delta \tilde{\mathcal{F}}}{\delta q_i} \equiv 0, \tag{136}
$$

where φ_i is a test function and the equation on the left of the mutual implication contains the definition of the functional derivative $\frac{\delta}{\delta q_i}$. The rest, as they say, is algebra:

$$
\begin{aligned}
&\left. \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \right|_{\varepsilon=0} \tilde{\mathcal{F}}[q_i + \varepsilon \varphi_i, q_{\setminus i}] = \\
&\quad = \left. \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \right|_{\varepsilon=0} \int \left(q(x_i) + \varepsilon \varphi(x_i) \right) \int q(x_{\setminus i}) \ln p(u, x|\chi) \, \mathrm{d}x_{\setminus i} \, \mathrm{d}x_i
\end{aligned}
\tag{137}
$$

$$
\begin{aligned}
&-\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0}\int\big(q(x_i)+\varepsilon\varphi(x_i)\big)\int q(x_{\setminus i})\ln\Big(\big(q(x_i)+\varepsilon\varphi(x_i)\big)q(x_{\setminus i})\Big)\,\mathrm{d}x_{\setminus i}\,\mathrm{d}x_i\\
&+\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0}\lambda_i\left(\int\Big(\big(q(x_i)+\varepsilon\varphi(x_i)\big)\Big)\mathrm{d}x_i-1\right)\\
&+\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0}\sum_{\setminus i}\lambda_{\setminus i}\left(\int q(x_{\setminus i})\mathrm{d}x_{\setminus i}-1\right)\\
=&\int\varphi(x_i)\int q(x_{\setminus i})\ln p(u,x|\chi)\,\mathrm{d}x_{\setminus i}\,\mathrm{d}x_i\\
&-\int\varphi(x_i)\int q(x_{\setminus i})\ln\Big(q(x_i)q(x_{\setminus i})\Big)\,\mathrm{d}x_{\setminus i}\,\mathrm{d}x_i\\
&-\int\varphi(x_i)\int q(x_{\setminus i})\frac{\varphi(x_i)q(x_{\setminus i})}{q(x_i)q(x_{\setminus i})}\mathrm{d}x_{\setminus i}\,\mathrm{d}x_i+\lambda_i\int\varphi(x_i)dx_i\\
=&\int\varphi(x_i)\langle\ln p(u,x|\chi)\rangle_{q_{\setminus i}}\mathrm{d}x_i-\int\varphi(x_i)\ln q(x_i)\underbrace{\int q(x_{\setminus i})\mathrm{d}x_{\setminus i}}_{=1}\mathrm{d}x_i\\
&-\int\varphi(x_i)\underbrace{\int q(x_{\setminus i})\ln q(x_{\setminus i})\,\mathrm{d}x_{\setminus i}}_{\cong S[q_{\setminus i}]}\mathrm{d}x_i-\int\varphi(x_i)\underbrace{\int q(x_{\setminus i})\mathrm{d}x_{\setminus i}}_{=1}\mathrm{d}x_i\\
&+\lambda_i\int\varphi(x_i)dx_i\\
=&\int\varphi(x_i)\left(\underbrace{\langle\ln p(u,x|\chi)\rangle_{q_{\setminus i}}}_{\cong I(x_i)}-\ln q(x_i)\underbrace{-S[q_{\setminus i}]-1+\lambda_i}_{\cong-\ln Z_i=\text{const.}(x_i)}\right)dx_i\\
=&\langle I(x_i)-\ln q(x_i)-\ln Z_i\rangle_{\varphi_i}=\langle\frac{\delta\tilde{\mathcal{F}}}{\delta q_i}\rangle_{\varphi_i}
\end{aligned}
$$

The optimal q_i is then characterized by

$$
\begin{aligned}
&\frac{\delta\tilde{\mathcal{F}}}{\delta q_i}=I(x_i)-ln\,q(x_i)-ln\,Z_i\equiv 0\\
&\Longrightarrow\quad q(x_i)=\frac{1}{Z_i}\exp\big(I(x_i)\big).
\end{aligned}
\tag{138}
$$

C. Calculation of the variational energies

As in the main text, I first deal with the three-level example model before turning to the general case. Using the notation

$$\langle f(x)\rangle_{q_j} \stackrel{\text{def}}{=} \int q(x_j) f(x) \mathrm{d}x_j \tag{139}$$

for the expectation of $f(x)$ under $q(x_j)$ together with the definition of the model described graphically in Figure 2, we can rewrite Eq. (27) as a sum of expectations

$$\begin{aligned} I\left(x_i^{(k)}\right) &= \langle \ln p\left(u^{(k)}, x^{(k)} \middle| \chi\right)\rangle_{q_{\setminus i}} \\ &= \langle \ln p\left(u^{(k)} \middle| x_1^{(k)}\right)\rangle_{q_{\setminus i}} + \langle \ln p\left(x_1^{(k)} \middle| x_2^{(k)}\right)\rangle_{q_{\setminus i}} \\ &\quad + \langle \ln p\left(x_2^{(k)} \middle| x_3^{(k)}, \kappa, \omega\right)\rangle_{q_{\setminus i}} + \langle \ln p\left(x_3^{(k)} \middle| \vartheta\right)\rangle_{q_{\setminus i}} \end{aligned} \tag{140}$$

The term $\ln p\left(u^{(k)} \middle| x_1^{(k)}\right)$ is included here to cover also models with sensory uncertainty as discussed in Section 5.1. In cases without such uncertainty (sc. $x_1 = u$), the term vanishes.

$p\left(x_1^{(k)} \middle| x_2^{(k)}\right)$ can be taken directly from Eq. (10), while

$$
\begin{aligned}
p\left(x_2^{(k)}\middle|x_3^{(k)},\kappa,\omega\right) &= \int p\left(x_2^{(k)}\middle|x_2^{(k-1)},x_3^{(k)},\kappa,\omega\right)p\left(x_2^{(k-1)}\right)\mathrm{d}x_2^{(k-1)} \\
&\approx \int p\left(x_2^{(k)}\middle|x_2^{(k-1)},x_3^{(k)},\kappa,\omega\right)q\left(x_2^{(k-1)}\right)\mathrm{d}x_2^{(k-1)} \\
&= \int \mathcal{N}\left(x_2^{(k)};x_2^{(k-1)},\exp\left(\kappa x_3^{(k)}+\omega\right)\right)\mathcal{N}\left(x_2^{(k-1)};\mu_2^{(k-1)},\sigma_2^{(k-1)}\right)\mathrm{d}x_2^{(k-1)} \\
&= \mathcal{N}\left(x_2^{(k)};\mu_2^{(k-1)},\sigma_2^{(k-1)}+\exp\left(\kappa x_3^{(k)}+\omega\right)\right)
\end{aligned}
\tag{141}
$$

and

$$
\begin{aligned}
p\left(x_3^{(k)}\middle|\vartheta\right) &= \int p\left(x_3^{(k)}\middle|x_3^{(k-1)},\vartheta\right)p\left(x_3^{(k-1)}\right)\mathrm{d}x_3^{(k-1)} \\
&\approx \int p\left(x_3^{(k)}\middle|x_3^{(k-1)},\vartheta\right)q\left(x_3^{(k-1)}\right)\mathrm{d}x_3^{(k-1)} \\
&= \int \mathcal{N}\left(x_3^{(k)};x_3^{(k-1)},\vartheta\right)\mathcal{N}\left(x_3^{(k-1)};\mu_3^{(k-1)},\sigma_3^{(k-1)}\right)\mathrm{d}x_3^{(k-1)} \\
&= \mathcal{N}\left(x_3^{(k)};\mu_3^{(k-1)},\sigma_3^{(k-1)}+\vartheta\right)
\end{aligned}
\tag{142}
$$

This is akin to the "prediction step" in the Kalman filter literature since it gives us the predictive densities for $x_i^{(k)}$ given inputs up to time $k-1$; it is this point where the mean field approximation unfolds its effect.

