Supporting Information

Wunderlich et al. 10.1073/pnas.0901077106

SI Text

Reinforcement Learning Model. Reinforcement Learning (RL) is concerned with learning the value of taking particular actions in different states of the world in which subjects do not have complete knowledge about the underlying reward generating process. Thus, it is ideally suited to model how subjects learn the value of taking the different actions over time.

We used a version of RL called Q-learning, where action values are updated using a simple Rescorla-Wagner rule. If an action is not selected in a trial its value is not updated. In contrast, if action a is selected on trial t, its value is updated via a prediction error, δ, as follows: $V_a(t+1) = V_a(t) + \eta(t)$, where η is a learning rate between 0 and 1. The prediction error $\delta(t)$ is calculated by comparing the actual reward received, $r(t)$, with the reward that the subject expected to receive from that action in that trial; that is, $\delta(t) = r(t) - V_a(t)$. Probabilistic rewards were delivered in free choice and forced choice trials and these trials were included in updating the value predictions. In null trials, subjects neither expected nor got any reward; hence no learning occurred and values were not updated.

To generate choices, we first used a soft-max procedure where in every trial, the probability (P) of choosing action a is given by: $P_{at} = \sigma(\beta(V_a(t) - V_b(t)) - \alpha$, where $\sigma(z) = 1/(1 + e^{-z})$ is the Luce choice rule or logistic sigmoid, $\alpha = 0$ denotes the indecision point (at which both actions are selected with equal probability), and β determines the degree of stochasticity involved in making decisions. In the paper we refer to V_e as the action value of the eye movement and V_h as the action value of the button press.

The model decision probabilities P_e and P_h were fitted against the discrete behavioral data B_e and B_h to estimate the free parameters (α and β). This was done using maximum likelihood estimation and a log likelihood function given by:

$$\log L = \frac{\Sigma B_e \log P_e}{N_e} + \frac{\Sigma B_h \log P_h}{N_h},$$

where N_e and N_h denote, respectively, the number of trials in which *eye* and *hand* were chosen, and B_e (B_h) equals one if *eye* (*hand*) was chosen in that trial, and zero otherwise.

We also fitted a model with an additional parameter that allowed the unchosen value to decay toward 0.5. However, we found that in our rather simple task with only two choice options the BIC corrected fit of this model was not significantly better than that of our simple learning rule.

The Competition Difference Model of the Decision Process. Based on our finding of a robust correlation between activity in the anterior cingulated cortex (ACC) and a variable equal to the difference between the value of the unchosen and chosen actions, we propose a simple conceptual model for how this value difference might be implemented in the brain to guide value-based choice.

Importantly, the model that we propose has two key properties: (1) it leads to stochastic choices, and (2) its output is sensitive to both the choice that is made and to the action values of the two alternatives.

The model consists of a neural network with N 'neurons.' Each neuron could take on either an ON or OFF state at every particular instant. These neurons were split into two discrete populations of N/2 neurons each: one population was associated with the value of an eye movement, the other with the value of a finger movement. In each trial the comparison process is initialized by turning ON a fraction of neurons in each population that is proportional to the action value of the associated action. Thus, for example, if $V_e = 0.56$ and $n = 200$, then 56 out of the 100 eye neurons were set to the ON state (see Fig. S4 for an illustration).

The network was then allowed to evolve in discrete steps as follows: 1. In every step every active neuron for one of the two actions is paired with a randomly chosen neuron (with replacement) for the other action. Once the assignment is made for all neurons the following rule is implemented: if the matching unit is ON, the neuron is switched OFF, otherwise no change is made on the state of the neuron. (Note: this rule is implemented simultaneously for all of the neurons, so there are no order effects).

2. Noise is injected after every iteration as follows: the state of every unit in the network is flipped to its opposite state with a probability that is given by the product of a noise parameter σ and the number of active units encoding the value of the same action.

The basic idea behind the CDM is that the decision process works by virtue of a stochastic mutual inhibitory competition between the two distinct populations of neurons encoding the value of the two actions. A "winner" is declared when one of the two populations reaches zero. At this point the population that has a positive number of ON neurons is declared the winner. Note that the model incorporates two desirable features: (1) higher value actions have a higher chance to win the competition process, which means that the better action is chosen with higher probability; and (2) the change in activity in every step scales with the amount of existing activity in the network.

We then added an additional layer (with constant positive input from which the previous result is subtracted) to the model to invert the output to the value difference between the action not chosen and the action chosen.

We simulated the model using a population of $n = 200$ as follows. First, we simulated the stochastic comparison process 1,000 times for each possible value difference between the two actions. Second, after the model converged in each simulation (which always occurred in less than 50 steps) we computed the number of ON units in the population that won the competition. Note that, since the model is stochastic, in some simulation it converged to the action with the larger value, but in others it converged toward the action with the smaller value. Third, we averaged the 1,000 simulations for each possible action value difference to estimate a reference output value for later use in the comparison regressor in the general linear models of the fMRI data described below. As depicted in Fig. S6, the averaging was done conditional on whether the optimal choice was made or not. We constructed a trial-by-trial parametric modulator by retrieving the stored values from this analysis in each trial for the current value difference and dependant on whether the subject chose optimally (action with the higher action value) or the action with the lower action value from either the red or blue curve in Fig. S6.

