

# A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex

Gerhard Jocham<sup>1</sup>, Laurence T Hunt<sup>1,2</sup>, Jamie Near<sup>1</sup> & Timothy E J Behrens<sup>1,3</sup>

**Although the ventromedial prefrontal cortex (vmPFC) has long been implicated in reward-guided decision making, its exact role in this process has remained an unresolved issue. Here we show that, in accordance with models of decision making, vmPFC concentrations of GABA and glutamate in human volunteers predict both behavioral performance and the dynamics of a neural value comparison signal. These data provide evidence for a neural competition mechanism in vmPFC that supports value-guided choice.**

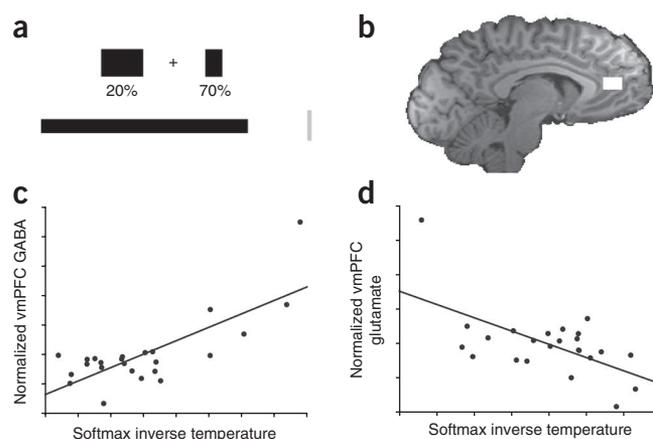
Organisms are constantly required to choose between options that differ in terms of their expected reward values. Although neural signals reflecting these values are widespread throughout the brain<sup>1–5</sup>, the vmPFC has attracted particular interest. Neural signals recorded in this region seem to indicate choice<sup>1,5</sup>, and damage to these areas in humans leads to deficits in decision making<sup>6</sup>. However, the nature of the computations underlying these decisions has remained elusive.

One popular mechanism for decision making is competition by mutual inhibition; this comprises a class of models in which representations of each available option inhibit one another until activity remains in only one. Such models can be implemented in abstract<sup>7</sup> or biophysical<sup>8</sup> forms, and dynamic neural signals consistent with these models can be found in the vmPFC<sup>9</sup>. A crucial prediction is that performance will depend heavily on the degree of inhibition relative to excitation in the network. If the vmPFC implements a decision-making process based on such an inhibitory competition mechanism, then both behavioral performance and the neural dynamics of the vmPFC value comparison signal should depend on the levels of the major excitatory and inhibitory neurotransmitters, glutamate and GABA, in the vmPFC. Such a finding would tie a neurochemical underpinning to our computational understanding of value-guided choice and provide a mechanistic explanation for interindividual variability in choice behavior.

We tested these predictions by using magnetic resonance spectroscopy (MRS) to obtain measures of each individual's basal GABA and glutamate concentrations from the vmPFC and from a right parietal region in the intraparietal sulcus (IPS) of 25 healthy male volunteers (Online Methods). Note that these neurochemical data are not time resolved

or choice related. They reflect the baseline neurotransmitter concentrations in each subject at rest, and we collected each from a single voxel in the vmPFC and a single voxel in the IPS. We chose the IPS as a control region because it has also been shown to encode value- and decision-related parameters in a number of studies<sup>2,10</sup>. Time constraints and methodological considerations precluded us from using further control regions such as the lateral orbitofrontal cortex (Online Methods). After MRS acquisition, subjects underwent a short version of a reward-guided decision-making task (Fig. 1 and Supplementary Fig. 1) during scanning with functional magnetic resonance imaging (fMRI). After scanning, participants completed a longer version of the decision-making task. During the task, subjects repeatedly made choices between two options of differing reward magnitude and reward probability (Fig. 1a).

