

## Supplemental Materials

### 1. Log model evidence for behavioural data

This appendix describes how we evaluated the log-evidence for our linear observation model of reaction times, which has the form:

$$Y = X\beta + \varepsilon \Rightarrow$$

$$p(Y, \beta, \sigma^2 | m) = (2\pi)^{-d/2} \sigma^2 \exp\left(-\frac{(Y - X\beta)^T (Y - X\beta)}{2\sigma^2}\right), \quad (\text{A1})$$

Here,  $Y$  represents the data (response speeds),  $X$  is a design matrix and  $\varepsilon \sim N(0, \sigma^2)$  are normally distributed errors. Using Jeffreys' (non-informative) priors for  $\beta$  and  $\sigma$  (i.e.  $\beta \sim 1, \sigma \sim 1/\sigma$ ), the evidence is given by

$$p(Y | m) = \int \int p(Y, \beta, \sigma | m) d\beta d\sigma$$

$$= (2\pi)^{(r-d)/2} |X^T X|^{-r/2} \Gamma(d - r - 1) (\lambda / 2)^{r+1-d} \quad (\text{A2})$$

where  $r$  is the number of parameters,  $d$  is the number of data-points and

$$\lambda = Y^T \left( I - X(X^T X)^{-1} X^T \right) Y \quad (\text{A3})$$

is the sum of squared residuals. Therefore the log model evidence is

$$\log(p(Y | m)) = \frac{r-d}{2} \log(2\pi) - r/2 \log(|X^T X|) + \log(\Gamma(d - r - 1)) + (r+1-d) \log(\lambda/2)$$

$$(\text{A4})$$

This is an exact expression for the log evidence of this model. It can be generalized to include observation models whose design matrix is informed by the trial-b-trial estimates of an underlying learning model with parameters  $\theta$  (e.g. the learning rate in the Rescorla-Wagner model):

$$y = X(\theta)\beta + \varepsilon \quad (\text{A5})$$

Assuming that the residuals are Gaussian ( $\varepsilon \sim N(0, \sigma^2 I_d)$ ) yields the likelihood function, i.e.:

$$p(y | \theta, \beta, \sigma) = N(X(\theta)\beta, \sigma^2 I_d) \quad (\text{A6})$$

We now can integrate over both  $\beta$  and  $\sigma$  to yield the restricted data likelihood  $p(y|\theta, m)$ :

$$\begin{aligned} p(y|\theta, m) &= \int p(y|\theta, \beta, \sigma, m) \underbrace{p(\beta|m)}_{\propto 1} \underbrace{p(\sigma|m)}_{\propto 1/\sigma} d\beta d\sigma \\ &= (2\pi)^{(r-d)/2} \Gamma(d-r-1) |X(\theta)^T X(\theta)|^{-r/2} \left(\frac{\lambda(\theta)}{2}\right)^{r+1-d} \end{aligned} \quad (\text{A7})$$

where we used Jeffreys' (non-informative) priors for  $\beta$  and  $\sigma$ , and the sum of squared estimated residual error  $\lambda(\theta) = \hat{\varepsilon}^T \hat{\varepsilon}$  is given by:

$$\lambda(\theta) = y^T \left( I_d - X(\theta) \left( X(\theta)^T X(\theta) \right)^{-1} X(\theta)^T \right) y \quad (\text{A8})$$

This (restricted) likelihood function is obviously not conjugate to, for example, simple Gaussian priors on  $\theta$ . This means that there is no analytical expression for the model evidence  $p(y|m)$ :

$$p(y|m) = \int p(y|\theta, m) p(\theta|m) d\theta \quad (\text{A9})$$

However, it is possible to use the so-called Laplace approximation (e.g. see (Friston et al., 2007)) to finesse this difficult integration problem. First, let us derive a second-order Taylor expansion to the log restricted likelihood (LReL)  $t(\theta)$ :

$$\begin{aligned} t(\theta) &= \ln p(y|\theta, m) \\ &\approx t(\hat{\theta}) + \frac{1}{2} (\theta - \hat{\theta})^T \underbrace{\frac{\partial^2 t}{\partial \theta^2} \Big|_{\hat{\theta}}}_{-H(\hat{\theta})} (\theta - \hat{\theta}) \end{aligned} \quad (\text{A10})$$