We only need to determine the $I\left(x_i^{(k)}\right)$ up to a constant because any constant term can always be absorbed into $\mathcal{Z}_i$ when forming $\hat{q}\left(x_i^{(k)}\right)$ according to Eq. (26). For the three levels of our example model, this means

$$
I\left(x_1^{(k)}\right) = \left\langle \ln p\left(u^{(k)}\middle|x_1^{(k)}\right)\right\rangle_{q_{\setminus 1}} + \left\langle \ln p\left(x_1^{(k)}\middle|x_2^{(k)}\right)\right\rangle_{q_{\setminus 1}} + \text{const.} \tag{143}
$$

$$I\left(x_2^{(k)}\right) = \langle \ln p\left(x_1^{(k)} \middle| x_2^{(k)}\right)\rangle_{q_{\setminus 2}} + \langle \ln p\left(x_2^{(k)} \middle| x_3^{(k)}, \kappa, \omega\right)\rangle_{q_{\setminus 2}} + \text{const.} \tag{144}$$

$$I\left(x_3^{(k)}\right) = \langle \ln p\left(x_2^{(k)} \middle| x_3^{(k)}, \kappa, \omega\right)\rangle_{q_{\setminus 3}} + \langle \ln p\left(x_3^{(k)} \middle| \vartheta\right)\rangle_{q_{\setminus 3}} + \text{const.} \tag{145}$$

With two exceptions, all integrals on the right-hand sides above can be solved analytically in all cases considered here, including sensory uncertainty and inference on continuous-valued states.

The two exceptions are the following: first, to solve $\langle \ln p\left(x_1^{(k)} \middle| x_2^{(k)}\right)\rangle_{q_{\setminus 1}}$, we expand $\ln s\left(x_2^{(k)}\right)$ to second order around the prior expectation $\mu_2^{(k-1)}$of $x_2^{(k)}$:

$$\begin{aligned} \ln s\left(x_2^{(k)}\right) \approx\ & \ln s\left(\mu_2^{(k-1)}\right) + \left(1 - s\left(\mu_2^{(k-1)}\right)\right)\left(x_2^{(k)} - \mu_2^{(k-1)}\right) \\ & + \frac{1}{2}\left(s\left(\mu_2^{(k-1)}\right)^2 - s\left(\mu_2^{(k-1)}\right)\right)\left(x_2^{(k)} - \mu_2^{(k-1)}\right)^2 \end{aligned} \tag{146}$$

Second, to solve $\langle \ln p\left(x_2^{(k)} \middle| x_3^{(k)}, \kappa, \omega\right)\rangle_{q_{\setminus 2}}$, we take

$$\left\langle \frac{1}{\sigma_2^{(k-1)} + \exp\left(\kappa \cdot x_3^{(k)} + \omega\right)} \right\rangle_{q_{\setminus 2}} \approx \frac{1}{\sigma_2^{(k-1)} + \exp\left(\kappa \cdot \mu_3^{(k-1)} + \omega\right)} \tag{147}$$

The result of doing the integrals in Eqs (143)-(145) is given in Eqs (28)-(30).

In the general case, I will give the calculation in full detail. For convenience, I restate the relevant definitions:

$$I\left(x_i^{(k)}\right) \stackrel{\text{def}}{=} \int q\left(x_{\setminus i}^{(k)}\right) \ln p\left(x_i^{(k)}, x_{\setminus i}^{(k)}, \chi \middle| u^{(1\ldots k)}\right) \mathrm{d}x_{\setminus i}^{(k)},$$

$$\begin{aligned} q\left(x_{\setminus i}^{(k)}\right) &\stackrel{\text{def}}{=} \prod_{j \neq i} q(x_j), \\ x_{\setminus i} &\stackrel{\text{def}}{=} \left\{x_j\right\}_{j \neq i}, \\ \chi &\stackrel{\text{def}}{=} \{\kappa_i, \omega_i, \vartheta\}_{1<i<n-1}. \end{aligned} \tag{148}$$

Unpacking this, we see

$$\begin{aligned} p\left(x_i^{(k)}, x_{\setminus i}^{(k)}, \chi \middle| u^{(1\ldots k)}\right) &= \int \prod_{j=1}^{n} p\left(x_j^{(k)} \middle| x_j^{(k-1)}, x_{\setminus j}^{(k)}, \chi\right) p\left(x_j^{(k-1)} \middle| u^{(1\ldots k)}\right) \mathrm{d}x_j^{(k-1)}, \\ p\left(x_j^{(k)} \middle| x_j^{(k-1)}, x_{\setminus j}^{(k)}, \chi\right) &= \mathcal{N}\left(x_j^{(k)}; x_j^{(k-1)}, \exp\left(\kappa_j x_{j+1}^{(k)} + \omega_j\right)\right), \\ p\left(x_j^{(k-1)} \middle| u^{(1\ldots k)}\right) &\approx \mathcal{N}\left(x_j^{(k-1)}; \mu_j^{(k-1)}, \sigma_j^{(k-1)}\right), \end{aligned} \tag{149}$$

where the last line reflects the mean field approximation. Performing this integral, we obtain

$$\begin{aligned} p\left(x_i^{(k)}, x_{\setminus i}^{(k)}, \chi \middle| \mu_1^{(1\ldots k)}\right) &= \prod_{j=1}^{n} \mathcal{N}\left(x_j^{(k)}; \mu_j^{(k-1)}, \sigma_j^{(k-1)} + \exp\left(\kappa_j x_{j+1}^{(k)} + \omega_j\right)\right), \\ \exp\left(\kappa_n x_{n+1}^{(k)} + \omega_n\right) &\equiv \vartheta. \end{aligned} \tag{150}$$

