

## Neural signals encoding shifts in beliefs



Philipp Schwartenbeck<sup>a,b,c,d,\*</sup>, Thomas H.B. FitzGerald<sup>a,d,1</sup>, Ray Dolan<sup>a,d</sup>

<sup>a</sup> The Wellcome Trust Centre for Neuroimaging, UCL, 12 Queen Square, London WC1N 3BG, UK

<sup>b</sup> Centre for Cognitive Neuroscience, University of Salzburg, Salzburg 5020, Austria

<sup>c</sup> Neuroscience Institute, Christian-Doppler-Klinik, Paracelsus Medical University Salzburg, Salzburg 5020, Austria

<sup>d</sup> Max Planck University College London Centre for Computational Psychiatry and Ageing Research, London WC1B 5EH, UK

### ARTICLE INFO

#### Article history:

Received 16 July 2015

Accepted 23 October 2015

Available online 28 October 2015

#### Keywords:

Dopamine

Information theory

Epistemic value

Belief formation

Psychosis

### ABSTRACT

Dopamine is implicated in a diverse range of cognitive functions including cognitive flexibility, task switching, signalling novel or unexpected stimuli as well as advance information. There is also longstanding line of thought that links dopamine with belief formation and, crucially, aberrant belief formation in psychosis. Integrating these strands of evidence would suggest that dopamine plays a central role in belief updating and more specifically in encoding of meaningful information content in observations. The precise nature of this relationship has remained unclear. To directly address this question we developed a paradigm that allowed us to decompose two distinct types of information content, information-theoretic surprise that reflects the unexpectedness of an observation, and epistemic value that induces shifts in beliefs or, more formally, Bayesian surprise. Using functional magnetic-resonance imaging in humans we show that dopamine-rich midbrain regions encode shifts in beliefs whereas surprise is encoded in prefrontal regions, including the pre-supplementary motor area and dorsal cingulate cortex. By linking putative dopaminergic activity to belief updating these data provide a link to false belief formation that characterises hyperdopaminergic states associated with idiopathic and drug induced psychosis.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### Introduction

Adaptive behaviour mandates agents build models of their environment and use these to appropriately update their beliefs about the causes of sensory information, a process formalised within Bayesian brain theories (Dolan and Dayan, 2013; Friston et al., 2014b; Friston, 2010; Körding and Wolpert, 2004; Tenenbaum et al., 2006). It has long been proposed that dopamine plays an important role in belief formation and, crucially, in aberrant belief formation in psychosis (Corlett et al., 2009; Howes et al., 2009; Howes et al., 2011a, 2011b; Kapur, 2003; Roiser et al., 2013; Shaner, 1999). This view is supported by studies indicating that dopamine is critical for cognitive flexibility (Bestmann et al., 2014; van Holstein et al., 2011) and task switching (Cools et al., 2009; Stelzel et al., 2010). It also resonates with observations showing dopamine neurons respond to both novel and unexpected stimuli (Bromberg-Martin et al., 2010; Bunzeck and Düzel, 2006; Horvitz, 2000; Iglesias et al., 2013) and reflect preferences for advance information (Bromberg-Martin and Hikosaka, 2009). These findings suggest that dopamine plays a central role in belief updating and, therefore, in

encoding the meaningful information content of observations. Despite these lines of suggestive evidence, a relationship between dopamine signalling and belief-updating in humans has not been demonstrated so far.

To examine this question we defined two related, but distinct, measures of information content that could be signalled by dopamine. The first quantity is surprise or surprisal (Barto et al., 2013) and reflects the information content of a stimulus in a purely information-theoretic sense. For an observation  $o$  given a model  $m$  with parameters  $\theta$  this is defined as:

$$-\ln P(o|\theta, m).$$

This is the negative log probability of an observation, which reflects the simple unexpectedness of an observation given an agent's model of the world. Importantly, this does not necessarily index the 'usefulness' of information to an agent, leading to the so-called 'white noise' paradox, where maximally informative stimuli are those that are entirely random and therefore meaningless for updating an agent's model about the world (Barto et al., 2013). Consequently, a surprising observation is not necessarily associated with improving an agent's beliefs about the environment.

This motivates an alternative definition of information in terms of actual shifts in beliefs. Formally, this corresponds to Bayesian surprise

\* Corresponding author at: Centre for Cognitive Neuroscience, University of Salzburg, Hellbrunnerstraße 34, 5020 Salzburg, Austria.

E-mail address: [philipp.schwartenbeck@stud.sbg.ac.at](mailto:philipp.schwartenbeck@stud.sbg.ac.at) (P. Schwartenbeck).

<sup>1</sup> These authors contributed equally to this work.

or the Kullback–Leibler divergence between prior and posterior beliefs about the world:

$$KL[P(\theta|o, m)||P(\theta|m)]$$

Bayesian surprise can be understood as meaningful information or the ‘epistemic value’ of an observation (Friston et al., 2015). This quantity has previously been used successfully to model visual search behaviour (Itti and Baldi, 2009) and exploration (Sun et al., 2011).

Since model updating occurs in response to unexpected observations, in many situations information-theoretic surprise and shifts in beliefs are highly correlated. However, they are decorrelated in situations where surprising stimuli do not lead to significant model updating, as occurs when an observation is similarly unlikely under all possible hypotheses. We exploited this fact to develop a task in which subjects were asked to infer, on a trial-by-trial basis, whether a visual or auditory cue determined the current probabilities of observing different trial outcomes (see Fig. 1a). Subjects were instructed, truthfully, that on any given trial outcomes were exclusively determined by one of the cue types (i.e., sensory modalities), and that this contingency would switch periodically. At the end of each trial, they were asked to rate their belief about the current contingency, on a visual analogue scale ranging from 1 (complete certainty that visual cues determined outcomes) to 11 (complete certainty that auditory outcomes did) (see Fig. 1b for an example trial). To allow us to clearly separate BOLD responses engendered by reward processing from those relating to information content, we used monetary gains and losses as outcomes. These outcomes determined subject’s earnings at the end of the experiment, but were independent of performance on the rating task.