where  $H(\hat{\theta}) = -\frac{\partial^2 t}{\partial \theta^2} \Big|_{\hat{\theta}}$  is the negative Hessian of the LReL. Then, under Jeffrey's non-informative priors for  $\theta$ , the log model evidence can be approximated as:

$$\begin{aligned} \ln p(y|m) &= \ln \int \exp(t(\theta)) d\theta \\ &\approx t(\hat{\theta}) + \frac{r_\theta}{2} \ln 2\pi + \frac{1}{2} \ln |H(\hat{\theta})^{-1}| \\ &= \underbrace{\ln p(y|\hat{\theta}, m) + \frac{r_\theta}{2} \ln 2\pi - \frac{1}{2} \ln |H(\hat{\theta})|}_F \end{aligned} \quad (\text{A11})$$

where  $r_\theta$  is the dimensionality of parameter vector  $\theta$  (i.e. the number of free parameters in the underlying learning model).

In equation 7,  $F \approx \ln p(y|m)$  is referred to as the Laplace approximation to the model evidence. Note that the well-known Bayesian Information Criterion (BIC) is simply the asymptotic limit ( $d \rightarrow \infty$ ) to  $F$ , as given above:

$$\begin{aligned}
 F &= \underbrace{\ln p(y|\hat{\theta}, m)}_{o(d)} + \underbrace{\frac{r_\theta \ln 2\pi}{2}}_{o(1)} - \frac{1}{2} \underbrace{\ln |H(\hat{\theta})|}_{o(r_\theta \ln d)} \\
 &\xrightarrow[\text{y iid}]{d \rightarrow \infty} \underbrace{\ln p(y|\hat{\theta}, m) - \frac{r_\theta}{2} \ln d}_{\text{BIC}}
 \end{aligned}
 \tag{A12}$$

Note that when the underlying learning model has no free parameters then  $\frac{r_\theta}{2} \ln d = 0$  and Eq. A12 reduces to the exact expression for the log-evidence in Eq. A4. This means that Eq. A12 can be used to compare any observation models, even if they differ in the number of parameters of the underlying learning model.

## 2. Bayesian volatility-based associative learning model

We start with the premise that subjects represent or infer the causes of their sensory inputs and optimise their behaviour on the basis of this inference. From a Bayesian perspective, the brain is an *observer* of its own sensory signals. This means subjects invert some forward or generative model of sensory inputs to represent the unobserved (hidden) causes of that input. Any learning then relies strongly on the subject's model of the world (the perceptual model), which determines predictions, and hence, prediction error. This Bayesian perspective has already been used to model behavioural decisions (e.g. (Kording et al., 2007)).

In what follows, we describe the volatility-based perceptual model used in this study to estimate the volatility and probabilities of the observed events (i.e. cue-outcome pairs). This model is based on the proposal by Behrens et al. (Behrens et al., 2007) and subsumes the set of probabilistic assumptions the brain encodes in order to represent the causes of paired audio-visual stimuli. The perceptual model generates sensory input  $u$  (e.g., experimental stimuli) from hidden causes  $x$  (e.g., experimental factors or environmental states) and can be expressed in terms of a likelihood model  $p(u|x)$  and prior beliefs  $p(x)$ . The states of the world  $x$  are unknown to the subject but are under experimental control. In our example,  $u$  is a series of cue-outcome pairs, presented to the observer, and  $x$  encodes the probabilistic cue-outcome association that the subject has to learn in order to predict its environment adequately. The prior belief itself is decomposed into a hierarchy of conditional probability density functions, as will be described bellow.

Let  $u_t$  be the outcome at trial  $t$  be a multinomial random variate such that:

$$p(u_t | u_t^c, r_t) = \text{Mult}(u_t | r_t) \\ = \prod_{i=1}^n (r_t^i)^{u_t^i},$$

where  $(r_t^i)_{i=1, \dots, n}$  is a  $n \times 1$  vector of probabilities describing completely the distribution of the  $n$  possible outcomes. This forms the likelihood of our generative model. Note that from there on, we will consider that each of the cues  $u_t^c$  is associated with its own likelihood, and consequently, its own generative model. This means that everything we state below is conditional on the given cue. As a consequence, the Bayesian inversion of such a set of generative models is cue-specific, and has to be replicated for all different cues.

