ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 21 (2004) 733–743

Trend detection via temporal difference model predicts inferior prefrontal cortex activation during acquisition of advantageous action selection

Martin P. Paulus,[a,b,c,*] Justin S. Feinstein,[a] Susan F. Tapert,[b,c] and Thomas T. Liu[c,d]

[a] *Laboratory of Biological Dynamics and Theoretical Medicine, University of California San Diego, San Diego, CA, USA*
[b] *Department of Psychiatry, University of California San Diego, San Diego, CA, USA*
[c] *Veterans Affairs San Diego Healthcare System, San Diego, CA, USA*
[d] *Department of Radiology, University of California San Diego, San Diego, CA, USA*

Received 22 July 2003; revised 12 September 2003; accepted 25 September 2003

The process of accurately predicting which actions are associated with advantageous versus disadvantageous outcomes is an important function of daily life. An integral part of this process is being able to detect when the association between an action and an outcome changes. This investigation examined the hypothesis that the inferior prefrontal cortex is critical for the detection of trends and that a trend process derived from the temporal difference model accomplishes this detection. Nineteen normal right-handed volunteers completed 120 4-s trials of a Rock Paper Scissors (RPS) task during functional magnetic resonance imaging. Subjects acquired the selection of advantageous actions during the RPS task. Activations in the medial frontal gyrus (BA 10), left ventrolateral frontal gyrus (BA 11/47), and left pallidum were significantly higher during trials in which subjects acquired the advantageous action. The time course of individually derived trend detection functions was found to be time-locked to the hemodynamic changes in the inferior frontal gyrus. These findings are consistent with the hypothesis that the inferior prefrontal cortex computes a trend from previously experienced action–outcome sequences based on a value function derived from the temporal difference model.

Keywords: Temporal difference model; Inferior prefrontal cortex; Functional magnetic resonance imaging

Introduction

Decision-making is complex and involves several component processes. For effective decision-making, these processes have to be integrated functionally, that is, information needs to be exchanged across these processes and neural substrates. A key process in decision-making is the establishment of a value system that can guide the selection of a response. This value system may represent the integration of associations of the available action with advantageous or disadvantageous outcomes as well as the strength of these associations relative to other available outcomes. The neural processes and computations that may underlie this value system has been under intense investigation and several theoretical models (Egelman et al., 1998; Gold and Shadlen, 2002; Shadlen and Newsome, 1996; Suri and Schultz, 1999) have been proposed to explain how response selection takes place in a decision-making situation.

These theoretical models have in common a connection between specific brain systems and processes underlying the adjustment of the value system. For example, in the Log-likelihood model (Shadlen and Newsome, 1996), a Bayesian-like process (Pouget et al., 2003) is proposed to modify probabilities of hypotheses according to perceived stimuli (Gold and Shadlen, 2002), which is thought to occur in the parietal cortex (Platt and Glimcher, 1999; Shadlen and Newsome, 2001). In comparison, the prediction error, or temporal difference model (Schultz et al., 1997; Suri and Schultz, 1999) focuses on monoaminergic neurons in general, which appear to broadcast prediction errors, that is, the difference between the expected and observed reward, as global teaching signals to different areas of the brain (Schultz and Dickinson, 2000), and dopamine neurons in particular, which generate a short-latency, phasic reward signal (Schultz, 2002).

The targets of these adjustment processes involve a wide range of brain areas, which have been implicated in decision-making processes (Paulus et al., 2001). Among these processes are the establishment of a relative value system in the inferior prefrontal cortex (Damasio et al., 1996; Frith et al., 1991; Jueptner et al., 1997), which includes ventromedial, ventrolateral, and orbitofrontal cortex (O'Doherty et al., 2003a); the detection of changing values in medial prefrontal cortex (Knutson et al., 2003; Zysset et al., 2002) or the adjacent anterior cingulate cortex (Bush et al., 2002; Pochon et al., 2002); and the implementation of an anticipatory reward or punishment signal in the ventral striatum (Knutson et al., 2001) and amygdala (Baxter and Murray, 2002; Kahn et al., 2002; Pagnoni et al., 2002).