To validate our model behaviorally, we determined for each possible value difference (Ve − Vh) in 1,000 model runs the fraction of runs in which the model settled on the eye choice. For this purpose, a noise parameter σ was estimated for each subject using the maximum likelihood procedure described in the RL section above. The resulting psychometric choice functions (probability of the model to choose eye dependant on Ve-Vh) are compatible with subjects' observed behavior and model performance is very similar to the reinforcement learning soft-max procedure (Table S3).

Note a few things about the model. First, it leads to stochastic choices, consistent with the behavior in Fig. 1*D*. Second, unlike other models such as the drift diffusion model (see the discussion

in the next section), the output signal depends on the action value difference when the best item is chosen (and is constant otherwise).

The Drift Diffusion Model of the Decision Process. A very popular model of how the comparison is made is called the drift diffusion model (DDM, sometimes also called race-to-barrier model) (1–3). This model has proven extremely useful in explaining the psychophysics of perceptual choice as well as some aspects of neural activity in areas such as LIP during perceptual decision tasks (4, 5).

The basic idea of the model, as applied to value-based decision-making, is illustrated in Fig. S5. The process computes a net value signal (say $V_h - V_e$) that fluctuates between two barriers until a decision is made. A decision is reached when the net value signal crosses either of the two barriers. If the top barrier is crossed the hand action is chosen. If the bottom barrier is crossed the eye action is chosen. The net value signal climbs to the hand barrier with a slope proportional to $V_h - V_e$, but it is also affected by white Gaussian noise. In the simple version of the model the net value signal commences the integration process mid-way between the two barriers, which implies that there is no bias between the two options (i.e., when $V_h = V_e$ both options are chosen with equal probability).

Note that this is a "high-level" computational model, which is silent about how the brain might implement these computations. This question needs to be answered to be able to make predictions about how to identify areas that might implement this process using fMRI. Consider an extremely simple neural implementation of the DDM. There are two populations of neurons: one encodes for the net value of a hand movement ($V_h - V_e$), the other encodes for the net value of an eye movement ($V_e - V_h$). Both populations encode a signal with a dynamic range 0 to M. Both signals begin the competition process at M/2 and the decision process stops when one of the signals reaches M. The signal in the two populations evolves in discrete time until a choice is made. Each of the populations is connected to an output signal that encodes the selected motor movement, triggered once the integration threshold M is reached. Note a few interesting properties of the neural implementation of the DDM. First, the sum of activity in all neurons at every instant during the comparison equals M. Second, the sum of activity in both output signals is also equal to a constant, call it B, independent of V_e and V_h.

These properties imply that an area implementing the comparison should have a level of neural activity equal to M (independent of V_e and V_h) from the onset of the trial until a choice is made. They also imply that the output of the process is characterized by a constant level of activity B (again, independent of V_e and V_h) that is on from the moment the decision is made to the time the motor output is executed.

These properties mean that the comparator activity of the DDM should be modeled in the general linear models of BOLD activity described below as an unmodulated regressor that begins with the onset of a free trial and ends with the deployment of one of the two actions (i.e., it has a duration equal to the reaction time). In contrast, the output activity should be modeled as an unmodulated regressor at the time of (either) action execution with a duration of 0 seconds.

Although these regressors provide a full characterization of the neural activity associated with the DDM, and they are easily incorporated in the general linear models described below, they present a major problem for fMRI. Consider, for example, the regressor for the comparator process. The activity for this process is perfectly correlated with those of other processes that come on-line during the evaluation process, that are also unmodulated by value, and that also last until a choice is made. Given that a large number of such processes are likely to exist (and in fact a large number of distinct areas are robustly activated by this type of contrasts in decision-making tasks), it is difficult to isolate the location of the DDM comparator process using fMRI, particularly for the range of reaction times taken for decisions in a standard fast-paced decision task such as the one featured here. A similar problem holds for the output signal of the DDM, since it is perfectly correlated with motor activity that is not modulated by action values.

Given these issues, we concluded that the neural signatures of the DDM cannot be identified using the fMRI methods deployed in the present study. It is important to emphasize that these measurement problems are not present in single-unit electrophysiology since this technique permits the independent measurement of neural activity in a putative decision region with sufficiently high spatial and temporal resolution. Moreover, the output of the model does not resemble the value difference signal we observed in the ACC in the present study. Thus, while we cannot assess the relevance of the DDM model to value-based decision making in the present study, it is the case that such a model does not provide a good account for the value comparison signal we observed in the ACC.

FMRI Data Acquisition. Functional images were taken with a gradient echo T2*-weighted echo-planar sequence (TR = 2.65 s, flip angle = 90°, TE = 30 ms, 64 × 64 matrix). Whole brain coverage was achieved by taking 45 slices (3 mm thickness, no gap, in-plane resolution 3 × 3 mm), tilted in an oblique orientation at 30° to the AC-PC line to minimize signal dropout in OFC. Subjects' head was restrained with foam pads to limit head movement during acquisition. Functional imaging data were acquired in two separate 568-volume runs, each lasting about 24 min. A high-resolution T1-weighted anatomical scan of the whole brain (MPRAGE sequence, 1 × 1 × 1 mm resolution) was also acquired for each subject.

