

Observing the Observer (II): Deciding When to Decide

Jean Daunizeau^{1,3*}, Hanneke E. M. den Ouden⁵, Matthias Pessiglione², Stefan J. Kiebel⁴, Karl J. Friston¹, Klaas E. Stephan^{1,3}

1 Wellcome Trust Centre for Neuroimaging, University College of London, London, United Kingdom, **2** Brain and Spine Institute, Hôpital Pitié-Salpêtrière, Paris, France, **3** Laboratory for Social and Neural Systems Research, Institute of Empirical Research in Economics, University of Zurich, Zurich, Switzerland, **4** Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, **5** Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

Abstract

In a companion paper [1], we have presented a generic approach for inferring how subjects make optimal decisions under uncertainty. From a Bayesian decision theoretic perspective, uncertain representations correspond to “posterior” beliefs, which result from integrating (sensory) information with subjective “prior” beliefs. Preferences and goals are encoded through a “loss” (or “utility”) function, which measures the cost incurred by making any admissible decision for any given (hidden or unknown) state of the world. By assuming that subjects make optimal decisions on the basis of updated (posterior) beliefs and utility (loss) functions, one can evaluate the likelihood of observed behaviour. In this paper, we describe a concrete implementation of this *meta-Bayesian* approach (i.e. a Bayesian treatment of Bayesian decision theoretic predictions) and demonstrate its utility by applying it to both simulated and empirical reaction time data from an associative learning task. Here, inter-trial variability in reaction times is modelled as reflecting the dynamics of the subjects’ internal recognition process, i.e. the updating of representations (posterior densities) of hidden states over trials while subjects learn probabilistic audio-visual associations. We use this paradigm to demonstrate that our meta-Bayesian framework allows for (i) probabilistic inference on the dynamics of the subject’s representation of environmental states, and for (ii) model selection to disambiguate between alternative preferences (loss functions) human subjects could employ when dealing with trade-offs, such as between speed and accuracy. Finally, we illustrate how our approach can be used to quantify subjective beliefs and preferences that underlie inter-individual differences in behaviour.

Citation: Daunizeau J, den Ouden HEM, Pessiglione M, Kiebel SJ, Friston KJ, et al. (2010) Observing the Observer (II): Deciding When to Decide. PLoS ONE 5(12): e15555. doi:10.1371/journal.pone.0015555

Editor: Olaf Sporns, Indiana University, United States of America

Received: August 5, 2010; **Accepted:** November 12, 2010; **Published:** December 14, 2010

Copyright: © 2010 Daunizeau et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by the Wellcome Trust (HDO, KJF), SystemsX.ch (JD, KES) and NCCR “Neural Plasticity” (KES). The authors also gratefully acknowledge support by the University Research Priority Program “Foundations of Human Social Behaviour” at the University of Zurich (JD, KES). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: j.daunizeau@fil.ion.ucl.ac.uk

Introduction

How can we infer subjects’ beliefs and preferences from their observed decisions? Or in other terms, can we identify the internal mechanisms that led subjects to act, as a response to experimentally controlled stimuli? Numerous experimental and theoretical studies imply that subjective prior beliefs, acquired over previous experience, strongly impact on perception, learning and decision-making ([2–6]). We also know that preferences and goals can impact subjects’ decisions in a fashion which is highly context-dependent and which subjects may be unaware of ([7–8]). But how can we estimate and disentangle the relative contributions of these components to observed behaviour? This is the nature of the so-called Inverse Bayesian Decision Theory (IBDT) problem, which has been a difficult challenge for analytical treatments.

In a companion paper [1], we have described a variational Bayesian framework for approximating the solution to the IBDT problem in the context of perception, learning and decision-making studies. Subjects are assumed to act as Bayesian observers, whose recognition of the hidden causes of their sensory inputs depends on the inversion of a *perceptual model* with subject-specific priors. The Bayesian inversion of this perceptual model derives from a variational formulation, through the minimization of

sensory surprise (in a statistical sense). More precisely, the variational Bayesian approach minimizes the so-called “free energy”, which is a lower bound on (statistical) surprise about the sensory inputs. The ensuing probabilistic subjective representation of hidden states (the posterior belief) then enters a *response model* of measured behavioural responses. Critically, decisions are thought to minimize expected loss or risk, given the posterior belief and the subject-specific loss (or utility) function that encodes the subject’s preferences. The response model thus provides a complete mechanistic mapping from experimental stimuli to observed behaviour. Over time or trials, the response model has the form of a state-space model (e.g., [9]), with two components: (i) an evolution function that models perception and learning through surprise minimization and (ii) an observation function that models decision making through risk minimization.

Solving the IBDT problem, or *observing the observer*, then reduces to inverting this state-space response model, given experimentally measured behaviour. This meta-Bayesian approach (experimenters make Bayesian inferences about subject’s Bayesian inferences) provides an approximate solution to the IBDT problem in that it enables comparisons of competing (perceptual and response) models and inferences on the parameters of those models. This is important, since evaluating the evidence of, for example, different

response models in the light of behavioural responses means we can distinguish between different loss functions (and thus preferences) subjects might have.

This paper complements the theoretical account in the companion paper by demonstrating the practical applicability of our framework. Here, we use it to investigate what computational mechanisms operate during learning-induced motor facilitation. While it has often been found that (correct) expectations about sensory stimuli speed up responses to those stimuli (e.g. [10–11]), explaining this acceleration of reaction times in computationally mechanistic terms is *not* trivial. We argue that such an explanation must take into account the dynamics of subjective representations, such as posterior beliefs about the causes that generate stimuli, and their uncertainty, as learning unfolds over trials. Throughout the text, “representation” refers to posterior densities of states or parameters. We investigate these issues in the context of an audio-visual associative learning task [12], where subjects have to categorize visual stimuli as quickly as possible. We use this task as a paradigmatic example of what sort of statistical inference our model-based approach can provide. As explained in detail below, this task poses two interesting explananda for computational approaches: (i) it relies upon a hierarchical structure of causes in the world: visual stimuli depend probabilistically on preceding auditory cues whose predictive properties change over time (i.e., a volatile environment), and (ii) it introduces a conflict in decision making, i.e. a speed-accuracy trade-off.

We construct two Bayesian decision theoretic (BDT) response models based upon the same speed-accuracy trade-off (c.f. [13] or [14]), but differing in their underlying perceptual model. These two perceptual models induce different learning rules, and thus different predictions, leading to qualitatively different trial-by-trial variations in reaction times. We have chosen to focus on reaction time data to highlight the important role of the response model and to show that optimal responses are not just limited to categorical choices.

Of course, the validity of a model cannot be fully established by application to empirical data whose underlying mechanisms or “ground truth” are never known with certainty. However, by ensuring that only one of the competing models was fully consistent with the information given to the subjects, we established a reference point against which our model selection results could be compared, allowing us to assess the construct validity of our approach. Furthermore, we also performed a simulation study, assessing the veracity of parameter estimation and model comparison using synthetic data for which the ground truth was known.

Methods

How does learning modulate reaction times? In this section, we first describe the associative learning task, and then the perceptual and response models we have derived to model the reaction time data. We then recall briefly the elements of the variational Bayesian framework which is described in the companion paper in detail and which we use to invert the response model given reaction time data. Next, we describe the Monte-Carlo simulation series we have performed to demonstrate the validity of the approach. Finally, we summarize the analysis of real reaction time data, illustrating the sort of inference that can be derived from the scheme, and establishing the construct validity of the approach.

The associative learning task

The experimental data and procedures have been reported previously as part of a functional magnetic resonance imaging study of audio-visual associative learning [12]. We briefly summarize the main points. Healthy volunteers were presented

visual stimuli (faces or houses) following an auditory cue. The subjects performed a speeded discrimination task on the visual stimuli. On each trial, one of two possible auditory cues was presented (simple tones of different frequencies; C_1 and C_2), each predicting the subsequent visual cue with a different probability. The subjects were told that the relationship between auditory and visual stimuli was probabilistic and would change over time but that these changes were random and not related to any underlying rule. The reaction-time (from onset of visual cue to button press) was measured on each trial.

The probability of a given visual outcome or response cue, say *face*, given C_1 was always the same as the probability of the alternative (*house*) given C_2 : $p(\text{face}|C_1) = 1 - p(\text{face}|C_2)$. Moreover, since the two auditory cues occurred with equal frequency, the marginal probability of a face (or house) on any given trial was always 50%. This ensured that subjects could not be biased by *a priori* expectations about the outcome. In the original regression analyses in [12] no differences were found between high and low tone cues, nor any interactions between cue type and other experimental factors; here, we therefore consider the trials cued by C_1 and C_2 as two separate (intermingled, but non-interacting) sequences. This allows us to treat the two sequences as replications of the experiment, under two different auditory cues. We hoped to see that the results were consistent under the high and low tone cues.

A critical manipulation of the experiment was that the probabilistic cue-outcome association pseudorandomly varied over blocks of trials, from strong $p(\text{face}|C) = 0.9$, and moderate $p(\text{face}|C) = 0.7$, to random $p(\text{face}|C) = 0.5$. Our subjects were informed about the existence of this volatility without specifying the structure of these changes (timing and probability levels). We prevented any explicit search for systematic relationships by varying the length of the blocks and by presenting predictive and random blocks in alternation. In one session, each block lasted for 28–40 trials, within which the order of auditory cues was randomized. Each of five sessions lasted approximately seven minutes. On each trial, an auditory cue was presented for 300 ms, followed by a brief (150 ms) presentation of the visual outcome. In order to prevent anticipatory responses or guesses, both the inter-trial interval (2000 ± 650 ms) and visual stimulus onset latency (150 ± 50 ms) were jittered randomly.

The conventional analysis of variance (ANOVA) of the behavioural measures presented in [12] demonstrated that subjects learned the cue-outcome association: reaction times to the visual stimuli decreased significantly with increasing predictive strengths of the auditory cues. In what follows, we try to better understand the nature of this learning and the implicit perceptual models the subjects were using.

Perceptual and response models

The first step is to define the candidate response models that we wish to consider. In what follows, we will restrict ourselves to two qualitatively different perceptual models, which rest on different prior beliefs and lead to different learning rules (i.e. posterior belief update rules or recognition processes). To establish the validity of our meta-Bayesian framework, the two models used for the analysis of the empirical data were deliberately chosen such that one of them was considerably less plausible than the other: whereas a “dynamic” model exploited the information given to the subjects about the task, the other (“static”) model ignored this information. This established a reference point for our model comparisons (akin to the “ground truth” scenario used for validating models by simulated data). These perceptual models were combined with a loss-function embodying the task instructions to form a complete BDT response model. This loss-function

had two opposing terms, representing categorization errors and the decision time, respectively, and thus inducing the speed-accuracy trade-off of the task. We now describe the form of these probabilistic models and their inversion.

Perceptual models. The sensory signals (visual outcomes) u presented to the subjects were random samples from two sets of images, composed of eight different faces and eight different houses, respectively. A two-dimensional projection of these images onto their two first principal eigenvectors clearly shows how faces and houses cluster around two centres that can be thought of as an “average” face and house, respectively (see Figure 1). We therefore assumed the sensory inputs u to be as a univariate variable (following some appropriate dimension reduction), whose expectation depends upon the hidden state (*face* or *house*). This can be expressed as a likelihood that is a mixture of Gaussians:

$$p(u_k | x_k^{(1)}) = [N(\eta_1, \alpha^2)]^{x_k^{(1)}} [N(\eta_2, \alpha^2)]^{1-x_k^{(1)}} \quad (1)$$

Here (η_1, η_2) are the expected sensory signals caused by houses and faces (the “average” face and house images), k is a trial index,

$x_k^{(1)} \in \{0, 1\}$ is an indicator state that signals the category ($x_k^{(1)} = 1$: *house*, $x_k^{(1)} = 0$: *face*), and α is the standard deviation of visual outcomes around the average face/house images. During perceptual categorization, subjects have to recognize $x_k^{(1)}$, given all the sensory information to date. As faces and houses are well-known objects for whose categorisation subjects have a life-long experience, it is reasonable to assume that (η_1, η_2) and α are known to the subjects. The hidden category states $x_k^{(1)}$ have a prior Bernoulli distribution conditioned on the cue-outcome associative strength $x_k^{(2)}$:

$$p(x_k^{(1)} | x_k^{(2)}) = \text{Bernoulli}(s(x_k^{(2)})) = s(x_k^{(2)})^{x_k^{(1)}} (1 - s(x_k^{(2)}))^{1-x_k^{(1)}} \quad (2)$$

