# Precision weighting of cortical unsigned prediction errors is mediated by dopamine and benefits learning

*J. Haarsma[1], P.C. Fletcher[1,2], H. Ziauddeen[1,2,3], T.J. Spencer[4], K.M.J. Diederen[1,4]*, G.K Murray[1,3]**

[1]Department of Psychiatry, University of Cambridge, United Kingdom, [2]Wellcome Trust MRC Institute of Metabolic Science, Cambridge Biomedical Campus, United Kingdom, [3]Cambridgeshire and Peterborough Foundation Trust, Fulbourn Hospital, United Kingdom, [4]Department of Psychosis studies, Institute of Psychiatry, Psychology and Neuroscience King's College London, De Crespigny Park, United Kingdom

***= joint last author/equal contributions**

**Correspondence to:**

Dr Diederen, Department of Psychosis studies, Institute of Psychiatry, Psychology and Neuroscience King's College London, De Crespigny Park, London, SE5 8AF, United Kingdom (kelly.diederen@kcl.ac.uk) **or**

Dr Murray, Department of Psychiatry, University of Cambridge, Douglas House, Trumpington Road 18b, Cambridge, CB2 8AH, United Kingdom (gm285@cam.ac.uk)

## *Abstract (150 words)*

The predictive coding framework construes the brain as performing a specific form of hierarchical Bayesian inference. In this framework the precision of cortical unsigned prediction error (surprise) signals is proposed to play a key role in learning and decision-making, and to be controlled by dopamine. To test this hypothesis, we re-analysed an existing data-set from healthy individuals who received a dopamine agonist, antagonist or placebo and who performed an associative learning task under different levels of outcome precision. Computational reinforcement-learning modelling of behaviour provided support for precision-weighting of unsigned prediction errors. Functional MRI revealed coding of unsigned prediction errors relative to their precision in bilateral superior frontal gyri and dorsal anterior cingulate. Cortical precision-weighting was (i) perturbed by the dopamine antagonist sulpiride, and (ii) associated with task performance. These findings have important implications for understanding the role of dopamine in reinforcement learning and predictive coding in health and illness.

***Introduction***

Recent decades have witnessed an increased interest in Bayesian theories of neural function, where neural signals representing prediction errors (expected – predicted outcome) play a key role (Rao & Ballard, 1999; den Ouden et al., 2008 & 2012; Bastos et al., 2012). These models, including predictive coding frameworks and the free-energy principle, conceptualize the brain as performing a specific form of hierarchical Bayesian inference, where each layer in the cortex aims to best predict its input, and updates its prediction by virtue of the prediction error, to achieve an accurate model of the world (Rao & Ballard, 1999; Bar, 2009; Friston, 2009; Mathys et al., 2011; Bastos et al., 2012; Clark, 2013 & 2015; Hohwy, 2013). A core idea included in these models is that prediction errors are scaled in a precision-weighted fashion, i.e. neural systems encoding predictions errors respond more strongly when new information is more reliable and hence more informative. Neural implementations of these theories have hypothesized the importance of the dopamine system in controlling the precision-weighted coding of prediction errors (Friston, 2009; Bastos et al., 2012; Adams et al., 2013). However, to the best of our knowledge, there is no direct evidence to support the hypothesis of a dopamine-mediated precision-weighted prediction error signalling in the cortex.

As with Bayesian theories of neural function, reinforcement-learning models also highlight the importance of the minimization of prediction error (Sutton & Barto, 1998). However, in this context a distinction is usually made between signed and unsigned prediction errors, which play different roles (Roesch et al., 2012). The signed prediction error signals the extent to which an outcome is better or worse than expected and changes the value of a stimulus accordingly (Rescorla & Wagner et al., 1972; Sutton, 1988). The unsigned prediction error signals the degree to which an outcome is unexpected, independent of its sign, and thereby controls the rate of learning (Pearce & Hall 1980; Hall & Pearce et al., 1982). In Bayesian reinforcement learning models unsigned prediction errors increase (or decrease) the uncertainty regarding prediction estimates and thereby increase learning in a fashion similar to reinforcement learning (Courville et al., 2006). Critically, influential Bayesian models also posit the presence of a cortical precision-weighted unsigned prediction error signal, potentially mediated by dopamine (Friston, 2009; Bastos et al., 2012; Adams et al., 2013).

Where in the brain can we expect to find evidence of a precision weighted unsigned prediction error signal that is mediate by dopamine? During reinforcement learning, neural

signals representing signed reward prediction errors and motivational salience prediction errors can be found in various locations (Roesch et al., 2012). For example, signed reward prediction errors are coded in a variety of brain regions, most notably the dopaminergic midbrain and ventral striatum (Schultz et al., 1997; O'Doherty et al., 2003 & 2004; Pessiglione et al., 2006; D'Ardenne et al., 2008; Diederen et al., 2016 & 2017, Tian et al., 2016). In contrast, neurophysiological evidence in experimental animals suggests that unsigned prediction errors, which are the main focus of interest here, are coded in the cortex, including the dorsal anterior cingulate cortex (dACC) (Hayden et al., 2011), as well as subcortically, in the basolateral amygdala (Li et al., 2011); both of these structures are innervated by dopamine (Esber et al., 2012; Paus et al., 2001). Evidence from human fMRI studies also points to prefrontal representations of unsigned prediction error, especially in medial and superior aspects of the prefrontal cortex, including the dACC (Ide et al., 2013; Metereau et al., 2013; Fouragnan et al., 2017).

Given the findings of an unsigned prediction error signal in the dorsal anterior cingulate cortex, and given that this area receives dense dopaminergic projections (Lewis et al., 1987; Berger et al., 1991), we hypothesize the following. Firstly, we expect that unsigned prediction errors are coded in the cortex, specifically in the dACC. Secondly, we expect that the unsigned prediction error signal is coded relative to the precision of environmental outcomes. Third, we expect the aforementioned process to be mediated by dopamine. We test these hypotheses with novel analyses on an existing dataset (Diederen et al 2017). In this study participants were randomised to receive either placebo, or the D2 receptor antagonist sulpiride, or the dopamine receptor agonist, bromocriptine. Participants undertook a cognitive task where they were presented with rewards/ monetary outcomes on each trial that were drawn from distributions with varying degrees of uncertainty. The participants were instructed to best predict drawn rewards, with the most accurate prediction corresponding to the mean of the reward distribution. Crucially, we manipulated the reliability of reward information by including distributions with varying standard deviations. Regression analyses clearly indicated a role for unsigned prediction errors in learning. Further computational modelling confirmed that behaviour was best explained by a model that included precision-weighted signed and unsigned prediction errors that drove value updating and learning rates respectively. Functional MRI (fMRI) in conjunction with a pharmacological manipulation allowed us to explore whether dopamine perturbs the

precision-weighting of prediction errors in the brain, and whether this perturbation was associated with diminished task performance. Our previous analysis of this dataset focussed exclusively on analysis of signed prediction errors and was largely focussed on subcortical region of interest analyses (Diederen et al 2017). Specifically, these analyses showed precision-weighting of signed prediction error signals in the dopaminergic midbrain and striatum, the extent of which was modulated by D2 receptor antagonism. However, in the current paper, we are concerned primarily with cortical representations of unsigned prediction error, and their dopaminergic modulation. In a novel analysis, we find precision-weighting of unsigned prediction error signals in the superior frontal cortex and dACC, which are modulated by the dopamine D2 receptor antagonist sulpiride. The findings provide evidence in favour of recent computational models that hypothesize a role for dopamine in modulating precision-weighting of cortical prediction error signals (Bastos et al., 2012; Adams et al., 2013).