FMRI Data Analysis. Image analysis was performed using SPM5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, U.K.). Images were first slice time corrected to TR/2, realigned to the first volume to correct for subject motion, spatially normalized to a standard T2* template with a voxel size of 3 mm, and spatially smoothed with a Gaussian kernel of 8-mm FWHM. Intensity normalization and high pass temporal filtering (using a filter width of 128 s) were also applied to the data.

We estimated several general linear models (GLM) for each individual.

GLM 1. Two events were modeled in each trial: the time of the choice cue, parametrically modulated by the trial-by-trial action values V_e and V_h, and the time of the presentation of the outcome, modulated by the prediction error δ. Trials in which subjects chose the eye action and trials in which subjects chose the hand action were modeled as separate regressors. Trials were further split to build separate regressors for each trial type: free choice, forced choice, and null trials. Choice and forced trials were modulated by the estimated action values to find neural representations of those signals. In null trials there were no modulators. The model also included an orthogonalized version of the parametric action value modulators described in the previous paragraph during the inter-trial interval. The rationale behind this last set of regressors was to allow for the possibility that participants might already be considering which option to choose next after receiving the feedback on the previous trial. In such a case, the ITI window would be part of the decision process. However, we did not find any correlates of value related signals during the ITI, which led us to focus our analysis on the time of the choice cue. All regressors were convolved with the canonical hemodynamic response function. In addition, the 6 scan-to-scan motion parameters produced during realignment and two session constants were included as additional regressors of no interest. We then computed contrasts of interest at the individual level using linear combinations of the regressors: Value chosen: Ve|eye_chosen + Vh|hand_hosen; Value difference ($V_{unchosen} - V_{chosen}$): Ve|hand_chosen + Vh|eye_chosen − Ve|eye_chosen −

Vh|hand_chosen. This model was used to generate the statistics reported in Figs. 2, 3, and 4*A*.

GLM 2. This model was identical to the first GLM except for the addition of following two regressors:

1. A Dirac delta function 700 ms into every trial (which is equal to the average response across subjects) modulated by the estimated output signal of the DCM given the values of V_h and V_e and the optimality of the choice made. The values of these modulators are depicted in Fig. S6, dependant on the action value difference and whether the subject chose optimally (red curve) or non-optimally (blue curve).

2. An indicator for the time of cue presentation modulated by a decision difficulty measure given by $-\mid V_e - V_h \mid$. Note that this modulator takes a maximum value when $V_e = V_h$. We also tested alternatively for subject specific conflict by taking into account individual choice biases in calculating the value difference $-\mid V_e - V_h - \alpha \mid$. For example, if a participant had a slight overall bias toward saccading, the difficulty would be centered on this subject's individual point of equilibrium. For this purpose we estimated a subject specific indecision point α by fitting the RL model with a third free parameter that allowed for horizontal shifts of the sigmoidal choice function.

As before, note that the value of the output signal of the CDM model used in every trial computed is obtained by averaging over 1,000 simulations, and that, due to the stochasticity of the CDM, it is a noisy measure of the actual activity during the trial.

The goal of this second GLM was to look for regions in which activity correlated with the output signal of the DCM. The results of this second GLM were used to generate the statistics reported in Fig. 4 *B* and *C*.

GLM 3 and 4. We carried out two further analyses to rule out the possibility that action-value signals observed in SMA and pre-SEF could be attributed to motor preparation. In the first such additional analysis, we estimated a GLM in which trials involving hand or eye movements were entered as separate indicator variables, and reaction times (RTs) for those hand and eye movements were used as parametric modulators around those indicator variables. We then tested for areas correlating separately with RTs for eye and hand movements (as a proxy for motor preparation).

In a fourth GLM, we re-ran the same analysis as in GLM 1, except this time with the inclusion of additional parametric modulators of RTs for hand and eye movements, to establish whether motor preparation as indexed by RTs could account even in part for the regions found to correlate with action-values.

To enable inference at the group level, we calculated second-level group contrasts using a one-sample *t* test. Results are reported at $P < 0.001$ uncorrected in the entire brain and tested in areas of interest at $P < 0.05$ after small volume correction (SVC) for multiple comparisons.

ROI Analyses. The effect size plots in Fig. 2*B* were computed by averaging GLM's beta values across subjects. To ensure the independence of the data used to compute the effect sizes from the data used to select the ROI we performed the following steps. First, for each subject we randomly selected half of the choice trials across the entire experiment and then created another design matrix in which we modeled the selected 50% of the trials (T1) and the remaining 50% of the trials (T2) as separate regressors. Similar to the GLM 1 (described above), regressors T1 and T2 each consisted of an onset time regressor and parametric modulators for Ve and Vh. Second, to define our ROIs we performed a whole-brain SPM analysis similar to the one shown in Fig. 1*A*, but this time restricted to the T1 trials only. Note that the activation map produced by this step (with a threshold set at $P < 0.005$ unc.) looks very similar to the one shown in Fig. 2*A*. Third, we defined the SMA/preSEF ROIs using a 6-mm sphere around the individual subject peak voxel within the activated cluster for T1. Finally, we extracted average effect sizes within these spheres from the remaining T2 regressor. Very similar results were obtained if instead of splitting the data into two groups of trials within subject, data from 50% of the subjects were used to define the ROIs, while data from the remaining 50% were used for extracting the effect sizes (from the co-ordinates defined in the first group of subjects). The effect size plots for vmPFC shown in Fig. 3*B* were calculated using an identical procedure.