$$s : x \rightarrow \frac{\exp(x)}{1 + \exp(x)}$$

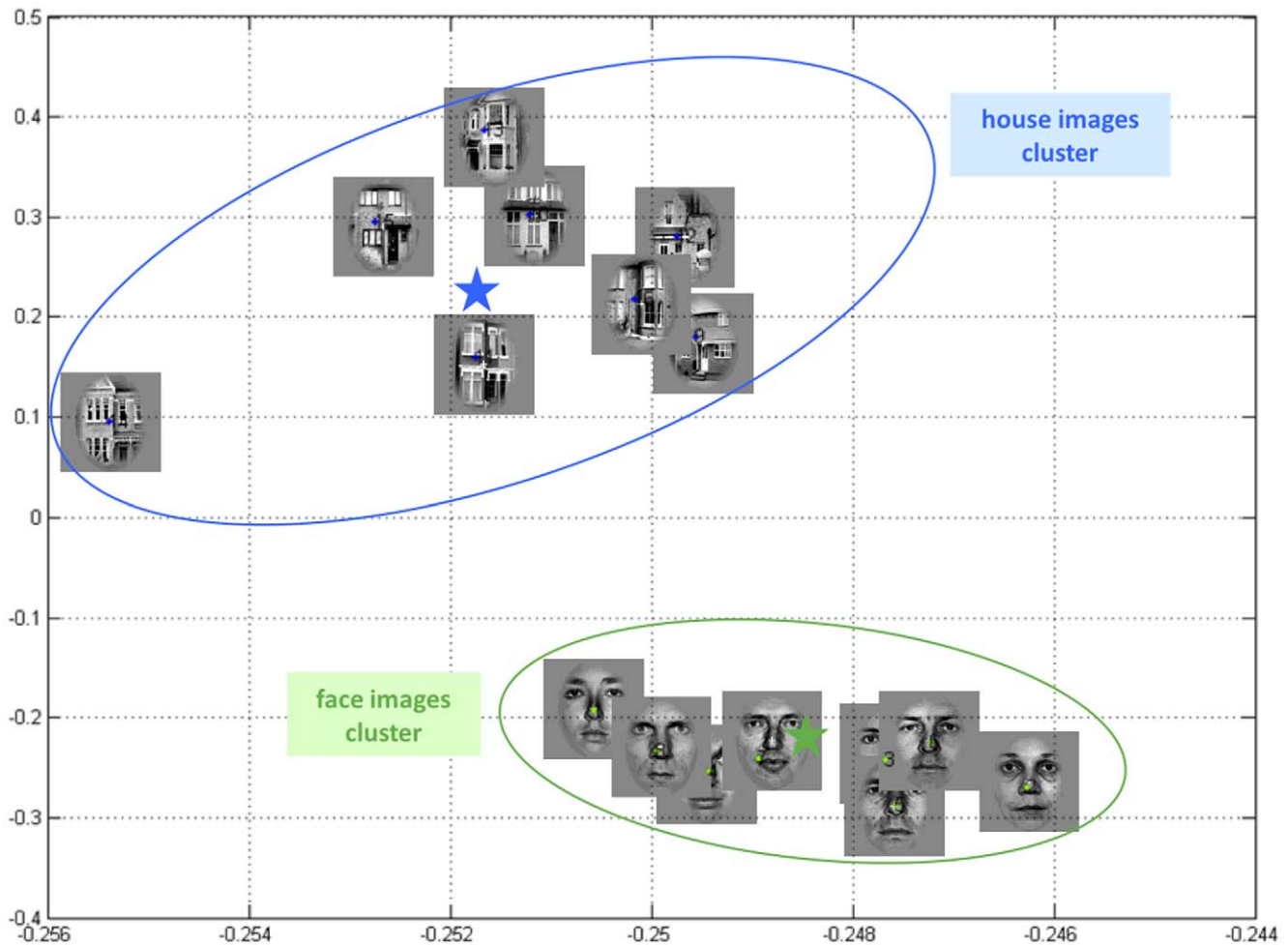


Figure 1. 2D projection of the visual stimuli that were presented to the subjects (two sets of eight face images and eight house images, respectively). X-axis: first principal component, y-axis: second principal component. On this 2D projection, house and face images clearly cluster (green and blue ellipses) around “average” face and house (green and blue stars), respectively. One might argue that these ellipses approximate the relative ranges of variations of faces and houses, as perceived by the visual system. doi:10.1371/journal.pone.0015555.g001

The sigmoid function $s(x_k^{(2)}) = p(x_k^{(1)} = 1 | C_i)$ maps the associative strength $x_k^{(2)}$ to the probability of seeing a house given the present auditory cue $C_i (i \in \{1, 2\})$. Figure 2 summarises the general structure of the perceptual models of associative learning in this paper.

We considered two perceptual models that differed only in terms of prior beliefs about the associative strength. Although both models have a prior expectation of zero for the associative strength, they differ profoundly in their predictions about how that associative strength changes over time. This is reflected by the different roles of the perceptual parameter ϑ in the two models:

- The *static perceptual model* $m_1^{(p)}$: Subjects were assumed to ignore the possibility of changes in associative strength and treat it as stationary. Under this model, subjects assume that the

associative strength has a constant value, $x_0^{(2)}$, across trials and is sampled from a Gaussian prior; *i.e.*:

$$\begin{aligned} x_k^{(2)} &= x_0^{(2)} : \forall k \\ p_{\vartheta}(x_0^{(2)}) &= N(0, \vartheta^{-1}) \end{aligned} \tag{3}$$

where ϑ is its (fixed) prior precision. Here, the perceptual parameter ϑ effectively acts as an (unknown) initial condition for the state-space formulation of the problem (see Equation 13 below).

- The *dynamic perceptual model* $m_2^{(p)}$: Subjects assumed *a priori* that the associative strength $x_k^{(2)}$ varied smoothly over time, according to

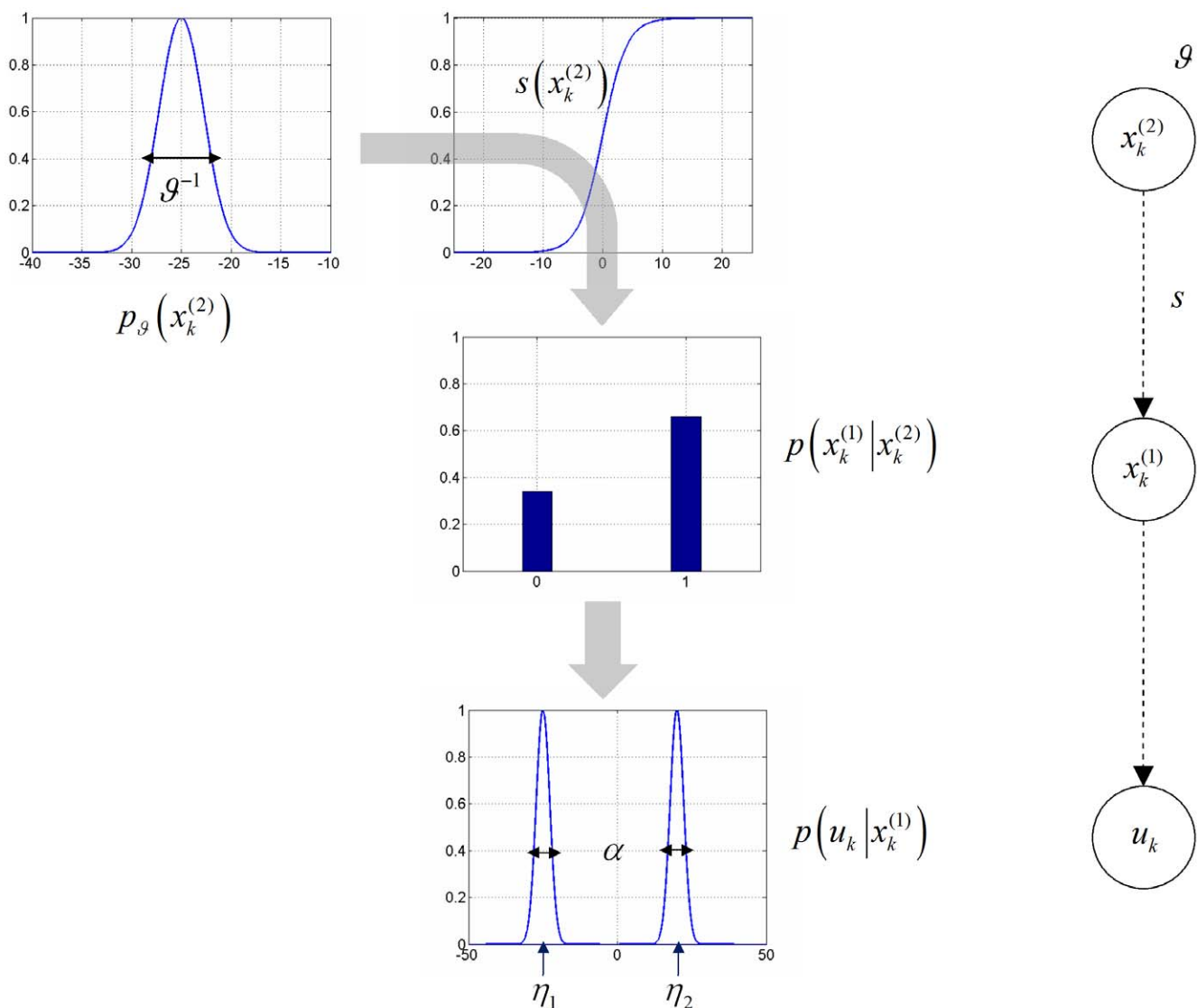


Figure 2. Conditional dependencies in perceptual models of associative learning. **Left:** cascade of events leading to the sensory outcomes. A Gaussian prior (with variance ϑ) is defined at the level of the cue-outcome association $x^{(2)}$. Passed through a sigmoid mapping, this determines the probability of getting a house ($x^{(1)} = 1$) or a face ($x^{(1)} = 0$). Finally, this determines the visual outcome u within the natural range of variation (α) of house/face images. **Right:** Equivalent graphical model. doi:10.1371/journal.pone.0015555.g002

a first-order Markov process. This is modelled as a random walk with a Gaussian transition density:

$$x_0^{(2)} = 0$$

$$p_{\mathcal{G}}(x_{k+1}^{(2)} | x_k^{(2)}) = N(x_k^{(2)}, \mathcal{G}^{-1}) \quad (4)$$

Here, \mathcal{G} is the precision hyperparameter which represents the roughness (inverse smoothness) of changes in associative strength (i.e., its volatility).

Note that the task information given to subjects did highlight the possibility of changes in cue strength. Therefore, from the point of view of the experimenter, it is more likely that the subjects relied upon the dynamic model to form their prior predictions. The choice of these two models was deliberate as it allowed for a clear prediction: we hoped to see that model comparison would show a pronounced superiority of the dynamic model (see section ‘Inverting the response model below’).

Recognition: the variational Bayesian inversion of the perceptual model. Given the perceptual models described above, we can now specify the recognition process in terms of their variational Bayesian inversion. The generic derivation of the recognition process is detailed in the companion paper [1]. In brief, subjects update their belief on-line, using successive stimuli to optimise $\lambda_k = \{\mu_k^{(1)}, \mu_k^{(2)}, \sigma_k^{(2)}\}$, the sufficient statistics of the posterior density on the k -th trial. Under a mean-field/Laplace approximation to the joint posterior, these sufficient statistics are (i) $\mu^{(1)}$, the first-order moment of the Bernoulli posterior $q(x_k^{(1)})$ about the outcome category $x^{(1)}$, and (ii) $(\mu^{(2)}, \sigma^{(2)})$, the first- and second- order moments of the Gaussian posterior $q(x_k^{(2)})$ about the associative strength $x^{(2)}$. The recognition process derives from the minimization of the surprise conveyed by sensory stimuli at each trial. Within a variational Bayesian framework, negative surprise is measured (or, more precisely, lower-bounded) via the so-called perceptual free-energy $F_k^{(p)}$ [15]:

$$F_k^{(p)} = E \left[\ln p(u_k | x_k^{(1)}) + \ln p(x_k^{(1)} | x_k^{(2)}) + \ln p(x_k^{(2)}) \right]$$

$$+ S(q(x_k^{(1)})) + S(q(x_k^{(2)}))$$

$$= -\frac{1}{2\alpha^2} (\mu_k^{(1)}(u_k - \eta_1)^2 + (1 - \mu_k^{(1)})(u_k - \eta_1)^2) - \ln \alpha - \frac{1}{2} \ln 2\pi$$

$$+ \mu_k^{(2)} (\mu_k^{(1)} - 1) + \ln s(\mu_k^{(1)}) + \frac{1}{2} (s(\mu_k^{(2)})^2 - s(\mu_k^{(2)})) \sigma_k^{(2)} \quad (5)$$

$$- \frac{1}{2\sigma_0^{(2)}} \left((\mu_k^{(2)} - \mu_{k-1}^{(2)})^2 + \sigma_k^{(2)} \right) - \frac{1}{2} \ln \sigma_0^{(2)} - \frac{1}{2} \ln 2\pi$$

$$- \mu_k^{(1)} \ln \mu_k^{(1)} - (1 - \mu_k^{(1)}) \ln (1 - \mu_k^{(1)}) + \frac{1}{2} \ln \sigma_k^{(2)} + \frac{1}{2} \ln 2\pi e$$

where the expectation is taken under the approximate posterior densities (representations) $q(x_k^{(1)})$ and $q(x_k^{(2)})$ and $S(\cdot)$ denotes the Shannon entropy. Note that the variance parameter $\sigma_0^{(2)}$ depends on the perceptual model; i.e.

$$\sigma_0^{(2)} = \sigma_{k-1}^{(2)} \quad \text{for the static model} \quad (6)$$

$$\sigma_0^{(2)} = \sigma_{k-1}^{(2)} + \mathcal{G} \quad \text{for the dynamic model.}$$

Note also that the perceptual free energy $F_k^{(p)}$ of the k -th trial depends on the representation of associative strength at the

previous trial, through the sufficient statistics $\mu_{k-1}^{(2)}$ and $\sigma_{k-1}^{(2)}$. Therefore, these affect the current optimal sufficient statistics λ_k (including that of the outcome category), allowing learning to be expressed over trials. Optimizing the perceptual free energy $F_k^{(p)}$ with respect to $q(x_k^{(1)})$ and $q(x_k^{(2)})$ yields the updated posterior densities of both the outcome category (*face or house*)

$$q(x_k^{(1)}) = \text{Bernoulli}(\mu_k^{(1)})$$

$$\mu_k^{(1)} = \frac{\pi_1}{\pi_1 + \pi_2}$$

$$\pi_1 = \exp\left(-\frac{1}{2\alpha^2} (u_k - \eta_1)^2 + \ln s(\mu_k^{(2)})\right)$$

$$+ \left(s(\mu_k^{(2)})^2 - s(\mu_k^{(2)})\right) \sigma_k^{(2)} \quad (7)$$

$$\pi_2 = \exp\left(-\frac{1}{2\alpha^2} (u_k - \eta_2)^2 + \ln(1 - s(\mu_k^{(2)})) - \mu_k^{(2)}\right)$$

$$+ \left(s(\mu_k^{(2)})^2 - s(\mu_k^{(2)})\right) \sigma_k^{(2)}$$

and of the associative strength

$$q(x_k^{(2)}) = N(\mu_k^{(2)}, \sigma_k^{(2)})$$

$$\mu_k^{(2)} = \arg \max_x I(x)$$

$$\sigma_k^{(2)} = -\left[\frac{\partial^2 I}{\partial x^2} \Big|_{\mu_k^{(2)}}\right]^{-1} = \left[\sigma_0^{(2)} + s(\mu_k^{(2)}) - s(\mu_k^{(2)})^2\right]^{-1} \quad (8)$$

$$I(x) = \ln s(x) + x(\mu_k^{(1)} - 1) - \frac{1}{2\sigma_0^{(2)}} (x - \mu_k^{(2)})^2$$

Note that functional form of the sufficient statistics above depends upon the perceptual model, through the variance parameter $\sigma_0^{(2)}$, which in turn depends upon the precision parameter \mathcal{G} (see Equation 6). This dependence is important, since it strongly affects the recognition process. Under the static perceptual model, equation 8 tells us that the subject’s posterior variance $\sigma_k^{(2)}$ about the associative strength is a monotonically decreasing function of trial index k . This means that observed cue-outcome stimuli will have less and less influence onto the associative strength representation, which will quickly converge. Under the dynamic perceptual model however, \mathcal{G} scales the influence the past representation has onto the current one. In other words, it determines the subject’s speed of forgetting (discounting): the more volatile the environment, the less weight is assigned to the previous belief (and thus past stimuli) in the current representation. The key difference between the two perceptual models thus reduces to their effective memory.