**Materials and methods**

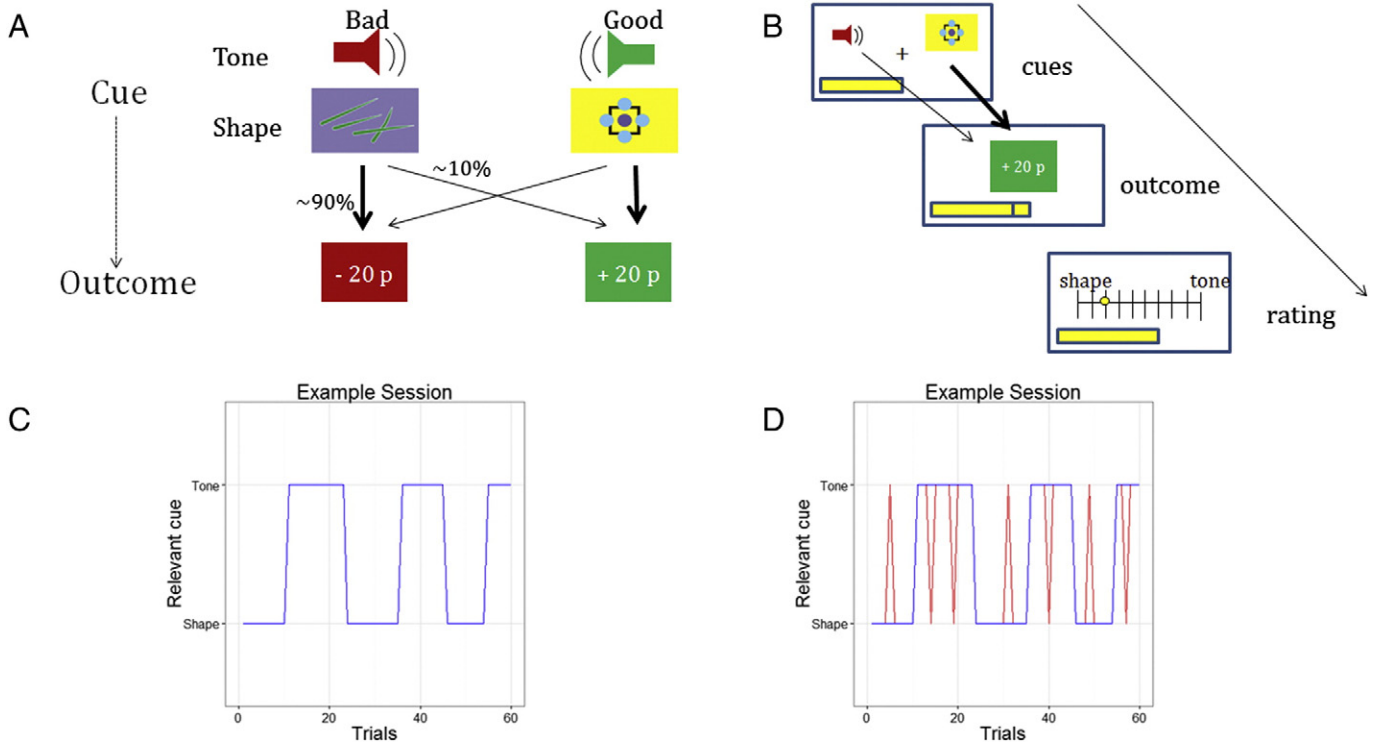
*Participants*

27 subjects (16 females) with no known history of psychiatric or neurological disorders provided written informed consent and participated in the study. Participants had a mean age of 24.9 (standard deviation = 5.28, range = [19,43]) and were recruited from the University College London Psychology Subject Pool. Subjects were instructed that they would be paid between £25 and £35, determined by the sum of their earnings in one behavioural training session and three sessions in the scanner. The average payment of subjects was £33.78. The study was approved by the UCL Research Ethics Committee.

*Procedure*

Participants underwent a 1-h training period followed by one and a half hours of scanning. In the experiment, subjects had to infer on a trial-by-trial basis whether auditory or visual cues currently determined probabilities of seeing a specific trial outcome, where the cue contingencies switched periodically (about 5 to 6 times in one session). Trial outcomes were wins or losses varying randomly between 10 and 30 pence.

Subjects were instructed that on any given trial only one of the two cue types (visual or auditory) predicted outcomes while the other had no predictive validity. In an initial training session subjects learned that in both the visual and auditory domains there were one ‘good’ cue that was predictive of a monetary win and one ‘bad’ cue that was predictive of a monetary loss, where cues had a validity of 90%. Subjects were instructed that these contingencies would remain constant throughout the experiment, such that the identity of the ‘bad’ tone and shape, predictive of a monetary loss, and the ‘good’ tone and shape, predictive of a monetary win, remained the same. Crucially,



**Fig. 1.** Task. (A) Subjects were instructed that throughout the experiment one ‘good’ shape and tone predicted a monetary win and one ‘bad’ shape and tone predicted a loss with 90% cue-validity. Monetary outcomes varied between ±10 and ±30 pence. (B) Each trial started with a simultaneous presentation of one shape (good or bad) and one tone (good or bad), followed by a trial outcome (win or loss). Based on this observation, subjects were asked to infer whether visual or auditory cues currently predict trial outcomes and indicate their belief on a rating bar. At any given trial, only one modality predicted the outcome. (C) Subjects were instructed that cue contingencies did not change on a trial-by-trial basis but periodically, with about 5–6 reversal of contingencies during one session (consisting of 60 trials). (D) Because cues had only 90% validity, sometimes unexpected trial-outcomes were observed (red line) without reflecting an actual shift in cue contingencies (blue line).

therefore, participants were asked to perform inference on cue contingencies, not on the valence of cues (good or bad).

In the actual task, each trial started with a simultaneous presentation of one visual (good or bad) and one auditory (good or bad) cue, followed by a trial outcome (monetary win or loss). Based on this observation, subjects had to infer whether auditory or visual information was currently predictive of trial outcomes and rate their belief on a scale ranging from 1 (complete certainty that one modality determined outcomes) to 11 (complete certainty that the other modality did), where the sides on which ‘Shape’ and ‘Tone’ would appear on the rating bar were counterbalanced between subjects. Importantly, half of the trials were informative (one cue being predictive of a win whereas the other being predictive of a loss) whereas the other half was uninformative (both cues being good or bad). This introduced a crucial dissociation between unexpected events (i.e., information-theoretic surprise) and actual shifts in beliefs (Bayesian surprise) that could take place in informative trials only, while decorrelating these two measures of information content from actual reward.

To thoroughly train subjects on the contingencies of the task, one full training session was performed prior to the actual experiment in the scanner, preceded by a training session on the valence of the cues (i.e., learning which cue was predictive of a win/loss) and a session in which cues had 100% validity to familiarise subjects with the general structure of the task.

Subjects performed three sessions consisting of 60 trials in the MRI Scanner. Each trial started with the presentation of the cues lasting for 2000 ms followed by a gap jittered between 2000 and 8000 ms. Subsequently, an outcome was shown for 2000 ms. After the outcome, a rating bar appeared for 4000 ms where subjects had to indicate their beliefs using a MRI-compatible button box. This was followed by an interstimulus-interval (fixation cross) jittered between 1000 and 3000 ms.

### Computational modelling

To derive trial-by-trial regressors for information-theoretic surprise and belief updating, we developed a simple model that performs Bayesian belief updating at each trial. We assume that subjects’ prior beliefs at the beginning of a session regard both hidden states to be equally likely:

$$t = 1: P(\theta|m) = [\text{vision} \quad \text{sound}] = [0.5 \quad 0.5] . \quad (1)$$

The likelihood of a given outcome depended on the observed auditory and visual cues ( $C_v$ ,  $C_a$ ), which could be good or bad, and a subject’s individual belief about the validity of a cue  $\tau$  (free parameter):

$$P(o_t|\theta, m) = \begin{bmatrix} o_t \cdot \tau \cdot C_v & o_t \cdot (1-\tau) \cdot C_a \\ o_t \cdot (1-\tau) \cdot C_v & o_t \cdot \tau \cdot C_a \end{bmatrix} . \quad (2)$$

The posterior belief about cue contingencies is obtained by an application of Bayes rule at any given trial:

$$P(\theta|o_t, m) = \frac{P(\theta|m) \cdot P(o_t|\theta, m)}{\sum (P(\theta|m) \cdot P(o_t|\theta, m))} . \quad (3)$$

The posterior belief at a given trial then serves as a prior belief at the subsequent trial:

$$t + 1: P(\theta|m) = P(\theta|o_t, m) \cdot E \quad (4)$$

where  $E$  encodes subjects’ expectations about the frequency of contingency reversal depending on reversal probability  $\delta$ :

$$E = \begin{bmatrix} \delta & 1-\delta \\ 1-\delta & \delta \end{bmatrix} . \quad (5)$$

We estimated the two free parameters  $\tau$  and  $\delta$  using (constrained) maximum-likelihood estimation, while constraining  $\tau$  between 0.5 and 1 and  $\delta$  between 0 and 1. To fit our model to behaviour, we used the estimated current belief about the cue contingencies as a predictor for the observed rating and specified the fitting algorithm (using the Matlab-routine `fmincon`) to minimise the residuals in that linear model (or, equivalently, maximise the explained variance  $R^2$  in observed behaviour).

Based on this model-fitting, we derived trial-by-trial measures of shifts in beliefs formalised as Bayesian surprise, i.e., the Kullback–Leibler divergence from prior to posterior belief  $KL[P(\theta|o_t, m)||P(\theta|m)] = \sum P(\theta|o_t, m) \cdot \log\left(\frac{P(\theta|o_t, m)}{P(\theta|m)}\right)$  as well as surprise defined as  $-\ln P(o|\theta, m)$ .

### Imaging data acquisition and analysis

We acquired  $T_2^*$ -weighted echo planar images (EPIs) by using a 32-channel head coil at a 3-T Trio Siemens scanner at the Wellcome Trust Centre for Neuroimaging. We acquired a partial volume consisting of 42 3-mm slices in descending order (echo time: 0.065 ms, repetition time: 2.940 ms) at an angle of 30° in the anterior–posterior axis to optimize sensitivity to prefrontal regions (Deichmann et al., 2003). Images were acquired with a field of view of 192 × 192 mm (matrix 64 × 64) resulting in a notional in-plane resolution of 3 × 3 × 3 mm. In each session, at least 245 volumes were acquired. An MR-compatible button box recorded right index and middle finger presses to move the cursor on the rating bar. Auditory cues were presented using MRI-compatible headphones. Foam head-restraint pads were used to minimise head movement and respiratory and cardiac activities were measured and used as covariates in the imaging analysis. Whole-brain 1 mm × 1 mm × 1 mm  $T_1$ -weighted structural images were acquired and coregistered with mean EPI images.