Small Volume Corrections. Seed region coordinates for small volume correction were defined by two alternative methods: First, we used an anatomical definition for supplementary motor cortex provided by the AAL human brain atlas (6), and we corrected for small volume within the entire area of supplementary motor cortex defined by this atlas (comprising both SEF and SMA), superimposed on the normalized average structural scan from our study. Second, we took the average peak co-ordinate from 16 previous fMRI studies identifying activation in SMA and defined a sphere of 12 mm around that averaged peak co-ordinate in which to perform the small volume correction (7–9). The size of the sphere in the functionally defined seed region was set to 1.5 times the size of the smoothing kernel used during preprocessing of the fMRI dataset. Using each and every one of these criteria, our effects survived correction for small volume with family wise error at $P < 0.05$.

The structural T1 images were co-registered to the mean functional EPI images for each subject and normalized using the parameters derived from the EPI images. Anatomical localization was carried out by overlaying the t-maps on a normalized structural image averaged across subjects, and with reference to an anatomical atlas (10).

1. Usher M, McClelland JL (2001) The time course of perceptual choice: The leaky, competing accumulator model. *Psychol Rev* 108:550–592.
2. Smith PL, Ratcliff R (2004) Psychology and neurobiology of simple decisions. *Trends Neurosci* 27:161–168.
3. Busemeyer JR, Johnson JG (2004) Computational models of decision making. In *Handbook of Judgement and Decision Making*, eds Koehler D, Narvey N (Blackwell Publishing Co, New York), pp 133–154.
4. Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* 86:1916–1936.
5. Shadlen MN, Britten KH, Newsome WT, Movshon JA (1996) A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci* 16:1486–1510.
6. Tzourio-Mazoyer N, et al. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
7. Lau HC, Rogers RD, Ramnani N, Passingham RE (2004) Willed action and attention to the selection of action. *Neuroimage* 21:1407–1415.
8. Lau HC, Rogers RD, Haggard P, Passingham RE (2004) Attention to intention. *Science* 303:1208–1210.
9. Picard N, Strick PL (2003) Activation of the supplementary motor area (SMA) during performance of visually guided movements. *Cereb Cortex* 13:977–986.
10. Duvernoy HM (1999) in *The Human Brain. Surface, Blood Supply and Three-Dimensional Section Anatomy* (Springer, NewYork), 2nd Ed.