We (experimentally) estimate the parameter \mathcal{G} through inversion of the response model $m^{(p)}$, as summarized in the next section. This means the optimisation of perceptual representations has to be repeated for every value of \mathcal{G} that is considered when observing the observer, i.e. during inversion of the response model. This is an important operational aspect of meta-Bayesian inference, where

inversion of the response model entails a nested inversion of the perceptual model.

Response model: deciding when to decide. Following the description of the perceptual models, we now define the BDT mapping from representations to behaviour. We assume that subjects decide on the basis of an implicit cost that ranks possible decisions in terms of *what* decision is taken and *when* it is made. This cost is encoded by a loss-function

$$\ell_\theta(x^{(1)}, c, t) = (x^{(1)} - c)^2 + \theta_1 t \quad (9)$$

where $c \in \{0, 1\}$ is the subject's choice (*face* or *house*) and $t \in \mathbb{R}^+$ is the decision time. The first term makes a categorisation error costly, whereas the second penalizes decision time. This loss-function creates a speed-accuracy conflict, whose optimal solution depends on the loss parameter θ_1 . Since the categorization error is binary, the loss parameter θ_1 can be understood as the number of errors subjects are willing to trade against one second delay. It is formally an error rate that controls the subject-dependent speed-accuracy trade-off. This can lead to an interaction between observed reaction times and choices, of the sort that explains why people make mistakes when in a hurry (see below).

This loss function is critical for defining optimal decisions: $\ell_\theta(x^{(1)}, c, t)$ returns the cost incurred by making choice c at time t while the outcome category is $x^{(1)}$. Because subjects experience perceptual uncertainty about the outcome category, the optimal decision (c^*, τ^*) minimizes the *expected loss*, which is also referred to as *posterior risk* Q_θ (this is discussed in more detail in the companion paper):

$$(c^*, \tau^*) = \arg \min_{c, t} Q_\theta(c, t) \quad (10)$$

$$Q_\theta(c, t) = \int \ell_\theta(x^{(1)}, c, t) q(x^{(1)}) dx^{(1)}$$

Note that because the expectation is taken with regard to the posterior density on the hidden states (i.e., the belief about stimulus identity), optimal decisions (concerning both choice c and response time t) do not only depend on the loss-function ℓ , but also on the perceptual model $m^{(p)}$.

To derive how posterior risk evolves over time within a trial, we make the representation of outcome category a function of within-trial peristimulus time t (dropping the trial-specific subscript k for clarity): $\mu^{(1)} \rightarrow \mu^{(1)}(t)$. We can motivate the form of $\mu^{(1)}(t)$ by assuming that the within-trial recognition dynamics derive from a gradient ascent on the perceptual free-energy $F_k^{(p)}$. This has been recently suggested as a neurophysiologically plausible implementation of the variational Bayesian approach to perception ([3, 16; 17]). Put simply, this means that we account for the fact that optimizing the perceptual surprise with respect to the representation takes time.

At each trial, the subject's representation is initialized at her prior prediction $\mu_0^{(1)} \equiv \mu^{(1)}(0)$, and asymptotically converges to the optimum perceptual free energy $\mu_\infty^{(1)} \equiv \lim_{t \rightarrow \infty} \mu^{(1)}(t)$. (Note that the prior prediction at the beginning of a trial, $\mu_0^{(1)}$, changes over trials due to learning the predictive properties of the auditory cue; see Equations 5–8 above). It turns out (see Appendix S1) that, the posterior risk in Equation 10 can be rewritten as a function of within-trial peristimulus time t and the difference $\Delta\mu_0^{(1)} = \mu_\infty^{(1)} - \mu_0^{(1)}$ between the posterior representation and the prior prediction of the outcome category (which can thus be thought of as a *post-hoc* prediction error):

$$Q_\theta(c, t) = (1 - 2c)\mu_0^{(1)} + (1 - 2c)\Delta\mu_0^{(1)}(1 - \exp(-2\theta_2 t)) + c + \theta_1 t \quad (11)$$

where the second response parameter θ_2 is an unknown scaling factor that controls the sensitivity to *post-hoc* prediction error.

Note that in the present experimental context, the sensory evidence in favour of the outcome category is very strong. Hence, at convergence of the recognition process, there is almost no perceptual uncertainty about the outcome ($\mu_\infty^{(1)} \approx x^{(1)}$). Thus, regardless of the prior prediction $\mu_0^{(1)}$, the *post-hoc* prediction error $\Delta\mu_0^{(1)}$ is always positive when a house is presented ($x^{(1)} = 1$) and always negative when a face is shown ($x^{(1)} = 0$). This means that categorization errors occur if: (C1) $\Delta\mu_0^{(1)} < 0$ and $c = 1$, or (C2) $\Delta\mu_0^{(1)} > 0$ and $c = 0$. These conditions can be unified by rewriting them as $\Delta\mu_0^{(1)}(2c - 1) < 0$ (see the Appendix for further mathematical details). An interesting consequence is that categorization errors can be interpreted as reflecting optimal decision-making: they occur whenever the (learned) prior prediction of the visual outcome is incorrect (e.g. $\mu_0^{(1)} \approx 0$ despite $x^{(1)} = 1$) and the delay cost is high enough. In other words, categorization errors are optimal decisions if the risk of committing an error quickly is smaller than responding correctly after a longer period.

Note that when $\Delta\mu_0^{(1)}(2c - 1) > 0$ (no categorization error), the posterior risk given in equation 11 is a convex function of decision time t . The shape of this convex function is controlled by both the error rate parameter θ_1 and the sensitivity θ_2 to *post-hoc* prediction error. Finally, Equation 11 yields the optimal reaction time:

$$\tau(\lambda, \theta, c) = \arg \min_t Q_\theta(c, t)$$

$$= \begin{cases} \frac{1}{2\theta_2} \ln \frac{2\theta_2 \Delta\mu_0^{(1)}(2c - 1)}{\theta_1} & \text{if } \frac{2\theta_2 \Delta\mu_0^{(1)}(2c - 1)}{\theta_1} > 1 \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

Note that this equation has two major implications. First, as one would intuit, optimal reaction times and *post-hoc* prediction error show inverse behaviour: as the latter decreases, the former increases. Second, and perhaps less intuitive, the optimal reaction time when committing perceptual categorization errors is zero, because in this case the *post-hoc* prediction error is such that: $\Delta\mu_0^{(1)}(2c - 1) < 0$. The reader may wonder at this stage whether predicted RTs of zero are at all sensible. It should be noted that this prediction arises from the deterministic nature of Equation 12. When combined with a forward model accounting for random processes like motor noise (see Equation 13 below), non-zero predicted RTs result. Put simply, Equation 12 states that the cost of an error is reduced, if the decision time is very short.

Inverting the response model

Together with equations 7 and 8, equations 11 and 12 specify the state-space form of our response model $m^{(r)}$:

$$\begin{cases} y_k = \tau(\lambda_k, \theta, c) + \varepsilon_k \\ \lambda_k = \arg \max F_k^{(p)}(\lambda_{k-1}, \vartheta, u_k) \end{cases} \quad (13)$$

where y_k is the observed reaction time at trial k and the residuals $\varepsilon_k \sim N(0, \Upsilon)$, with precision $\Upsilon^{-1} = \theta_3$, account for (i.i.d. Gaussian) random variability in behavioural responses (e.g. motor noise). The

second (evolution) equation models recognition through the minimization of perceptual free energy (or negative sensory surprise) and the first (observation) equation models decision making through the minimization of posterior risk. The functional form of the optimal decision time is given in equation 12 (evaluating the *post-hoc* prediction error $\Delta\mu_0^{(1)}$ at the current trial) and that of the perceptual free energy is given in equation 5 (recall that learning effects are modulated by the perceptual parameter ϑ). Equation 13 basically implies that the current reaction time y_k is a nonlinear function of both the response parameters θ and the perceptual parameter ϑ , through the history of representations $\lambda_1, \lambda_2, \dots, \lambda_k$. The trial-to-trial variation of reaction times y_1, y_2, \dots, y_k therefore informs us about both the hidden loss and the belief structures of the observer.

The complete formulation of the probabilistic response model involves the definition of the likelihood function (directly derived from equation 13) and the prior density over the unknown model parameters (ϑ, θ) . Here, we use weakly informative log-normal priors (see [18]) on the perceptual parameter ϑ and the response parameters $\{\theta_1, \theta_2\}$ to enforce positivity. These are given in table 1. In addition, the variational Bayesian inversion of the response model makes use of a mean field approximation $p(\theta, \vartheta | y, m^{(r)}) \approx r(\theta_1, \theta_2, \vartheta) r(\theta_3)$ that separates the noise precision parameter θ_3 from the remaining parameters. Lastly, we relied on a Laplace approximation to the marginal posterior $r(\theta_1, \theta_2, \vartheta)$. This reduces the Bayesian inversion to finding the first- and second-order moments of the marginal posterior (see equations 13 and 14 in the companion paper [1] for a complete treatment).

The algorithmic implementation of the variational Bayesian inversion of the response model is formally identical to that of a Dynamic Causal Model (DCM, see e.g. [19] for a recent review). The variational Bayesian scheme furnishes the approximate marginal posteriors and a lower bound on the response model evidence (via the response free energy $F^{(r)} \approx p(y | m^{(r)})$), which is used for model comparison. One can also recover the representations since these are a function of the perceptual parameter ϑ , for which we obtain a posterior density $r(\vartheta)$ (see equations 12 and 14 in the companion paper [1]):

$$\begin{aligned} \hat{q}(x^1) &\approx \text{Binom}(\hat{\mu}^{(1)}) \\ \hat{\mu}^{(1)} &= \mu^{(1)} \Big|_{\hat{\vartheta}} \\ \hat{q}(x^2) &\approx N(\hat{\mu}^{(2)}, \hat{\sigma}^{(2)}) \\ \hat{\mu}^{(2)} &= \mu^{(2)} \Big|_{\hat{\vartheta}} \\ \hat{\sigma}^{(2)} &= \sigma^{(2)} \Big|_{\hat{\vartheta}} + \frac{\partial \mu^{(2)}}{\partial \vartheta} \Big|_{\hat{\vartheta}} \text{Var}[\vartheta | y] \frac{\partial \mu^{(2)}}{\partial \vartheta} \Big|_{\hat{\vartheta}}^T \end{aligned} \tag{14}$$

Table 1. First and second order moments of the prior density over perceptual and response parameters (under both static and dynamical perceptual models).

parameter	prior mean	prior variance
log ϑ (dynamic perceptual model)	0	10^2
(static perceptual model)	2	10^2
log $\theta_{1,2}$	$[0 \ 0]^T$	$10^2 I_2$
θ_3	10^4	10^6

Note that we used log-normal priors for ϑ and $\theta_{1,2}$, and a Gamma prior for the residuals' precision θ_3 .

doi:10.1371/journal.pone.0015555.t001

where all sufficient statistics and gradients are evaluated at the mode $\hat{\vartheta} \equiv \int \vartheta r(\vartheta) d\vartheta$ of the approximate posterior $r(\vartheta)$ and $\text{Var}[\vartheta | y] \equiv \int (\vartheta - \hat{\vartheta})^2 r(\vartheta) d\vartheta$ is the experimenter's posterior variance about the perceptual parameter.

Results

In what follows, we first apply our approach to simulated data in order to establish the face validity of the scheme, both in terms of model comparison and parameter estimation. We then present an analysis of the empirical reaction-time data from the audio-visual associative learning task in [12].

Monte-Carlo evaluation of model comparison and parameter estimation

We conducted two series of Monte-Carlo simulations (sample size = 50), under the static (series A) and dynamic perceptual models (series B). In each series, the (log) perceptual parameters were sampled from the intervals $[0, 3]$ for series A and $[-2, 2]$ for series B. For both series, the first two (log) response parameters were sampled from the interval $[-2, 2]$. As an additional and orthogonal manipulation, we systematically varied the noise on reaction times across several orders of magnitude: $\theta_3 \in \{1, 10^{-2}, 10^{-4}\}$. Each simulated experiment comprised a hundred trials and the sequence of stimuli was identical to that used in the real audio-visual associative learning study. We chose the parameters (α, μ) of the perceptual likelihood such that the discrimination ratio $(\alpha / |\mu_1 - \mu_2| \approx 10)$ was approximately similar to that of the natural images (see Figure 1). We did not simulate any categorization error. For each synthetic data set, we used both static and dynamic perceptual models for inversion of the response model and evaluated the relative evidence of the perceptual models. Since we knew the ground truth (i.e., which model had generated the data) this allowed us to assess the veracity of model comparison.