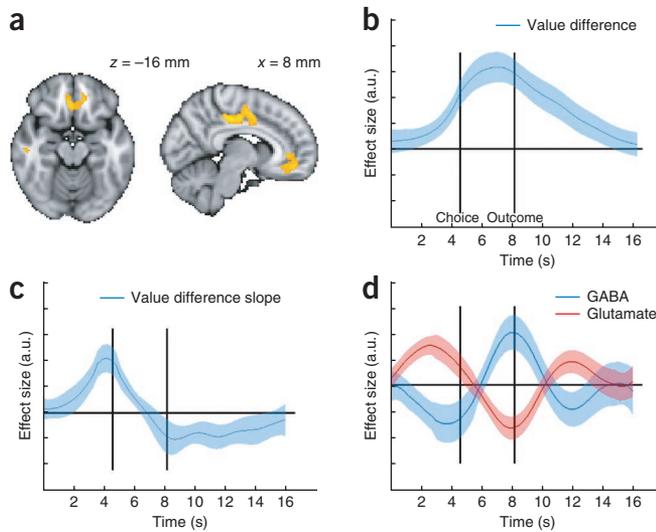
Subject performance in this task can be characterized by using a standard prospect theory model (Online Methods). Whereas statistically optimal behavior on the task would be to multiply magnitude and probability and to choose the option with the highest Pascalian value, the model has two parameters that warp probability and reward space to match subject behavior. A third, crucial parameter (the softmax inverse temperature,  $\tau$ ), indicates the accuracy of subject decisions. Subjects with a low  $\tau$  value require a substantial value difference to select the option with the higher subjective value reliably (Supplementary Fig. 1d). If decisions are made by mutual inhibition,



**Figure 1** Experimental task and correlation of spectroscopy data with behavior. (a) Example trial task schematic. The width of the bars represents reward magnitudes, and the percentages underneath specify the reward probability. The black bar at the bottom represents participant's cumulative earnings, and the gray bar at right is a target that volunteers try to reach. (b) Location of the MRS voxel in the vmPFC. (c, d) Performance was best in subjects with high GABA concentrations (c) and low glutamate concentrations (d). Values represent arbitrary units, as glutamate (c) and GABA (d), respectively, were regressed out of both variables to show effects that are orthogonal with respect to the other neurotransmitter.

<sup>1</sup>Oxford Centre for Functional MRI of the Brain (FMRIB Centre), University of Oxford, John Radcliffe Hospital, Oxford, UK. <sup>2</sup>Department of Experimental Psychology, University of Oxford, Oxford, UK. <sup>3</sup>Wellcome Trust Centre for Neuroimaging, University College London, London, UK. Correspondence should be addressed to G.J. (g.jocham@fmrrib.ox.ac.uk).

Received 21 February; accepted 17 May; published online 17 June 2012; doi:10.1038/nn.3140



**Figure 2** Relationship between GABA and glutamate and value difference slope. (a) Whole-brain analysis showing the effects of value difference. (b) Time-resolved regression of value difference against the fMRI signal from the vmPFC. (c) Slope (temporal derivative) of the value-difference signal in b. In b and c, the solid line represents the mean across subjects, the shaded area indicates the s.e.m. (d) Regression of vmPFC concentrations of GABA and glutamate against the slope of the vmPFC value difference signal (c) at each time point in a trial. Solid lines represent regression coefficients, shaded area represents the s.e.m. High concentrations of glutamate and low concentrations of GABA lead to faster ramping up of the value difference signal. a.u., arbitrary units.

$\tau$  should increase with higher GABA and lower glutamate concentrations (see **Supplementary Fig. 2** for this relation in one instantiation of one mutual-inhibition model<sup>9</sup>).