By substituting this into the definition of I in Eq. (148), we get

$$
\begin{aligned}
I\left(x_i^{(k)}\right) &= \int q\left(x_{\setminus i}^{(k)}\right) \ln p\left(x_i^{(k)}, x_{\setminus i}^{(k)}, \chi \middle| \mu_1^{(1\ldots k)}\right) \mathrm{d}x_{\setminus i}^{(k)} \\
&= \int q\left(x_{\setminus i}^{(k)}\right) \ln\left(\prod_{j=1}^{n} \mathcal{N}\left(x_j^{(k)}; \mu_j^{(k-1)}, \sigma_j^{(k-1)} + \exp\left(\kappa_j x_{j+1}^{(k)} + \omega_j\right)\right)\right) \mathrm{d}x_{\setminus i}^{(k)} \\
&= \sum_{j=1}^{n} \int q\left(x_{\setminus i}^{(k)}\right) \ln \mathcal{N}\left(x_j^{(k)}; \mu_j^{(k-1)}, \sigma_j^{(k-1)} + \exp\left(\kappa_j x_{j+1}^{(k)} + \omega_j\right)\right) \mathrm{d}x_{\setminus i}^{(k)} \\
&= \int q\left(x_{\setminus i}^{(k)}\right) \ln \mathcal{N}\left(x_{i-1}^{(k)}; \mu_{i-1}^{(k-1)}, \sigma_{i-1}^{(k-1)} + \exp\left(\kappa_{i-1} x_i^{(k)} + \omega_{i-1}\right)\right) \mathrm{d}x_{\setminus i}^{(k)} \\
&\quad + \int q\left(x_{\setminus i}^{(k)}\right) \ln \mathcal{N}\left(x_i^{(k)}; \mu_i^{(k-1)}, \sigma_i^{(k-1)} + \exp\left(\kappa_i x_{i+1}^{(k)} + \omega_i\right)\right) \mathrm{d}x_{\setminus i}^{(k)} \\
&\qquad + \text{const.} \\
&= \int q\left(x_{i-1}^{(k)}\right) \ln \mathcal{N}\left(x_{i-1}^{(k)}; \mu_{i-1}^{(k-1)}, \sigma_{i-1}^{(k-1)} + \exp\left(\kappa_{i-1} x_i^{(k)} + \omega_{i-1}\right)\right) \mathrm{d}x_{i-1}^{(k)} + \\
&\quad + \int q\left(x_{i+1}^{(k)}\right) \ln \mathcal{N}\left(x_i^{(k)}; \mu_i^{(k-1)}, \sigma_i^{(k-1)} + \exp\left(\kappa_i x_{i+1}^{(k)} + \omega_i\right)\right) \mathrm{d}x_{i+1}^{(k)} \\
&\qquad + \text{const.}
\end{aligned}
\tag{151}
$$

We thus have to solve these last two integrals. The first one can be solved analytically:

$$
\begin{aligned}
&\int q\left(x_{i-1}^{(k)}\right) \ln \mathcal{N}\left(x_{i-1}^{(k)}; \mu_{i-1}^{(k-1)}, \sigma_{i-1}^{(k-1)} + \exp\left(\kappa_{i-1} x_i^{(k)} + \omega_{i-1}\right)\right) \mathrm{d}x_{i-1}^{(k)} \\
&= -\frac{1}{2} \ln\left(\sigma_{i-1}^{(k-1)} + \exp\left(\kappa_{i-1} x_i^{(k)} + \omega_{i-1}\right)\right) \\
&\qquad - \frac{1}{2} \frac{\sigma_i^{(k)} + \left(\mu_i^{(k)} - \mu_i^{(k-1)}\right)^2}{\sigma_{i-1}^{(k-1)} + \exp\left(\kappa_{i-1} x_i^{(k)} + \omega_{i-1}\right)}
\end{aligned}
\tag{152}
$$

To solve the second integral, we take

$$\left\langle \frac{1}{\sigma_i^{(k-1)} + \exp\left(\kappa_i x_{i+1}^{(k)} + \omega_i\right)} \right\rangle_{q_{i+1}} \approx \frac{1}{\sigma_i^{(k-1)} + \exp\left(\kappa_i \mu_{i+1}^{(k-1)} + \omega_i\right)}, \tag{153}$$

yielding

$$\begin{aligned} &\int q\left(x_{i+1}^{(k)}\right) \ln \mathcal{N}\left(x_i^{(k)}; \mu_i^{(k-1)}, \sigma_i^{(k-1)} + \exp\left(\kappa_i x_{i+1}^{(k)} + \omega_i\right)\right) \mathrm{d}x_{i+1}^{(k)} \\ &\approx -\frac{1}{2} \frac{1}{\sigma_i^{(k-1)} + \exp\left(\kappa_i \mu_{i+1}^{(k-1)} + \omega_i\right)} \left(x_i^{(k)} - \mu_i^{(k-1)}\right)^2. \end{aligned} \tag{154}$$

The sum of Eqs (152) and (154) then is the variational energy of Eq. (31), up to a constant term that can be absorbed into the normalization constant $\mathcal{Z}_i$ (cf. Eq. (26)):

$$\begin{aligned} I\left(x_i^{(k)}\right) = &-\frac{1}{2} \ln\left(\sigma_{i-1}^{(k-1)} + \exp\left(\kappa_{i-1} x_i^{(k)} + \omega_{i-1}\right)\right) \\ &- \frac{1}{2} \frac{\sigma_i^{(k)} + \left(\mu_i^{(k)} - \mu_i^{(k-1)}\right)^2}{\sigma_{i-1}^{(k-1)} + \exp\left(\kappa_{i-1} x_i^{(k)} + \omega_{i-1}\right)} \\ &- \frac{1}{2} \frac{1}{\sigma_i^{(k-1)} + \exp\left(\kappa_i \mu_{i+1}^{(k-1)} + \omega_i\right)} \left(x_i^{(k)} - \mu_i^{(k-1)}\right)^2. \end{aligned} \tag{155}$$

D. The negative free energy under the Laplace approximation

The Laplace approximation is the approximation of a probability distribution by a Gaussian distribution centered around the approximated distribution's mode. Originally introduced for beta distributions on the unit interval, the Laplace approximation can be extended to many other posterior distributions since in many cases, for moderate to large samples, the posterior may be expected to be sharply peaked around its mode (Kass and Raftery, 1995).

In what follows, we take the posterior $p(\xi|u)$ to be approximated by the Gaussian $q(\xi|\mu,\Sigma)$, that is

$$\begin{aligned} p(\xi|u) &\approx q(\xi|\mu,\Sigma) \\ &= \mathcal{N}(\xi;\mu,\Sigma) \\ &= (2\pi)^{-\frac{d}{2}}|\Sigma|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}(\xi-\mu)^{\mathrm{T}}\Sigma^{-1}(\xi-\mu)\right), \end{aligned} \tag{156}$$

where $\xi, \mu \in \mathbb{R}^d$, Σ a positive definite $d \times d$-Matrix, and $|\Sigma|$ its determinant.

Our goal here will be to determine those values of μ (the *conditional mode*) and Σ (the *conditional variance*) that maximize the negative free energy $\mathcal{F}$ and thus make q approximate the true posterior density most closely (cf. Appendix A). To simplify the notation, we use $f_\xi \stackrel{\text{def}}{=} \frac{\partial f}{\partial \xi}$, $f_{\xi\xi} \stackrel{\text{def}}{=} \frac{\partial^2 f}{\partial \xi^2}$, etc. for partial derivatives and set

$$\mathcal{L}(\xi) \stackrel{\text{def}}{=} \ln p(u,\xi). \tag{157}$$

For what follows, it will also be useful to note that for a constant $d \times d$-Matrix A

$$\begin{aligned}\langle(\xi-\mu)^{\mathrm{T}}A(\xi-\mu)\rangle_q &= \langle\sum_{i,j}(\xi_i-\mu_i)(A)_{ij}\big(\xi_j-\mu_j\big)\rangle_q \\ &= \sum_{i,j}\underbrace{\langle(\xi_i-\mu_i)\big(\xi_j-\mu_j\big)\rangle_q}_{\stackrel{\text{def}}{=}(\Sigma)_{ij}}(A)_{ij} \\ &= \mathrm{tr}(\Sigma A),\end{aligned} \tag{158}$$

and specifically

$$\langle(\xi-\mu)^{\mathrm{T}}\Sigma^{-1}(\xi-\mu)\rangle_q = \mathrm{tr}(\Sigma\Sigma^{-1}) = \mathrm{tr}\big(\mathrm{id}_{\mathbb{R}^d}\big) = d. \tag{159}$$