This vector of cue-outcome association probabilities obeys *a priori* the following Dirichlet distribution:

$$p(r_t | r_{t-1}, v_t) = \text{Dir}(r_t | a_t) \\ = \frac{\Gamma(a_t^0)}{\prod_{i=1}^n \Gamma(a_t^i)} \prod_{i=1}^n (r_t^i)^{a_t^i - 1}$$

This transition density is actually a martingale; i.e. it is a first order Markov process whose current first order moment is equal to its previous realization:

$$\langle r_t \rangle = r_{t-1}.$$

Furthermore, the precision (i.e. inverse variance) of the transition from  $r_{t-1}$  to  $r_t$  is parameterized by a scalar quantity  $v_t$ , which measures the volatility of the environment:

$$\sum_{i=1}^n a_t^i = \exp(-v_t) + 1$$

The volatility itself is assumed to vary over time as a martingale, and the above parameterization makes a simple AR(1) model possible:

$$p(v_t | v_{t-1}, K) = N(v_t | v_{t-1}, K) \\ = \frac{1}{\sqrt{2\pi K}} \exp\left(-\frac{1}{2K} (v_t - v_{t-1})^2\right),$$

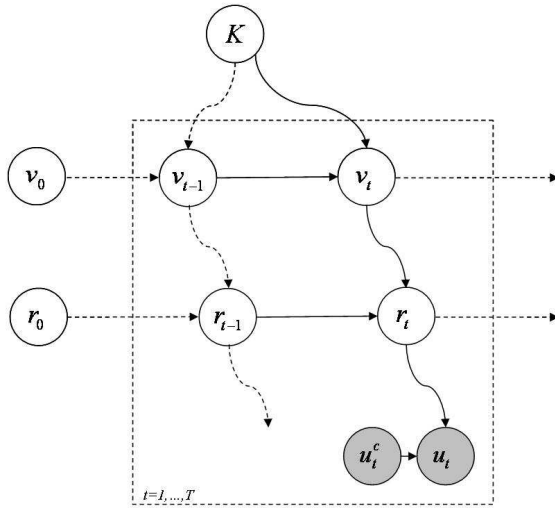
where  $K$  is the prior variance of the volatility, i.e. the volatility's volatility.

The prior on  $K$  itself is supposed to be non-informative, i.e.:

$$p(K) \propto 1.$$

To summarize, the generative model assumes the following cascade of events (illustrated in the graph in Figure S1):

- 1- A value for the volatility variance  $K$  is randomly drawn from its prior pdf  $p(K)$ . Then, at each trial  $t$ :
- 2- This value determines the transition pdf of the volatility. Then, a first value  $v_t$  is randomly drawn from  $p(v_t|v_{t-1}, K)$ .
- 3- Knowing the volatility  $v_t$  then allow us to derive the transition density for  $r_t$ . Then, a value for the cue-outcome association probability is drawn from  $p(r_t|r_{t-1}, v_t)$ .
- 4- This finally defines the likelihood of the outcome itself: the first outcome  $u_t$  is then drawn randomly from  $p(u_t|u_t^c, r_t)$ .
- 5- The steps 2, 3 and 4 are repeated in time, giving rise to three time series for the volatility  $v_t$ , the cue-outcome association probability  $r_t$  and the observed outcomes  $u_t$ .



**Figure S1. Graph illustration of the volatility model.**  $u_t$ = observed outcome at trial  $t$ ;  $r_t$  = cue-outcome association probability;  $v_t$ = volatility;  $K$  = variance of the volatility.

The model assumes that the observer updates its posterior belief on-line, in the light of incoming data, in a Kalman filter-like manner. The joint posterior pdf over the full set of unknown variables, namely  $x = \{K, v, r\}$ , then follows the following prediction and update steps:

$$\text{prediction: } p(r_t, v_t, K | u_{1:t-1}) = \iint p(r_t | r_{t-1}, v_{t-1}) p(v_t | v_{t-1}, K) p(r_{t-1}, v_{t-1}, K | u_{1:t-1}) dr_{t-1} dv_{t-1}$$

$$\text{update: } p(r_t, v_t, K | u_{1:t}) = \frac{p(r_t, v_t, K | u_{1:t-1}) p(u_t | r_t)}{\iiint p(r_t, v_t, K | u_{1:t-1}) p(u_t | r_t) dr_t dv_t dK}$$

These two steps are iterated as long as new data are measured and, after each cue-outcome observation, yield estimates of both the current cue-outcome association probability  $r_t$  and the environmental volatility  $v_t$ , as well as an estimator of the static volatility's variance  $K$ , given all previously observed data. In the present study, the trajectory of these estimates as a function of time (trial  $t$ ) served as predictors for behavioural data (response speeds) and neuroimaging data (fMRI data in SPM and DCM analyses).