***Results***

*Environmental precision and D2 antagonism modulate performance.*

First, we explored whether performance increases when environmental precision is high (i.e. when the standard deviation of rewarding outcomes is low). We measured performance by investigating how close participants predictions were to the actual mean of the distribution in the final trials of each task block (i.e., the average difference between the mean of the reward distribution and the final three predictions the participants made). We found that performance increased when the standard deviation decreased ($F\{1,56\}=11.3$, $p=.001$). We found no interaction with group ($F\{2,56\}=1.3$, $p=.27$). However, there was a trend-level effect of group on average final performance ($F\{2,56\}=2.5$, $p=.094$). Post-hoc tests suggested reduced performance in the sulpiride group compared to Placebo ($T\{38\}=2.25$, $p=.029$), but no significant decrease in performance in the bromocriptine group ($T\{37\}=1.41$, $p=.16$; Figure 1A). However we should be cautious interpreting this result in the absence of a significant effect of group in the one-way ANOVA.

Since dopamine plays an important role in motivation, we explored whether the dopaminergic manipulation influenced measures of motivation, such as reaction times (Crespi, 1942; Niv, 2007), how far the participants scrolled the mouse in order to come to

state their prediction (Fig. 1A; please see methods section for more detail), and the number of missed trials (Fig. 1D). Using one-way ANOVAs, we found that only reaction times were significantly different across medication groups, with the sulpiride group being faster than placebo (RT: F{2,55}=4.21, *p*=.019; Misses: F{2,55}=0.33, *p*=.72; Scrolling distance: F{2,55}=0.56, *p*=.57; Fig. 1C). Differences in RT remained significant when controlling for scrolling distance: F{2,55}=3.32, *p*=.044. More rapid reactions in the sulpiride condition makes it unlikely that any performance deficit secondary to sulpiride could be driven by motivational impairments.

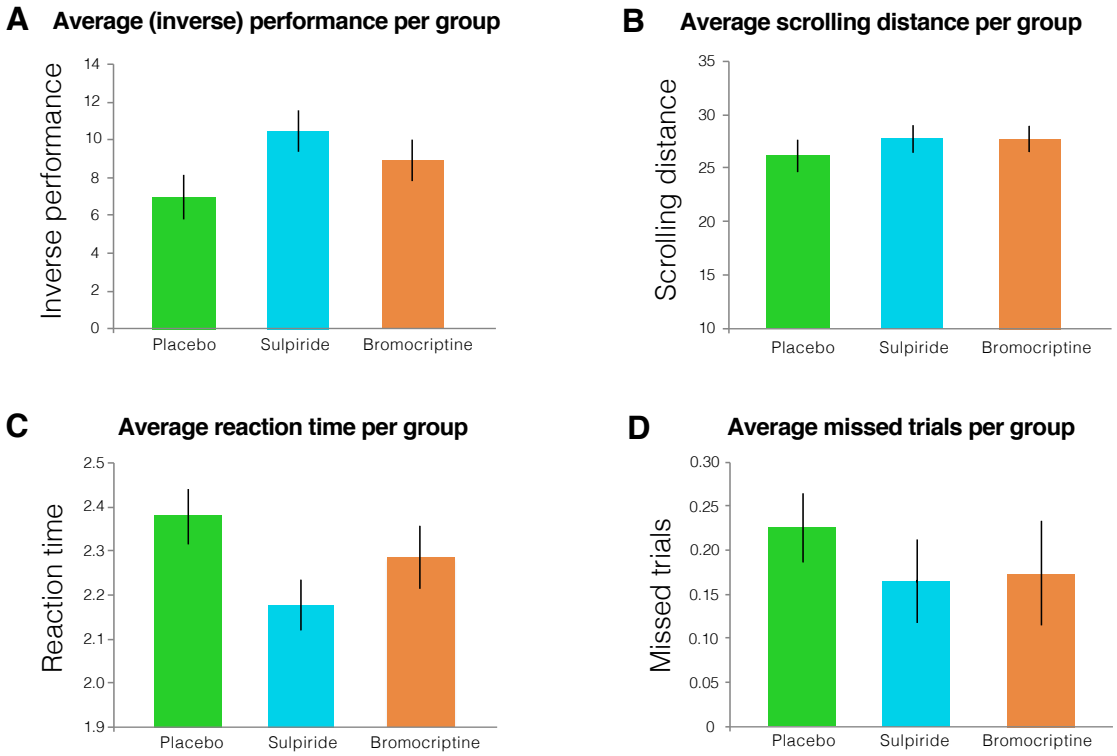


*Figure 1: A. Average (inverse) performance, B. scrolling distance, C. reaction time and D. missed trials per group. Error bars represent standard error of the mean. Participant performed worse after receiving sulpiride, and they responded significantly faster.*

*Unsigned and signed prediction errors contribute to belief updating*

In order to explore whether signed and unsigned prediction errors contribute to belief updating, we fitted a simple regression model to the behavioural data, where belief updating was the dependent variable and signed and unsigned prediction errors the independent variables. Belief updating here is defined as the change in predicted reward from $t_n$ to $t_{n+1}$. We expected a main effect of signed prediction error, but not unsigned prediction as the former informs the participants about the direction of belief updating,

whereas the latter does not. However, as reinforcement learning models like the Pearce-Hall model suggest that unsigned prediction errors increase the amount of attention devoted to a stimulus, we expect there to be an interaction between unsigned and signed prediction errors. As expected we indeed found an effect of signed prediction errors ($F\{1,10763\}=246.95$, $p<.0001$), but not unsigned prediction errors ($F\{1,10763\}=.95$, $p=.33$). Most importantly we found an interaction effect between signed and unsigned prediction errors ($F\{1,10763\}=51.7$, $p<.0001$), showing that the influence of signed prediction error on belief updating is a function of both the signed and unsigned prediction error. We furthermore tested whether SD condition had an interaction with signed prediction error on belief updating. Although there was no main-effect of SD ($F\{2,10761\}=.25$, $p=.78$) there was indeed a significant interaction ($F\{2,10761\}=25.8$, $p<.0001$). In order to visualise the interaction between signed and unsigned prediction error we binned signed prediction errors and plotted belief updating on the y-axis (Fig.2A). The significant interaction predicts a logit-function, amplifying the contribution of signed prediction error on belief updating when the unsigned prediction errors are highest (i.e. when signed prediction errors are highly positive or highly negative). We thus conclude unsigned prediction error contribute to learning in addition to signed prediction errors. We furthermore visualised the effect of SD, where we see that signed prediction errors have a stronger effect on learning in the SD5 condition compared to SD15 (Fig.2B). The significant interaction predicts a stronger effect of signed prediction error on belief updating in the SD5 condition compared to the SD15 condition. We thus conclude signed prediction errors contribute more to learning when reward information is reliable.