In the decision-making task, GABA and glutamate concentrations indeed predicted average choice accuracy  $\tau$ . Because GABA is synthesized from glutamate, the concentrations of the two neurotransmitters are highly correlated within each brain region. It is therefore crucial to orthogonalize all effects with respect to the other neurotransmitter<sup>11</sup>; that is, it is necessary to compute partial correlations. By doing so, we found that vmPFC GABA and glutamate had opposing effects:  $\tau$  was highest in subjects with high GABA concentrations ( $r = 0.76$ ,  $P < 0.00001$ ) and low glutamate concentrations ( $r = -0.598$ ,  $P = 0.001$ ; **Fig. 1**). This pattern was specific to the vmPFC, as no such relation was found with IPS GABA ( $r = 0.25$ ,  $P = 0.13$ ) and glutamate levels ( $r = -0.29$ ,  $P = 0.1$ ). Furthermore, when the two pairs of neurotransmitter concentrations were formally compared in a single linear model, vmPFC GABA had a significantly greater positive effect than IPS GABA ( $t = 2.05$ ,  $P = 0.027$ ), and vmPFC glutamate had a significantly greater negative effect than IPS glutamate ( $t = 2.29$ ,  $P = 0.017$ ).

In agreement with previous studies<sup>5</sup>, blood oxygenation level-dependent (BOLD) activity in the vmPFC recorded in the fMRI session correlated positively with the value difference between chosen and unchosen options on each trial (**Fig. 2a,b**). If the evolution of this value-difference signal is indeed dependent on a balance between mutual inhibition and recurrent excitation, then it should ramp up faster and ramp down earlier in subjects with relatively high ratios of excitation to inhibition. We therefore computed the temporal derivative of the value-difference signal throughout the trial (**Fig. 2c**) and examined its correlation with vmPFC GABA and glutamate concentrations. Early in the trial, individuals with high glutamate ( $t = 4.96$ ,  $P < 0.00005$ ; **Fig. 2d**) and low GABA ( $t = -3.05$ ,  $P = 0.0029$ ) concentrations showed higher derivatives, indicating faster ramping up. Late in the trial, the

same individuals had the most negative derivatives, indicating faster ramping down (GABA:  $t = 3.42$ ,  $P = 0.0012$ ; glutamate:  $t = -4.25$ ,  $P < 0.0002$ ). Notably, this effect was specific to the value-difference correlation; no such pattern was found in the raw BOLD signal, thus making it unlikely that our finding is due to a general effect on the BOLD signal (**Supplementary Fig. 3**).

We have shown that interindividual variability in vmPFC GABA and glutamate concentrations explains variability in choice behavior and in fMRI signals recorded during value-guided choice. Taken together, these findings indicate that value-guided choice is governed by a competition by mutual inhibition that is mediated by a balance between GABAergic inhibition and glutamatergic excitation in the vmPFC. We have recently shown using magnetoencephalography that vmPFC shows dynamics predicted by neural competition<sup>9</sup>. The findings presented here further support the idea that vmPFC has a central role not only in valuation<sup>12</sup> but also in choice<sup>13</sup>. However, in showing that this competitive process is predictably dependent on GABA and glutamate concentrations, our findings provide a clear link from neurochemical to computational mechanisms and thus to economic choice. Such an understanding of neurochemical mechanisms has potential clinical relevance. For example, it is noteworthy that altered prefrontal levels of GABA and glutamate have been reported in individuals with major depressive disorder<sup>14</sup>, a condition that has impairments in decision making as one of its diagnostic criteria.

## METHODS

Methods and any associated references are available in the online version of the paper.

*Note: Supplementary information is available in the online version of the paper.*

## ACKNOWLEDGMENTS

This work was supported by a Wellcome Trust Research Career Development fellowship to T.E.J.B. (WT088312AIA). L.T.H. was supported by a 4-year D.Phil. studentship from the Wellcome Trust (WT080540MA). J.N. is supported by the UK Medical Research Council. We are very grateful to A. Soltani for providing the MATLAB code for the biophysical model. We thank S. Knight for his valuable help in data acquisition.