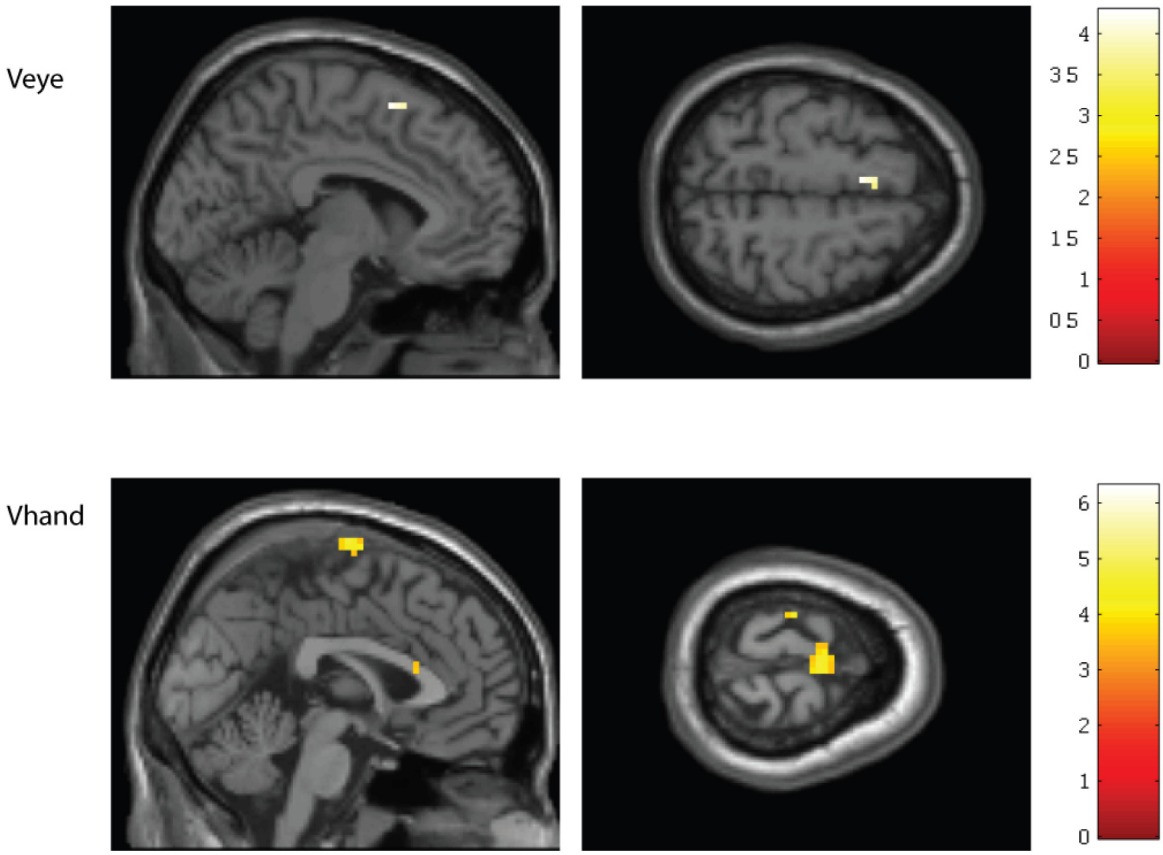

Fig. S1. A region of pre-SEF showing correlations with action-values for eye movements (Ve, top) and a region of supplementary motor area showing correlations with action-values for hand movement (Vh, bottom). T-maps are shown from a whole brain analysis thresholded at $P < 0.001$ uncorrected. The color bars indicate the magnitude of the t-scores. These same two contrasts are also shown combined in Fig. 2*A*.

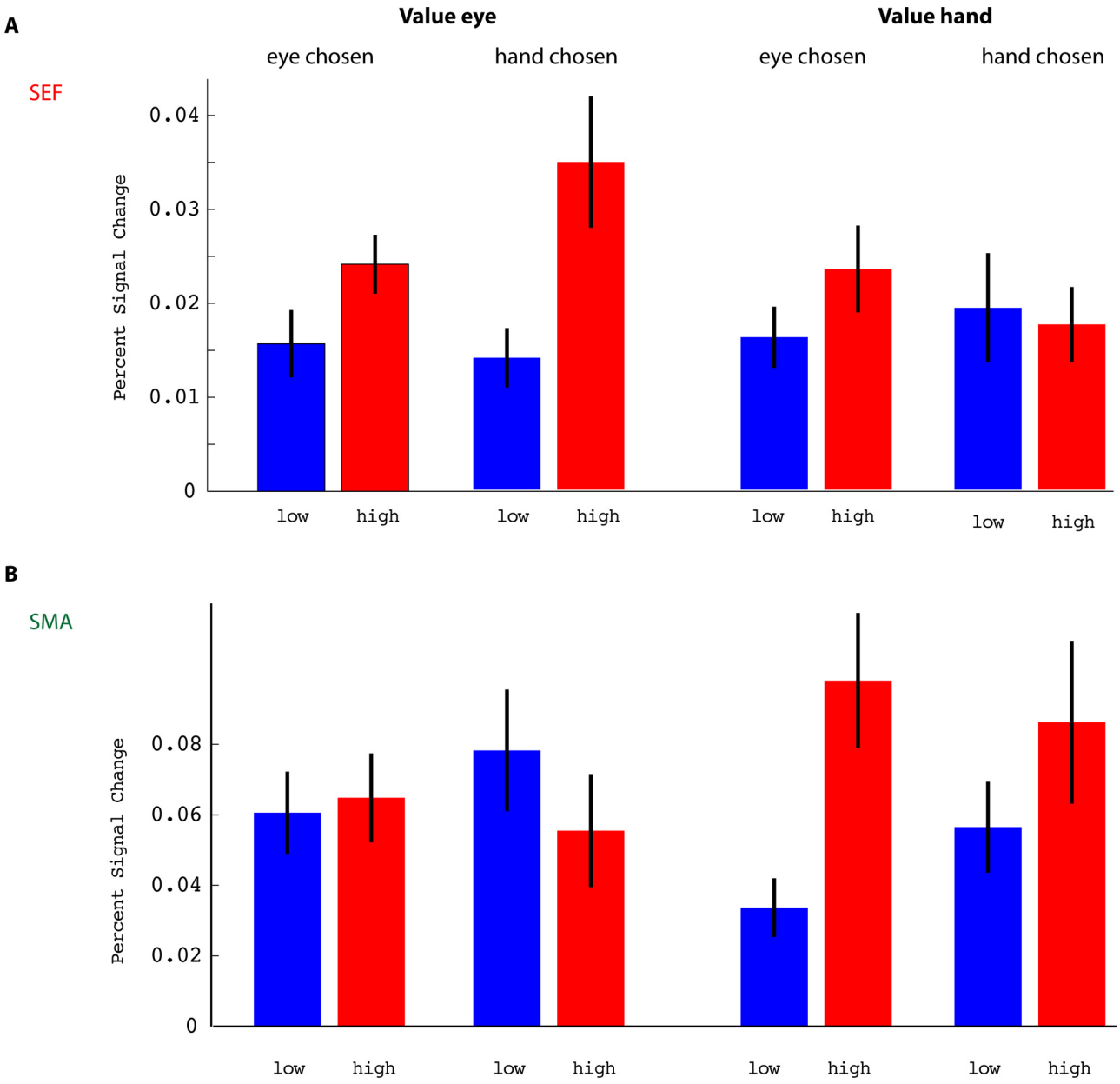

Fig. S2. Posthoc plots of effect sizes (expressed as percent signal change) averaged across all voxels in the activated clusters at the group level for each subject and then averaged across subjects separately for the pre-SEF (shown in *A*) and SMA (shown in *B*). (*A*). The graph on the left hand shows average % signal change extracted from pre-SEF separately for trials in which the action value of the eye movement is low (Veye$<=$0.5 percentile) and high (Veye $>$0.5 percentile), further separated by trials in which either the eye movement or hand movement was actually chosen. As expected (given this area correlates with action-values for eye movements), the % change plots discriminate high and low eye values irrespective of whether that action is chosen on that trial. Importantly however, when activity within the same area is plotted as a function of the action value for the hand movement (shown on the right hand side), the signal change on high and low hand value trials does not discriminate between high and low hand values. Although for eye_chosen trials the value hand signal does appear to separate in the direction of and high and low hand values, this difference is not statistically significant [paired *t* test t $=$ 1.8; $P < 0.08$; note this post-hoc comparison for the value of hand is independent of the parametric contrast used to select the voxels (value_eye)]. (*B*). Similar plot for SMA showing that this region distinguishes high and low action_values for hand movements irrespective of whether that action is chosen, but does not distinguish the value of eye movements.