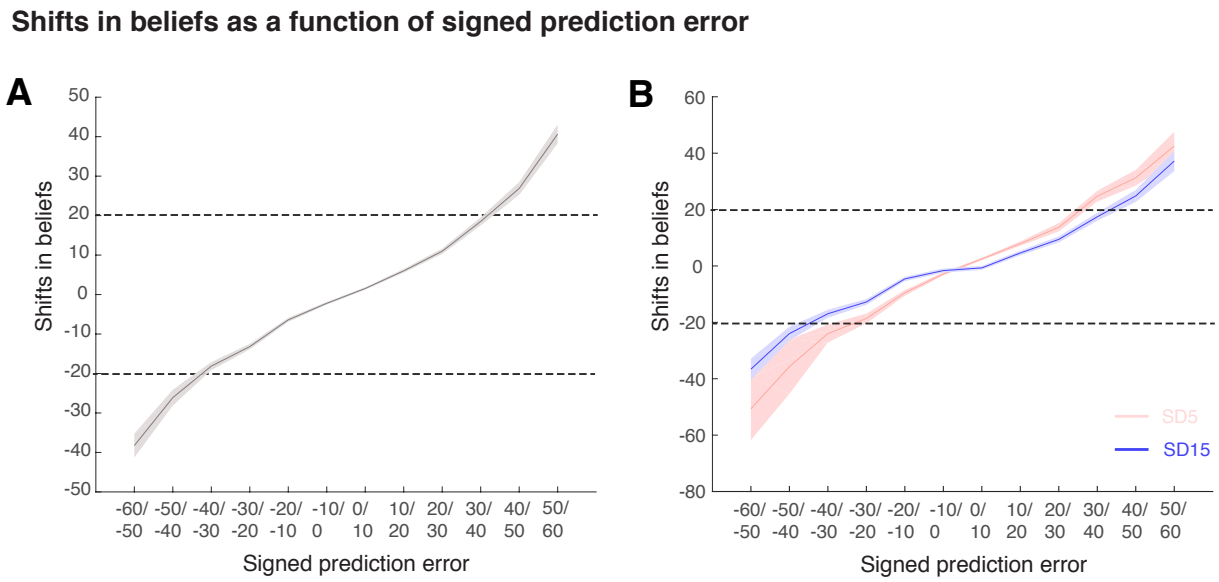


*Figure 2: Here we show the relationship between signed prediction errors (x-axis) and belief-updating (y-axis) collapsed for SD5 SD10 & SD15 (A) and separate for SD5 & SD15 (B) (SD10 is not plotted for clarity purposes). The shaded area is one standard error of the mean. In fig2A the interaction effect between unsigned prediction error and signed prediction is shown (above the top and below the bottom dotted line unsigned prediction errors are strongest, amplifying the effect of the signed prediction error). In figure B the effect of SD is shown, revealing a stronger effect of signed prediction error in the SD5 condition compared to SD15.*

*Reinforcement Learning modelling of the behavioural data supports a precision-weighted unsigned and signed prediction error*

We fitted different reinforcement learning models to the participants' trial-wise predictions to investigate whether participants were likely to make us of precision weighted unsigned and signed prediction errors. These models included a simple Rescorla-Wagner (RW, Rescorla & Wagner, 1972) model without precision weighting of prediction errors, as well as a precision weighted RW model. In addition, as previous work using similar tasks showed that learning rates tend to decrease as trials progress as a function of unsigned prediction error size in uncertain environments we included four Pearce-Hall (PH, Pearce & Hall, 1980) models that allow the learning rate to decrease as trials progress: 1) without precision weighting of prediction errors, 2) precision weighting of signed and unsigned prediction errors 3) estimated precision weighting of unsigned and signed prediction errors 4) separate estimation of precision weighting of unsigned and signed prediction errors. On average across all groups, the Pearce-Hall model with estimated precision weighted signed and unsigned prediction errors performed the best in explaining the behavioural data (model comparison using AIC to penalize models for the number of free parameters; see table 2 for an overview of the model (free) parameters and models fits and comparisons). These results

indicate that both unsigned and signed prediction errors are precision-weighted to facilitate efficient learning when the environment is uncertain/ variable.

Table 1: Overview AIC for each reinforcement learning model per medication group. The lowest AIC for each group is printed in bold.

| | Free parameters | AIC overall |
|---|---|---|
| RW | 1 | 894.7 |
| RW scaled | 1 | 903.2 |
| PH not scaled | 2 | 884.2 |
| PH both scaled | 2 | 883.6 |
| PH-estimated SD | 3 | 882.7 |
| PH-estimated SD separate for signed and unsigned PE | 4 | 879.4 |

*Unsigned prediction errors are coded in the SFC/dACC*

Our previously published analysis of these neuroimaging data focussed exclusively on signed prediction errors (Diederen et al., 2017), and was largely confined to subcortical region of interest analysis. To investigate the coding of unsigned prediction errors in the brain, here we tested for a main effect on the unsigned prediction errors parametric modulators (regressors whose magnitude is proportional to the degree of unsigned prediction error). We conducted analyses across the whole brain using a high threshold: $p_{FWE}$ <.01. Unsigned prediction errors were significantly coded in a right superior frontal cluster, spanning the superior and middle frontal gyri (peak: 27 8 54, cluster size: 145 voxels, T=8.43, Z=7.68, $p_{FWE}$<.001), a left superior frontal cluster spanning the superior and middle frontal gyri (peak: -26 0 54, cluster size: 32 voxels, T=7.66, Z=7.09, $p_{FWE}$<.001), and a dorsal anterior cingulate cortex cluster (peak: -7 12 50, cluster size: 6 voxels, T=5.32, Z=5.1 $p_{FWE}$=.002 Fig.3A). We also found a right superior occipital cluster (peak: 24 -74 34, cluster size 21 voxels, T=7.24, Z=6.75, $p_{FWE}$<.001) and two parietal clusters: (peak: 34 -48 38, cluster size: 18 voxels, T=5.69, Z=5.43, $p_{FWE}$<.001) (peak: 20 -60 26, cluster size: 4 voxels, T=5.24, Z=5.03, $p_{FWE}$=.003). We used the left and right SFC and dACC clusters as ROIs with which to take forward our analysis of dopaminergic effects in our primary analysis.

*Precision-weighting of unsigned prediction errors is mediated by dopamine in the SFC/dACC*

In order to test for an effect of outcome precision (i.e., the SD of reward distributions) and pharmacological perturbations on precision-weighting of unsigned prediction errors, we extracted the parameter estimates (betas) of the unsigned prediction error parametric modulators from the left and right superior frontal cortex (SFC) and dACC cluster that showed a main effect of unsigned prediction error coding at whole brain corrected *p*FWE<.01. Using a two-factor-mixed ANOVA with three medication groups (including placebo, sulpiride and bromocriptine), and three SD conditions (SD5, SD10, SD15) we tested first for both the left and right frontal clusters and the dACC whether there were significant interactions across SD condition and medication group. In the case of a significant effect we explored further which medication group was driving the effect of decreased precision-weighting.

First, we examined the left SFC cluster, and we found a significant interaction across SD conditions and medication group, suggesting that medication group had a significant effect on precision-weighting of unsigned prediction errors (F{2,56}=4.025 *p*=.023; Fig.3B). We then explored whether the effect was driven by differing effects of bromocriptine or sulpiride compared to placebo. In the analysis of differential effects of sulpiride and placebo, we found a significant interaction between medication group and SD condition (F{2,37} = 5.44, *p*=.025), which indicates that sulpiride dampens precision-weighting of unsigned prediction errors. Comparing the placebo and bromocriptine group, there was a significant effect of SD (F{1,37} = 14.93, *p*<.001), but no significant effect of group (F{1,37} = 2.781, *p*=.104) or interaction between medication group and SD (F{2,36} = 0.02, *p*=.894). This suggests that SFC unsigned prediction error signals are precision-weighted, but unaffected by bromocriptine.