We performed preprocessing and all statistical analyses using SPM12b (Wellcome Trust Centre for Neuroimaging, London, UK, <http://www.fil.ion.ucl.ac.uk/spm>). To allow for  $T_1$  relaxation effects the first 6 functional images were discarded. fMRI time series were realigned to the mean image and unwarped using fieldmaps generated by the Fieldmap toolbox as implemented in SPM12b (Hutton et al., 2002). EPIs and structural scans were normalised and smoothed using a 6 mm<sup>2</sup> full-width at half-maximum (FWHM) Gaussian kernel using the DARTEL toolbox (Ashburner, 2007), allowing the comparison of data between subjects in common anatomical space.

To investigate for effects in the dopaminergic midbrain, we used a hand drawn region of interest based on previous literature (Schwartenbeck et al., 2014a) for small volume correction of our fMRI analyses.

We conducted a standard mass-univariate analysis to interrogate the data for model-based effects. We created a GLM with separate events for fixation crosses, cue onset, outcome presentation and rating bar onset, modelled as stick functions with duration 0. The outcome regressor was modulated by five additional parametric regressors, namely surprise (z-scored), shifts in beliefs (KL-divergence, z-scored), the difference between estimated shifts in beliefs and the behavioural shift in rating (z-scored), current monetary return and reward prediction error, i.e., a mismatch between expected and observed reward, where the former was defined as the sum of the valence of the auditory and visual cues multiplied by the current belief about cue contingencies as indicated by subjects’ rating at the previous trial. The full correlation matrix for surprise, shifts in beliefs, monetary outcome and reward prediction errors is displayed in Table 1. To control for any confounding effects of reward anticipation or confidence on effects at outcome presentation, we included subjects’ current belief about the relevant modality (as indicated by the rating at the previous trial), the confidence of these beliefs as indicated by the cursor position at the previous trial, expected (relevant) value (as described above) and expected *irrelevant*

**Table 1**  
Correlation matrix for surprise, shifts in beliefs, monetary outcome and reward prediction errors (RPE).

	Surprise	Shifts in beliefs	Monetary outcome	RPE
Surprise		0.581	−0.001	0.008
Shifts in beliefs			−0.006	−0.003
Monetary outcome				0.589

value (i.e., the valence of cues multiplied with subjects' beliefs that the cue is *not* relevant, as indicated by the rating at the previous trial) as parametric regressors at cue onset. Furthermore, we included the number of button presses as parametric regressor on the rating regressor. Note that we did not serially orthogonalise these regressors, thus removing all shared variance from the analysis.

We additionally included motion, respiratory and cardiac information as physiological covariates of no interest and included temporal derivatives (first-order differences) to account for slice-timing effects. Furthermore, an AR (1) model was used to account for serial correlations in the fMRI time series and we applied a 128-s cutoff high pass filter.

A standard summary statistic approach was used to test for second-level effects using one-sample t-tests on the estimated responses of the first-level (within-subject) analysis. Random field theory was used to correct for multiple comparisons.

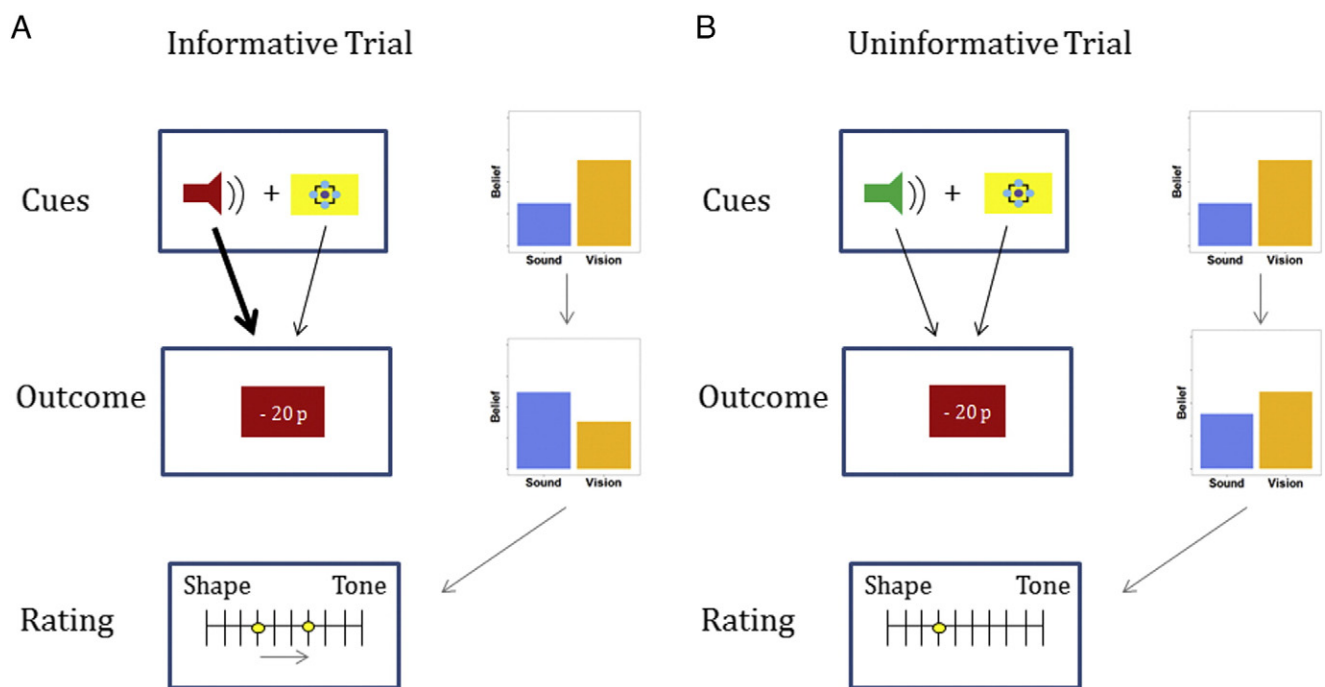
## Results

In a training period prior to the experiment, subjects learned that in both visual and auditory domains one 'bad' cue predicted a monetary loss whereas one 'good' cue predicted a monetary win, where cues had a validity of 90%. Subjects were instructed that these contingencies remained constant throughout the entire experiment. Thus, for the

entire experiment subjects knew the identity of the 'bad' shape and tone that predicted a probable loss, as well as the 'good' shape and tone that predicted a probable win (stimulus contingencies are depicted in Fig. 1a). Each trial started with a simultaneous presentation of one of the two cues from the visual and auditory domains, followed by an outcome, which was either a monetary win or loss (varying randomly between 10 and 30 pence). Current earnings were indicated by a progress bar at the bottom of the screen.