The free energy is $\mathcal{F} = \langle\mathcal{L}(\xi)\rangle_q + S[q]$ (cf. Eq. (128)). I will deal with those terms separately, beginning by expanding the first term in powers to second order around μ:

$$\begin{aligned}\langle\mathcal{L}(\xi)\rangle_q &\approx \langle\mathcal{L}(\mu)+\mathcal{L}_\xi(\mu)^{\mathrm{T}}(\xi-\mu)+\frac{1}{2}(\xi-\mu)^{\mathrm{T}}\mathcal{L}_{\xi\xi}(\mu)(\xi-\mu)\rangle_q \\ &= \mathcal{L}(\mu)+\mathcal{L}_\xi(\mu)^{\mathrm{T}}\underbrace{\langle(\xi-\mu)\rangle_q}_{=0}+\frac{1}{2}\underbrace{\langle(\xi-\mu)^{\mathrm{T}}\mathcal{L}_{\xi\xi}(\mu)(\xi-\mu)\rangle_q}_{=\mathrm{tr}\left(\Sigma\mathcal{L}_{\xi\xi}(\mu)\right)} \\ &= \mathcal{L}(\mu)+\mathrm{tr}\left(\Sigma\mathcal{L}_{\xi\xi}(\mu)\right).\end{aligned} \tag{160}$$

The second term in $\mathcal{F}$, the entropy $S[q]$, can be calculated exactly:

$$\begin{aligned} S[q] &= -\langle \ln q(\xi) \rangle_q \\ &= \langle \frac{d}{2}\ln(2\pi) + \frac{1}{2}\ln|\Sigma| + \frac{1}{2}(\xi-\mu)^{\mathrm{T}}\Sigma^{-1}(\xi-\mu)\rangle_q \\ &= \frac{d}{2}\ln(2\pi) + \frac{1}{2}\ln|\Sigma| + \frac{1}{2}\underbrace{\langle (\xi-\mu)^{\mathrm{T}}\Sigma^{-1}(\xi-\mu)\rangle_q}_{=d} \\ &= \frac{d}{2}\ln(2\pi\mathrm{e}) + \frac{1}{2}\ln|\Sigma|. \end{aligned} \tag{161}$$

Taking the two terms together, we have

$$\mathcal{F} \approx \mathcal{L}(\mu) + \mathrm{tr}\left(\Sigma \mathcal{L}_{\xi\xi}(\mu)\right) + \frac{d}{2}\ln(2\pi\mathrm{e}) + \frac{1}{2}\ln|\Sigma|. \tag{162}$$

The conditional variance that maximizes this approximation to $\mathcal{F}$ can now be found by differentiating with respect to Σ, setting to zero, and solving for Σ. We take the vectors $\{e_1, e_2, \dots, e_d\}$ to be an orthonormal basis of $\mathbb{R}^d$. We do two preliminary calculations. First, (with A as above):

$$\begin{aligned} \partial_\Sigma \mathrm{tr}(\Sigma A) &= \partial_\Sigma \sum_{k.l} (\Sigma)_{kl}(A)_{lk} \\ &= \sum_{i,j} \frac{\partial}{\partial(\Sigma)_{ij}} \left(\sum_{k.l} (\Sigma)_{kl}(A)_{lk} \right) e_i \otimes e_j \\ &= \sum_{i,j} \delta_{ik}\delta_{jl}(A)_{lk} e_i \otimes e_j \\ &= \sum_{i,j} (A)_{ji} e_i \otimes e_j = A^{\mathrm{T}}, \end{aligned} \tag{163}$$

where δ_{ij} (0 for $i \neq j$; 1 for $i = j$) is the Kronecker symbol.

Since Σ is symmetric and positive definite, we may assume without loss of generality that it is diagonal with respect to the basis $\{e_1, e_2, \dots, e_d\}$, so that $|\Sigma| = \prod_k (\Sigma)_{kk}$. Then

$$\begin{aligned}\partial_\Sigma \ln|\Sigma| &= \partial_\Sigma \ln \prod_k (\Sigma)_{kk} = \partial_\Sigma \sum_k \ln(\Sigma)_{kk} \\ &= \sum_{i,j} \frac{\partial}{\partial(\Sigma)_{ij}} \left(\sum_k \ln(\Sigma)_{kk} \right) e_i \otimes e_j \\ &= \sum_{i,j} \delta_{ik}\delta_{jk} \frac{1}{(\Sigma)_{kk}} e_i \otimes e_j = \Sigma^{-1}.\end{aligned} \tag{164}$$

Applying this to Eq. (162) and setting to zero, we have

$$\begin{aligned}&\partial_\Sigma \mathcal{F} = \frac{1}{2}\left(\mathcal{L}_{\xi\xi}(\mu) + \Sigma^{-1}\right) = 0 \\ &\implies \Sigma = -\mathcal{L}_{\xi\xi}(\mu)^{-1}\end{aligned} \tag{165}$$

Inserting this into Eq. (162) gives us

$$\mathcal{F} \approx \mathcal{L}(\mu) + \frac{1}{2}\ln|\Sigma| + \frac{d}{2}\ln(2\pi). \tag{166}$$

To evaluate this, we need the mode μ of $\mathcal{L}$. In general, this cannot be found analytically but has to be approximated numerically.

E. Response models for saccadic reaction times

In the following, I explain the functional form of our three response models in more detail. All models assume a linear relationship between α and RS, parameterized by the two parameters ζ_1 and ζ_2 (cf. Eq. (120) and Figure 22). α represents the proportion of total attentional capacity that is allocated to the cued location (and therefore lies in the unit interval) and should amount to 0.5 if both target locations are equally likely. These constraints, which all response models conform to, can be summarized as:

$$\begin{aligned} &\text{C1: } 0 \leq \alpha \leq 1 \\ &\text{C2: } \alpha = 0.5 \text{ for } \hat{\mu}_1 = 0.5 \end{aligned} \tag{167}$$

Given these constraints, our response models differ in which attribute of the predicted validity of the cue maps to the attentional factor α (and thus determines RS in Eq. (120)). The functional forms of these models are motivated in the following and are depicted graphically in Figure 34. (Note that the vertical axis in Figure 34 is attention to outcome location. For valid trials, this is equal to attention to cued location α, while for invalid trials it is $1 - \alpha$.)

The 'precision' model (Eq. (121)) links attention to the precision of predictions as suggested by Feldman and Friston (2010). In our specific case, the precision of the prediction at the 1st level ($\hat{\pi}_1$) has a minimal value of 4 when $\hat{\mu}_1 = 0.5$ and approaches infinity as $\hat{\mu}_1$ approaches 1 (cf. Eq. (38)). The most parsimonious way to meet the above constraints C1 and C2 is to define α as the logistic sigmoid of $\hat{\pi}_1$, minus its minimum (cf. Eq. (121)):

$$\alpha = s(\hat{\pi}_1 - 4). \tag{168}$$

Note that since the cue becomes a counter-indication of outcome location when μ_2 falls below 0 (or equivalently, when $\hat{\mu}_1$ drops below 0.5), a suitable definition of α for the whole range of $\hat{\mu}_1$ is

$$\alpha^{(t)} = s\left(\text{sign}\left(\mu_2^{(t-1)}\right)\left(\hat{\pi}_1^{(t)} - 4\right)\right). \tag{169}$$

This ensures that attention to the cued location falls to 0 as $\hat{\mu}_1$ approaches 0.