In the right SFC we did not find a significant interaction (F{2,56}=1.70, *p*=.193; Fig. 3C), and therefore did not perform any further post-hoc tests. However, the brain response pattern is largely the same as in the left SFC (see figure 3).

In the dACC we found a “trend-level” marginally significant interaction (F{2,56}=2.81*, p*=.069; Fig. 3D). Post-hoc tests between the placebo and sulpiride group revealed a trend-level interaction (F{1,38}=3.043, *p*=.089). Testing the placebo and bromocriptine group revealed a significant effect of SD (F{1,37}=9.32, *p*=.004), but no effect of group (F{1,37}=1.17, *p*=.29)

interaction (F{1,37}=.172, *p*=.68). A similar pattern was thus found in the dACC as in the left SFC (see figure (3).

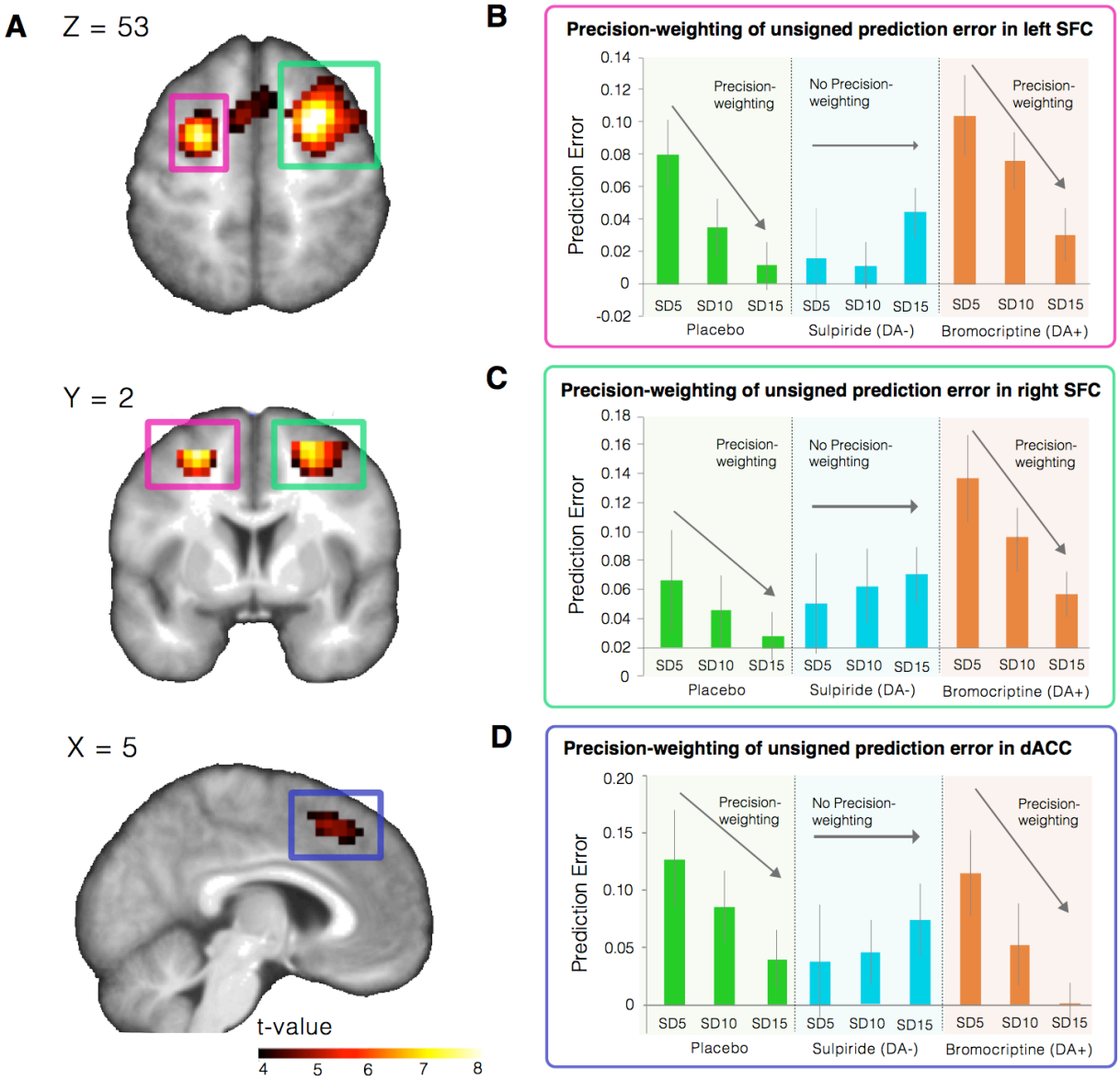


*Figure 3: A. Unsigned prediction errors were coded in a region spanning SFC/dACC. B-D. When exploring these regions further, we find that unsigned prediction errors are coded in a precision-weighted fashion as indicated by the strong unsigned prediction error signal in the SD5 condition which declines over the SD10 & SD15 condition in the placebo and bromocriptine group. Importantly, sulpiride perturbed precision-weighting in this region. Error bars represent standard error of the mean.*

*Performance correlates with the precision-weighting of unsigned prediction errors*

As neural precision-weighting is thought to facilitate task performance, we computed Spearman correlations between the degree of precision-weighting (quantified as the average parameter estimates (betas) of the unsigned prediction error in the SD5 condition

minus the SD15 condition) and performance (quantified as the mean performance error (*abs*(actual mean – predicted mean)) on the last three trials in each session. Thus a lower performance (error) error equals higher accuracy (performance)). We found that both precision-weighting in the left SFC/dACC (Rho=-.43, *p*=.001; Fig. 4A) and the right superior frontal cortex (Rho=-.369, *p*=.004; Fig. 4B) was significantly related to performance, in such a way that increased precision-weighting resulted in more accurate predictions (smaller difference between prediction and actual mean). In a general linear model controlling for medication group this effect remained significant (F{57}=7.53, *p*=.008).