### 3. Alternative learning models

In addition to the two learning models described in the main text (i.e. a linear model based on the true probabilities generating the stimulus sequence and a hierarchical Bayesian observer), we tested three further models, following suggestions by our reviewers. Firstly we employed a simple "model-free" reinforcement learning approach using a classical Rescorla-Wagner learning model. As there were no significant differences in behaviour for the two cues (see main text), the model fitted one joint learning rate for both cue types. The associative strengths  $V_t$  of the cues to the outcomes (face = 1, house = 0) were modelled according to the equation

$$V_t = V_{t-1} + \alpha(\lambda_{t-1} - V_{t-1})$$

where  $\lambda$  is the outcome (face or house) and  $\alpha$  denotes the learning rate, which is fitted for each individual subject.

The Rescorla Wagner model, like the Bayesian learning model, has no explicit knowledge of the task structure. It is conceivable, however, that over the course of the experiment the subjects learned to recognise the discrete levels of predictability of the cues. Therefore, we tested two additional models that did represent the underlying structure of the task. These models consisted of two variants of a first order hidden Markov model (HMM), which were used to model the sequences of observed cue-outcome combinations (Rabiner, 1989). An HMM is a set of hidden states, each of which is probabilistically associated with an observable output (in our case a cue-outcome combination). Transitions between states occur stochastically; the conditional probabilities (transition probabilities) are such that the state at time  $t$  depends only on the state at time  $t-1$  (Markov property). In the HMMs used here, five hidden states were used to represent the five discrete levels of associative cue strengths. This knowledge about the levels of associative strengths in the experiment enabled the model to represent the subject's potential "metalearning" about the levels that the strengths could change to, allowing for sudden jumps between levels when associations changed rather than having to learn them anew each time.

We used the Baum-Welch algorithm (also known as forward-backward algorithm (Baum et al., 1970)) to find the transition matrix that best explained the observed cue-outcome combinations. In a first version of the HMM ("HMM fixed" in Fig. S1 and Table S1), the optimal transition matrix was learned from the entire observed sequence. However, this equates inference with learning, and assumes that subjects know the structure of the task from the start of the experiment. An alternative scheme ("HMM learn") is to update the transition matrix each time an observation is made, that is, to run the Baum-Welch algorithm at each trial anew, using the trial sequences  $\{[1, 2], [1, 2, 3], \dots, [1 \dots N]\}$ . A priori, i.e. at the start of the experiment, there was an equal belief to be in each of the 5 states.

Figure S2 shows the estimated probability of observing a face following one of the CS in block 3, (cf. figure 1C main text) as computed by each of the four learning models and juxtaposes these trial-by-trial estimates to the true probabilities. For each of these models the log model evidence was calculated as described in section S1, taking into account the additional parameter of the RW model (i.e. the learning rate) by means of the Bayesian Information Criterion.

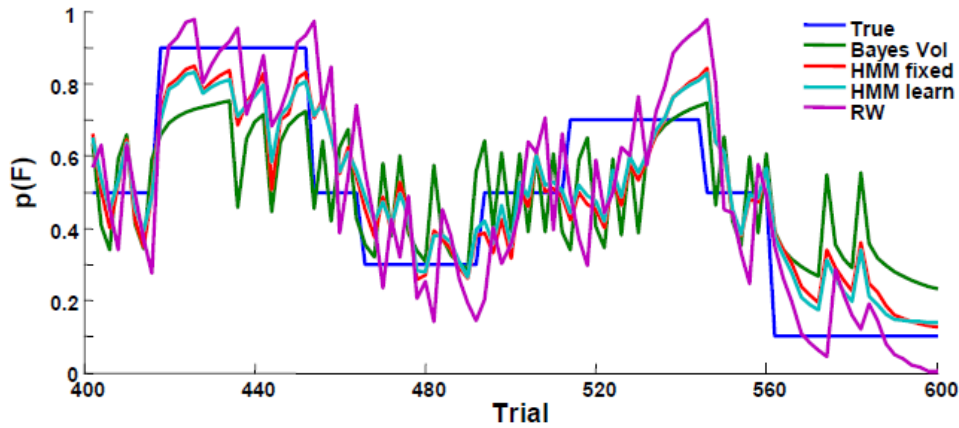


Figure S2.  $p(F|CS)$  as estimated by the various models.