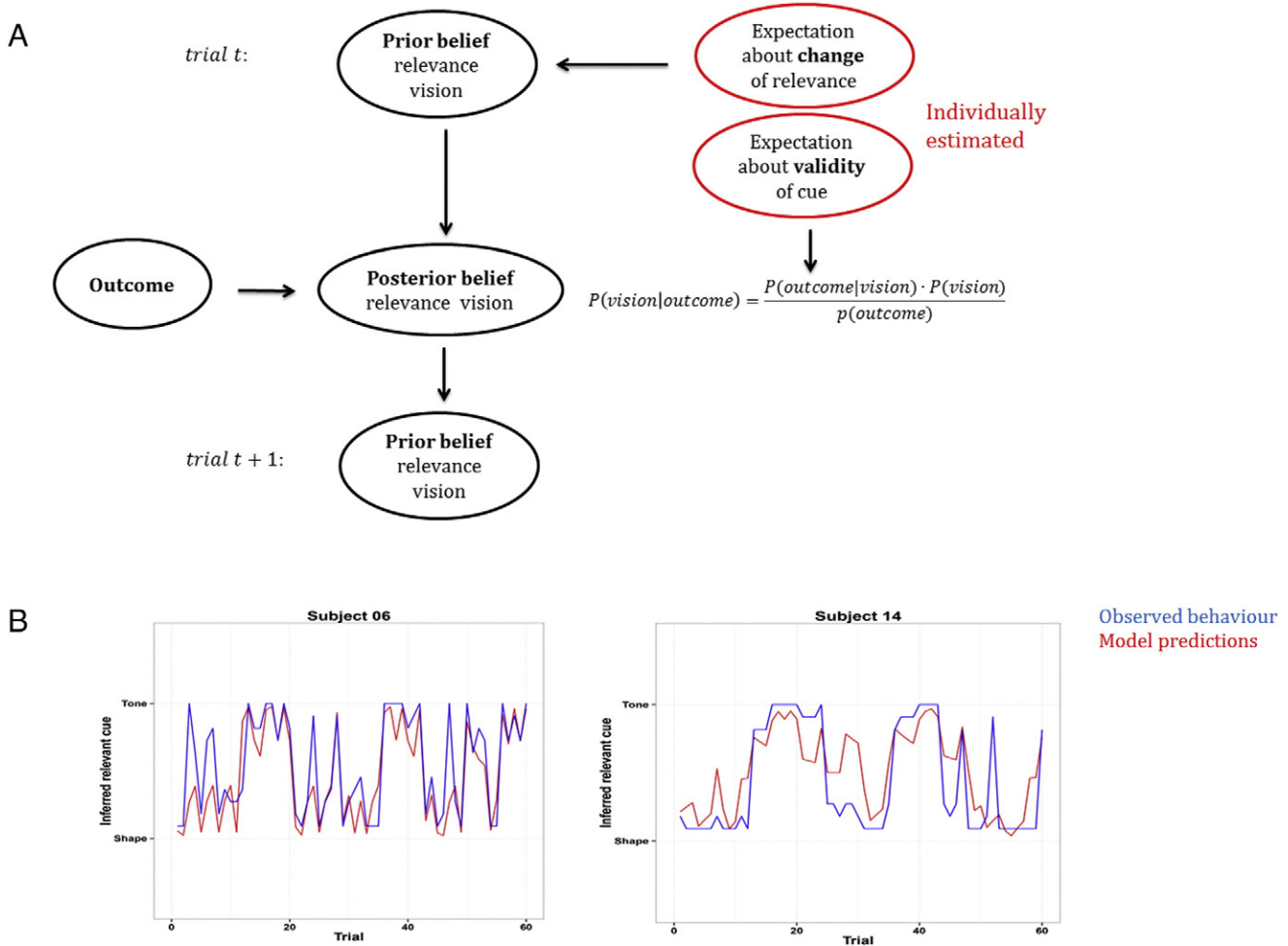
The experiment was divided into 60 trial sessions, with reversals occurring between five and six times in a session. Importantly, in half of the trials the cues were uninformative, and both cues were either good or bad (congruent valence of cues). In the other half of the trials cues were informative, i.e. one cue was good and the other bad (incongruent valence of cues), providing a decorrelation of information-theoretic and Bayesian surprise (illustrated in Fig. 2).

Subjects underwent an extensive training period to familiarise themselves with the task contingencies, in particular the approximate frequency of reversals (see [Methods](#) for details). Subsequently, they performed three sessions of the task during acquisition of fMRI data. To analyse performance on the task, we developed a simple computational model that performs Bayesian belief-updating on a trial-by-trial basis about which modality was currently relevant (model details are depicted in Fig. 3a). To derive trial-by-trial regressors for the use in the imaging analysis we fit individual subjects' data using constrained maximum-likelihood estimation, with subjects' beliefs about the validity of a cue and probability of contingency reversal as free parameters (see [Methods](#) for details). We found that our simple model had high accuracy in predicting subject's contingency-ratings ( $R^2 = 0.82$ , 95%-confidence interval: [0.97, 0.76]) (see examples in Fig. 3b and Table 2 for a full list of all estimated parameters). Furthermore, we found a high correlation between Bayesian surprise and observed shifts in beliefs ( $r = 0.58$ , 95%-confidence interval: [0.51, 0.64]). Note that the aim of our computational model was to derive imaging regressors for



**Fig. 2.** Difference between Bayesian and information-theoretic surprise. (A) Illustration of an informative trial with a shape predicting a win and a tone predicting a loss, assuming that the agent believes that visual information is relevant (thus expecting a win as a trial outcome). Observing a monetary loss will contradict the agent's expectations and induce a shift in beliefs about cue contingencies, since the trial outcome was correctly predicted by the tone but not by the shape. This is an illustration of a trial with high Bayesian surprise (i.e., a shift in beliefs). (B) Conversely, two 'good' cues followed by a (unexpected) negative trial outcome will not induce a shift in beliefs because the outcome was neither predicted by the tone nor by the shape. Thus, despite the observation of an unexpected outcome the trial provides no meaningful information content and is thus completely uninformative. Note that both examples yield high information-theoretic surprise (i.e., an unexpected outcome is observed), but only in informative trials actual shifts in beliefs can occur. Crucially, each session consisted of the same number of informative and uninformative trials, thus providing a decorrelation between information-theoretic and Bayesian surprise.





**Fig. 3.** Behavioural modelling. (A) We tested whether subjects updated their beliefs optimally, i.e. updating prior beliefs about cue contingencies to posterior beliefs according to Bayes rule. Note that  $P(\text{vision}) = 1 - P(\text{sound})$ . We performed (constrained) maximum-likelihood estimation of subjects' expectation about the frequency of contingency reversal and about the validity of cues. Posterior beliefs at one trial served as prior beliefs at the subsequent trial. (B) We found that our model accounted well for observed behaviour ( $R^2 = 0.82$ ), indicating that subjects updated their beliefs at least close to optimally (difference of predicted and observed ratings is illustrated in two example subjects).

belief-updates rather than optimally capture behaviour, and as such, we did not compare a range of different models.

To check that subjects were able to adequately perform the task, we correlated subjects' rating at the end of each trial with the true contingencies and found moderate relationship contingencies ( $r = 0.54$ , 95%-confidence interval: [0.61, 0.45]). This level of performance is in line with the challenging nature of our task. The fact that half of all trials were uninformative, implying that only in half of the trials subjects could actually infer on contingency reversals, naturally induces a delay in tracking changes in current cue-contingencies. To account for this, we additionally compared performance with that of an ideal Bayesian observer (an instantiation of our behavioural model using the true parameters of the task). We found a high correlation between the ideal observer and subjects' performance ( $r = 0.78$ , 95%-confidence interval: [0.83, 0.71]), confirming that subjects were able to perform the task successfully.

We next derived, based on the model-fitting, trial-by-trial regressors encoding information-theoretic surprise, epistemic value and reward prediction errors, and examined their neuronal correlates. In particular we were interested in whether one of these quantities was encoded in midbrain regions commonly associated with dopaminergic function, namely the Substantia nigra pars compacta (SNc) and Ventral Tegmental Area (VTA) (Düzel et al., 2009).

In keeping with our hypothesis, we found that dopamine-rich midbrain regions encoded Bayesian surprise or epistemic value

(small-volume corrected, peak MNI  $-6 - 15 - 10$ ,  $P_{\text{peak}} < 0.001$ ,  $T_{\text{peak}} = 6.24$  and peak MNI  $10 - 18 - 10$ ,  $P_{\text{peak}} = 0.049$ ,  $T_{\text{peak}} = 3.64$  see Fig. 4a). To control for additional effects of belief updating beyond estimated Bayesian surprise, we used the observed change in rating at the end of each trial as an overt behavioural proxy for belief updating. We included the difference between estimated Bayesian surprise and the observed change in rating as an additional regressor to test whether midbrain activity predicted shifts in beliefs over and above those predicted by our model. We detected additional effects in the dopaminergic midbrain for belief updating over and beyond our model predictions (small-volume corrected, peak MNI  $-10 - 20 - 10$ ,  $P_{\text{peak}} = 0.006$ ,  $T_{\text{peak}} = 4.65$ ), suggesting that midbrain activity predicted fluctuations in behaviour over and above those predicted by the model. Furthermore, we replicated these effects in the midbrain in an additional analysis, where we used the actual observed shifts in rating as opposed to the belief-updates as estimated by our model (small-volume corrected, peak MNI  $-6 - 15 - 10$ ,  $P_{\text{peak}} = 0.03$ ,  $T_{\text{peak}} = 3.87$  and, reaching trend-significance, peak MNI  $12 - 18 - 10$ ,  $P_{\text{peak}} = 0.065$ ,  $T_{\text{peak}} = 3.48$ ). Taken together, both sets of findings provide convergent evidence that putative dopaminergic midbrain regions encode epistemic value, or meaningful information, of a stimulus.