## AUTHOR CONTRIBUTIONS

G.J. designed the research, and acquired and analyzed the data. L.T.H. modified the biophysical model and analyzed the biophysical model predictions. J.N. acquired and analyzed MRS data. T.E.J.B. designed the research and analyzed the data. All authors were involved in writing the manuscript.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/doi/10.1038/nn.3140>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Wunderlich, K., Rangel, A. & O'Doherty, J.P. *Proc. Natl. Acad. Sci. USA* **106**, 17199–17204 (2009).
2. Gershman, S.J., Pesaran, B. & Daw, N.D. *J. Neurosci.* **29**, 13524–13531 (2009).
3. Padoa-Schioppa, C. & Assad, J.A. *Nature* **441**, 223–226 (2006).
4. Cai, X., Kim, S. & Lee, D. *Neuron* **69**, 170–182 (2011).
5. Boorman, E.D., Behrens, T.E., Woolrich, M.W. & Rushworth, M.F. *Neuron* **62**, 733–743 (2009).
6. Bechara, A., Tranel, D. & Damasio, H. *Brain* **123**, 2189–2202 (2000).
7. Ratcliff, R. & Rouder, J.N. *Psychol. Sci.* **9**, 347–356 (1998).
8. Wong, K.F. & Wang, X.J. *J. Neurosci.* **26**, 1314–1328 (2006).
9. Hunt, L.T. *et al. Nat. Neurosci.* **15**, 470–476 (2012).
10. Hare, T.A., Schultz, W., Camerer, C.F., O'Doherty, J.P. & Rangel, A. *Proc. Natl. Acad. Sci. USA* **108**, 18120–18125 (2011).
11. Stagg, C.J., Bachtari, V. & Johansen-Berg, H. *Curr. Biol.* **21**, 480–484 (2011).
12. Schoenbaum, G., Roesch, M.R., Stalnaker, T.A. & Takahashi, Y.K. *Nat. Rev. Neurosci.* **10**, 885–892 (2009).
13. Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E. & Behrens, T.E. *Neuron* **70**, 1054–1069 (2011).
14. Hasler, G. *et al. Arch. Gen. Psychiatry* **64**, 193–200 (2007).

## ONLINE METHODS

**Participants.** Twenty-five healthy male participants (18 to 35 years) participated in the experiment after written informed consent was obtained. All experimental procedures were approved by the Milton Keynes Ethics Committee.

**Magnetic resonance (MR) imaging.** MR data were acquired at 3 T on a Siemens Trio using a 32-channel coil. First, a high-resolution T1-weighted scan was acquired using an MPRAGE sequence<sup>15</sup>. Spectroscopy voxel placement was based on this scan. MRS data were acquired as described previously<sup>15</sup>, with repetition time (TR) = 3,200 ms. The vmPFC voxel (anteroposterior 1.5 × mediolateral 3.0 × dorsoventral 1.0 cm) was mediolaterally centered on the midline and dorsoventrally on the genu of the corpus callosum, with the posterior voxel boundary just rostral to the genu. The parietal voxel (2.0 × 2.0 × 2.0 cm) was centered on the right IPS. We acquired 180 averages for the parietal and 360 averages for the vmPFC voxel to compensate for the reduced signal-to-noise ratio. A rather thin (dorsoventrally) voxel was used for the vmPFC to reduce the effect of field inhomogeneities in this brain area. The presence of field inhomogeneities was also the reason for choosing IPS as the control site rather than lateral orbitofrontal cortex. Temporal constraints precluded the acquisition of more than two voxels. MRS was followed by fMRI, during which subjects performed the task. Forty-five slices with a voxel resolution of 3 mm isotropic were obtained using a sequence optimized for the orbitofrontal cortex<sup>16</sup>. Field maps were acquired using a dual echo 2D gradient echo sequence with TR = 488 ms and TE of 7.65 ms and 5.19 ms on a 64 × 64 × 40 grid.

**Behavioral task.** The task involved repeatedly choosing between two options to obtain a monetary reward (Supplementary Fig. 1a). Each option consisted of one horizontal bar (reward magnitude) and a percentage written underneath it (reward probability). On a subset of trials ('no-brainer' trials), the magnitude and probability of one option were higher than those of the alternative option. The reward schedule was designed to minimize the correlation between chosen and unchosen value (mean  $r = 0.21$  and  $0.33$  across subjects for the fMRI and post-scanning task). One hundred trials were presented during fMRI. After scanning, participants undertook another 400 trials of this task (without the initial viewing period and with quicker timing) on a laptop outside the scanner.