The model evidences were then used for a random effects Bayesian model selection across subjects (see main text). The hierarchical Bayesian learner turned out to be superior to any of the other tested models (Table S1), with an exceedance probability of 66% that this model was more likely than any other model considered.

Table S1. BMS results for behavioural model comparison

	Dirichlet parameters $\alpha$	Exceedance probability $\varphi$
<i>True categorical model</i>	1.00	0.00
<i>Bayesian volatility model</i>	8.99	0.66
<i>HMM (fixed)</i>	6.53	0.22
<i>HMM (learn)</i>	3.39	0.02
<i>Rescorla Wagner</i>	5.11	0.09

#### 4. Additional SPM results

The main text deals only with key questions of interest for this study, namely characterization of stimulus-independent and stimulus-specific surprise responses and connectivity changes. For completeness, the results of additional analyses are reported here; these include a detailed analysis of the main effects of the stimuli as well as an analysis of regional responses associated with the estimated volatility of the probabilistic associations.

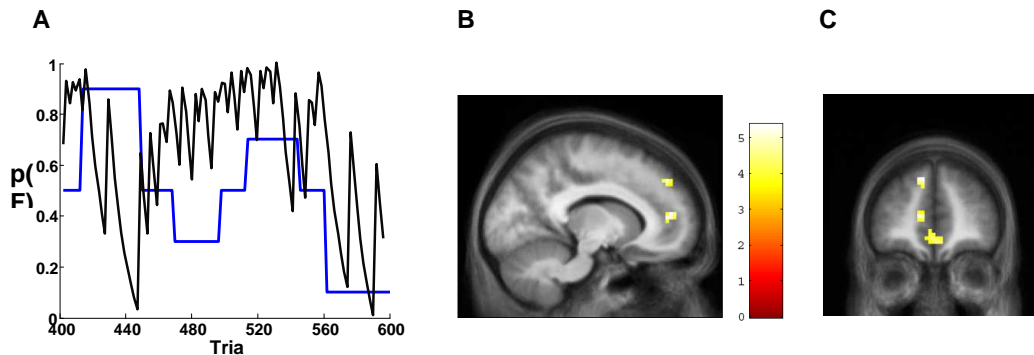
### Stimulus main effects in FFA and PPA

As expected, the mid fusiform gyrus was activated more strongly to *face* stimuli than to *house* stimuli (FFA, Table S1), and the parahippocampal gyrus showed the opposite effect (PPA, Table S1). At the group level the FFA activation was significant at whole brain corrected level only in the right hemisphere, but the left FFA activation was significant within an anatomically defined ROI for the fusiform gyrus (table S1).

### Volatility dependent brain activations

Although this was not the focus of this study, for completion we also tested in which areas activity increased or decreased with the trial-by-trial volatility estimates. Following the results by Behrens et al., (2007), who demonstrated that ACC activity correlated with volatility estimates during reward learning, we tested whether volatility encoding in the ACC would also be present in our purely perceptual paradigm which did not include any rewards. Figure S3A shows the trial-by-trial estimates of the volatility for session 3 (compare figure 1C in the main text for the parallel probability estimates). Indeed, activity in the dorsal and rostral ACC and the ventromedial prefrontal cortex correlated significantly with the volatility estimates (Table S2 and Fig. S3B-C).

Although the use of a volatile environment was not a phenomenon of primary interest for this study, but merely a means of enforcing continuous learning (and thus maximising induction of synaptic plasticity and hence changes in connectivity), it is noteworthy that our analysis of volatility effects replicated previous results (Behrens et al., 2007).



**Figure S3. Volatility effects.** **A)** Trial-by-trial estimates of the volatility. Blue line = true probabilities, black = volatility. **(B,C)** The anterior cingulate cortex and ventromedial prefrontal cortex show a positive correlation with the volatility estimate of the Bayesian observer model, here shown in a sagittal **(A)** and axial **(B)** slice

**Table S2. MNI coordinates and Z-values for significantly activated regions**

Foci of activation	MNI coords.			Z score
	x	y	Z	
<b>Main effects of sensory stimulation</b>				
<i>House&gt;Face</i>				
R parahippocampal gyrus*	30	-51	12	7.01
L parahippocampal gyrus*	-24	-57	-18	6.70
<i>Face&gt;House</i>				
R mid fusiform gyrus*	45	-57	-24	5.42
L amygdala*	-21	-12	-9	4.31
L mid fusiform gyrus**	-45	-54	-21	3.47
<b>Volatility effects</b>				
<i>positive correlation</i>				
Ventromedial prefrontal ctx*	3	48	-9	3.64
ACC**	-12	45	9	4.11
Ventral ACC / subgenual ctx**	-6	36	-3	3.57
L caudate/thalamus*	-21	-9	9	4.32
<b>negative correlation</b>				
No significant activations.				