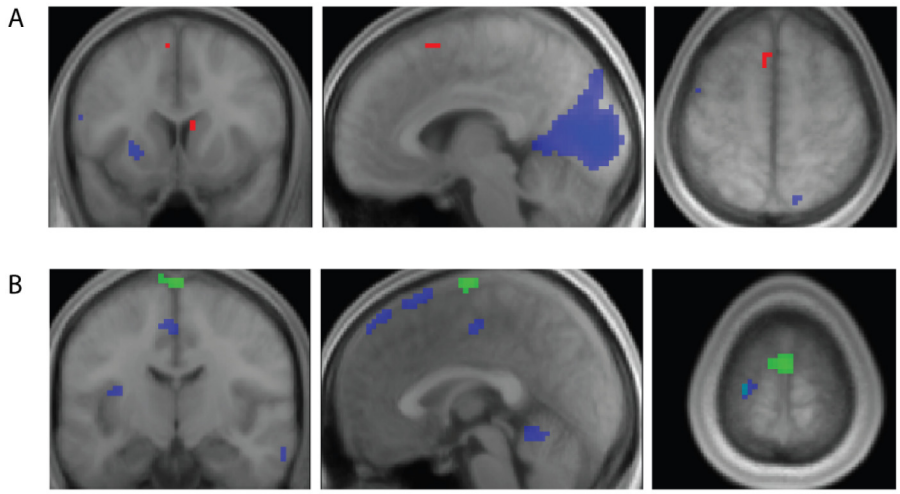

Fig. S3. (*A*) Correlations with eye action value V_{eye} (red, same as in Fig. 2*A*) don't overlap with correlations of the local predominant choice bias for eye (blue). Areas in blue indicate regions that are significantly more active during periods in which the subject predominantly chooses eye. (*B*) Similarly, the correlation with hand action value V_{hand} in SMA (green, same as displayed in Fig. 2*A*) is not overlapping with areas that are significantly more active at times when the subject predominantly chooses hand actions (blue).

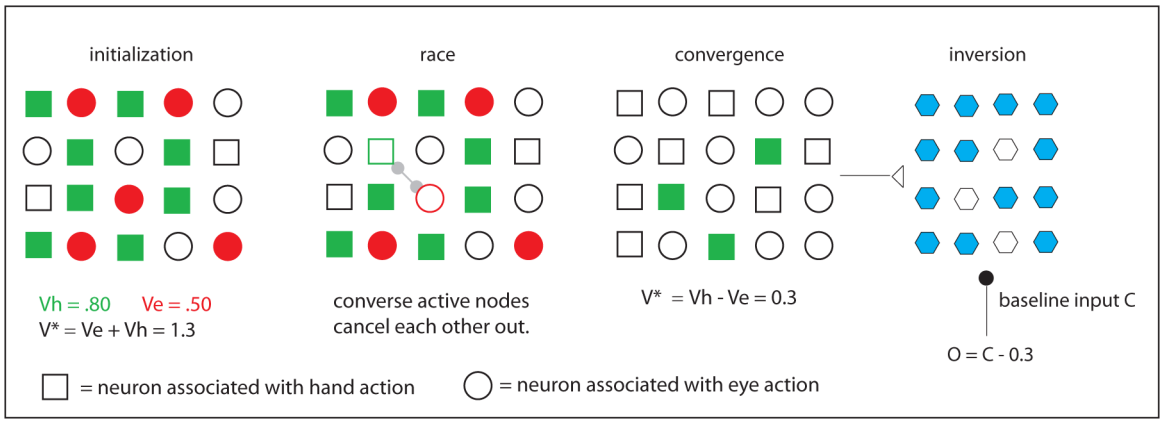

Fig. S4. Illustration of the competition difference model. The decision process was modeled as an iterative algorithm of mutual competition between the neuronal populations associated with the valuation of eye and hand actions. Model inputs were action-values for eye and hand movements. The evolution of the model output over time within a trial is illustrated in a hypothetical case with action-values of Vh = 0.8 and Ve = 0.5 at the time of model initialization (left), during competition between populations (middle), and after convergence (right). The model includes an additional final layer (with constant positive input from which the previous result is subtracted) that inverts the output of the network to compute the value difference between the action not chosen and the action chosen.