A simpler model of attention allocation given a cue-induced belief about outcome is that attention is proportional to predicted probability of outcome: if the agent believes that the probability of seeing outcome 'left' is p (e.g., 80%), then it will allocate proportion p (i.e., 80%) of its attentional resources to location 'left'. I call this the 'belief' model (cf. Eq. (122)). In terms of our perceptual model, the predicted probability of a valid trial is simply $\hat{\mu}_1$:

$$\alpha = \hat{\mu}_1. \tag{170}$$

Finally, the 'surprise' model describes the attentional factor α as a function of the Shannon surprise (the negative logarithm of the probability of the outcome being the cued location given the prediction $\hat{\mu}_1$) (cf. Bestmann et al., 2008). For a predicted probability $\hat{\mu}_1 = 1$ of a valid trial, surprise is zero, whereas for $\hat{\mu}_1 = 0$, it is infinite. In the first case, attention is therefore allocated exclusively to the cued location (i.e., $\alpha = 1$), while in the second case, attention is allocated exclusively to the non-cued location (i.e., $\alpha = 0$). The simplest way this can be achieved under consideration of constraints C1 and C2 is (cf. Eq. (123)):

$$\alpha^{(t)} = \frac{1}{\left(1 + surprise\left(\hat{\mu}_1^{(t)}\right)\right)} \tag{171}$$

where $surprise\left(\hat{\mu}_1^{(t)}\right) = -\log_2 p\left(x_1^{(t)} = 1 \middle| \hat{\mu}_1^{(t)}\right) = -\log_2\left(\hat{\mu}_1^{(t)}\right)$

Note that we make use of $\log_2 x = \ln x \,/\, \ln 2$ to ensure we meet constraint C2.

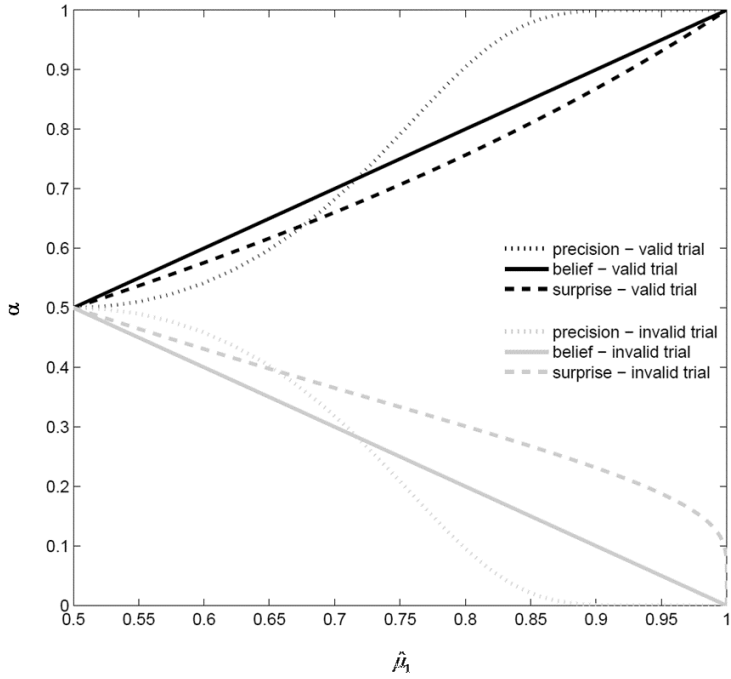

Figure 34. ***Attention as a function of belief.*** *Illustration of the amount of attentional resources α for the three different theoretical response models as a function of* $\hat{\mu}_1$.

REFERENCES

Anderson, A. J., and Carpenter, R. H. S. (2006). Changes in expectation consequent on experience, modeled by a simple, forgetful neural circuit. *J Vis* 6. Available at: http://www.journalofvision.org/content/6/8/5 [Accessed November 15, 2012].

Bastos, A., Moran, R. J., Litvak, V., Fries, P., and Friston, K. (2011). A Dynamic Causal Model of how inter-areal synchronization is achieved in canonical microcircuits. *SFN Abstract No. 622.16*.

Beal, M. J. (2003). Variational Algorithms for Approximate Bayesian Inference (Ph.D. thesis, UCL). Available at: http://www.cse.buffalo.edu/faculty/mbeal/thesis/ [Accessed October 7, 2009].

Beauchamp, M. S., Petit, L., Ellmore, T. M., Ingeholm, J., and Haxby, J. V. (2001). A Parametric fMRI Study of Overt and Covert Shifts of Visuospatial Attention. *NeuroImage* 14, 310–321.

Beck, J., Ma, W., Kiani, R., Hanks, T., Churchland, A., Roitman, J., Shadlen, M., Latham, P., and Pouget, A. (2008). Probabilistic Population Codes for Bayesian Decision Making. *Neuron* 60, 1142–1152.

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., and Rushworth, M. F. S. (2008). Associative learning of social value. *Nature* 456, 245–249.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience* 10, 1214–1221.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior* 10, 122–142.

Bestmann, S., Harrison, L. M., Blankenburg, F., Mars, R. B., Haggard, P., Friston, K. J., and Rothwell, J. C. (2008). Influence of Uncertainty and Surprise on Human Corticospinal Excitability during Preparation for Action. *Current Biology* 18, 775–780.

Bohnet, I., and Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, 467–484.

Brefczynski, J. A., and DeYoe, E. A. (1999). A physiological correlate of the "spotlight" of visual attention. *Nature Neuroscience* 2, 370–374.

Bresciani, J.-P., Dammeier, F., and Ernst, M. O. (2006). Vision and touch are automatically integrated for the perception of sequences of events. *Journal of Vision* 6, 554–564.

Brodersen, K. H., Penny, W. D., Harrison, L. M., Daunizeau, J., Ruff, C. C., Duzel, E., Friston, K. J., and Stephan, K. E. (2008). Integrated Bayesian models of learning and decision making for saccadic eye movements. *Neural Networks* 21, 1247–1260.

Brown, H., and Friston, K. J. (2012). Free-energy and illusions: the Cornsweet effect. *Front. Psychology* 3, 43.

Broyden, C. G. (1970). The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA J Appl Math* 6, 76–90.

Bruce, C. J., and Goldberg, M. E. (1985). Primate frontal eye fields. I. Single neurons discharging before saccades. *J Neurophysiol* 53, 603–635.

Buskens, V., Raub, W., and van der Veer, J. (2010). Trust in triads: An experimental study. *Social Networks* 32, 301–312.

Carpenter, R. H. S., and Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature* 377, 59–62.

Carver, C. S., and White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology* 67, 319–333.

Chiau, H.-Y., Tseng, P., Su, J.-H., Tzeng, O. J. L., Hung, D. L., Muggleton, N. G., and Juan, C.-H. (2011). Trial type probability modulates the cost of antisaccades. *J Neurophysiol* 106, 515–526.

Chikkerur, S., Serre, T., Tan, C., and Poggio, T. (2010). What and where: A Bayesian inference theory of attention. *Vision Research* 50, 2233–2247.