\* significant at  $p < 0.05$  FEW cluster-level corrected across the whole-brain

\*\* significant at  $p < 0.05$  cluster-level corrected for a priori region of interest

## 5. Optimisation of fixed connections (DCM)

In order to optimise the fixed connections, a basic model was defined that included the minimal number of connections necessary to test the hypothesis outlined in the main text. The endogenous connectivity of this ‘minimal’ model was then optimised by systematically adding connections. A random effects Bayesian model selection (BMS) procedure was then used to select the optimal model at the group level (Stephan et al., 2009). This procedure quantifies the relative goodness of models in terms of exceedance probabilities  $\phi_i$  which denote the probability that model  $i$  is superior to any other model considered, given the data from all subjects. Note that exceedance probabilities are a function of model space; for example, because they sum to unity over all models considered, they decrease monotonically when increasing the set of alternative models. They are thus to be interpreted in relative, not absolute terms.

A minimal model (Figure S4,  $m_1$ ) included only the endogenous connections from the sensory areas to the PMd, and these connections were modulated by the activity from the putamen. An additional six models ( $m_2$ - $m_7$ ) were derived from this basic architecture in two steps. In a first step, we compared all combinations of endogenous

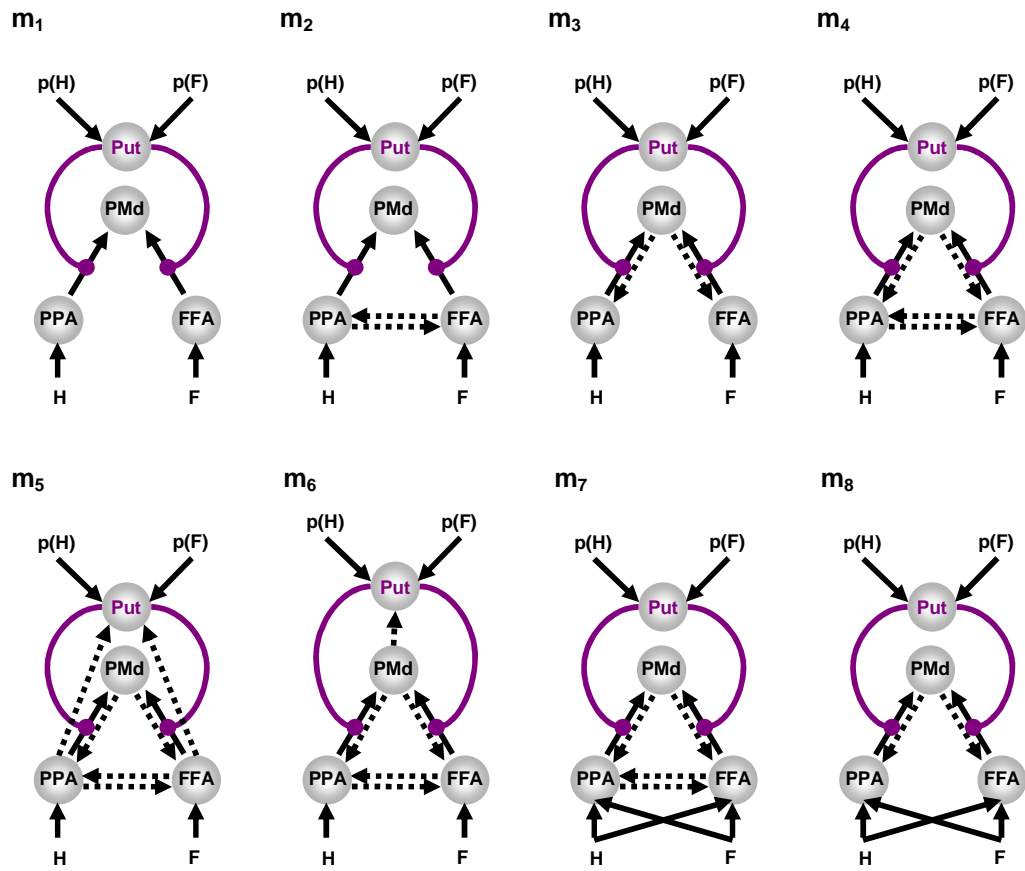
connections between PPA, FFA and PMd ( $m_1$ - $m_4$ ) and two models ( $m_5$ ,  $m_6$ ) with connections to the putamen from the visual and premotor cortex, respectively. Model  $m_5$  tested whether there was any direct influence of FFA and PPA on the putamen. Model  $m_6$  included a direct connection from the PMd to the putamen (Fig. S4,  $m_6$ ) because there exist direct projections from the premotor cortex to the putamen (Takada et al., 1998;Leh et al., 2007). Based on a suggestion by one of our reviewers we also included 2 models in which Face and House stimuli directly entered PPA and FFA, respectively. For this, we tested a model with ( $m_7$ ) and without ( $m_8$ ) reciprocal connections between the PPA and FFA.