In this investigation, we combined a computational approach based on the prediction error or temporal difference model that establishes values for available actions in a decision-making situation with functional neuroimaging to elucidate the functions of

* Corresponding author. Department of Psychiatry, UCSD, La Jolla CA 92093-0603. Fax: +1-858-534-7653.
E-mail address: martin@mag.ucsd.edu (M.P. Paulus).
Available online on ScienceDirect (www.sciencedirect.com.)

1053-8119/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2003.09.060

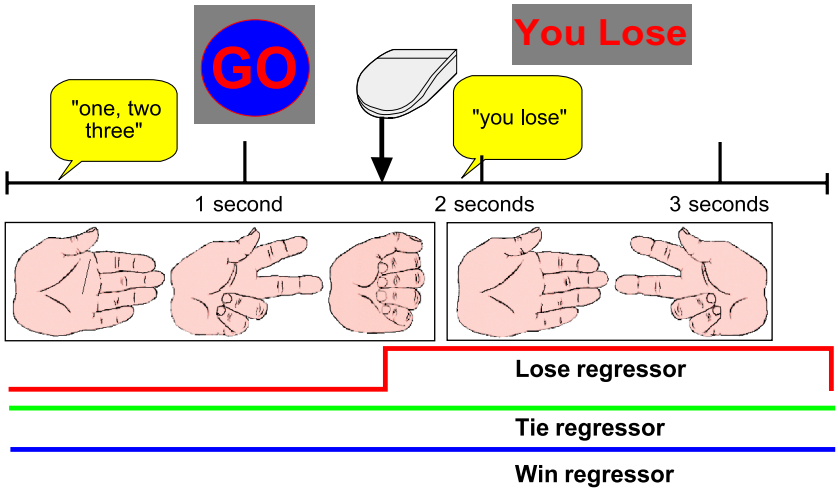

Fig. 1. Task design and an example of a subject specific outcome regressor.

neural substrates in adjusting the value system when actions become advantageous or disadvantageous. Although the associations between available actions and advantageous or disadvantageous outcomes were changed suddenly, subjects were not informed about this change and, thus, needed to detect these changes empirically to select the best action. The goal was to determine whether an 'online' implementation of the association between actions and outcome values and a signal that would indicate a change in this association (trend detection) was consistent with functional magnetic resonance imaging (fMRI)-related activation changes in the inferior prefrontal cortex. Therefore, this investigation examined the hypothesis that the inferior prefrontal cortex is critical for the detection of trends, and that a trend process derived from the temporal difference model accomplishes this detection. To investigate this hypothesis, a computerized version of the Rock Paper Scissors (RPS) task was used with pre-determined and varying contingencies.

Methods

Nineteen normal right-handed volunteers (14 males, 5 females, age = 36.3 ± 6.2, mean ± SD) gave informed consent for a