In a whole brain analysis, we found effects for the difference between behavioural rating and modelled shifts in beliefs in the inferior frontal cortex (peak MNI  $45 8 21$ ,  $P_{\text{cluster}} < 0.001$ ,  $T_{\text{peak}} = 4.84$ ), the posterior parietal cortex (peak MNI  $40 - 72 32$ ,  $P_{\text{cluster}} = 0.006$ ,  $T_{\text{peak}} =$

**Table 2**

Model parameters. Individual beliefs about cue validity and expectations of reversal probability per session estimated individually based on observed behaviour.

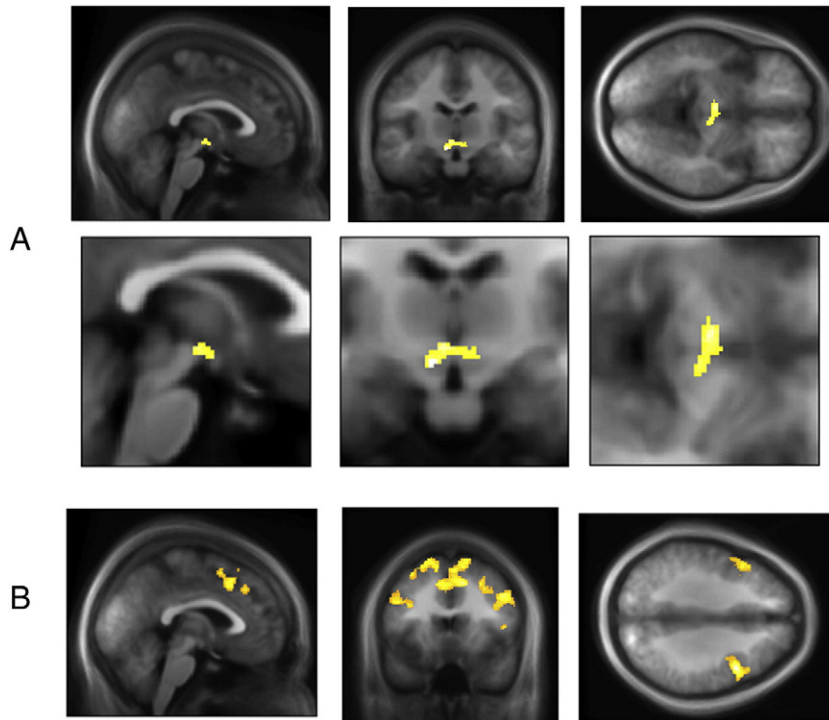
Subject	$\tau_{\text{sess } 1}$	$\tau_{\text{sess } 2}$	$\tau_{\text{sess } 3}$	$\delta_{\text{sess } 1}$	$\delta_{\text{sess } 2}$	$\delta_{\text{sess } 3}$
1	0.99	0.99	0.77	0.99	0.77	0.68
2	0.95	0.98	0.92	0.98	0.99	0.94
3	0.69	0.50	0.91	0.72	0.71	0.94
4	0.50	0.7	0.94	0.94	0.99	0.99
5	0.92	0.84	0.86	0.99	0.98	0.98
6	0.98	0.94	0.93	0.83	0.74	0.75
7	0.99	0.79	0.94	0.98	0.83	0.99
8	0.99	0.97	1.00	0.99	0.96	0.93
9	0.83	0.66	0.84	0.96	0.95	0.91
10	0.99	0.99	0.98	0.99	0.99	0.99
11	0.92	0.89	0.77	0.99	0.98	0.88
12	0.82	0.84	0.85	0.97	0.97	0.96
13	1.00	0.99	0.50	0.97	0.72	0.97
14	0.99	0.79	0.95	0.97	0.94	0.96
15	0.83	0.72	0.80	0.91	0.89	0.87
16	0.96	0.86	0.99	0.99	0.97	1.00
17	0.68	0.78	0.91	0.92	0.91	0.94
18	0.88	0.95	0.86	0.96	0.96	0.90
19	0.95	0.96	0.67	0.97	0.97	0.89
20	0.55	0.50	0.97	0.99	0.96	0.99
21	0.97	0.72	0.97	0.96	0.84	0.95
22	0.82	0.89	0.99	0.97	0.98	1.00
23	1.00	1.00	1.00	0.99	1.00	0.99
24	0.92	0.87	0.50	0.99	0.99	0.89
25	0.96	1.00	0.90	0.94	0.88	0.86
26	0.89	0.89	0.84	0.98	0.98	0.95
27	0.74	0.73	0.81	0.97	0.95	0.97

5.05) and in the anterior cingulate cortex (ACC) (peak MNI 6 27 27,  $P_{\text{cluster}} < 0.001$ ,  $T_{\text{Peak}} = 5.27$ ). This is in keeping with previous work which also implicates these areas in model updating (O'Reilly et al., 2013), and suggests possible cortical targets of, or inputs to, midbrain neurons encoding epistemic value.

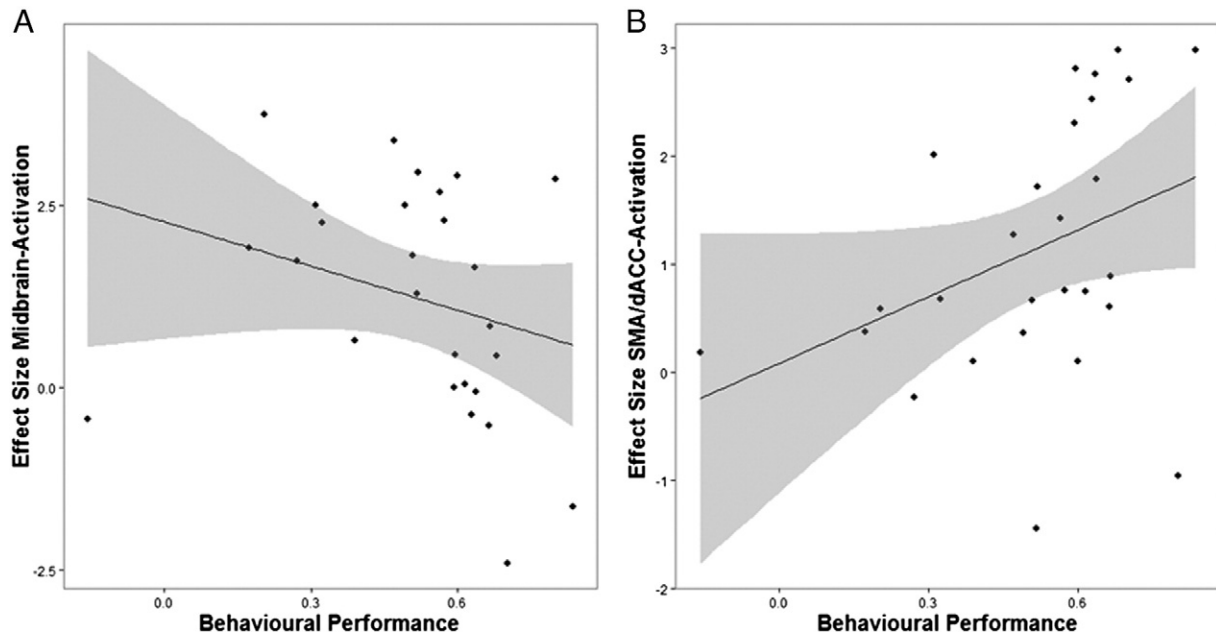
Next, we investigated whether individual differences in performance would relate to the effect size of midbrain activity for meaningful information. As a measure of behavioural performance, we correlated each subject's rating with the true contingencies at any given trial. As a proxy for the effect size of the midbrain activation, we extracted individual betas from the peak-voxel of the combined effects for estimated and observed shifts in beliefs (peak MNI  $-9 - 20 - 12$ ,  $P_{\text{Peak}} = 0.02$ ,  $T_{\text{Peak}} = 4.10$ ). We found that this measure of performance correlated negatively with the effect size of activation in midbrain regions for shifts in beliefs (Spearman's  $\rho = -0.44$ ,  $p = 0.02$ , see Fig. 5a), suggesting that subjects who performed comparatively well in tracking the true contingencies displayed decreased effects in midbrain regions for shifts in beliefs. A similar result (reaching trend-significance) was found when we tested the relationship between performance and the effect size of midbrain activation using a correlation between subjects' behaviour and the ideal observer model as measure of performance (Spearman's  $\rho = -0.36$ ,  $p = 0.07$ ). This raises the intriguing possibility that individual differences in dopaminergic sensitivity to informative cues might affect learning performance, which might explain the abnormal belief formation that has been associated with dopaminergic abnormalities (Howes and Kapur, 2009).