Figure 3 shows a single example of simulated recognition, in terms of the subject's belief about both the stimulus and the cue-outcome association. For this simulation, the volatility of the association was set to $\log \vartheta = -2$ (emulating a subject who assumes a low volatile environment), both for generating stimuli and recognition. We found that the variational Bayesian recognition recovers the stimulus categories perfectly (see blue line in upper-right panel of figure 3) and the cue-outcome association strength well (see lower-left panel and green lines in upper-right panels). This demonstrates that variational recognition is a close approximation to optimal Bayesian inference.

Figure 4 shows the inversion of the response model, given the synthetic reaction time data in Figure 3 which were corrupted with unit noise ($\theta_3 = 1$). Adding this observation noise yielded a very low signal-to-noise ratio (SNR = 0 dB, see Figure 4), where by definition: $\text{SNR} = 10 \log_{10} \langle \tau \rangle^2 / Y$. We deliberately used this high noise level because it corresponded roughly to that seen in the empirical data reported below. Table 1 lists the priors we placed on the parameters for this example and for all subsequent inversions with the dynamic perceptual model. Despite the low SNR of the synthetic data, the posterior estimates of the response parameters (grey bars) were very close to the true values (green circles), albeit with a slight overconfidence (upper left panel in Figure 4). Furthermore, the posterior correlation matrix shows that the perceptual and the response parameters are identifiable and separable (upper centre panel). The non-diagonal elements in the posterior covariance matrix measure the degree to which any pair of parameters is non-identifiable (see appendix in the companion paper [1]). Note that the model fit looks rather poor

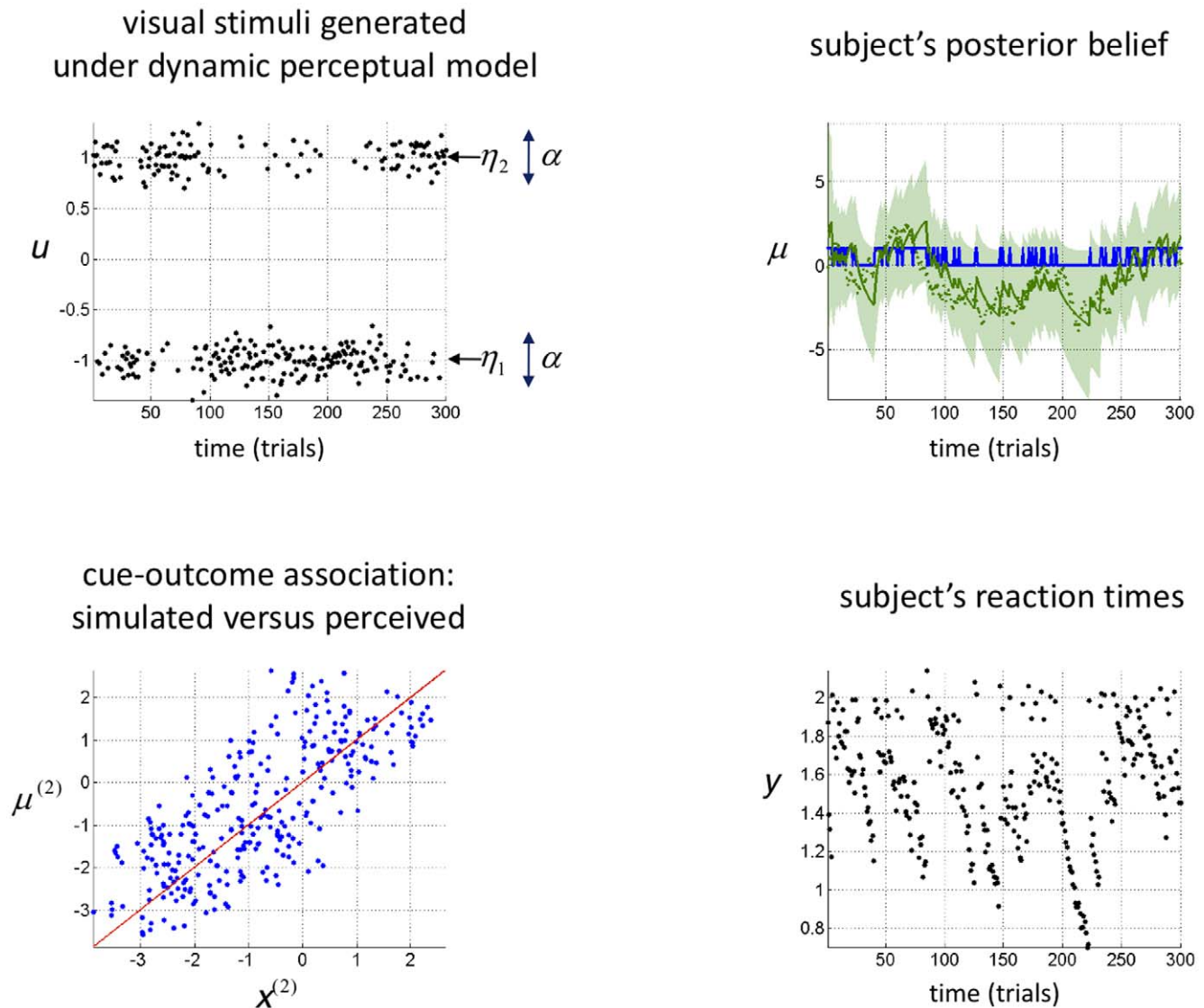


Figure 3. Variational Bayesian recognition of visual stimuli: Upper Left: time series of sensory cues, sampled from the generative model summarized in Figure 2. Note that the discrimination ratio ($\alpha/|\mu_1 - \mu_2|$) is approximately similar to that of the natural images (see Figure 2). **Upper Right:** Subject's posterior belief, as obtained using the inversion of the perceptual model given observed sensory cues (green: cue-outcome association, blue: visual stimulus category; solid line: posterior mean $\mu^{(2)}$, shaded area: 99% posterior confidence interval, dots: sampled hidden states). Note that on each trial, the category of the visual stimuli was recognized perfectly. **Lower Left:** scatter plot comparing the simulated (sampled, x-axis) versus perceived (estimated, y-axis) cue-outcome associative strength. **Lower right:** simulated reaction times. doi:10.1371/journal.pone.0015555.g003

and gives the impression that the RT data are systematically “under-fitted” (lower right and lower centre panels of Fig. 4). This, however, is simply due to the high levels of observation noise: In contrast, the estimation of the true subjective beliefs is precise and accurate (see upper right and lower left panels of Fig. 4). This means that the variational Bayesian model inversion has accurately separated the “observed” reaction time data into noise and signal components. In other words, the estimation of the deterministic trial-by-trial variations of reaction times is not confounded by high levels of observation noise. This result (using simulated data) is important because it lends confidence to subsequent analyses of empirical reaction time data.

Figure 5 shows the results of the model comparison based on series A and B. This figure shows the Monte-Carlo empirical distribution of the response free-energy differences $F_{\text{static}}^{(r)} - F_{\text{dynamic}}^{(r)}$,

where $F_{\text{static}}^{(r)}$ (respectively $F_{\text{dynamic}}^{(r)}$) is the approximate log-evidence for the response model under the static (respectively dynamic) perceptual model. This relative log-evidence is the approximate log-Bayes factor or log odds ratio of the two models. It can be seen from the graphs in Figure 5 that model comparison identifies the correct perceptual model with only few exceptions for the static model (left panel) and always for the dynamic model (note that a log-evidence difference of zero corresponds to identical evidence for both models). Table 2 provides the average free-energy differences over simulations as a function of the true model (simulation series) and SNR.

It is interesting that the free-energy differences are two orders of magnitude larger for series B, relative to series A. In other words, when the data-generating model is the dynamic one, it is easier to identify the true model from reaction times than when the static model generated the data. This might be due to the fact that the

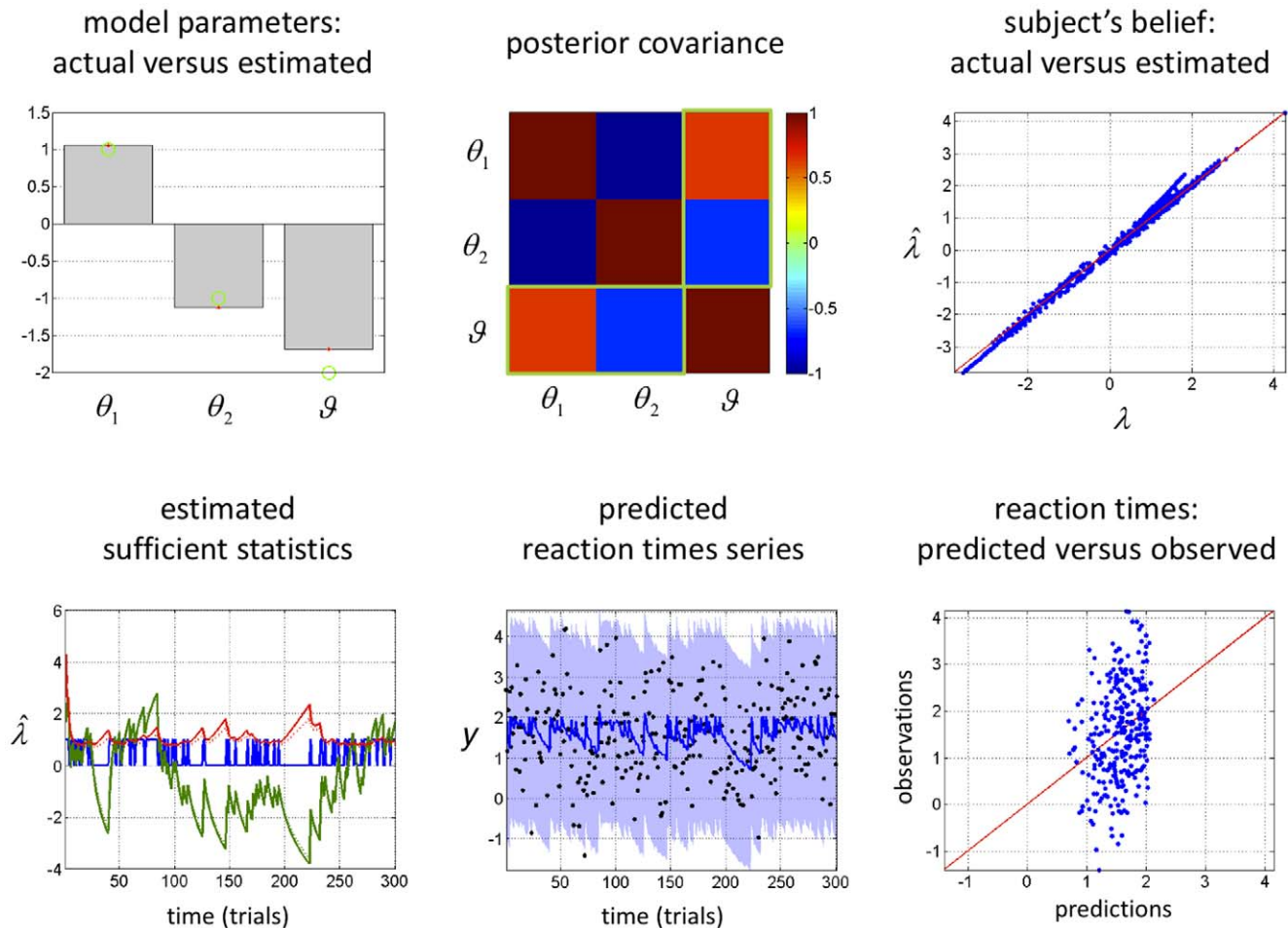


Figure 4. Observing the observer: follow-up example from Figure 3. **Upper left:** Comparison between the estimated (grey bars) and actual (green dots) perceptual and response parameters. Note that simulated and estimated parameters are shown in log-space. **Upper centre:** posterior joint correlation matrix of the perceptual and response parameters (the green rectangles depict the correlation between the perceptual parameter ϑ and the response parameters $\theta_{1,2}$). **Upper Right:** scatter plot comparing the simulated (x-axis) and estimated (y-axis) sufficient statistics λ of the approximate subject's posterior. **Lower left:** time series of estimated (solid lines) and simulated (dotted lines) sufficient statistics λ of the approximate subject's posterior (blue: cue identity, green: expected association, red: posterior variance of the associative strength). **Lower centre:** time series of the simulated (black dots) and predicted (solid line: posterior expectation, shaded area: 99% confidence interval) reaction times. **Lower right:** scatter plot comparing the simulated (y-axis) versus predicted (x-axis) reaction times.
doi:10.1371/journal.pone.0015555.g004

static model is a limiting case of the dynamic model; i.e. when the volatility ϑ tends to zero the dynamical perceptual model can account for the variability in reaction times generated using the static model. However, note that this difference in model complexity does not distort or bias our model comparisons since the free energy approximation to the model evidence accounts for such differences in complexity [20].

As expected there is also a clear effect of noise: the higher the SNR, the larger the relative log-evidences. This means that model comparison will disambiguate models more easily the more precise the experimental data.

We next characterised the accuracy of parameter estimation under the best (correct) model, using the sum of squared error (SSE), in relation to the true values. We computed the Monte-Carlo empirical distribution of the SSE for each set of (perceptual and response) parameters, for each simulation series (A and B) and SNR. Figure 6 shows these distributions and Table 3 provides the Monte-Carlo averages.

Quantitatively, the parameters are estimated reasonably accurately, except for the perceptual parameter (prior precision

on associative strength) of the static model. This reflects the fact that the prior on association strength has little impact on the long-term behaviour of beliefs, and hence on reaction times. This is because within the static perceptual model, ϑ acts as an initial condition for the dynamics of the representation λ , which are driven by a fixed point attractor that is asymptotically independent of ϑ . Thus, only the first trials are sensitive to ϑ . The ensuing weak identifiability of ϑ expressed itself as a high estimation error (high SSE). Again, there is a clear effect of noise, such that the estimation becomes more accurate when SNR increases. Also, consistent with the model comparison results above, parameter estimates are more precise for the dynamic model than for the static one.