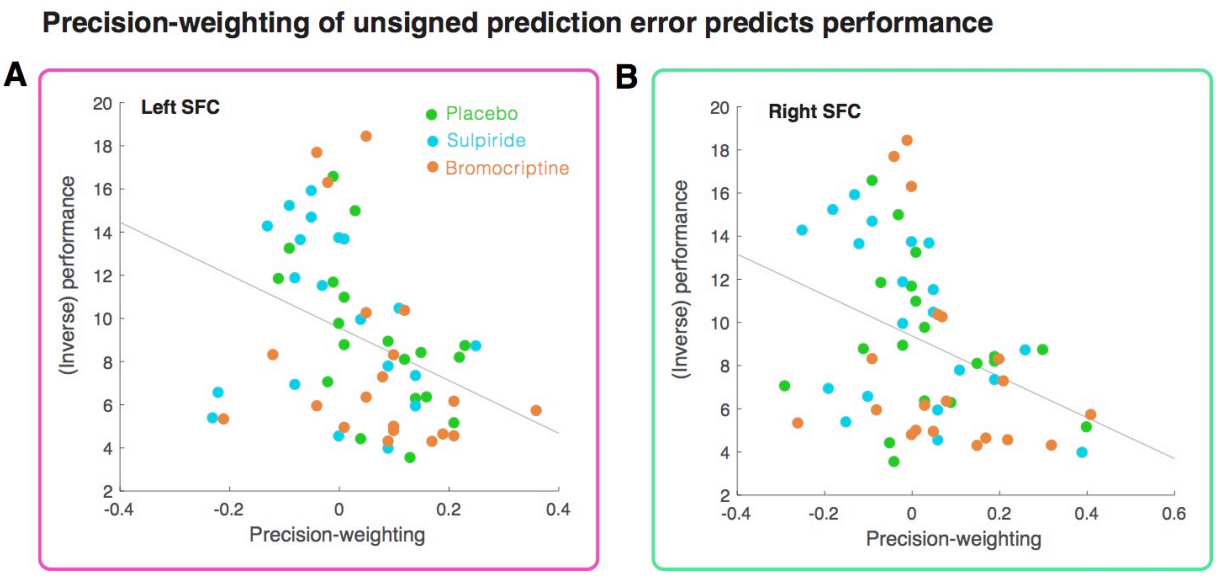


*Figure 4: Precision-weighting of unsigned prediction errors in the A. left SFC/dACC and B. right SFC/dACC predicts performance (i.e. difference between mean of the reward distribution and predicted mean) on the task.*

*Secondary outcome variable analyses*

In addition to the analyses we conducted in our primary regions of interest, i.e. SFC and dACC, we also test for an interaction between SD conditions and medication in the remainder of clusters that coded unsigned prediction errors. We first tested the superior occipital cortex cluster. Here we didn't find an interaction (F{2,56}=1.60, *p*=.21). In the bigger parietal cluster we didn't find a significant interaction either (F{2,56}=.82, *p*=.45), and neither in the smaller parietal cluster (F{2,56}=1.53, *p*=.23).

***Discussion***

We first confirmed that unsigned prediction errors play a role in learning using a regression analysis. We subsequently confirm using computational modelling on the behavioural level that a separately estimated precision-weighted signed and unsigned prediction errors provided the best description of the behavioural data, thus suggesting precision-weighted signed and unsigned prediction errors in the brain (the former confirmed previously by fMRI in humans (Diederen et al 2016), including in a prior analysis of these data (Diederen et al., 2017). In the present study, we tested the prediction of brain representation of unsigned prediction errors. We found evidence supporting this prediction, by showing that unsigned prediction errors are coded relative to the SD in an associative learning paradigm (i.e. precision-weighted) in the bilateral superior frontal cortex and dorsal anterior cingulate cortex (SFC/dACC). Importantly, this mechanism was significantly diminished in the sulpiride (D2 receptor antagonism) group in comparison to the other groups in the right SFC, with marginal evidence of a medication effect in the dACC. This result suggests that the dopamine D2 receptor plays a key role in the mechanisms underlying precision-weighting of unsigned prediction errors. We furthermore found that a decrease in precision-weighting of unsigned prediction error was significantly correlated to performance on the task, where an increase in precision-weighting resulted in more accurate predictions of upcoming rewards. These results confirm the hypothesis that there exist cortical unsigned prediction error signals, which influence performance and are precision weighted by dopamine.

The coding of unsigned prediction errors in the SFC/dACC is in line with earlier findings by Hayden (et al., 2011) who found unsigned prediction errors in the dACC of monkeys, and with prior fMRI studies in humans (Fouragnan et al., 2017; Metereau et al., 2013; Ide et al., 2013). However, the finding that cortical unsigned prediction error signals are precision-weighted in humans is a novel finding to the best of our knowledge. The fact that such a precision weighted signal exists in the dACC is notable given that, aside from the motor cortex, dopaminergic innervation of cortex is greatest in ACC (Lewis et al., 1987; Berger et al., 1991; Paus, 2001), compatible with the hypothesis that precision weighting is influenced here by dopaminergic input.

The observation that dopamine plays a role in the precision-weighting of unsigned prediction errors extends previous work showing the precision-weighting of signed (reward)

prediction errors (Diederen et al., 2016 & 2017), and fits well with contemporary computational neuroscience models which suggest that the brain is a hierarchically organized system which aims to minimize prediction error. According to these accounts, neural modulatory systems like dopamine play a key role in controlling the precision of prediction errors (Bastos et al., 2012; Adams et al., 2013; Friston et al., 2014a,b). Whilst our sulpiride results accord with such models, we did not find evidence that bromocriptine modulated precision-weighting of prediction error signals. However, in both the left and the right SFC, there is an overall amplification of the prediction error signal in the bromocriptine condition. This effect wasn't significant in the left SFC. However, in the right SFC there was a significant main effect of group ($p$=.032), and a direct comparison between placebo and bromocriptine did reveal a significant effect of group ($p$=.013). This effect is likely due to dopaminergic effects on the excitability of neuronal populations irrespective of environmental precision (Wang et al., 2001). However, these results weren't the main focus of the paper and should therefore be interpreted with caution.

If the precision-weighting of prediction errors is important in learning, we can expect aberrant learning to occur when prediction errors are not scaled optimally to the precision of information. This mechanism could be of importance to psychosis, which is characterized by delusional beliefs and hallucinatory perception (Fletcher & Frith et al., 2009). Previous work showed aberrant prediction error coding in people with psychosis (Murray et al., 2008; Corlett et al., 2007a,b). As psychosis has consistently been associated with dopamine dysfunction (Howes et al., 2009), it is conceivable that the precision-weighting process is impaired in psychosis. Indeed, it has been suggested that dopamine dysregulation causes psychosis due to affecting the brains capacity to precision-weight prediction error (Adams et al., 2013). That is, if unreliable prediction errors were given excessive weight, they could have an exaggerated influence on driving changes in the brain's model of the world, thereby contributing to the formation of abnormal beliefs. As such, future studies should aim to explore the relationship between precision-weighting of prediction error and psychosis.

In conclusion, we found evidence of precision-weighted unsigned prediction errors in the superior frontal and dorsal anterior cingulate cortices. Furthermore, we found that the precision-weighting of prediction errors was modulated by the dopaminergic antagonist

sulpiride, and we found that the degree of precision-weighting in this area was correlated to performance on the task, providing evidence for the first time that dopamine plays a role in precision-weighting of unsigned prediction error signals during learning.


## *Acknowledgements*

We would like to thank Wolfram Schultz for his important role in designing the experiment and acquiring the data analysed here, and to the staff of the Clinical Research Facility and Wolfson Brain Imaging Centre, Cambridge Biomedical Research Centre for assistance in data collection.

Author roles: J.H.: Conceptualization, methodology, formal analysis, writing (original draft preparation, review and editing); P.C.F.: Conceptualization, project administration, funding acquisition, writing (review and editing). T.S. and H.Z. investigation, project administration, writing (review and editing). K.M.J.D. Conceptualization, methodology, project administration, implementation, formal analysis, supervision, writing (original draft preparation, review and editing) G.K.M., conceptualization, project administration, methodology, supervision, writing (original draft preparation, review and editing)

Conflicts of Interest: P.C.F. has received payments in the past for ad hoc consultancy services to GlaxoSmithKline All other authors declare no competing interests.

Funding: This work was supported by the NSPN Welcome Trust Strategic Grant to J.H., Wellcome Trust to P.C.F. and W.S., Bernard Wolfe Health Neuroscience Fund to P.C.F. and H.Z., and the Niels Stensen Foundation to K.M.J.D. and the NSPN Welcome Trust.