Christie, R., and Geis, F. L. (1970). *Studies in Machiavellianism*. Academic Press.

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., Linenweber, M. R., Petersen, S. E., Raichle, M. E., Essen, D. C. V., et al. (1998). A Common Network of Functional Areas for Attention and Eye Movements. *Neuron* 21, 761–773.

Corrado, G. S., Sugrue, L. P., Sebastian Seung, H., and Newsome, W. T. (2005). Linear-Nonlinear-Poisson Models of Primate Choice Dynamics. *J Exp Anal Behav* 84, 581–617.

Cox, R. T. (1946). Probability, Frequency and Reasonable Expectation. *Am. J. Phys.* 14, 1–13.

Daunizeau, J., David, O., and Stephan, K. E. (2011). Dynamic causal modelling: A critical review of the biophysical and statistical foundations. *NeuroImage* 58, 312–322.

Daunizeau, J., Friston, K. J., and Kiebel, S. J. (2009). Variational Bayesian identification and prediction of stochastic nonlinear dynamic causal models. *Physica D: Nonlinear Phenomena* 238, 2089–2118.

Daunizeau, J., den Ouden, H. E. M., Pessiglione, M., Kiebel, S. J., Friston, K. J., and Stephan, K. E. (2010a). Observing the Observer (II): Deciding When to Decide. *PLoS ONE* 5, e15555.

Daunizeau, J., den Ouden, H. E. M., Pessiglione, M., Kiebel, S. J., Stephan, K. E., and Friston, K. J. (2010b). Observing the Observer (I): Meta-Bayesian Models of Learning and Decision-Making. *PLoS ONE* 5, e15554.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology* 10, 85.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.

Dayan, P., Hinton, G. E., Neal, R. M., and Zemel, R. S. (1995). The Helmholtz Machine. *Neural Computation* 7, 889–904.

Dayan, P., and Huys, Q. J. M. (2009). Serotonin in Affective Control. *Annu. Rev. Neurosci.* 32, 95–126.

Dayan, P., Kakade, S., and Montague, P. R. (2000). Learning and selective attention. *Nature Neuroscience* 3, 1218–1223.

Dayan, P., and Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Curr. Opin. Neurobiol* 18, 185–196.

Deneve, S. (2008). Bayesian Spiking Neurons II: Learning. *Neural Computation* 20, 118–145.

Desimone, R., and Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience* 18, 193–222.

Deubel, H. (2008). The time course of presaccadic attention shifts. *Psychological Research* 72, 630–640.

Deubel, H., and Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research* 36, 1827–1837.

Ditterich, J., Mazurek, M. E., and Shadlen, M. N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. *Nature Neuroscience* 6, 891–898.

Doré-Mazars, K., Pouget, P., and Beauvillain, C. (2004). Attentional selection during preparation of eye movements. *Psychological Research* 69, 67–76.

Doya, K. (2008). Modulators of decision making. *Nat Neurosci* 11, 410–416.

Eriksen, C. W., and James, J. D. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics* 40, 225–240.

Eriksen, C. W., and Yeh, Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance* 11, 583–597.

Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience* 9, 292–303.

Farrell, S., Ludwig, C. J. H., Ellis, L. A., and Gilchrist, I. D. (2010). Influence of environmental statistics on inhibition of saccadic return. *PNAS* 107, 929–934.

Fearnhead, P., and Liu, Z. (2007). On-line inference for multiple changepoint problems. *J Royal Statistical Soc B* 69, 589–605.

Fehr, E. (2009). On the Economics and Biology of Trust. *Journal of the European Economic Association* 7, 235–266.

Feldman, H., and Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Front. Hum. Neurosci* 4, 215.

Feynman, R. P. (1972). *Statistical mechanics: a set of lectures*. W. A. Benjamin.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Exp Econ* 10, 171–178.

Fischer, B., Biscaldi, M., and Otto, P. (1993). Saccadic eye movements of dyslexic adult subjects. *Neuropsychologia* 31, 887–906.

Fletcher, R. (1970). A new approach to variable metric algorithms. *The Computer Journal* 13, 317–322.

Frank, M. J. (2008). Schizophrenia: A Computational Reinforcement Learning Perspective. *Schizophrenia Bulletin* 34, 1008 –1011.

Frank, M. J., Doll, B. B., Oas-Terpstra, J., and Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in

exploration and exploitation. *Nature Neuroscience* 12, 1062–1068.

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., and Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *PNAS* 104, 16311–16316.

Friston, K. (2008). Hierarchical Models in the Brain. *PLoS Comput Biol* 4, e1000211.

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences* 13, 293–301.

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

Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *NeuroImage* 19, 1273–1302.

Friston, K. J., and Stephan, K. E. (2007). Free-energy and the brain. *Synthese.* 159, 417–458.

Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris* 100, 70–87.

Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., and Penny, W. (2007). Variational free energy and the Laplace approximation. *NeuroImage* 34, 220–234.

Geisler, W. S., and Diehl, R. L. (2002). Bayesian natural selection and the evolution of perceptual systems. *Philos Trans R Soc Lond B Biol Sci* 357, 419–448.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis*. Chapman & Hall/CRC.

Gershman, S. J., and Niv, Y. (2010). Learning latent structure: carving nature at its joints. *Current Opinion in Neurobiology* 20, 251–256.

Giessing, C., Thiel, C. M., Rösler, F., and Fink, G. R. (2006). The modulatory effects of nicotine on parietal cortex activity in a cued target

detection task depend on cue reliability. *Neuroscience* 137, 853–864.

Gitelman, D. R. (2002). ILAB: A program for postexperimental eye movement analysis. *Behavior Research Methods, Instruments, & Computers* 34, 605–612.

Glimcher, P. W. (2001). Making choices: the neurophysiology of visual-saccadic decision making. *Trends in Neurosciences* 24, 654–659.

Glimcher, P. W. (2003). The Neurobiology of Visual-Saccadic Decision Making. *Annual Review of Neuroscience* 26, 133–179.

Gluck, M. A., Shohamy, D., and Myers, C. (2002). How do People Solve the Weather Prediction Task?: Individual Variability in Strategies for Probabilistic Category Learning. *Learning & Memory* 9, 408–418.

Godijn, R., and Theeuwes, J. (2003). Parallel allocation of attention prior to the execution of saccade sequences. *Journal of Experimental Psychology: Human Perception and Performance* 29, 882–896.

Goldfarb, D. (1970). A family of variable-metric methods derived by variational means. *Mathematics of Computation* 24, 23–26.

Gu, Q. (2002). Neuromodulatory transmitter systems in the cortex and their role in cortical plasticity. *Neuroscience* 111, 815–835.

de Haan, B., Morgan, P. S., and Rorden, C. (2008). Covert orienting of attention and overt eye movements activate identical brain regions. *Brain Research* 1204, 102–111.

Häfner, H., Maurer, K., Löffler, W., an der Heiden, W., Munk-Jørgensen, P., Hambrecht, M., and Riecher-Rössler, A. (1998). The ABC schizophrenia study: a preliminary overview of the results. *Social Psychiatry and Psychiatric Epidemiology* 33, 380–386.

Herz, A. V. M., Gollisch, T., Machens, C. K., and Jaeger, D. (2006). Modeling Single-Neuron Dynamics and Computations: A Balance of Detail and Abstraction. *Science* 314, 80–85.