Application to empirical reaction times

The Monte-Carlo simulations above demonstrate the face validity of the method, in the sense that one obtains veridical model comparisons and parameter estimates, given reaction time data with realistic SNR. We now apply the same analysis to empirical reaction times from nineteen subjects [12]. Specifically,

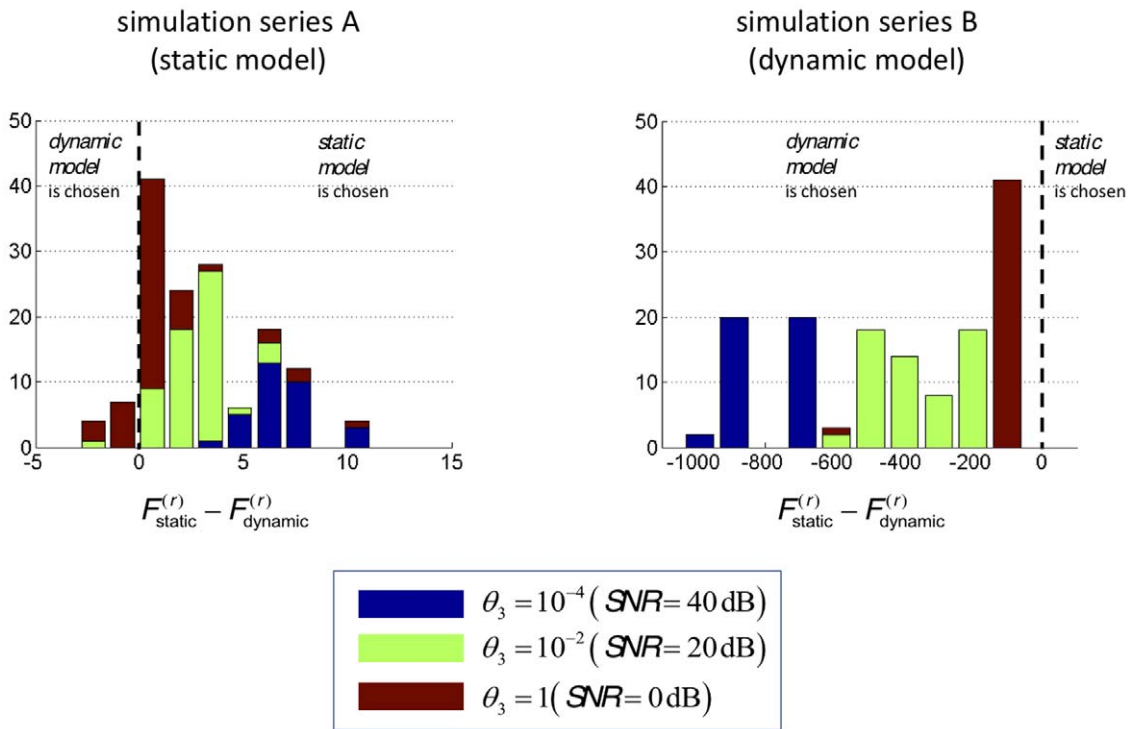


Figure 5. Monte-Carlo empirical distributions of model log-evidence differences ($F_{static}^{(r)} - F_{dynamic}^{(r)}$). Blue: SNR = 40 dB, green: SNR = 20 dB, red: SNR = 0 dB. **Left:** Monte-Carlo simulation series A (under the static perceptual model). **Right:** Monte-Carlo simulation series B (under the dynamic perceptual model). doi:10.1371/journal.pone.0015555.g005

we hoped to show two things to provide evidence for the construct validity of our approach: first, that the dynamic model (which was consistent with the information given to the subjects) would have higher evidence than the static model (which was not), and secondly, that our results would generalise over both auditory cues, both in terms of model comparison and parameter estimates (as explained above, we treated reaction times for the two cues as separate sequences).

We conducted a hierarchical (two-level) analysis of the data from the nineteen subjects. Note that the original study by [12] contained twenty subjects. For experimental reasons, one of these subjects experienced a different stimulus sequence than the rest of the group. Even though it would have been perfectly possible to analyze this subject with the present approach, we decided, for reasons of homogeneity in the inter-subject comparison, to focus on subjects with identical stimulus sequence. In a first-level analysis, we inverted both dynamic and static models on both type I cues (high pitch tones) and type II cues (low pitched tones)

separately, for each subject. As in the simulations above, the parameters (α, μ) of the perceptual likelihood (equation 1) were chosen such that stimulus discriminability ($\alpha/|\mu_1 - \mu_2| \approx 10$) was similar to that of the natural images (see Figure 1). Also, categorization errors were assigned a response time of zero (see histograms in upper right panels of Figs. 12–13) and a very low precision Υ , relative to the other trials. This allowed us to effectively remove these trials from the data without affecting the trial-to-trial learning effects.

Figure 7 summarizes the model comparison results for each subject, showing the difference in log-evidence for both auditory cues. A log-evidence difference of three (and higher) is commonly considered as providing strong evidence for the superiority of one model over another [21]. Using this conventional threshold, we found that in 13 subjects out of 19 the competing perceptual models could be disambiguated clearly for at least one cue type. It can be seen that for all of these subjects except one the dynamic perceptual model was favoured. Also, it was reassuring to find that the variability of response model evidences across cue types was much lower than its variability across subjects. In particular, in 10 out of the 13 subjects where the perceptual models could be distinguished clearly, the model evidences were consistent across cue types.

In a second step, we performed a Bayesian group-level random effect analysis of model evidences [20]. Assuming that each subject might have randomly chosen any of the two perceptual models, but consistently so for both cues, we used the sum of the subject-specific log-evidences over both cues for model comparison at the group level. Figure 8 shows the ensuing posterior Dirichlet distribution of the probability $\mathcal{G}_{dynamic}$ of the dynamic perceptual model across the group, given all datasets. Its posterior expectation was approximately $E[\mathcal{G}_{dynamic}|y] \approx 0.82$. This indicates how

Table 2. Monte-Carlo averages of log-evidence differences as a function of simulation series (A: static and B: dynamic) and SNR.

	Series A (static) $F_{static}^{(r)} - F_{dynamic}^{(r)}$	Series B (dynamic) $F_{dynamic}^{(r)} - F_{static}^{(r)}$
$\theta_3 = 10^{-4} \Rightarrow SNR = 40dB$	7.52	752.9
$\theta_3 = 10^{-2} \Rightarrow SNR = 20dB$	3.22	320.7
$\theta_3 = 1 \Rightarrow SNR = 0dB$	1.86	28.6

doi:10.1371/journal.pone.0015555.t002

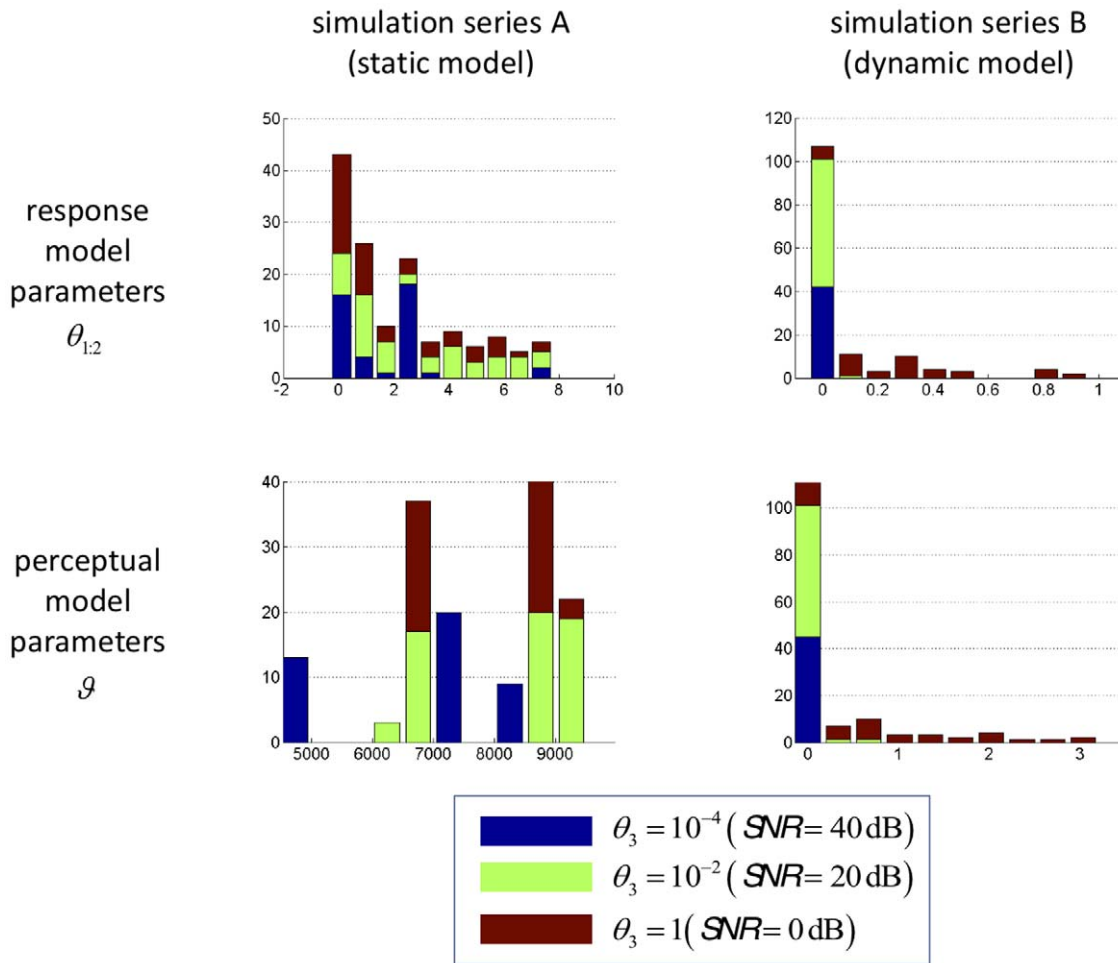


Figure 6. Monte-Carlo empirical distributions of the parameter estimation error (SSE score). Blue: SNR = 40 dB, green: SNR = 20 dB, red: SNR = 0 dB. **Upper left:** response model parameters, Monte-Carlo simulation series A. **Upper right:** response model parameters, Monte-Carlo simulation series B. **Lower left:** perceptual model parameters, Monte-Carlo simulation series A. **Lower right:** perceptual model parameters, Monte-Carlo simulation series B. Note that the SSE score was evaluated in log-space. doi:10.1371/journal.pone.0015555.g006

frequently the dynamic model won the model comparison within the group, taking into account how discernable these models were. We also report the so-called “exceedance probability” of the dynamic model being more likely than the static model, given all

datasets: $P(\vartheta_{dynamic} \geq \vartheta_{static} | y) = 0.999$. This measures the overall strength of evidence in favour of the dynamic perceptual model, at the group level. This is a pleasing result because, as described above, the dynamic model (where subjects assume *a priori* that the cue-outcome association is varying in time) was consistent with the information delivered to the subjects (whereas the static model was not).

Having established the dynamic model as the more likely model of reaction time data at the group level, we now focus on the actual estimates of both response and perceptual parameters. First, we tested for the reliability of the parameter estimates, that is, we asked whether the subject-dependent posterior densities $r(\theta, \vartheta)$ of the perceptual and response parameters were reproducible across both types of cues. Specifically, we hoped to see that the variability across both types of cues was smaller than the variability across subjects. For the three parameters $[\vartheta, \theta_1, \theta_2]$, figures 9, 10 and 11 display the variability of the posterior densities across both cues and all subjects, taking into account the posterior uncertainty $Var[\theta, \vartheta | y]$ (see equation 14). First, it can be seen that there is a consistent relationship between cue-dependent parameter estimates. Second, there is a comparatively higher dispersion of parameter estimates across subjects than across cues. Taken together, this demonstrates the reliability of parameter estimates in

Table 3. Monte-Carlo averages of the SSE as a function of simulation series (A or B) and SNR, for perceptual and response parameters.