***Methods***

*Demographics*

| | N | Mean Age (SD) | Male (Female) |
|---|---|---|---|
| **Medication study** | 59 | 24.19 (4.46) | 31 (28) |
| *Placebo* | *20* | *23.9 (4.76)* | *9 (11)* |
| *Sulpiride* | *20* | *24.8 (4.51)* | *12 (8)* |
| *Bromocriptine* | *19* | *23.7 (4.3)* | *10 (9)* |

*Participants*

We recruited 63 participants, of which 4 were excluded due to feeling unwell during scanning and/or a failure to complete the paradigm, or side-effects of the medication (see Diederen et al., 2017 for a detailed description). All participants were recruited via the distribution of flyers in Cambridge and advertisements on Gumtree. Drug screenings were all negative. After receiving detailed information about the study, all participants gave written informed consent.

Table 2: Participants

| | N Sessions | N Trials | SDs | EVs | N participants |
|---|---|---|---|---|---|
| **Dopamine Study** | 3 | 62 | 5, 10 & 15 | 35 & 65 | 60 (20 per group) |

*Pharmacological dopaminergic challenge*

Participants received a single dose of the D2-antagonist sulpiride (600mg), the dopamine agonist Bromocriptine (2.5 mg), or placebo, in a double-blind fashion. Domperidone was added to prevent potential nausea, as it would be an indicative factor of having taken a drug, and could thus influence the blinding procedure, as well as making people feel unwell (see Diederen et al., 2017 for a detailed description of the dopaminergic manipulation).

*Task design*

Durin functional (f)MRI data acquisition, participants predicted the magnitude of upcoming rewards. Optimal performance on this task required the participant to estimate the mean of the distribution from which the rewards were drawn. The task consisted of three sessions of 10 minutes each. The rewards were drawn from six different pseudo-Gaussian distributions that differed with respect to their standard deviation (SD) and expected value (i.e. mean of the distribution; EV). See Table 1 for an overview of the SD's and EV's. Distributions were counterbalanced to ensure that the two conditions within each session differed with respect to the EV and SD. The conditions were presented in short blocks, each including 4-6 trials. Each distribution consisted of 21 trials, which resulted in a total number of trials of 42 per block. The participants were informed beforehand that each distribution (of which two per session) had a different level of precision, which could be one of three levels: low, medium or high precision, corresponding to SD5, 10 & 15, although the exact SD's were not revealed to the participants. Each trial started with a fixation cross that was presented between 2100 and 4200ms. The fixation cross was followed by a cue that was presented for 500ms. This cue informed the participant from which of two distributions (high, medium or low) in that block the upcoming reward was being drawn. After the cue was presented the participants were required to predict the magnitude of the upcoming reward. They had 3500ms to make this prediction. Making the prediction involved scrolling a mouse ball with their fingers up and down, and clicking the left mouse button to state their prediction. This resulted in moving a bar up and down the screen on a scale that ranged from 0 to 100. The starting point on the bar was randomized throughout the experiment, to ensure that scrolling distance did not correlate with participants predictions. The participants were instructed to minimize the prediction error, i.e. the difference between the expected and the actual obtained reward. After the prediction, another fixation cross was presented for a duration between 2100 and 5250ms. Thereafter, the obtained reward was presented in addition to the prediction and the reward prediction error for 1000ms (see figure 5 for an example trial).

Pay-off

If the participants would always receive the amount of money drawn by the computer, this might have reduced their motivation to reveal their true prediction. Therefore, the participants were rewarded according to their accuracy (i.e., proximity to the mean of the

reward distribution) on 20% of the trials in order to keep them motivated to predict as well as possible.

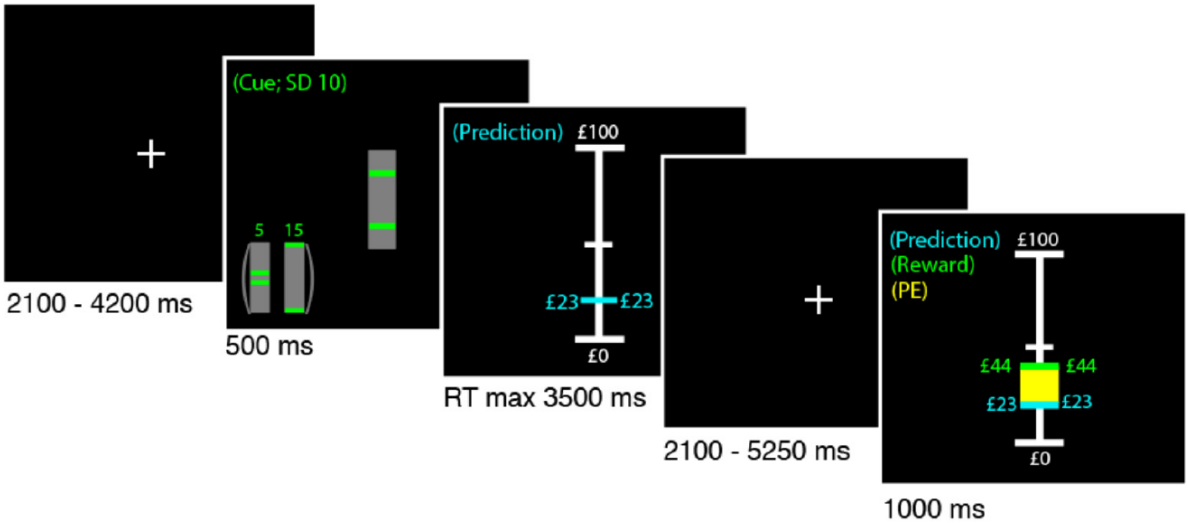


*Figure 5: Example of a trial. The participants were instructed to learn the mean of a reward distribution. First a fixation cross was presented after which the participants were informed about the standard deviation of the reward distribution. Subsequently the participants were asked to make a prediction regarding the upcoming reward, which was presented to the participant in combination with the prediction error (in yellow) after an anticipation period.*

*Behavioural analyses*

Behavioural data was analysed in Matlab. In order to test the performance of the participants, we analysed whether the predictions of the participants approached the mean of the distribution (or expected value (EV)). We calculated the absolute difference between the EV and the prediction, which we refer to as the (inverse) performance. We reasoned that if participants' prediction became closer to the EV of the reward distribution from which rewards were drawn, the participants were learning. We tested this hypothesis by computing a paired t-test on the performance at the start of a task block (condition), and at the end of that block for each of the six conditions. In addition, we analysed differences in scrolling distance to exclude the possibility that drug effects influenced willingness to do the task, resulting in a decrease in scrolling distance. We also analysed differences in reaction times and missed trials for the same reason.

*Linear regression model*

In order to explore the contribution of signed and unsigned prediction error to learning we used a linear regression model with belief-updating as dependent variable and both

unsigned and signed prediction error as independent variables. Belief updating here is defined as the change in value estimate from $t_n$ to $t_{n+1}$. That is, when the participant expects a value of 20 on $t_1$ and expect as value of 30 on $t_2$ there was a belief update of 10, which is expected to be explained by a prediction error experienced at $t_1$. Since signed prediction errors, but not unsigned prediction errors inform the participant about the direction of learning, we expect the former but not the latter to have a main effect on belief-updating. Furthermore, since formal learning models like the Pearce-Hall model (Pearce & Hall, 1980), suggest that unsigned prediction errors increase learning, we expect an interaction between unsigned and signed prediction error on belief-updating. In order to visualise the effect of signed and unsigned prediction errors on belief-updating, we will create bins on the basis of signed prediction errors and plot belief-updating on the y-axis. If unsigned prediction errors interact with signed prediction error in predicting belief-updating we should expect to see a logit or sigmoid relationship between signed prediction error and belief updating.