Consistent with previous findings, effects of information-theoretic surprise were observed in a network of frontal regions including the ventral pre-supplementary motor cortex (pre-SMA) and dorsal anterior cingulate cortex (peak MNI 12 16 42,  $P_{\text{cluster}} < 0.001$ ,  $T_{\text{Peak}} = 4.86$ , see Fig. 4b). This is in keeping with the general idea of the brain as a predictive organ (Dayan et al., 1995; Rao and Ballard, 1999), and supports a distinction between the neuronal processing of meaningful information and simple surprise.

We again tested whether individual differences in the effect size of SMA/dorsal cingulate activation would relate to behavioural performance. Intriguingly, we found a strong positive relationship between behavioural performance and individual effect size of activation for surprise (betas extracted from peak voxel MNI 12 16 42, see Fig. 5b),



**Fig. 4.** Imaging results. (A) Shifts in beliefs were encoded in dopamine-rich midbrain regions including the Substantia Nigra compacta and Ventral Tegmental Area (small-volume corrected, peak MNI  $-6 - 15 - 10$ ,  $P_{\text{Peak}} < 0.001$ ,  $T_{\text{Peak}} = 6.24$  and peak MNI  $10 - 18 - 10$ ,  $P_{\text{Peak}} = 0.049$ ,  $T_{\text{Peak}} = 3.64$ ; voxel extend threshold of 100 for display purposes). (B) Information-theoretic surprise was encoded in frontal regions including the dorsal cingulate cortex (peak MNI 12 16 42,  $P_{\text{cluster}} < 0.001$ ,  $T_{\text{Peak}} = 4.86$ ; image thresholded at  $p = 0.01$  and voxel extend threshold of 392 for display purposes).



**Fig. 5.** Relationship between performance and midbrain as well as SMA/dACC activation. A) Subjects who were comparatively better (performance on x axis) at inferring the actual cue contingencies displayed relatively less midbrain activation for shifts in beliefs compared to subjects who were worse at tracking these contingencies during the task. B) In contrast, subjects who were comparatively good at tracking the cue contingencies showed higher activity in the SMA/dACC for information-theoretic surprise.

both for performance with respect to the true task contingencies (Spearman's Rho  $\rho = 0.48$ ,  $p = 0.01$ ) and the ideal observer (Spearman's Rho  $\rho = 0.59$ ,  $p = 0.001$ ). Put simply, this implies that subjects who were comparatively good at tracking the task contingencies showed increased effect sizes of activation in the SMA/dACC for information-theoretic surprise.

Finally, we tested for monetary outcome and reward prediction error effects at outcome time, which are typically associated with activity in the dopaminergic midbrain and its targets such as the ventral striatum (O'Doherty et al., 2004; Schultz, 2012; Schultz et al., 1997). Even at a very liberal threshold of  $p < 0.05$  uncorrected we did not detect effects for those regressors in dopaminergic midbrain or striatal regions (see Table 3 for details). While the absence of significant effects cannot be taken as evidence against the neuronal encoding of reward prediction errors, these results are in line with the general structure of our paradigm wherein subjects were over-trained on the task contingencies. Specifically, subjects had to perform inference on the relevance of cues based on trial-outcomes without a need to perform any kind of reward learning. Our results suggest that, across subjects at least, the dopamine system was if at all only weakly engaged by the reward component of the task, compared with its epistemic aspects, most likely due to the fact that the value per se of outcomes was irrelevant for guiding future behaviour. This finding is in line with evidence that prediction responses to reward in regions such as the striatum are most prominent in situations where they have relevance for on-going behaviour (FitzGerald et al., 2014; Klein-Flügge et al., 2011).

**Table 3**  
Effects of reward prediction error on the whole brain level (reported for  $p < 0.05$  uncorrected).

Region	Cluster size	FWE-corrected p-value	T-value peak voxel	Peak coordinates (MNI)		
				X	Y	Z
Precentral gyrus	317	0.001	5.22	33	-12	58

## Discussion

We provide evidence that activity in dopamine-rich midbrain regions encodes the quantity of meaningful information imparted by a stimulus. This is often referred to as its epistemic value (Friston et al., 2015), defined as the Kullback–Leibler divergence between beliefs before and after observing the stimulus. These results support suggestions that dopaminergic rich regions play a key role in belief updating, consistent with the broader notion that abnormal belief formation is associated with hyperdopaminergic function (Kapur, 2003). It also highlights the importance of dopamine for the expression of cognitive flexibility (Bestmann et al., 2014; van Holstein et al., 2011). We also found evidence for a dissociation between information-theoretic surprise, which is mainly encoded in frontal regions, and actual shifts in beliefs which are associated with activity in putatively dopaminergic midbrain regions. Thus, in keeping with prior work, we show that information-theoretic surprise was encoded in regions typically associated with internal conflict and task difficulty, such as the pre-SMA and dorsal cingulate cortex (Duncan, 2010; Kool and Botvinick, 2013).

In a post-hoc analysis, we detected a between-subject effect, such that participants who were better at tracking the true contingencies during the task displayed smaller effect sizes in terms of shifts in beliefs induced activation within dopamine-rich midbrain regions. Speculatively, this might relate to the effort of tracking the true contingencies or indeed the amount of meaningful information imparted by an observation, such that incoming information is more 'meaningful' to subjects who perform poorly on tracking the true contingencies, requiring a stronger update of beliefs compared to subjects whose beliefs in the contingencies are more accurate in the first place. Alternatively, subjects who perform poorly in the task may be hypersensitive to unexpected events and shift their beliefs too often, where this hypersensitivity is reflected in stronger midbrain activation. In either case, our results raise the possibility that inter-individual differences in dopaminergic function drive altered belief formation, over and above any effects of the dopamine system in signalling reward (Schultz et al., 1997). Intriguingly, we found the opposite pattern for effect sizes of activation for information-theoretic surprise in the supplementary motor area and dorsal anterior cingulate cortex: subjects who were comparatively better at tracking the task contingencies showed higher activation for

surprising outcomes. While one has to be cautious with post-hoc explanations, this may indeed speak for a between-subjects anticipation effect: subjects who were better at anticipating and tracking reversals of the task contingencies displayed attenuated effects in the midbrain, which may reflect less requirement of an update or, put differently, observations may have become less meaningful with respect to the subject's beliefs. Surprising events, however, cannot be anticipated – and thus may elicit stronger effects in subjects who are better at tracking the true contingencies. Furthermore, this might also indicate that subjects who performed poorly at tracking the true contingencies found it harder to dissociate between surprising and meaningful information, which is a promising topic for future investigations and translational approaches.

Existing work reports a dopaminergic response to novel and unexpected stimuli (Bromberg-Martin et al., 2010; Bunzeck and Düzel, 2006; DeYoung, 2013; Horvitz, 2000). Since these stimuli will, in general, lead to significant belief-updating (Barto et al., 2013) they should also, on an epistemic value account of dopaminergic firing, occasion significant responses. A key contribution of our study is to show, by separating surprise from actual shifts in beliefs, that dopamine is sensitive only to meaningful information, rather than the simple unexpectedness of a stimulus. This highlights an important role for dopamine in belief formation over and above other proposed roles such as reward learning (Schultz, 2012; Schultz et al., 1997), motivational salience (Berridge, 2007), exploration (DeYoung, 2013; Redgrave and Gurney, 2006) and the precision of action selection (Friston et al., 2014a). Further, the dissociation between surprise and meaningful information has also been reported in the context of memory reconsolidation, such that prediction error or surprisal has been shown to control the balance between memory reconsolidation and new memory formation (Sevenster et al., 2014), highlighting the important interplay between memory reconsolidation and dopamine function.

Our findings also resonate with previous work showing that dopamine neurons encode reward uncertainty, displaying highest sensitivity for maximal uncertainty, i.e. a probability of 0.5 for receiving a reward (Fiorillo et al., 2003). Intuitively, it seems plausible that a shift in beliefs relates to higher uncertainty about the environment and thus higher uncertainty about the delivery of a reward. However, as displayed in Fig. 2a, shifts of beliefs can also occur in informative trials when there was relatively little uncertainty about the outcome. Furthermore, beliefs can also shift towards confirming current expectations about the cue contingencies. In tasks where surprise and model updating are highly correlated a surprising outcome will necessarily induce a shift in beliefs as well as higher uncertainty about the task contingencies. The crucial contribution of this study, however, is a dissociation of surprise from belief updates, showing that only the latter elicits midbrain activation whereas the former is associated with effects in the SMA and dorsal anterior cingulate cortex.