**Analysis of behavioral data.** Subjective reward magnitudes and probabilities were derived by fitting utility functions according to prospect theory<sup>17</sup>

$$r_s = r_o^\alpha$$

$$p_s = \frac{p_o^\gamma}{(p_o^\gamma + (1 - p_o)^\gamma)^{1/\gamma}}$$

where the objective reward magnitude and probability  $r_o$  and  $p_o$  are transformed into subjective magnitude and probability,  $r_s$  and  $p_s$ , respectively. From these values, subjective expected values can be calculated as

$$sEV = r_s \times p_s$$

The modeled probability to choose either of the two options was given by a softmax rule

$$P_{(C=K)} = \frac{e^{sEV_K \cdot \tau}}{\sum_{i=1}^n e^{sEV_i \cdot \tau}}$$

where  $K$  = choice made by subject,  $n$  = number of options and  $\tau$  = softmax inverse temperature. We also fitted two models with only two free parameters, where either  $\alpha$  or  $\gamma$  were fixed at 1. Calculation of the Bayes information criterion (BIC) showed that model fits under the two alternative models were significantly worse than under the full model ( $P < 0.008$ , Supplementary Table 1).

We custom-implemented a Bayesian estimation procedure in MATLAB (MathWorks) to obtain the best-fitting parameters  $\alpha$ ,  $\gamma$  and  $\tau$  (Supplementary Fig. 1b,c). Choice probabilities as a function of  $\tau$  are shown in Supplementary

Figure 1d. To test whether subjects integrated probabilities and magnitudes, we performed a logistic regression analysis. Reward probability, magnitude, choice on the previous trial and outcome on the previous trial were entered as independent variables  $X$  to predict the binary outcome  $Y$  (choices, 0 or 1 for left and right choices, respectively, Supplementary Fig. 1e). A further linear regression tested the effect of value difference, value sum and no-brainer trials on subjects' log reaction times (Supplementary Fig. 1f).

**Processing and analysis of MRS data.** A semi-automated MATLAB-based preprocessing routine was applied to all spectra before analysis. Motion-corrupted signal averages were identified and removed, and then frequency and phase-drift corrections were performed to ensure exact alignment of the remaining averages; these steps were followed by signal averaging. Fully processed spectra were then analyzed as in ref. 15. GABA and glutamate values are reported as a ratio to creatine. GABA and glutamate were successfully detected in 24 (vmPFC) and 22 (parietal cortex) volunteers. Only two spectra from vmPFC had a water line-width slightly  $\geq 10$  Hz (11.5 and 11.66). As reported in the supplements, excluding them did not change the pattern of results, which is why the data reported in the main text include those two vmPFC data sets. The T1-weighted anatomical scans were segmented into gray and white matter using FAST (FMRIB's automated segmentation tool)<sup>18</sup> to calculate relative volumes of gray matter, white matter and cerebrospinal fluid in the MRS voxels. The reported concentrations of GABA and glutamate are corrected for the relative gray matter volumes (Supplementary Table 2).

**Analysis of fMRI data.** Analysis of fMRI data was performed using tools from the FMRIB Software Library (FSL<sup>19</sup>), using the same routine as in ref. 20, with the spatial filter set to 6 mm full-width at half maximum. To investigate activity related to the value difference between the chosen and unchosen options, we set up a general linear model (GLM) that contained the following nine regressors: value difference, value sum, outcome value (reward versus no reward obtained), reward magnitude obtained, main effect from stimulus presentation to response onset, main effect from response onset to outcome delivery, main effect of outcome phase and two stick functions (modeling left and right button presses, respectively). In addition, the six motion parameters from motion correction were included. Contrast images from the first level were then taken to group level using a random effects analysis. Results are reported at an activation-level threshold of  $P < 0.005$  ( $z > 2.58$ ) combined with a cluster-forming threshold of  $P < 0.05$ .