*Behavioural modelling*

To investigate learning, we fitted several reinforcement learning models to participants' prediction sequences. Each model used a common updating rule in which predictions on trial *n* depended on the prediction error and the learning rate on trial *n-1:*

$$\mathbf{y_n = y_{n-1} + k_n\, \delta_n} \qquad \textbf{Equation 1}$$

Here, $y_n$ is the prediction made on trial *n*, $k_n$ refers to the learning rate and $\delta_n$ denotes the size of the prediction error. This is a standard reinforcement learning model, and allows to estimate to which degree prediction errors are being used by the participant by estimating the learning rate parameter (Sutton & Barto et al., 1998).

The first model consisted of a Rescorla-Wagner (RW) reinforcement learning model with a fixed learning rate. A fixed learning rate prescribes that each prediction error is weighted equally during learning (e.g., independent of whether a prediction error occurred at the start, or the end of a session):

$$\mathbf{k_n = \alpha} \qquad \textbf{Equation 2}$$

The second model consisted of a Pearce-Hall (PH) model with a trial-wise, dynamic, learning rate, which prescribes higher weighting of prediction errors (i.e., more learning) at the start of a task session compared to later trials (Pearce et al., 1980). In uncertain environments, it is more optimal to decrease the weighting of prediction errors as learning progresses (once participants become more certain of their predictions) as prediction errors will continue to occur as a result of the imposed uncertainty. The PH learning rate decays as trial progress, as a function of the previous learning rate and the experienced prediction error, which allows for fast updating of predictions at the beginning of each task session:

$$k_n = \gamma C|\delta_{n-1}| + (1 - \gamma)\, k_{n-1} \qquad \textbf{Equation 3}$$

in which $|\delta|$ is the absolute prediction error, C is an arbitrary scaling coefficient and $\gamma$ is the learning rate decay.

We additionally explored whether scaling prediction error to the reliability of the environment (i.e., precision –weighting) benefitted learning by comparing models that scaled the prediction error term and models that did not. We both explored models that simply precision-weighted the prediction error by a constant term $\omega$ (equation 4&5), and models that estimated the degree $\upsilon$ of precision-weighting for each individual (equation 6). We estimated both models that had a single estimated $\upsilon$ for both signed and unsigned prediction errors and a model that had a separate variable for signed and unsigned prediction errors. In contrast to earlier studies (Diederen et al., 2015, 2016 & 2017) we used a more simplified model that utilised a linear precision-weighted model to ease interpretation.

$$y_n = y_{n-1} + k_n\, (\delta/\, \omega)_n \qquad \textbf{Equation 4}$$

$$k_n = \gamma C|(\delta_{n-1}/\, \omega)| + (1 - \gamma)\, k_{n-1} \qquad \textbf{Equation 5}$$

$$\omega = 1 - \upsilon + \upsilon * (SD) \qquad \textbf{Equation 6}$$

For the Pearce-Hall model we explored the scaling of the signed and unsigned prediction error separately, as well as combined.

We used the Akaike Information Criterion (AIC) to estimate the model evidence in favour of each of the two models. The model with the lowest AIC was used for further analyses.

*fMRI pre-processing steps*

All pre-processing steps were performed using SPM8 (available at http://www.fil.ion.ucl.ac.uk; Wellcome Department of Cognitive Neurology, London, England) in combination with the Donders matlab (dmb) toolbox for combining data from different echo acquisition times. Data was acquired at 4 different echo times (TE's) of 12, 27,91, 43,82 and 59,73 ms. Corrections for slice timing was not applied considering the relatively fast repetition times (TR) of 2100ms. Using SPM8 the functional images were realigned first, after which the four different echoes were summed and averaged in order to minimize signal intensity inhomogeneity due to differences in T2* relaxation times across the brain (Poser et al., 2006). The functional images were subsequently coregistered to the T1-weighted anatomical image in native-space by first coregistering one averaged functional image, after which the coregistration parameters of this registration were applied to the all the functional images in order to registered to native space. Unified segmentation was used in order to achieve the normalization of the functional and anatomical images to MNI space (Ashburner et al., 2005). Unified segmentation is a single iterative model that combines segmentation, bias correction and normalization in order to achieve optimal results. The segmentation step was omitted for one of the participants; since it gave erroneous segmentation parameters resulting in distortions in the normalized anatomical T1 scan. The segmentation step was later conducted using a different template, namely the East-Asian template, which resulted in a correct segmentation. Spatial smoothing was performed using 8mm Gaussian kernels. The time series in each session were high-pas filtered at 128 Hz.

*1st-level analysis*

fMRI analyses were conducted using SPM8. We used a single statistical linear regression model for our main analyses. We modelled for each trial the onset of the cue as an event (i.e. a delta function of zero duration), the onset of the prediction event (i.e., when participant could start making their prediction) as a single epoch lasting until they indicated their prediction. Each event was convolved by the standard canonical haemodynamic response function in SPM8. Importantly, we used parametric modulation to identify neural

correlates of unsigned prediction error responses by specifying for all outcome events the reward prediction errors. The reward events were separately modelled for the different SD conditions in order to test for differences in precision-weighting of prediction errors as evidences by different sizes of slopes for the coding of unsigned prediction errors under different levels of certainty. Contrast were created on the 1$^{st}$-level which combined the two EV conditions for each SD in order to leave a single contrast for each SD conditions which could be taken to the 2$^{nd}$-level (i.e., group-level).

*2$^{nd}$-level analysis*

For the analysis of the main effect of unsigned prediction errors we first created 1$^{st}$-level contrast in which we combined the 2 EV's for each separate SD condition, resulting in 3 contrasts corresponding to the 3 different SD conditions (SD5, SD10 & SD15). We modelled these 3 SD contrasts for each pharmacological condition separately specifying the effects to be dependent, and having equal variance. We subsequently tested for a main effect of prediction error by doing a positive 2$^{nd}$-level contrast on the parametric modulators for the unsigned prediction error.

To test for an effect of SD on the coding of unsigned prediction error and the effect of pharmacological group, we extracted the parameter estimates (beta's) for all SD conditions for each medication group separately from the left and right superior frontal cortex clusters and dACC cluster surviving FWE<.01 coding unsigned prediction errors; for each cluster we used the mean values per cluster per participants in a mixed-2-factor-ANOVA with a 3-level within subject factor (SD5, SD10, SD15) and a 3-level between subject factor (placebo vs sulpiride, or placebo vs bromocriptine), and tested for main effect of SD and interaction between SD conditions and medication using a linear contrast in SPSS (version 21). When a significant interaction was found we explored this further to see which group was driving the interaction.

**References**


Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., & Friston, K. J. (2013). The computational anatomy of psychosis. *Frontiers in Psychiatry*, *4*(May), 47. http://doi.org/10.3389/fpsyt.2013.00047

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *Neuroimage*, *26*(3), 839-851.

Bar, M. (2009). Predictions: a universal principle in the operation of the human brain.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695-711.