Our paradigm also relates to what has been reported as the 'gambler's fallacy' or the 'Perruchet effect'. In brief, this fallacy reflects a failure in acknowledging the independence of events in decision-making paradigms, leading to the overestimation of the likelihood for winning a gamble or receiving a shock as expressed in subjects' beliefs. In contrast, previous work has provided evidence for a second, more automated learning system expressed in the galvanic skin response, which correctly tracks the true contingencies (McAndrew et al., 2012; Weidemann et al., 2009). In our paradigm, however, subjects were asked to track cue contingencies on a trial-by-trial basis where the independence assumption was explicitly violated; the likelihood for a reversal of cue contingencies increased with ongoing time. Therefore, the question of interest becomes how subjects actually represent the 'non-independence' of trials in such situations. While this is an interesting question, testing this would be beyond the scope of the present study, which aimed at dissociating belief updates due to meaningful information and the pure unexpectedness of an observation as reflected in information-theoretic surprise. Testing for the representation of those

dependencies between trials and its potential physiological correlates, such as skin conductance or pupil diameter, however, remains an interesting area for future research.

In a previous study using a saccadic eye movement task, surprise has been associated with activity in the inferior posterior parietal cortex, whereas model updating was associated with the anterior cingulate cortex (ACC) (O'Reilly et al., 2013). While we partly replicated the findings of an ACC-involvement in model updating, we did not detect any parietal activity for the unexpectedness of an observation. These differences are likely to reflect the lack of a clear motor planning component to our task. Conversely, O'Reilly et al. do not report activity in the midbrain relating to model-updating. While it is dangerous to over-interpret a negative finding, one crucial difference is that in O'Reilly et al. 'surprise' and 'model-update' trials were explicitly indicated, whereas in our paradigm the epistemic value of feedback cues had to be inferred. The aspect of inferring the epistemic value (meaningful information content) of a stimulus may be the crucial factor determining an involvement of the dopaminergic system. Also the fact that we only detected a weak signal in the ACC is likely to be due to the different nature of tasks, such the ACC typically is more strongly involved in instrumental tasks (Matsumoto et al., 2007; O'Reilly et al., 2013; Rudebeck et al., 2008) whereas our task only required subjects to make inferences based on observations.

Our data provide the first association between information signalling and dopamine putative function in humans. Understanding the link between dopamine and belief updating that implicates a selective sensitivity for meaningful information also sheds new light on findings that associate dopamine with signalling advance information (Bromberg-Martin and Hikosaka, 2009). In the light of our results, dopamine-rich regions might be thought of in terms of encoding advance information not because information is inherently rewarding per se. Instead advance information has epistemic value for an agent in terms of inducing a shift in its beliefs. Our data, therefore, speak to a direct association between dopamine and an agent's beliefs about the world, predicting dopaminergic function will reflect a sensitivity for meaningful information that affects beliefs.

Finally, understanding the link between dopamine and belief formation can provide crucial insight into how dopaminergic abnormalities in conditions such as in schizophrenia, obsessive compulsive disorder and addiction (Friston et al., 2014b; Howes et al., 2009; Howes et al., 2011a, 2011b; Kapur, 2003; Roiser et al., 2013; Schwartenbeck et al., 2014b; Shaner, 1999). In particular, it is widely hypothesised that dysregulated dopaminergic function leads to increased salience being conferred on sensory stimuli, a step in a path leading to psychosis (Howes and Kapur, 2009) (though the precise mechanisms by which this might occur remain to be mapped out). Our results nuance this hypothesis by suggesting that the increased salience conferred by dysregulated dopaminergic activity may be best understood in terms of the pathological assignment of meaningfulness or epistemic value to sensations. This is remarkably consistent with the phenomenology of prodromal psychosis (the so-called 'Truman sign') (Fusar-Poli et al., 2008), and provides an intriguing and potentially important topic for future research. By implicating dopamine in signalling the meaningful information imparted by an observation (or, equivalently, the degree of belief updating that is necessary in response to it), our findings suggest new directions for future theoretical and empirical work on the expression of aberrant beliefs.

## Acknowledgments

Philipp Schwartenbeck is a recipient of a DOC Fellowship of the Austrian Academy of Sciences at the Centre for Cognitive Neuroscience, University of Salzburg. This work was supported by the Wellcome Trust (Ray Dolan Senior Investigator Award 098362/Z/12/Z). The Wellcome Trust Centre for Neuroimaging is supported by core funding from the Wellcome Trust 091593/Z/10/Z.