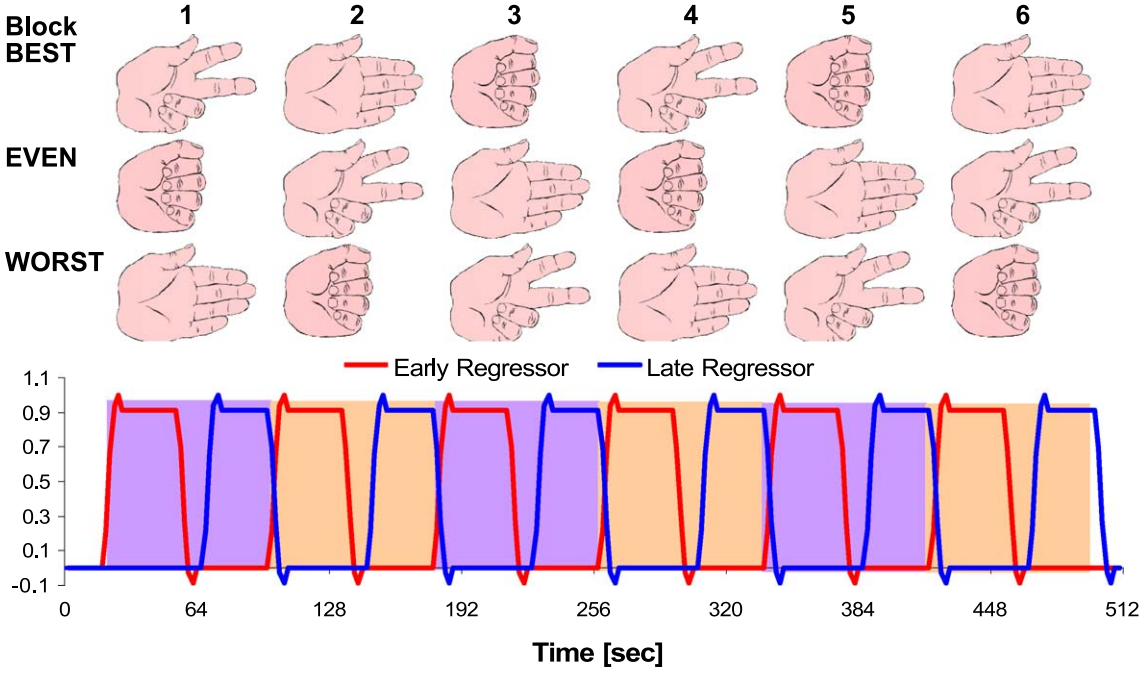

Fig. 2. One hundred twenty trials of the Rock Paper Scissors task were divided into six blocks of 20 trials (80 s) during which the preferred response was kept constant. Selecting the preferred response won 9/10 times, selecting the even response won "5/10" times, and selecting the worst response won "1/10" times. The preferred, even, and worse response were switched every 20 trials. For example, when the subject selects scissors during the first block, the computer will select paper in 9/10 encounters. When the subject selects paper, the computer will select scissors in 9/10 encounters and when the subject selects rock, the computer will select rock in 8/10 encounters and scissors or paper in 1/10 encounters. Each block was divided into an "early" (red line) and "late" (blue line) phase and a regressor was used for each phase to estimate the hemodynamic response.

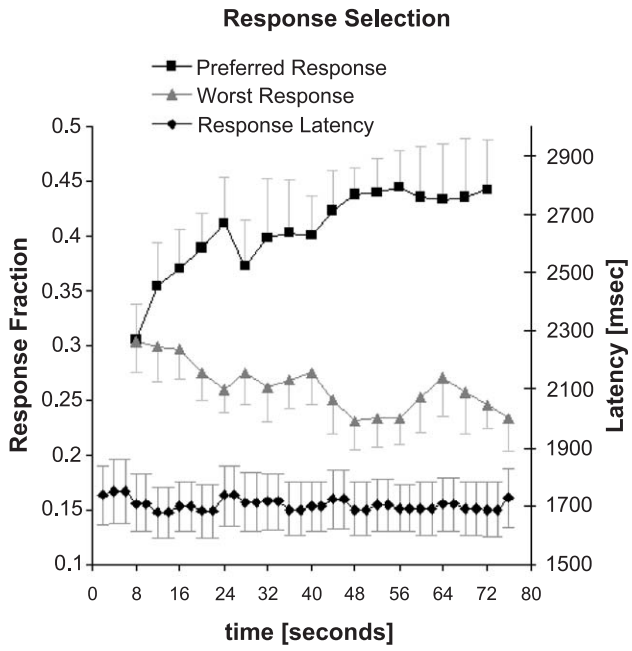

Fig. 3. Acquisition curves for the selection of the preferred (black) over the worst (grey) response during the course of a trial block and average response latencies with SEM across trial blocks.

protocol approved by the UCSD Human Research Protections Program. Subjects had no history of psychiatric disorders or medical problems. Before entering the MRI scanner, subjects completed a 5-min practice version of the RPS task (Fig. 1). Once in the scanner, subjects completed 120 4-s trials of the RPS task flanked by 16 s resting periods in the beginning and end during an 8 min and 32 s functional scan acquisition. After the scanning session, structured post-task questionnaires were obtained from 13/19 subjects; the remainder of the subjects gave general verbal reports.