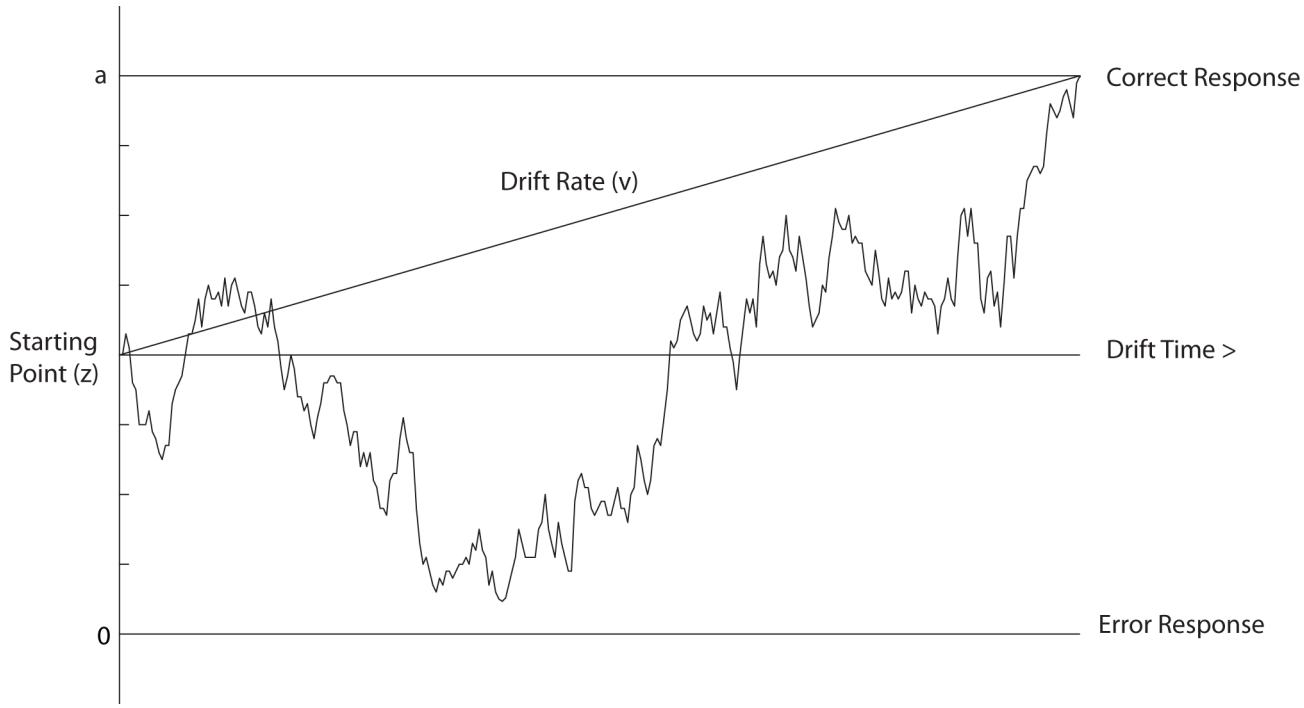

Fig. S5. Illustration of the drift diffusion model. During the decision making process the model computes a net value signal (say $V_h - V_e$) that fluctuates between two barriers. A decision is reached when the net value signal crosses either of the two barriers. If the top barrier is crossed the hand action is chosen. If the bottom barrier is crossed the eye action is chosen. The net value signal climbs to the hand barrier with a slope proportional to $V_h - V_e$, but it is also affected by white Gaussian noise. In the case depicted in the figure, $V_h > V_e$ so that hand is the correct choice.

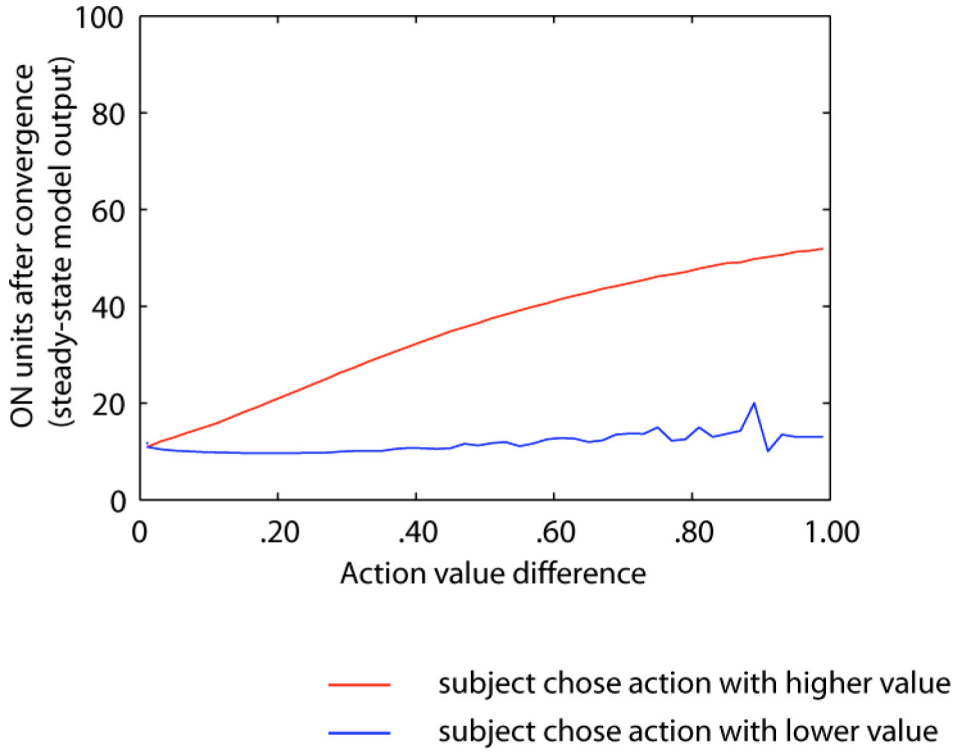

Fig. S6. Average steady state output of the computational decision model of the choice process as a function of value difference between the two action values. The remaining number of ON units after 50 iterations was averaged across 1,000 model runs. The red curve displays the average remaining total ON units in the model in trial that the model converged toward the action with the larger action value. In contrast, the blue curve displays the same statistic for trials in which the model selected the action with the lower action value.