## References

- Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. *NeuroImage* 38, 95–113. <http://dx.doi.org/10.1016/j.neuroimage.2007.07.007>.
- Barto, A., Mirrolli, M., Baldassarre, G., 2013. Novelty or surprise? *Front. Psychol.* 4, 907. <http://dx.doi.org/10.3389/fpsyg.2013.00907>.
- Berridge, K.C., 2007. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)* 191, 391–431. <http://dx.doi.org/10.1007/s00213-006-0578-x>.
- Bestmann, S., Ruge, D., Rothwell, J., Galea, J.M., 2014. The role of dopamine in motor flexibility. *J. Cogn. Neurosci.* 1–12. [http://dx.doi.org/10.1162/jocn\\_a.00706](http://dx.doi.org/10.1162/jocn_a.00706).
- Bromberg-Martin, E., Hikosaka, O., 2009. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* 63, 119–126. <http://dx.doi.org/10.1016/j.neuron.2009.06.009>.
- Bromberg-Martin, E.S., Matsumoto, M., Hikosaka, O., 2010. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834. <http://dx.doi.org/10.1016/j.neuron.2010.11.022>.
- Bunzeck, N., Düzel, E., 2006. Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron* 51, 369–379. <http://dx.doi.org/10.1016/j.neuron.2006.06.021>.
- Cools, R., Frank, M.J., Gibbs, S.E., Miyakawa, A., Jagust, W., D'Esposito, M., 2009. Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *J. Neurosci.* 29, 1538–1543. <http://dx.doi.org/10.1523/JNEUROSCI.4467-08.2009>.
- Corlett, P.R., Frith, C.D., Fletcher, P.C., 2009. From drugs to deprivation: a Bayesian framework for understanding models of psychosis. *Psychopharmacology (Berl)* 206, 515–530. <http://dx.doi.org/10.1007/s00213-009-1561-0>.
- Dayan, P., Hinton, G.E., Neal, R.M., Zemel, R.S., 1995. The Helmholtz machine. *Neural Comput.* 7, 889–904.
- Deichmann, R., Gottfried, J.A., Hutton, C., Turner, R., 2003. Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage* 19, 430–441.
- DeYoung, C.G., 2013. The neuromodulator of exploration: a unifying theory of the role of dopamine in personality. *Front. Hum. Neurosci.* 7. <http://dx.doi.org/10.3389/fnhum.2013.00762>.
- Dolan, R.J., Dayan, P., 2013. Goals and habits in the brain. *Neuron* 80, 312–325. <http://dx.doi.org/10.1016/j.neuron.2013.09.007>.
- Duncan, J., 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn. Sci.* 14, 172–179. <http://dx.doi.org/10.1016/j.tics.2010.01.004>.
- Düzel, E., Bunzeck, N., Guitart-Masip, M., Wittmann, B., Schott, B.H., Tobler, P.N., 2009. Functional imaging of the human dopaminergic midbrain. *Trends Neurosci.* 32, 321–328. <http://dx.doi.org/10.1016/j.tins.2009.02.005>.
- Fiorillo, C.D., Tobler, P.N., Schultz, W., 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299 (80–), 1898–1902. <http://dx.doi.org/10.1126/science.1077349>.
- FitzGerald, T.H.B., Schwartenbeck, P., Dolan, R.J., 2014. Reward-related activity in ventral striatum is action contingent and modulated by behavioral relevance. *J. Neurosci.* 34, 1271–1279. <http://dx.doi.org/10.1523/JNEUROSCI.4389-13.2014>.
- Friston, K.J., 2010. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. <http://dx.doi.org/10.1038/nrn2787>.
- Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., Dolan, R.J., 2014a. The anatomy of choice: dopamine and decision-making. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369. <http://dx.doi.org/10.1098/rstb.2013.0481>.
- Friston, K., Stephan, K.E., Montague, R., Dolan, R.J., 2014b. Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatry* 1, 148–158. [http://dx.doi.org/10.1016/S2215-0366\(14\)70275-5](http://dx.doi.org/10.1016/S2215-0366(14)70275-5).
- Friston, K., Rigoli, F., Ognibene, D., Mathys, C., FitzGerald, T., Pezzulo, G., 2015. Active inference and epistemic value. *Cogn. Neurosci.* <http://dx.doi.org/10.1080/17588928.2015.1020053>.
- Fusar-Poli, P., Howes, O., Valmaggia, L., McGuire, P., 2008. "Truman" signs and vulnerability to psychosis. *Br. J. Psychiatry* 193, 168. <http://dx.doi.org/10.1192/bjp.193.2.168>.
- Horvitz, J.C., 2000. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96, 651–656.
- Howes, O.D., Kapur, S., 2009. The dopamine hypothesis of schizophrenia: version III – the final common pathway. *Schizophr. Bull.* 35, 549–562. <http://dx.doi.org/10.1093/schbul/sbp006>.
- Howes, O.D., Montgomery, A.J., Asselin, M.-C., Murray, R.M., Valli, I., Tabraham, P., Bramon-Bosch, E., Valmaggia, L., Johns, L., Broome, M., McGuire, P.K., Grasby, P.M., 2009. Elevated striatal dopamine function linked to prodromal signs of schizophrenia. *Arch. Gen. Psychiatry* 66, 13–20. <http://dx.doi.org/10.1001/archgenpsychiatry.2008.514>.
- Howes, O.D., Bose, S., Turkheimer, F., Valli, I., Egerton, A., Stahl, D., Valmaggia, L., Allen, P., Murray, R., McGuire, P., 2011a. Progressive increase in striatal dopamine synthesis capacity as patients develop psychosis: a PET study. *Mol. Psychiatry* 16, 885–886. <http://dx.doi.org/10.1038/mp.2011.20>.
- Howes, O.D., Bose, S.K., Turkheimer, F., Valli, I., Egerton, A., Valmaggia, L.R., Murray, R.M., McGuire, P., 2011b. Dopamine synthesis capacity before onset of psychosis: a prospective [18 F]-DOPA PET imaging study. *Psychiatry Online* 1311–1317.
- Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., Turner, R., 2002. Image distortion correction in fMRI: a quantitative evaluation. *NeuroImage* 16, 217–240. <http://dx.doi.org/10.1006/nimg.2001.1054>.
- Iglesias, S., Mathys, C., Brodersen, K.H., Kasper, L., Piccirelli, M., den Ouden, H.E.M., Stephan, K.E., 2013. Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron* 80, 519–530. <http://dx.doi.org/10.1016/j.neuron.2013.09.009>.
- Itti, L., Baldi, P., 2009. Bayesian surprise attracts human attention. *Vis. Res.* 49, 1295–1306. <http://dx.doi.org/10.1016/j.visres.2008.09.007>.
- Kapur, S., 2003. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry* 160, 13. <http://dx.doi.org/10.1176/appi.ajp.160.1.13>.
- Klein-Flügge, M.C., Hunt, L.T., Bach, D.R., Dolan, R.J., Behrens, T.E.J., 2011. Dissociable reward and timing signals in human midbrain and ventral striatum. *Neuron* 72, 654–664. <http://dx.doi.org/10.1016/j.neuron.2011.08.024>.
- Kool, W., Botvinick, M., 2013. The intrinsic cost of cognitive control. *Behav. Brain Sci.* 36, 697–726. <http://dx.doi.org/10.1017/S0140525X1300109X>.
- Körding, K., Wolpert, D., 2004. Bayesian integration in sensorimotor learning. *Nature* 427, 244–247. <http://dx.doi.org/10.1038/nature02169>.
- Matsumoto, M., Matsumoto, K., Abe, H., Tanaka, K., 2007. Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* 10, 647–656. <http://dx.doi.org/10.1038/nn1890>.
- McAndrew, A., Jones, F.W., McLaren, R.P., McLaren, I.P.L., 2012. Dissociating expectancy of shock and changes in skin conductance: an investigation of the Perruchet effect using an electrodermal paradigm. *J. Exp. Psychol.* 38, 203–208. <http://dx.doi.org/10.1037/a0026718>.
- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., Dolan, R.J., 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304 (80–), 452–454. <http://dx.doi.org/10.1126/science.1094285>.
- O'Reilly, J.X., Schüffelgen, U., Cuell, S.F., Behrens, T.E.J., Mars, R.B., Rushworth, M.F.S., 2013. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc. Natl. Acad. Sci. U. S. A.* 110, E3660–E3669. <http://dx.doi.org/10.1073/pnas.1305373110>.
- Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. <http://dx.doi.org/10.1038/4580>.
- Redgrave, P., Gurney, K., 2006. The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* 7, 967–975. <http://dx.doi.org/10.1038/nrn2022>.
- Roiser, J.P., Howes, O.D., Chaddock, C.A., Joyce, E.M., McGuire, P., 2013. Neural and behavioral correlates of aberrant salience in individuals at risk for psychosis. *Schizophr. Bull.* 39, 1328–1336. <http://dx.doi.org/10.1093/schbul/sbs147>.
- Rudebeck, P.H., Behrens, T.E., Kennerly, S.W., Baxter, M.G., Buckley, M.J., Walton, M.E., Rushworth, M.F.S., 2008. Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J. Neurosci.* 28, 13775–13785. <http://dx.doi.org/10.1523/JNEUROSCI.3541-08.2008>.
- Schultz, W., 2012. Updating dopamine reward signals. *Curr. Opin. Neurobiol.* 23, 229–238. <http://dx.doi.org/10.1016/j.conb.2012.11.012>.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science* 275 (80–), 1593–1599.
- Schwartenbeck, P., FitzGerald, T.H.B., Mathys, C., Dolan, R., Friston, K., 2014a. The dopaminergic midbrain encodes the expected certainty about desired outcomes. *Cereb. Cortex* <http://dx.doi.org/10.1093/cercor/bhu159>.
- Schwartenbeck, P., FitzGerald, T.H.B., Mathys, C., Dolan, R., Wurst, F., Kronbichler, M., Friston, K., 2014b. Optimal inference with suboptimal models: addiction and active Bayesian inference. *Med. Hypotheses* 84, 109–117. <http://dx.doi.org/10.1016/j.mehy.2014.12.007>.
- Sevenster, D., Beckers, T., Kindt, M., 2014. Prediction error demarcates the transition from retrieval, to reconsolidation, to new learning. *Learn. Mem.* 21, 580–584.
- Shaner, A., 1999. Delusions, superstitious conditioning and chaotic dopamine neurodynamics. *Med. Hypotheses* 52, 119–123. <http://dx.doi.org/10.1054/mehy.1997.0656>.
- Stelzel, C., Basten, U., Montag, C., Reuter, M., Fiebach, C.J., 2010. Frontostriatal involvement in task switching depends on genetic differences in d2 receptor density. *J. Neurosci.* 30, 14205–14212. <http://dx.doi.org/10.1523/JNEUROSCI.1062-10.2010>.
- Sun, Y., Gomez, F., Schmidhuber, J., 2011. Planning to be surprised: optimal Bayesian exploration in dynamic environments. *Artif. Intell.* 6830, 41–51.
- Tenenbaum, J.B., Griffiths, T.L., Kemp, C., 2006. Theory-based Bayesian models of inductive learning and reasoning. *Trends Cogn. Sci.* 10, 309–318. <http://dx.doi.org/10.1016/j.tics.2006.05.009>.
- Van Holstein, M., Aarts, E., van der Schaaf, M.E., Geurts, D.E.M., Verkes, R.J., Franke, B., van Schouwenburg, M.R., Cools, R., 2011. Human cognitive flexibility depends on dopamine D2 receptor signaling. *Psychopharmacology (Berl)* 218, 567–578. <http://dx.doi.org/10.1007/s00213-011-2340-2>.
- Weidemann, G., Tange, J.M., Lovibond, P.F., Mitchell, C.J., 2009. Is Perruchet's dissociation between eyeblink conditioned responding and outcome expectancy evidence for two learning systems? *J. Exp. Psychol.* 35, 169–176. <http://dx.doi.org/10.1037/a0013294>.