Berger, B., Gaspar, P., & Verney, C. (1991). Dopaminergic innervation of the cerebral cortex: unexpected differences between rodents and primates. *Trends in neurosciences*, *14*(1), 21-27.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. http://doi.org/10.1017/S0140525X12000477

Clark, A. (2015). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world, *10*(7). http://doi.org/10.1016/j.tics.2006.05.004

Corlett, P. R., Murray, G. K., Honey, G. D., Aitken, M. R. F., Shanks, D. R., Robbins, T. W., … Fletcher, P. C. (2007). Disrupted prediction-error signal in psychosis: Evidence for an associative account of delusions. *Brain*, *130*(9), 2387–2400. http://doi.org/10.1093/brain/awm173

Corlett, P. R., Honey, G. D., & Fletcher, P. C. (2007). From prediction error to psychosis: ketamine as a pharmacological model of delusions. *Journal of Psychopharmacology (Oxford, England)*, *21*(3), 238–252. http://doi.org/10.1177/0269881107077716

Crespi LP: Quantatative variation of incentive and performance in the white rat. American Journal of Psychology. 1942, 55: 467-517. 10.2307/1417120.

D'ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, *319*(5867), 1264-1267.

Den Ouden, H. E., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2008). A dual role for prediction error in associative learning. *Cerebral cortex*, *19*(5), 1175-1185.

Den Ouden, H. E., Kok, P., & De Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in psychology*, *3*.

Diederen, K. M. M. J., Spencer, T., Vestergaard, M. D. D., Fletcher, P. C. C., & Schultz, W. (2016). Adaptive Prediction Error Coding in the Human Midbrain and Striatum Facilitates Behavioral Adaptation and Learning Efficiency. *Neuron*, *90*(5), 1127–1138. http://doi.org/10.1016/j.neuron.2016.04.019

Diederen, K. M. J., & Schultz, W. (2015). Scaling prediction errors to reward variability benefits error-driven learning in humans. *Journal of Neurophysiology*, *114*(3), 1628–1640. http://doi.org/10.1152/jn.00483.2015

Diederen, K. M. J., Ziauddeen, H., Vestergaard, M. D., Spencer, T., Schultz, W., & Fletcher, P. C. (2017). Dopamine Modulates Adaptive Prediction Error Coding in the Human Midbrain and Striatum. *Journal of Neuroscience*, *37*(7), 1708–1720. http://doi.org/10.1523/JNEUROSCI.1979-16.2016

Esber, G. R., Roesch, M. R., Bali, S., Trageser, J., Bissonette, G. B., Puche, A. C., ... & Schoenbaum, G. (2012). Attention-related Pearce-Kaye-Hall signals in basolateral amygdala require the midbrain dopaminergic system. *Biological psychiatry*, *72*(12), 1012-1019.

Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews. Neuroscience*, *10*(1), 48–58. http://doi.org/10.1038/nrn2536

Fouragnan, E., Queirazza, F., Retzler, C., Mullinger, K. J., & Philiastides, M. G. (2017). Spatiotemporal neural characterization of prediction error valence and surprise during reward learning in humans, (May), 1–18. http://doi.org/10.1038/s41598-017-04507-w

Friston, K. (2009). The free-energy principle: a rough guide to the brain?. *Trends in cognitive sciences*, *13*(7), 293-301.

Friston, K. J., Stephan, K. E., Montague, R., & Dolan, R. J. (2014). Computational psychiatry: The brain as a phantastic organ. *The Lancet Psychiatry*. http://doi.org/10.1016/S2215-0366(14)70275-5

Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., & Dolan, R. J. (2014). The anatomy of choice: dopamine and decision-making. *Phil. Trans. R. Soc. B*, *369*(1655), 20130481.

Function, M., & Wiley, J. (1991). Dopaminergic innervation of the cerebral cortex: unexpeded differences between rodents and primates, (1966), 35–39.

Hall, G., & Pearce, J. M. (1982). Restoring the associability of a pre-exposed CS by a surprising event. *The Quarterly Journal of Experimental Psychology Section B*, *34*(3b), 127-140.

Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., & Platt, M. L. (2011). Surprise Signals in Anterior Cingulate Cortex : Neuronal Encoding of Unsigned Reward Prediction Errors Driving Adjustment in Behavior, *31*(11), 4178–4187. http://doi.org/10.1523/JNEUROSCI.4652-10.2011

Hohwy, J. (2013). *The predictive mind*. Oxford University Press.

Howes, O. D., & Kapur, S. (2009). The dopamine hypothesis of schizophrenia: version III—the final common pathway. *Schizophrenia bulletin*, *35*(3), 549-562.

IBM Corp. Released 2012. IBM SPSS Statistics for Macintosh, Version 21.0. Armonk, NY: IBM Corp.

Ide, J. S., Shenoy, P., Yu, A. J., & Li, C. R. (2013). Bayesian Prediction and Evaluation in the Anterior Cingulate Cortex, *33*(5), 2039–2047. http://doi.org/10.1523/JNEUROSCI.2201-12.2013

Lewis, D. A., Campbell, M. J., Foote, S. L., Goldstein, M., & Morrison, J. H. (1987). The distribution of tyrosine hydroxylase-immunoreactive fibers in primate neocortex is widespread but regionally specific. *Journal of Neuroscience*, *7*(1), 279-290.

Li, J., Schiller, D., Schoenbaum, G., Phelps, E. A., & Daw, N. D. (2011). Differential roles of human striatum and amygdala in associative learning. *Nature neuroscience*, *14*(10), 1250.

Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A Bayesian foundation for individual learning under uncertainty. *Frontiers in human neuroscience*, *5*, 39.

Metereau, E., Dreher, J., & Recherche, M. De. (2018). Cerebral Correlates of Salient Prediction Error for Different Rewards and Punishments, (February 2013), 477–487. http://doi.org/10.1093/cercor/bhs037

Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, a D., Honey, G., … Fletcher, P. C. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry*, *13*(3), 239, 267–276. http://doi.org/10.1038/sj.mp.4002058

Niv Y: Cost, benefit, tonic, phasic: what do response rates tell us about dopamine and motivation?. Ann N Y Acad Sci. 2007, 1104: 357-376. 10.1196/annals.1390.018.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329-337.

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*(5669), 452-454.

Paus, T. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nature reviews neuroscience*, *2*(6).

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review*, *87*(6), 532.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042.

Poser, B. a., Versluis, M. J., Hoogduin, J. M., & Norris, D. G. (2006). BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. *Magnetic Resonance in Medicine*, *55*(6), 1227–1235. http://doi.org/10.1002/mrm.20900

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. http://doi.org/10.1038/4580

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, *2*, 64-99.

Roesch, M. R., Esber, G. R., Li, J., Daw, N. D., & Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce–Hall and Rescorla–Wagner coexist within the brain. *European Journal of Neuroscience*, *35*(7), 1190-1200.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593-1599.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine learning*, *3*(1), 9-44.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1, No. 1). Cambridge: MIT press.

Tian, J., Huang, R., Cohen, J. Y., Osakada, F., Kobak, D., Machens, C. K., ... & Watabe-Uchida, M. (2016). Distributed and mixed information in monosynaptic inputs to dopamine neurons. *Neuron*, *91*(6), 1374-1389.

Wang, J., & O'donnell, P. (2001). D1 dopamine receptors potentiate NMDA-mediated excitability increase in layer V prefrontal cortical pyramidal neurons. *Cerebral Cortex*, *11*(5), 452-462.