Table S1. Characteristics of each value signal type in terms of the specific variable that activity in a given region should be correlated with as a function of the action performed (choice taken)

	Action performed	
Value signal categories	*Eye chosen*	*Hand chosen*
Action_value: eye	V_{eye}	V_{eye}
Action_value: hand	V_{hand}	V_{hand}
Value_chosen	V_{eye}	V_{hand}
Value_chosen: eye only	V_{eye}	—
Value_chosen: hand only	—	V_{hand}

Table S2. Locations of significant correlation with parametric contrasts in the fmri analysis (threshold $P < 0.001$) MNI coordinates denote the group peak voxel of each cluster

	x	*y*	*Z*	*T*	No. voxels	
				Vh (Fig 2A)		
01	−21	−90	−18	6.29	46	Left occipital cortex
02	57	−54	−09	5.68	65	Right inferior temporal gyrus
03	−30	−24	75	5.58	59	Left postcentral sulcus
04	00	−12	78	4.61	33	SMA*
05	39	−78	39	4.34	32	Left intraparietal sulcus
				Ve (Fig 2*A*)		
01	27	15	−03	4.27	5	ventral striatum
02	−06	09	60	4.16	4	preSMA*
				Vchosen (Fig 3*A*)		
01	−39	−36	69	6.37	142	Left postcentral gyrus
02	21	−48	60	5.84	133	Right postcentral sulcus
03	36	−21	39	5.59	24	Right central sulcus
04	−15	−36	45	5.37	34	Left cingulate sulcus
05	−48	−15	51	5.19	50	Left central sulcus
06	06	45	−18	4.76	51	Ventromedial prefrontal cortex*
07	−60	−09	−06	4.72	38	Sup. Temporal sulcus
08	−54	−30	03	4.43	64	Planum temporale
				Vunchosen-Vchosen (Fig 4*A*)		
01	−27	24	00	8.02	249	Left anterior insula
02	36	24	06	7.0	176	Right anterior insula
03	03	24	51	5.64	268	Dorsomedial frontal cortex & anterior cingulate*
04	−33	−54	42	5.47	48	Intraparietal sulcus
05	−48	12	36	5.1	214	Inferior frontal sulcus
				Hand bias (Fig S3)		
01	−39	−21	51	7.61	558	Left precentral gyrus
02	21	−45	−30	6.4	144	Right cerebellum
03	06	21	66	6.16	158	Right dorsal medial frontal cortex
04	−39	−66	−39	6.02	168	Left cerebellum
05	66	−27	−21	5.93	132	Right temporal lobe
06	21	−39	60	5.76	32	Right central sulcus
07	−39	−66	57	5.54	83	Left superior parietal gyrus
08	48	−51	54	5.26	320	Right superior parietal gyrus
09	30	15	45	5.02	58	Right middle frontal gyrus
10	33	−27	69	4.92	30	Right precentral gyrus
11	−12	−42	30	4.57	126	Left posterior cingulated gyrus
12	45	−72	−33	4.56	44	Right cerebellum
				Eye bias (Fig S3)		
01	12	−72	12	7.76	2,734	Occipital lobe (bilateral)
02	−27	12	00	5.26	25	Left striatum
03	24	−42	45	5.08	20	Right parietal cortex

*, $P < 0.05$ SVC corrected.

Table S3. Model performance in predicting individual subject choices of the soft-max model and the CDM model

	Soft-max model		CDM model	
Subject	R2	$P <$	R2	$P <$
1	0.57	0.000	0.54	0.000
2	0.65	0.000	0.65	0.000
3	0.20	0.000	0.15	0.000
4	0.20	0.000	0.20	0.000
5	0.18	0.000	0.15	0.000
6	0.04	0.019	0.02	0.118
7	0.13	0.000	0.11	0.000
8	0.54	0.000	0.54	0.000
9	0.04	0.011	0.05	0.009
10	0.18	0.000	0.17	0.000
11	0.64	0.000	0.62	0.000
12	0.06	0.003	0.05	0.004
13	0.36	0.000	0.35	0.000
14	0.25	0.000	0.22	0.000
15	0.56	0.000	0.53	0.000
16	0.58	0.000	0.57	0.000
17	0.31	0.000	0.27	0.000
18	0.38	0.000	0.39	0.000
19	0.46	0.000	0.44	0.000
20	0.56	0.000	0.57	0.000
21	0.61	0.000	0.62	0.000
22	0.19	0.000	0.18	0.000
23	0.06	0.003	0.06	0.004
	0.34		0.32	