Note that comparing DCMs with additional connections is not equivalent to testing whether these connections do or do not exist anatomically, but whether these connections play a functional role in the process modelled. Comparing all six models ( $m_1$ - $m_6$ ) against each other using random effects BMS, model  $m_4$  turned out to be the best model, i.e. a model with full reciprocal connectivity between PPA, FFA and PMd (but not direct connections from either visual areas nor PMd to the putamen). The exceedance probability for model  $m_4$  was  $\varphi_4 = 0.44$ , surpassing the exceedance probabilities of all other models (which ranged from 0.01 to 0.28; see Table S3).

Once we had identified the most likely pattern of connections among the areas, we constructed an additional model ( $m_{pm}$ ) and compared this to model  $m_4$  in order to verify the specificity of the modulatory influence exerted by the putamen (note that in the main text  $m_4$  is referred to as  $m_{pt}$ ). Since the putamen and the PMd showed similar prediction error related activations (Fig. 3, main text), we wished to demonstrate that putamen activity gated visuomotor connections (Fig. 4C, main text), instead of premotor activity gating visuostratial connections (Fig. 4B, main text). Indeed, BMS showed that the reversed model ( $m_7$ , with PMd as source of modulatory effects) was clearly inferior to the original model ( $m_4$ , with the putamen as source of modulatory effects), with an exceedance probability of 99% in favour of the latter (see Fig. 4D, main text).

**Table S3. BMS results for models  $m_1$ - $m_8$**

	<b>Dirichlet parameters <math>\alpha</math></b>	<b>Exceedance probability <math>\varphi</math></b>
$m_1$	1.68	0.01
$m_2$	4.11	0.17
$m_3$	1.70	0.01
$m_4$	5.63	0.44
$m_5$	3.06	0.07
$m_6$	4.81	0.28
$m_7$	1.02	0.00
$m_8$	1.00	0.00



**Figure S4. Alternative DCMs.**  $M_1$  includes the minimal connections needed to model the observed modulatory effects in the premotor cortex (PMd); this model includes endogenous connections from the PPA and FFA to the PMd, and these connections are modulated by output activity from the putamen. Models 2-4 then add or exclude connections between the sensory and premotor areas. Model 5 includes direct connections from the sensory areas to putamen, and model 6 includes a connection from PMd to the putamen. Model 7 and 8 include direct inputs of both visual stimulus types to both the PPA and FFA. In an additional model ( $m_{pm}$ ) the role of the putamen and the PMd were swapped such that PMd activity modulated visual afferents to the putamen; this model is shown in Figure 4C in the main text.

## Reference List

Baum LE, Petrie T, Soules G, Weiss N (1970) A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. *Ann Math Statistics* 1: 164-171.

Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10: 1214-1221.

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

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

Leh SE, Ptito A, Chakravarty MM, Strafella AP (2007) Fronto-striatal connections in the human brain: a probabilistic diffusion tractography study. *Neurosci Lett* 419: 113-118.

Rabiner LR (1989) *Proc IEEE* 77: 257-286.

Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. *Neuroimage* 46: 1004-1017.

Takada M, Tokuno H, Nambu A, Inase M (1998) Corticostriatal projections from the somatic motor areas of the frontal cortex in the macaque monkey: segregation versus overlap of input zones from the primary motor cortex, the supplementary motor area, and the premotor cortex. *Exp Brain Res* 120: 114-128.