		Series A (Static)	Series B (Dynamic)
perceptual parameters	SNR = 40dB	$7.8 \cdot 10^3$	$6.0 \cdot 10^{-4}$
	SNR = 20dB	$8.4 \cdot 10^3$	$6.1 \cdot 10^{-2}$
	SNR = 0dB	$8.2 \cdot 10^3$	1.09
response parameters	SNR = 40dB	2.20	$2.0 \cdot 10^{-4}$
	SNR = 20dB	3.39	$2.1 \cdot 10^{-2}$
	SNR = 0dB	2.38	$3.5 \cdot 10^{-1}$

doi:10.1371/journal.pone.0015555.t003

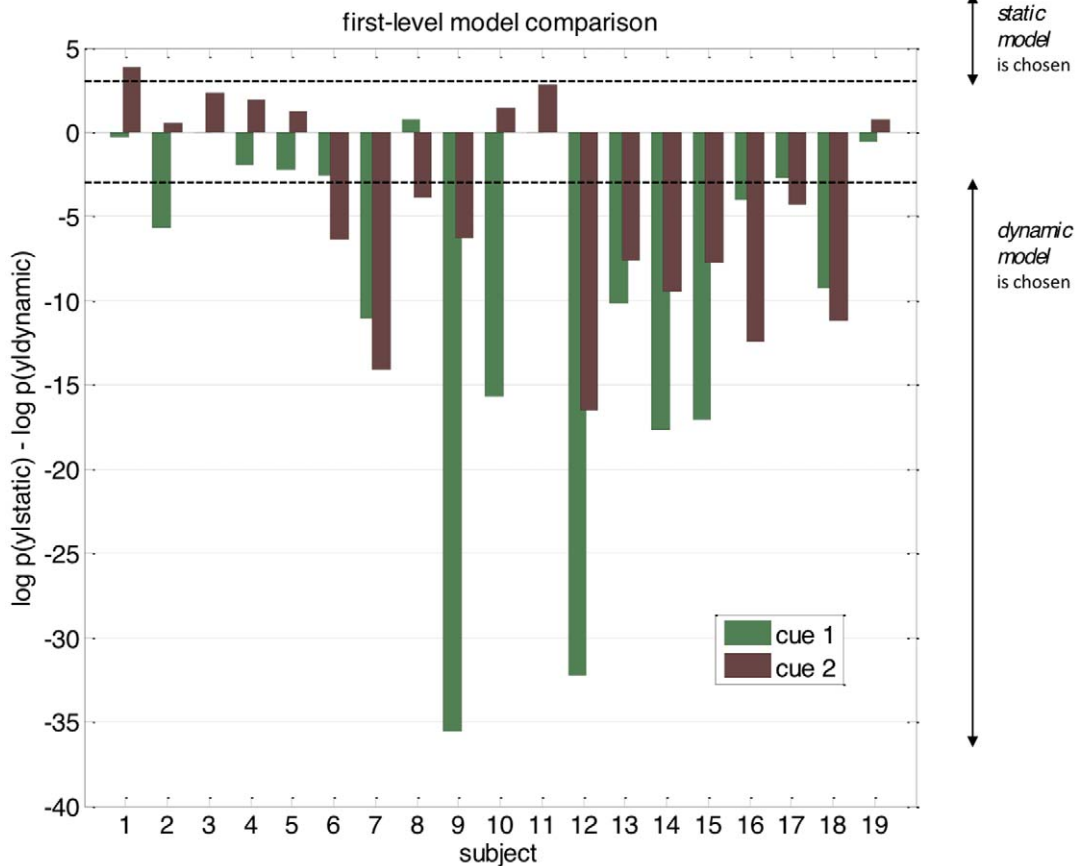


Figure 7. Subject-level model comparison. The graph is a bar plot of the difference in model evidence for the static model versus the dynamic model ($F_{\text{static}}^{(r)} - F_{\text{dynamic}}^{(r)}$), for each subject (along the x-axis) and each cue (green: type I cue, red: type II cue). doi:10.1371/journal.pone.0015555.g007

the context of empirically measured behavioural data with low SNR (i.e., reaction times). This implies that one can obtain robust and subject-specific inferences with our approach.

Such inferences concern both subject-specific a priori beliefs (e.g., about the stability of the environment; see equations 3 and 4) and preferences (as encoded by their individual loss function; see equation 9). To demonstrate the potential of our approach for characterizing inter-individual differences in beliefs and preferences, we present a subject-specific summary of the inverted response model (under the dynamic perceptual model) for two individuals (subjects 5 and 12). These results are summarized by Figures 12 and 13.

First, we would like to stress that, as for the group as a whole, the SNR of empirical data from these two subjects is similar to the SNR of the Monte Carlo simulation series described above (around 0 dB; see Figure 4). It is therefore not surprising that the model fit to the empirical looks similarly bad as in our simulations (compare upper left panels in Figs. 12-13 with lower right panel in Fig. 4). Note, however, that our simulations demonstrated that despite this poor fit the model parameters were estimated with high accuracy and precision; this instils confidence in the analysis of the empirical data.

Even though the two histograms of reaction time data from these two subjects were almost identical (compare upper right panels in figures 12 and 13), the trial-to-trial variations of reaction time data allowed us to identify rather different subject-specific structures of beliefs and preferences (loss functions). Concerning

the beliefs of these two individuals, the parameter estimates indicated that, a priori, subject 5 assumed a much more stable environment (i.e., had a much lower prior volatility \mathcal{G}) than subject 12; as a consequence, the dynamics of her estimates of the associative strength $\mu^{(2)}$ are considerably smoother across trials (compare lower left panels in figures 12 and 13). In other words, she averaged over more past cue-outcome samples when updating her posterior belief or representation than subject 12. Another consequence of this is the fact that subject 5 uncertainty $\sigma^{(2)}$ about the associative strength is much smaller and less “spiky” than subject 12’s. This has an intuitive interpretation: since subject 12 assumes a volatile environment, a series of predicted visual outcomes (approaching a nearly deterministic association) is highly surprising to her. This causes high perceptual uncertainty about the tracked associative strength whenever its trial-to-trial difference approaches zero.

As for the preferences (loss functions) that guided the actions of these two subjects, subject 12 displays a greater variability in her optimal decision times for very small post-hoc prediction errors ($\mu_{\infty}^{(1)} - \mu_0^{(1)}$, see equation 12). As a consequence, her optimal decision time is greater than that of subject 5, for any given magnitude of the *post-hoc* prediction error (compare lower right panels in Figures 12 and 13). This is because both subject 12’s error rate (i.e., θ_1) and sensitivity to *post-hoc* prediction error (i.e., θ_2) is smaller than subject 5’s.

In summary, subject 12 is assuming a more variable associative strength. This means that, when compared to subject 5, she

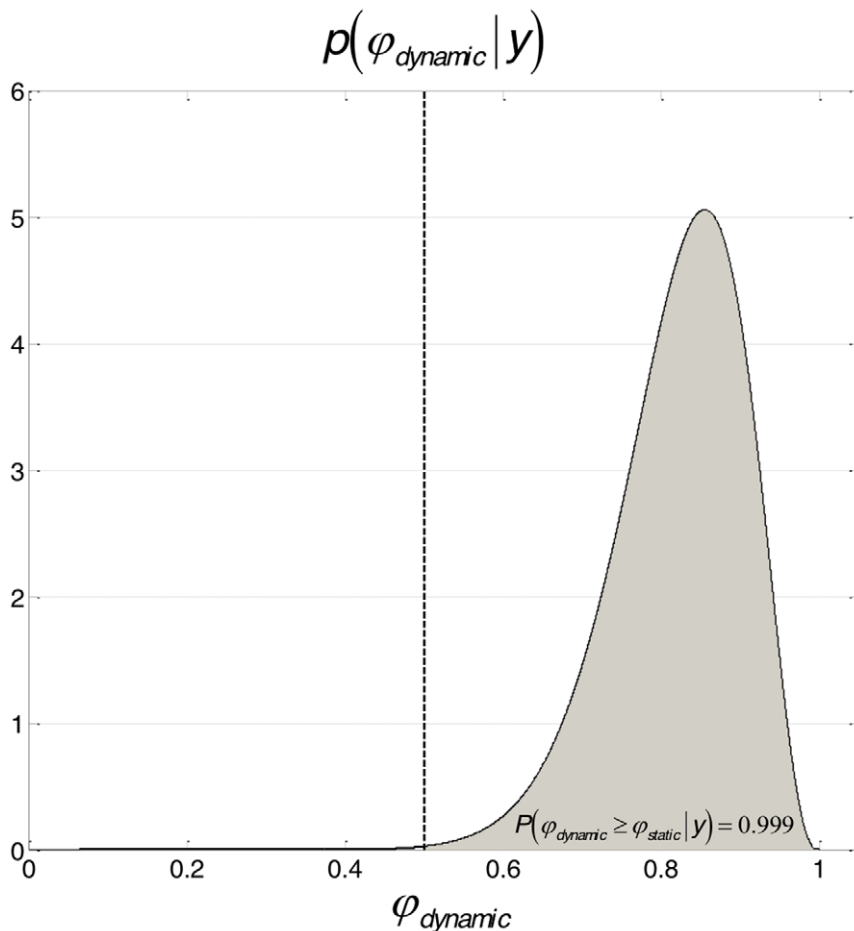


Figure 8. Group-level model comparison. Dirichlet posterior distribution of the frequency (within the group of subjects) $\vartheta_{dynamic}$ of the dynamic model, given all subjects data $p(\vartheta_{dynamic}|y)$. The grey area depicts the exceedance probability $P(\vartheta_{dynamic} \geq \vartheta_{static}|y)$, i.e. the probability that the dynamic model is more likely (within the group) than the static one.
doi:10.1371/journal.pone.0015555.g008

discards information about past cue-outcome associations more quickly and has more uncertain (prior) predictions about the next outcome. However, she is willing to make more categorization errors per second delay than subject 5. This is important, since she effectively needs more time to update her uncertain (i.e. potentially inaccurate) prior prediction to arrive at a correct representation. In contrast, subject 5 is more confident about her prior predictions and is more willing to risk categorization errors in order to gain time.

Discussion

In a companion paper [1], we have described a variational Bayesian framework for approximating the solution to the Inverse Bayesian Decision Theory (IBDT) problem in the context of perception, learning and decision-making studies. We propose a generic statistical framework for (i) comparing different combinations of perceptual and response models and (ii) estimating the posterior distributions of their parameters. Effectively, our approach represents a meta-Bayesian procedure which allows for Bayesian inferences about subject's Bayesian inferences. In this paper, we have demonstrated this approach by applying it to a simple perceptual categorization task that drew on audio-visual associative learning. We have focused on the problem of 'deciding when to decide', i.e. we have modelled reaction time data as

arising from subjective beliefs and preferences under the constraint of a speed-accuracy trade-off. This model is novel and quite different from classical evidence accumulation and 'race' models (e.g. [22,23,10]), in two ways. First, a reaction time is understood in terms of the convergence speed of an optimization process, i.e. perceptual recognition. This is because it takes time for a (variational) Bayesian observer to arrive at an optimal representation or belief. In this paper, the within-trial (peri-stimulus time) dynamics of the recognition process emerged from a gradient-ascent on the free-energy, where free-energy is a proxy for (negative) perceptual surprise under a given perceptual model. The resulting form of the response model is analytically tractable and easy to interpret. Second, the variability of reaction times across subjects is assumed to depend on individual differences in prior beliefs (e.g., about the stability of the environment) and preferences (i.e., loss or utility functions). Our approach thus provides insights into both within-trial mechanisms of perception as about inter-individual differences in beliefs and preferences.

In this work, we have chosen to focus on modelling reaction time data and have deliberately ignored categorization errors. This is because considering both reaction time and choice data at the same time would have required an extension of the response likelihood. The difficulty here is purely technical: the ensuing bivariate distribution is Bernoulli-Gaussian, whose sufficient statistics follow from the posterior risk (equation 11). Although

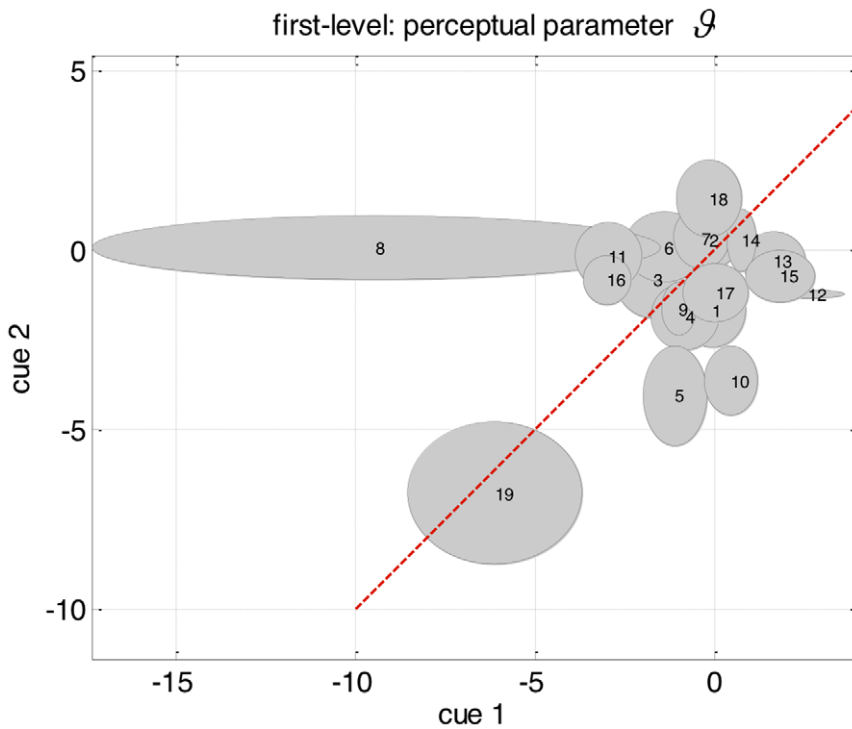


Figure 9. Plot of the reliability of perceptual parameter estimates $\hat{\vartheta}$ across cue types. The perceptual parameter estimate $\hat{\vartheta} = E[\vartheta|y]$ and its posterior variance $Var[\vartheta|y]$ are plotted as a function of cue types (on the x and y axis) and shown as an ellipse for each subject. The centre of the ellipse represents $\hat{\vartheta}$ for each cue, and its vertical and horizontal axis show one posterior standard deviation around it. The red line shows the ideal positions of the parameter estimates (the centre of the ellipses) if there was perfect reliability (i.e. no variability across cue types).
doi:10.1371/journal.pone.0015555.g009