Iglesias, S., Mathys, C., Brodersen, K. H., Kasper, L., Piccirelli, M., and Stephan, K. E. (2012). Generic and task-specific effects of uncertainty and prediction errors in associative learning. *HBM abstract*. Available at: http://ww4.aievolution.com/hbm1201/index.cfm?do=abs.viewAbs&abs=5697.

Itti, L., and Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research* 49, 1295–1306.

Jaynes, E. T. (1957). Information Theory and Statistical Mechanics. *Phys. Rev.* 106, 620.

Jaynes, E. T. (2003). *Probability theory: the logic of science*. Cambridge University Press.

Jeffreys, H. (1935). Some Tests of Significance, Treated by the Theory of Probability. *Mathematical Proceedings of the Cambridge Philosophical Society* 31, 203–222.

Jeffreys, S. H. (1961). *Theory of probability*. Clarendon Press.

Jongen, E. M. M., and Smulders, F. T. Y. (2007). Sequence effects in a spatial cueing task: Endogenous orienting is sensitive to orienting in the preceding trial. *Psychological Research* 71, 516–523.

Jonides, J. (1980). Towards a model of the mind's eye's movement. *Canadian Journal of Psychology/Revue canadienne de psychologie* 34, 103–112.

Kalman, R. E. (1960). A new approach to linear filtering and prediciton problems. *Journal of Basic Engineering* 82, 35–45.

Kass, R. E., and Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association* 90, 773–795.

Kastner, S., Pinsk, M. A., Weerd, P. D., Desimone, R., and Ungerleider, L. G. (1999). Increased Activity in Human Visual Cortex during Directed Attention in the Absence of Visual Stimulation. *Neuron* 22, 751–761.

King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., and Montague, P. R. (2008). The Rupture and Repair of Cooperation in Borderline Personality Disorder. *Science* 321, 806 –810.

Kording, K. P., and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature* 427, 244–247.

Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S.-C., and Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences* 106, 17951 –17956.

Laplace, P.-S. (1774). Mémoire sur la probabilité des causes par les évènemens. *Mém. Acad. Roy. Sci.* 6, 621–656.

Laplace, P.-S. (1812). *Théorie Analytique des Probabilités*. Paris: Courcier Imprimeur.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., and Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied* 8, 75–84.

Lomakina, E., Vezhnevets, A., Mathys, C., Brodersen, K. H., Stephan, K. E., and Buhmann, J. (2012). Bayesian global optimization for model-based neuroimaging. *HBM e-poster*. Available at: https://ww4.aievolution.com/hbm1201/index.cfm?do=abs.viewAbs&abs=6320.

London, M., and Hausser, M. (2005). Dendritic computation. *Annu. Rev. Neurosci.* 28, 503–532.

Luck, S. J., Chelazzi, L., Hillyard, S. A., and Desimone, R. (1997). Neural Mechanisms of Spatial Selective Attention in Areas V1, V2, and V4 of Macaque Visual Cortex. *J Neurophysiol* 77, 24–42.

Mathys, C., Daunizeau, J., Friston, K. J., and Stephan, K. E. (2011). A Bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* 5:, 39.

McCrae, R. R., and Costa Jr., P. T. (2004). A contemplated revision of the NEO Five-Factor Inventory. *Personality and Individual Differences* 36, 587–596.

Montague, P. R., Hyman, S. E., and Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature* 431, 760–767.

Murphy, R. O., Ackermann, K. A., and Handgraaf, M. (2011). Measuring Social Value Orientation. *SSRN eLibrary*. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1804189 [Accessed May 16, 2012].

Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., Jones, P. B., Bullmore, E. T., Robbins, T. W., and Fletcher, P. C. (2007). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry* 13, 267–276.

Nassar, M. R., Wilson, R. C., Heasly, B., and Gold, J. I. (2010). An Approximately Bayesian Delta-Rule Model Explains the Dynamics of Belief Updating in a Changing Environment. *J. Neurosci.* 30, 12366–12378.

Nelder, J. A., and Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal* 7, 308–313.

Nobre, A. C., Gitelman, D. R., Dias, E. C., and Mesulam, M. M. (2000). Covert Visual Spatial Orienting and Saccades: Overlapping Neural Systems. *NeuroImage* 11, 210–216.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. *Science* 304, 452–454.

Orbán, G., Fiser, J., Aslin, R. N., and Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences* 105, 2745 –2750.

den Ouden, H. E. M., Daunizeau, J., Roiser, J., Friston, K. J., and Stephan, K. E. (2010). Striatal Prediction Error Modulates Cortical Coupling. *J. Neurosci.* 30, 3210–3219.

Patton, J. H., Stanford, M. S., and Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *J Clin Psychol* 51, 768–774.

Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M., and Leff, A. P. (2010). Comparing Families of Dynamic Causal Models. *PLoS Comput Biol* 6, e1000709.

Perry, R. J., and Zeki, S. (2000). The neurology of saccades and covert shifts in spatial attention An event-related fMRI study. *Brain* 123, 2273–2288.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., and Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045.

Pestilli, F., Carrasco, M., Heeger, D. J., and Gardner, J. L. (2011). Attentional Enhancement via Selection and Pooling of Early Sensory Responses in Human Visual Cortex. *Neuron* 72, 832–846.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology* 32, 3–25.

Preuschoff, K., and Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. *Ann. N. Y. Acad. Sci* 1104, 135–146.

R Core Team (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing Available at: http://www.R-project.org.

Rao, R. (2005). Bayesian inference and attentional modulation in the visual cortex. [Miscellaneous Article]. *Neuroreport November 7, 2005* 16, 1843–1848.

Rescorla, R. A., and Wagner, A. R. (1972). "A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement," in Classical conditioning II: Current research and theory (New York: Appleton-Century-Crofts), 64–99.

Risko, E. F., and Stolz, J. A. (2010). The proportion valid effect in covert orienting: Strategic control or implicit learning? *Consciousness and Cognition* 19, 432–442.

Rizzolatti, G., Riggio, L., Dascola, I., and Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia* 25, 31–40.

Rushworth, M. F. S., and Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience* 11, 389–397.

Saproo, S., and Serences, J. T. (2010). Spatial Attention Improves the Quality of Population Codes in Human Visual Cortex. *J Neurophysiol* 104, 885–895.

Schall, J. D., and Hanes, D. P. (1993). Neural basis of saccade target selection in frontal eye field during visual search. *, Published online: 02 December 1993; | doi:10.1038/366467a0* 366, 467–469.

Schall, J. D., Purcell, B. A., Heitz, R. P., Logan, G. D., and Palmeri, T. J. (2011). Neural mechanisms of saccade target selection: gated accumulator model of the visual–motor cascade. *European Journal of Neuroscience* 33, 1991–2002.

Schneider, W. X. (1995). VAM: A neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Visual Cognition* 2, 331–376.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science* 275, 1593–1599.

Shanno, D. F. (1970). Conditioning of quasi-Newton methods for function minimization. *Mathematics of Computation* 24, 647–656.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27, 379–423,623–656.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., and Frith, C. D. (2004). Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science* 303, 1157–1162.

Smith, A., Li, M., Becker, S., and Kapur, S. (2006). Dopamine, prediction error and associative learning: A model-based account. *Network* 17, 61–84.

Stampe, D. M. (1993). Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behavior Research Methods, Instruments, & Computers* 25, 137–142.