**Region of interest (ROI) analyses.** The above whole-brain analysis yielded a positive effect of value difference in the vmPFC (Supplementary Fig. 4a). BOLD time courses were extracted from the resulting activation. Each volunteer's time series was cut into trials with a duration of 16.2 s (symbol presentation at 0 s, response onset at 4.58 s, outcome presentation at 8.12 s, corresponding to the mean onsets across subjects and trials). Time series were resampled to a resolution of 300 ms. A GLM containing the parameters of interest was then fitted at each time point for each volunteer. This resulted in a time course of effect size for each regressor and for each volunteer. Time courses were then averaged across participants (Fig. 2 and Supplementary Fig. 4b,c).

**Correlation analyses.** As mentioned in the main text, GABA and glutamate are correlated. Therefore, we performed partial correlation analyses. To test whether GABA and glutamate concentrations from vmPFC were better at predicting  $\tau$  than IPS concentrations of those transmitters, GABA and glutamate from both regions were entered as regressors in the design matrix (along with a constant term) to predict the data,  $\tau$ . The effects of vmPFC GABA and glutamate were directly contrasted with those of IPS GABA and glutamate.

For the time-resolved analysis of GABA and glutamate on the slope of the value-difference signal, both GABA and glutamate were entered into one single GLM such that again, the reported effects are exclusively on the orthogonal, non-shared variance between the two neurotransmitters. To test whether the pattern of results could be due to a general relation between GABA and glutamate levels and the BOLD response, we conducted the same analysis; however, this time we looked at the effects of the two neurotransmitters on the main effect rather than on the slope of the value difference (Supplementary Fig. 3).

**Modeling.** We implemented a mean-field reduction of the spiking neuronal network model described in ref. 21. Full details are given in ref. 9. It is notable that

a number of models exist that implement a neural competition on the basis of mutual inhibition, for example those described in refs. 21–23. We do not claim that our results are specific to this particular model; instead we choose this model as an example of the class. We have used this model successfully in a recent study<sup>9</sup> to predict local field potential data, which are more closely related to the fMRI BOLD signal than neuronal spiking activity, which is why we decided to base our predictions on this model.

**Model analysis.** For model predictions of cross-subject behavioral variation, we fit softmax functions to model choice behavior in exactly the same way as was done for individual subjects, choosing the softmax parameter that maximized the log-likelihood of each model instantiation's choices. The regression of these softmax parameters against the degree of recurrent excitation ( $w^+$ ) is plotted in **Supplementary Figure 2a**.

For model predictions of cross-subject neural activity variation, we first estimated a linear regression model for each instantiation of the model, with value difference and overall value as independent variables and with the network's

response as the dependent variable. As for the fMRI data, we calculated the first temporal derivative of the parameter estimate for value difference. We then ran a second-level linear regression analysis (*across* model instantiations) in which this temporal derivative was the dependent variable and inhibition balance was an independent explanatory variable. The *t* statistic from this second-level analysis is plotted in **Supplementary Figure 2b**.

15. Stagg, C.J. *et al.* *J. Physiol. (Lond.)* **589**, 5845–5855 (2011).
16. Deichmann, R., Gottfried, J.A., Hutton, C. & Turner, R. *Neuroimage* **19**, 430–441 (2003).
17. Tversky, A. & Kahneman, D. *J. Risk Uncertain.* **5**, 297–323 (1992).
18. Zhang, Y., Brady, M. & Smith, S. *IEEE Trans. Med. Imaging* **20**, 45–57 (2001).
19. Smith, S.M. *et al.* *Neuroimage* **23** (suppl. 1), S208–S219 (2004).
20. Boorman, E.D., Behrens, T.E., Woolrich, M.W. & Rushworth, M.F. *Neuron* **62**, 733–743 (2009).
21. Wang, X.J. *Neuron* **36**, 955–968 (2002).
22. Usher, M. & McClelland, J.L. *Psychol. Rev.* **108**, 550–592 (2001).
23. Ditterich, J., Mazurek, M.E. & Shadlen, M.N. *Nat. Neurosci.* **6**, 891–898 (2003).