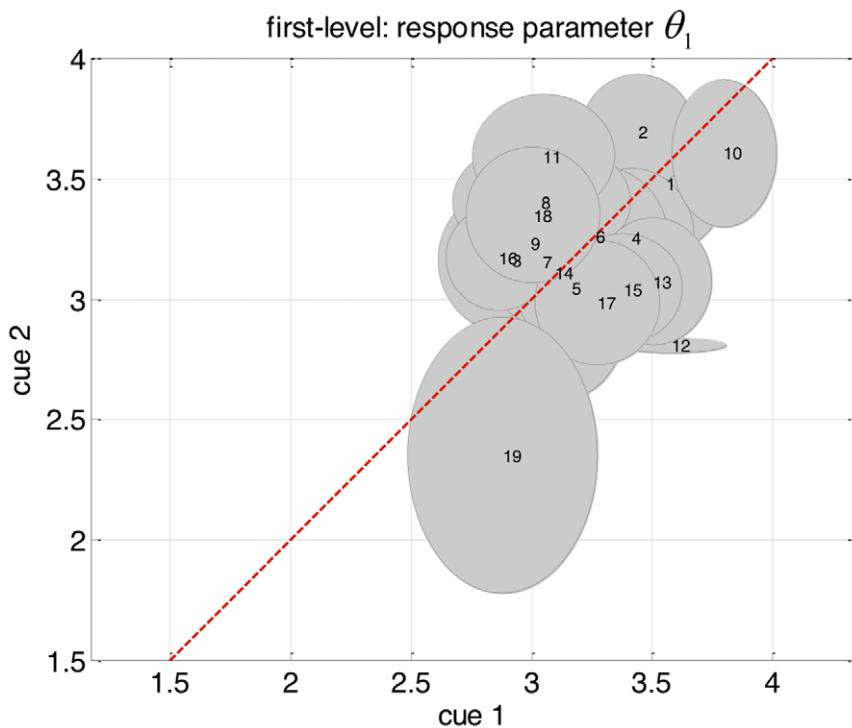
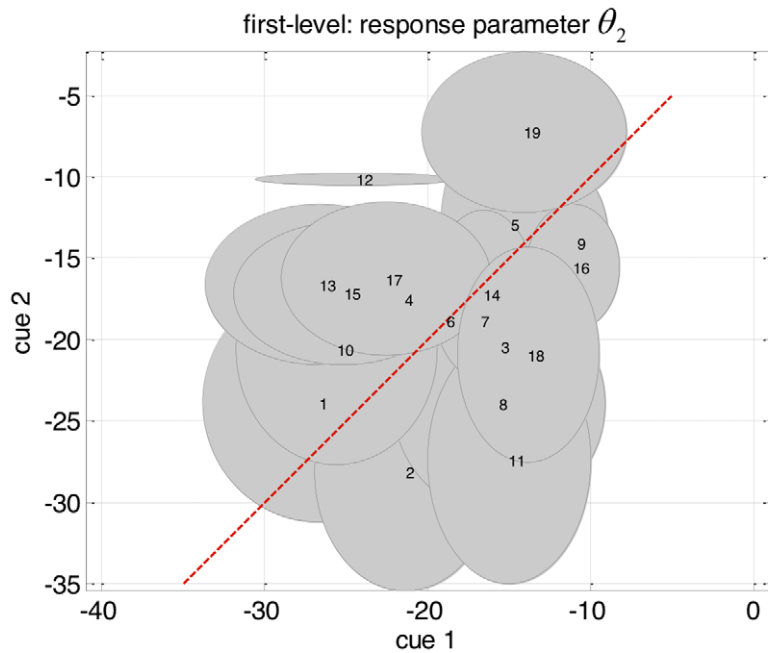


Figure 10. Plot of the reliability of response parameter estimates $\hat{\theta}_1$. See legend to Fig. 9 for explanations.
doi:10.1371/journal.pone.0015555.g010



1

Figure 11. Plot of the reliability of response parameter estimates $\hat{\theta}_2$. See legend to Fig. 9 for explanations.
doi:10.1371/journal.pone.0015555.g011

feasible, deriving this extended response model would have significantly increased its complexity. Since the focus of this article was to provide a straightforward demonstration of our theoretical framework (described in the companion paper), we decided not to include choice data in the response model. Clearly, this is a limitation as we are not fully exploiting the potential information about underlying beliefs and preferences that is provided by observed categorization errors. This extension will be considered in future work.

In our model, categorization errors arise when incorrect prior predictions coincide with high delay costs (see equations 11 and 12). One might think that there is an irreconcilable difference between this deterministic scheme and stochastic diffusion models of binary decisions ([24]; see also [25], for a related Bayesian treatment). However, there are several ways in which our scheme and stochastic diffusion models can be reconciled. For example, the trial-wise deterministic nature of our scheme can be obtained by choosing the initial condition of the stochastic process such that the probability of reaching the upper or lower decision threshold is systematically biased in a trial-by-trial fashion. Also, delay costs can be modelled by letting the distance between lower and upper diffusion bounds shrink over time. Alternatively, one could motivate the form of stochastic diffusion models by assuming that the brain performs a stochastic (ensemble) gradient ascent on the free energy. This would relate the frequency of categorization errors to task difficulty; for example, when a stimulus is highly ambiguous or uncertain, the perceptual free energy landscape is flat (perceptual uncertainty is related to the local curvature of perceptual free energy; see equations 5 and 6 of the companion paper). In summary, there are several ways in which our approach and stochastic diffusion models could be formally related. The utility of such hybrid models for explaining speed-accuracy trade-offs (cf. [26]) will be explored in future work.

We initially evaluated the method using Monte-Carlo simulations under different noise levels, focusing on model inversion

given synthetic data and on how well alternative models could be disambiguated. This enabled us to assess both the efficiency of parameter estimation and veracity of model comparison as a function of SNR. Importantly, we found that even under very high noise levels (SNR = 0dB, comparable to the SNR of our empirical data), and therefore poor model fit, the model nevertheless (i) yielded efficient estimates of parameters, enabling us to infer and track the trial-to-trial dynamics of subjective beliefs from reaction time data, and (ii) robustly disambiguated correct and wrong models. We then applied the approach to empirical reaction times from 19 subjects performing an associative learning task, demonstrating that both model selection results and parameter estimates could be replicated across different cue types. Reassuringly, the model selection results were consistent with the information available to the subjects. In addition, we have shown that subject-to-subject variability in reaction times can be captured by significant differences in parameter estimates (consistently again across cue types) where these parameters encode the prior beliefs and preferences (loss functions) of subjects,

Together, the simulations and empirical analyses establish the construct validity of our approach and illustrate the type of inference that can be made about subjects' priors and loss-functions. Our results suggest that the approach may be fairly efficient when it comes to comparing and identifying models of learning and decision-making on the basis of (noisy) behavioural data such as reaction times.

Some readers may wonder why we have used a relatively complicated criterion to evaluate the relative goodness of competing models; i.e., an approximation to the log-evidence, instead of simply comparing their relative fit. Generally, pure model fit indices are not appropriate for comparing models and should be avoided (cf. [27-29]). There are many reasons why a perfectly reasonable model may fit a particular data set poorly; for example, independent observation noise (see Figure 4 for an example). On the other hand, it is easy to construct complex

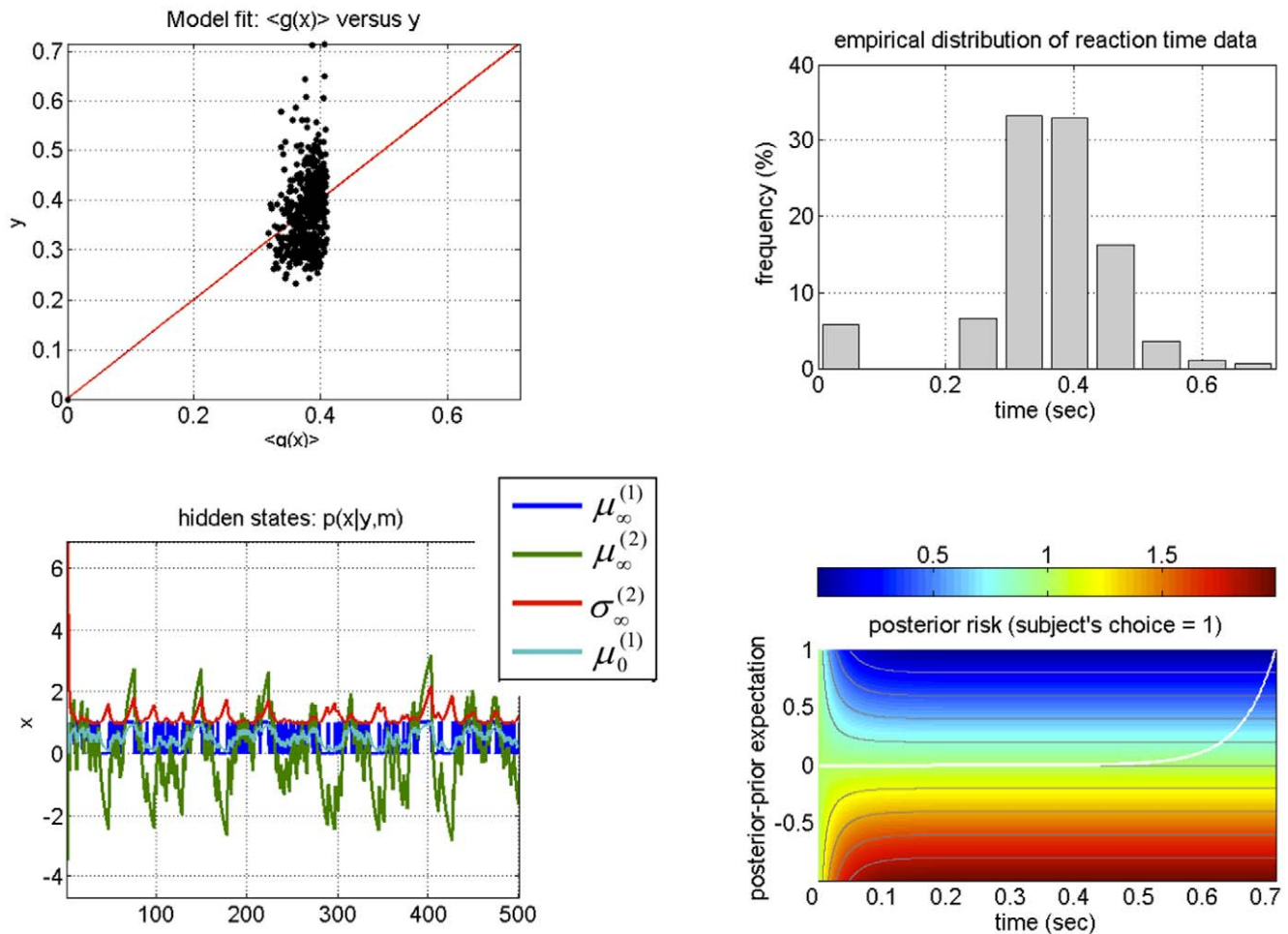


Figure 12. Results of inverting the response model for subject 5. Upper-left: predicted (x-axis) versus observed (y-axis) reaction time data. The red line indicates perfect agreement between the model and the data. Upper-right: observed reaction times empirical histogram. Note that incorrect decisions were assigned a response time of zero and did not influence model fit; see main text for details. Lower-left: time series (trial-by-trial) of the sufficient statistics of subject 5's representations of both the outcome category ($\mu^{(1)}$) and associative strength ($\mu^{(2)}$ and $\sigma^{(2)}$). See main text for the precise meaning of these variables. Lower-right: posterior risk as a function of *post-hoc* prediction error (y-axis), i.e. the difference between posterior and prior expectations, and decision time (x-axis). The posterior risk is evaluated at subject 5's response parameters estimate θ for 'house' decisions (i.e. $c=1$); it can be symmetrically derived for $c=0$. The white line shows the optimal decision time $\tau(c)$ for each level of *post-hoc* prediction error (see Equation 12 in the main text). Note that $\tau(c)$ is identically zero for all negative *post-hoc* prediction error. This signals a perceptual categorization error ($\Delta\mu_0^{(1)}(2c-1) > 0$, see Equation 11 in main text), which is emitted (at the limit) instantaneously.
doi:10.1371/journal.pone.0015555.g012

models with excellent or even perfect fit, which are mechanically meaningless and do not generalize (i.e., “over-fitting”). In brief, competing models cannot be compared on the basis of their fit alone; instead, their relative complexity must also be taken into account. This is exactly what is furnished by the (log) model evidence, which reports the balance between model fit and complexity (and can be approximated efficiently by the variational techniques used in this paper). This allows us to compare models of different complexity in an unbiased fashion. Crucially, our Bayesian model selection method does not require models to be nested and does not impose any other constraints on the sorts of model that can be compared ([30,20]). For example, alternative models compared within our framework could differ with regard to the mathematical form of the perceptual or the response model, the priors or the loss function – or any combination thereof. In principle, this makes it possible to investigate the relative plausibility of different explanations: For example, whether individual differences in behaviour are more likely to result from

individual differences in the perceptual or the response model. For clarity, however, the empirical example shown in this paper dealt with a very simple case, in which the perceptual model was varied while the response model was kept fixed.

As with all inverse problems, the identifiability of the BDT model parameters depends upon both the form of the model and the experimental design. In our example, we estimated only one parameter of the perceptual models we considered. One might argue that rather than fixing the sensory precision (α , see Equation 1) with infinitely precise priors, we should have tried to estimate it from the reaction time data. It turns out, however, that estimating (ϑ, α) and (θ_1, θ_2) together represents a badly conditioned problem; i.e. the parameters are not jointly identifiable because of posterior correlations among the estimates. This speaks to the utility of generative models for decision-making: the impact that their form and parameterisation has on posterior correlations can be identified before any data are acquired. Put simply, if two parameters affect the prediction of data in a similar way, their