Stephan, K. E., Friston, K. J., and Frith, C. D. (2009a). Dysconnection in Schizophrenia: From Abnormal Synaptic Plasticity to Failures of Self-monitoring. *Schizophrenia Bulletin* 35, 509 –527.

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., and Friston, K. J. (2009b). Bayesian model selection for group studies. *NeuroImage* 46, 1004–1017.

Steyvers, M., and Brown, S. (2006). "Prediction and change detection," in Adv. Neural Inf. Process Syst., 1281–1288.

Steyvers, M., Lee, M. D., and Wagenmakers, E.-J. (2009). A Bayesian analysis of human decision-making on bandit problems. *Journal of Mathematical Psychology* 53, 168–179.

Sutton, R. S., and Barto, A. G. (1998). *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press.

Thiel, C. M., Huston, J. P., and Schwarting, R. K. (1998). Hippocampal acetylcholine and habituation learning. *Neuroscience* 85, 1253–1262.

Vossel, S., Weidner, R., and Fink, G. R. (2011). Dynamic Coding of Events within the Inferior Frontal Gyrus in a Probabilistic Selective Attention Task. *Journal of Cognitive Neuroscience* 23, 414–424.

Whiteley, L., and Sahani, M. (2008). Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. *J Vis* 8. Available at: http://www.journalofvision.org/content/8/3/2 [Accessed November 16, 2012].

Wilson, R. C., Nassar, M. R., and Gold, J. I. (2010). Bayesian Online Learning of the Hazard Rate in Change-Point Problems. *Neural Computation* 22, 2452–2476.

Xu, F., and Tenenbaum, J. B. (2007). Sensitivity to sampling in Bayesian word learning. *Dev Sci* 10, 288–297.

Yamagishi, T., and Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motiv Emot* 18, 129–166.

Yang, T., and Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature* 447, 1075–1080.

Yu, A. J., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692.

Yuille, A., and Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends Cogn. Sci. (Regul. Ed.)* 10, 301–308.

ACKNOWLEDGMENTS

I thank my supervisors Klaas Enno Stephan and Klaas Prüssmann for the exceptional level of inspiration, support, and encouragement they have provided me. It has been a great privilege to experience their calmly confident belief in the value of my research. Klaas Enno especially has guided my work into a direction that will hopefully not only make its results stand, but also ensure their dissemination by various relevant applications and by a presentation that opens them up even to the less mathematically-minded.

Many colleagues have made significant contributions to the content of this thesis. Chief among them is Simone Vossel who was responsible for the experimental part of the study on saccadic reaction times. Ekaterina Lomakina-Rumyantseva and Sandra Iglesias were an essential part of the simulation study, and Saee Paliwal helped with the HGF toolbox. Jean Daunizeau and Karl Friston were a crucial inspiration and help with their deep knowledge of free energy, Bayesian inference, and decision theory. Tony Williams and Ernst Fehr provided an economic perspective, extensive behavioral experience and insights – as well as z-Tree programming skills - for the observed iterated trust game study. Lilian Weber and Andreea Diaconescu provided advice and wrote software for the same study. I am indebted to Lars Kasper and Marco Piccirelli for MRI sequence development, and to the staff at the SNS Lab – Karl Treiber, Adrian Etter, and Sebastian Grässli – for their professional support and expertise.

Both at the SNS Lab and at the IBT, especially at the nascent Translational Neuromodeling Unit (TNU), it was a great joy and inspiration to have such knowledgeable, insightful, and helpful colleagues as Justin Chumbley, Jan Engelmann, Jakob Heinzle, Quentin Huys, Falk Lieder, Friederike Meyer, Sudhir Shankar Raman, and Gabor Stefanics; Kay Brodersen, with his astounding dedication to science, his unfailing generosity, curiosity, energy, and friendship, has been the best officemate I could ever wish for.

For their administrative competence and support, I thank Mirjam Bebi, Marianne Berg, Eva and Sally Gschwend, Tamara Herz, Ursi Meier, Bettina Petralli, Tina Wentz, Suzanne Wilde, and Karin Zeder.

I thank my parents for equipping me early on with what it takes to pass this great marshmallow test of completing a thesis. My warmest thanks go to Vanessa, my fiancée, for her patience, encouragement, affection, and kindness – but also for help with devising the task in the observed iterated trust game study, for analyzing part of the data in that study, and for help with the preparation of this manuscript.

This thesis was supported by the NCCR “Neural plasticity and repair”.

KURZFASSUNG

Mathematische Modelle der Informationsverarbeitung sind entscheidend für ein Verständnis der Mechanismen, die situationsbedingtem Verhalten zugrundeliegen. Die zwei gegenwärtig auf diesem Gebiet dominanten Herangehensweisen, Verstärkungslernen (reinforcement learning – RL) und Bayes'sches Lernen, haben jedoch beide gewisse Nachteile. Bayes'sche Modelle kennen zum Beispiel oft keine interindividuellen Unterschiede und erfordern den Umgang mit analytisch nicht lösbaren Integralen, was zeitgerechtes Lernen erschwert, wenn nicht verunmöglicht. In der vorliegenden Dissertation führe ich den hierarchischen Gauss'schen Filter (HGF) ein, ein generisches hierarchisches Bayes'sches Modell für individuell angepasstes Lernen unter verschiedenen Formen von Unsicherheit (zum Beispiel eine schnell wechselnde Umwelt oder sensorische Unsicherheit). Der HGF verfolgt den Ansatz, verborgene Zustände der Umwelt als eine Hierarchie von Gauss'schen Zufallsprozessen aufzufassen, wobei jeweils die nächsthöhere Ebene die Varianz der darunterliegenden bestimmt. Die Kopplung zwischen den Ebenen wird von Parametern gesteuert, die den Einfluss von Unsicherheit auf die Informationsverarbeitung auf eine individuell verschiedene Weise prägen. Unter Rückgriff auf Variationsrechnung und Molekularfeldtheorie sowie mit Hilfe einer neuartigen Näherung für die Energiefunktion im informationstheoretischen Sinne leite ich Aktualisierungsgleichungen her, die (i) von geschlossener Form und daher extrem effizient sind, wodurch sie zeitgerechtes Lernen erlauben, (ii) sich auf natürliche Weise im Rahmen von RL deuten lassen und die (iii) Parameter enthalten, die Prozesse abbilden, welche in aktuellen Lerntheorien eine Schlüsselrolle spielen, zum Beispiel die Präzisionsgewichtung von Voraussagefehlern. Diese Parameter erlauben die Ausprägung individueller Unterschiede im Lernverhalten und könnten mit neuromodulatorischen Mechanismen im Hirn in Beziehung stehen. Der HGF ist sehr allgemein: er kann sowohl diskrete wie auch stetige Zustände beschreiben, und er umfasst sowohl deterministische als auch stochastische Beziehungen zwischen Umweltereignissen und Wahrnehmungszuständen. Diese Eigenschaften erläutere ich anhand von Simulationen und Analysen empirischer

Zeitreihen, die auf die Finanzmärkte zurückgehen, auf Verhaltensexperimente und auf funktionelle Magnetresonzmessungen. Insgesamt bietet der Rahmen des HGF eine neuartige mathematische Grundlage für das Verständnis normalen und pathologischen Lernens; er stellt RL in den Zusammenhang eines generischen Bayes'schen Ansatzes und verbindet es so mit Optimalitätsgrundsätzen aus der Wahrscheinlichkeitstheorie.