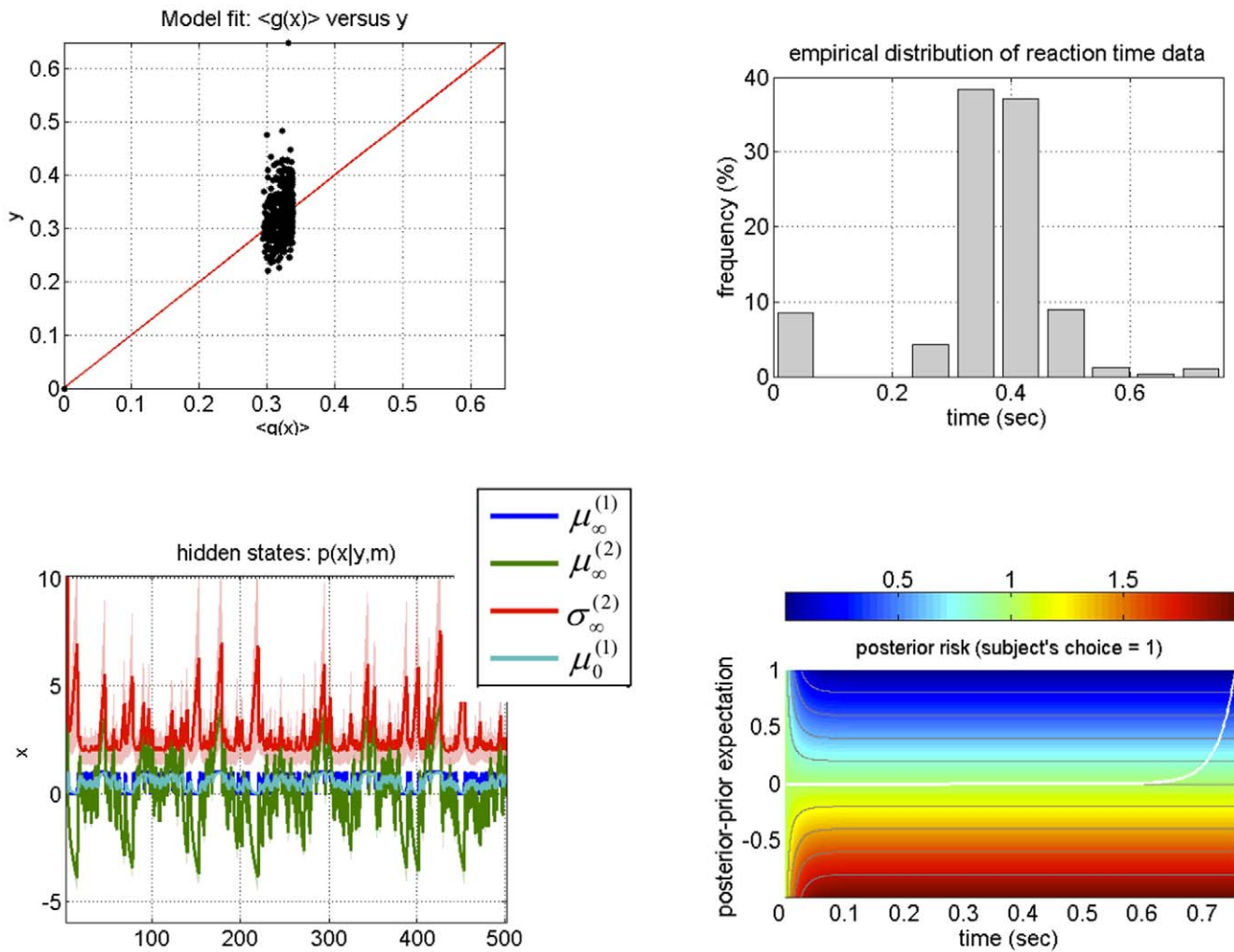


Figure 13. Results of inverting the response model for subject 12. See legend to Fig. 12 for details.
doi:10.1371/journal.pone.0015555.g013

unique estimation will be less efficient. In our particular example, there is no critical need to estimate α from the data. This is because faces and houses are well-known objects for whose categorisation subjects have a life-long experience. It is therefore reasonable to assume that α is known to the subjects, and its value can be chosen in correspondence with the statistics of the visual stimuli (see above). However, pronounced inter-individual differences can be observed empirically in face-house discrimination tasks, and this may result from differences in the individuals' history of exposure to faces throughout life. A limitation of our model is that it does not account for such inter-subject variability but assumes that α is fixed across subjects.

In contrast to α , which can (and should) be treated as a fixed parameter, it is necessary to estimate the perceptual parameter ϑ . Note that from the subject's perspective, ϑ (similar to α) is quasi-fixed (i.e., with nearly infinite precision) as this prior has been learnt throughout life. From the experimenter's perspective, however, ϑ is an unknown parameter which has to be inferred from the subject's behaviour. Estimating this parameter is critical for the experimenter as its value determines the subject's learning rate. This is best explained by highlighting the link between 'learning rates', as employed by reinforcement learning models, and Bayesian priors, or more precisely *prior precision parameters*. In the 'dynamic' perceptual model, the learning rule effectively

replaces the past history of sensory signals with a summary based on the previous representation (see Eq. 8). In turn, the perceptual representation discounts past sensory signals with an exponential weighting function, whose half-life is an affine transformation of the prior volatility ϑ . The link between ϑ and the subject's learning rate can be seen by considering the solution to equation 8 (at convergence):

$$\frac{\partial I}{\partial x} \Big|_{\mu_k^{(2)}} = 0 \Rightarrow \mu_k^{(2)} - \mu_{k-1}^{(2)} \propto \underbrace{\left(\sigma_k^{(2)} + \vartheta \right)}_{\text{effective learning rate}} \left(\mu_k^{(1)} - s\left(\mu_k^{(2)}\right) \right) \quad (15)$$

where $\mu_k^{(1)}$ is the belief about the auditory outcome category (face/house) and $s\left(\mu_k^{(2)}\right)$ is its posterior prediction based on the past history of sensory signals. Equation 15 gives the effective update rule for the perceived associative strength $\mu_k^{(2)}$ when the perceptual free energy has been optimized. Note that the form of Eq. 15 corresponds to the Rescorla-Wagner learning rule [31], in which the change in associative strength $\mu_k^{(2)} - \mu_{k-1}^{(2)}$ is proportional to the prediction error, i.e. $\delta = \mu_k^{(2)} - s\left(\mu_k^{(2)}\right)$.

In summary, for the model in the present paper, the subject's learning rate depends on the prior volatility \mathcal{G} of cue-outcome associations. Note, however, that there may not always be a quantitative relation between prior precision parameters and learning rates because this depends on the specificities of the perceptual model. There is a general qualitative relationship between the two quantities, however, because the prior precision of hidden causes within hierarchical perceptual models controls the relative weight of upcoming sensory information and prior (past) beliefs in forming the actual posterior representation. In short, this means the learning rate itself (and thus any 'forgetting' effect) emerges from optimal Bayesian recognition (see e.g., [32] for a nice example). A full treatment of these issues will be presented in forthcoming work [33].

Another analogy concerns the optimal decision time derived from the speed-accuracy trade-off given in Equation 12 which is similar in form to Hick's law. This law relates the reaction times to the amount of extracted information (c.f. [34]). In its simplest form, Hick's law is given by: $RT = a + b \log(n)$, where RT is the expected reaction time and n is the number of choice alternatives. Here, $\log(n)$ is the perceptual uncertainty (as measured by Shannon entropy). It turns out that when no categorization error is made, Equation 12 could be rewritten as $RT = a + b \log|\Delta\mu_0|$, where $\Delta\mu_0$ is the *post-hoc* prediction error, i.e. posterior minus prior expectation. Put simply, $\log|\Delta\mu_0|$ measures incoming information. There are obvious formal (information theoretic) differences between Equation 12 and Hick's law, but they capture similar intuitions about the mechanisms causing variations in reaction times.

This paper has demonstrated the practical application of the meta-Bayesian framework described in the companion paper, using empirical reaction time data from an audio-visual associative learning task reported in [12]. Authors presented several analyses of these data, including a formal comparison of alternative learning models. The results provided in the present article finesse the original comparisons and take us substantially beyond the previous report. First, the paper [12] did not provide any decision theoretic explanation for (learning induced) motor facilitation. In that paper, the behavioural comparison of different learning models was a precursor to using prediction error estimates in a model of fMRI data. It therefore only used a very simple response model assuming that (inverse) reaction times scale linearly with prediction error. In contrast, we have proposed a response model that is fully grounded in decision theory and does not assume a specific (e.g., logarithmic) relationship between prediction errors and motor facilitation. Second, we conducted a full two-level

analysis of the reaction time data, in order to assess inter-individual differences. This was made possible because, as opposed to the work in [12], we allowed for inter-individual differences in both the perceptual and response parameters (see above).

Finally, we wish to emphasize that the "observing the observer" (OTO) approach for inference on hidden states and parameters can be obtained in a subject-specific fashion, as demonstrated by our empirical analyses in this paper (see Figs. 9-13). This allows for analyses of inter-individual differences in the mechanisms that generate observed behaviour. Such quantitative inference on subject-specific mechanisms is not only crucial for characterizing inter-individual differences, an important theme in psychology and economics in general, but also holds promise for clinical applications. This is because spectrum diseases in psychiatry, such as schizophrenia or depression, display profound heterogeneity with regard to the underlying pathophysiological mechanisms, requiring the development of models that can infer subject-specific mechanisms from neurophysiological and/or behavioural data [35]. In this context, the approach presented in this paper can be seen as a complement to DCM: OTO may be useful for inference on subject-specific mechanisms expressed through behaviour, in a similar way as DCM is being used for inference on subject-specific mechanisms underlying neurophysiology.

Supporting Information

Appendix S1 Appendix S1 ('deciding when to decide') is included as 'supplementary material'. It summarizes the mathematical derivation of the optimal reaction times (as given in equation 12) from first principles, within the framework of Bayesian Decision Theory. (DOC)

Acknowledgments

We would like to thank the anonymous reviewers for their thorough comments, which have helped us improving the manuscript. The participants had no history of psychiatric or neurological disorders. Written informed consent was obtained from all volunteers before participation. The study was approved by the National Hospital for Neurology and Neurosurgery Ethics Committee (UK).

Author Contributions

Conceived and designed the experiments: KS HEMdO. Performed the experiments: HEMdO, JD. Analyzed the data: JD. Contributed reagents/materials/analysis tools: JD, KJF. Wrote the paper: JD, KJF, KS, SJK, MP.

References

- Daunizeau J, Den Ouden HEM, Pessiglione M, Stephan KE, Kiebel SJ, et al. (2010) Observing the Observer (I): Meta-Bayesian models of learning and decision making. *PLoS ONE* 5(12): e15554. doi:10.1371/journal.pone.0015554.
- Dayan P, Hinton GE, Neal RM (1995) The Helmholtz machine. *Neural Comput* 7: 889–904.
- Friston K, Kilner J, Harrison L (2006) A free-energy principle for the brain. *J of physiol Paris* 100: 70–87.
- Kersten D, Mamassian P, Yuille A (2004) Object perception as Bayesian inference. *Annu Rev Psychol* 55: 271–304.
- Lee TS, Mumford D (2003) Hierarchical Bayesian inference in the visual cortex. *J Opt Soc Am A Opt Image Sci Vis* 20: 1434–1448.
- Summerfield C, Koehlin E (2008) A neural representation of prior information during perceptual inference. *Neuron* 59: 336–347.
- Glimcher PW (2003) The neurobiology of visual-saccadic decision making. *Annu Rev Neurosci* 26: 133–79.
- Hare TA, Camerer CF, Rangel A (2009) Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324: 646–648.
- Durbin J, Koopman S (2001) Time series analysis by state-space methods. Dover (ISBN 0486442780).
- Carpenter RH, Williams ML (1995) Neural computation of log likelihood in control of saccadic eye movements. *Nature* 377: 59–62.
- Bestmann S, Harrison LM, Blankenburg F, Mars RB, Haggard P, et al. (2008) Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Curr Biol* 18: 775–780.
- Den Ouden HEM, Daunizeau J, Roiser J, Friston KJ, Stephan KE (2010) Striatal prediction error modulates cortical coupling. *J Neurosci* 30: 3210–3219.
- Cauraugh JH (1990) Speed-accuracy tradeoff during response preparation. *Res Q Exerc Sport* 61(4): 331–337.
- Usher M, Olami Z, McClelland JL (2002) Hick's Law in a stochastic Race Model with speed-accuracy tradeoff. *J Math Psychol* 46: 704–715.
- Beal M (2003) Variational algorithms for approximate Bayesian inference, PhD thesis, ION, UCL, UK.
- Friston KJ, Stephan KE (2007) Free-energy and the brain. *Synthese* 159: 417–458.
- Friston K, Kiebel S (2009) Cortical circuits for perceptual inference. *Neural Netw*, in press.
- Friston K, Mattout J, Trujillo-Barreto N, Ashburner J, Penny W (2007) Variational free-energy and the Laplace approximation. *NeuroImage* 1: 220–234.

19. Daunizeau J, David O, Stephan KE (2010) Dynamic Causal Modelling: a critical review of the biophysical and statistical foundations. *Neuroimage*, in press.
20. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. *NeuroImage* 46: 1004–1017.
21. Kass RE, Raftery AE (1995) Bayes factors. *J Am Stat Assoc* 90: 773–795.
22. Gold JJ, Shadlen MN (2001) Neural computations that underlie decisions about sensory stimuli. *Trends Cogn Sci* 5(1): 10–16.
23. Glimcher P, Fehr E, Camerer C, Poldrack E (2009) *Handbook of Neuroeconomics*. San Diego: Academic Press.
24. Ratcliff R, Smith PL (2004) A comparison of sequential sampling models for two-choice reaction time. *Psychol Rev* 111: 333–367.
25. Vandekerckhove L, Tuerlinckx F, Lee MD (2010) Hierarchical diffusion models for two-choice response times. *Psychological Methods* (in press).
26. Bogacz R, Hu PT, Holmes PJ, Cohen JD (2010) Do humans produce the speed-accuracy tradeoff that maximizes reward rate? *Q J Exp Psychol* 63: 863–891.
27. Bishop CM (2006) *Pattern recognition and machine learning*. Springer. ISBN: 978-0-387-31073-2.
28. MacKay DJC (2003) *Information Theory, Inference, and Learning Algorithms*. Cambridge: Cambridge University Press.
29. Pitt MA, Myung IJ (2002) When a good fit can be bad. *Trends in Cognitive Science* 6: 421–425.
30. Penny WD, Stephan KE, Mechelli A, Friston KJ (2004) Comparing Dynamic Causal Models. *NeuroImage* 22(3): 1157–1172.
31. Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement, *Classical Conditioning II*. A. H. Black, W. F. Prokasky Eds. Appleton-Century-Crofts. pp 64–69.
32. Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10(9): 1214–21.
33. Matthys C, Daunizeau J, Friston K, Stephan K (submitted) A Bayesian foundation for individual learning under uncertainty. *Plos Comp Biol* (submitted).
34. Usher M, Olami Z, McClelland JL (2002) Hick's law in a stochastic race model with speed-accuracy tradeoff. *J Math Psychol* 46: 704–715.
35. Stephan KE, Friston KJ, Frith CD (2009) Dysconnection in schizophrenia: From abnormal synaptic plasticity to failures of self-monitoring. *Schizophr Bull* 35: 509–527.