Rock paper scissors task

This task was used to determine whether subjects are able to select responses that are associated with wins and avoid responses that are associated with losses; whether subjects are able to change their response selection when the associated outcome contingencies change. This task is based on the well-known Rock Paper Scissors game. For the task, the typical rules apply: paper beats rock, rock beats scissors, and scissors beat paper. The subjects were instructed that they were playing against the computer and were told to maximize their total point account (1 point for a win, 0 points for a tie, and −1 point for a loss). The probability of reinforcement, that is, beating the computer (e.g. subject chooses paper, computer selects rock, subject gains one point), was pre-determined for each response within a trial block (see Fig. 2). Unbeknownst to the subject and without changing trial duration or inter-trial interval, the preferred, even, and worst response were switched every 20 trials. For example, when the subject selects scissors (preferred response) during the first block, the computer will select paper in 9/10 encounters. When the subject selects rock (even response), the computer will select rock in 8/10 encounters, paper in 1/10 encounters and scissors in 1/10 encounters. When the subject selects paper (worst response), the computer will select scissors in 9/10 encounters. A total of 120 trials are presented; therefore the "best response" block switches 6 times after 20 consecutive trials in the following order: scissors, paper, rock, scissors, rock, paper. The main dependent variable was the proportion of response that is associated with wins 9/10 times during each trial block.

During each trial, subjects saw pictures of a hand forming paper, scissors, and rock signs on the left, middle, and right side of the computer screen, respectively, for 1 s and heard the instruction "one, two, three" over a headphone (see Fig. 1). At 1 s into the trial, subjects were presented with a "Go" sign in the center of the screen, which provided the cue to select paper, scissors, or rock by pushing the left, middle, or right button with the index, middle, or ring finger of the right hand. Subjects had 2.5 s to respond, after which the trial timed out until the next trial. Immediately after selecting a response, the outcome was presented on the computer screen: the subject's selection was shown on the left and the selection of the computer was shown on the right side of the screen. At the same time, the subjects heard "you win", "you lose", or "a tie", and the score was incremented, reduced by one, or left unchanged, while the total score was displayed on the top of the screen.

Functional magnetic resonance imaging

During the RPS task, a functional imaging run sensitive to blood oxygenation level-dependent (BOLD) contrast was collected for each subject using a 1.5-T Siemens (Erlangen, Germany) scanner (T2*-weighted echo planar imaging, TR = 2000 ms, TE = 40 ms, 64 × 64 matrix, 20 4-mm axial slices, 256 scans). fMRI volume acquisitions were time-locked to the onset of each trial. During the same experimental session, a T1-weighted image (MPRAGE, TR = 11.4 ms, TE = 4.4 ms, flip angle = 10°, FOV = 256 × 256, 1 mm^3 voxels) was obtained for anatomical reference. For preprocessing, voxel time series were interpolated to correct for non-simultaneous slice acquisition within each volume and corrected for three-dimensional motion. Two subjects were excluded due to large movement artifacts apparent during systematic visual inspection of the voxel time series.

fMRI analysis pathway

The data were preprocessed and analyzed with the software AFNI (Cox, 1996). The echo planar images were realigned to the 128th acquired scan and time corrected for slice acquisition order. To exclude the voxels showing an artifact-related signal drop, a combined threshold/cluster-growing algorithm was applied to the mean of the functional images to compute a whole-brain mask. This screened out non-brain voxels and voxels falling within the artifact region. Three separate analyses were carried out.

First, two a priori regressors of interest were constructed to determine the activation during (1) early task-block phase and (2) late task-block phase. Preliminary behavioral studies using the RPS task revealed that subjects showed significant acquisition of the preferred response when comparing the first 10 trials to the second 10 trials within a trial block. However, for the regressors to be nearly orthogonal, we constructed a regressor for the first eight trials and a regressor for the last eight trials of each trial block, which allowed for a return of the regressor height to baseline when convolved with a prototypical hemodynamic response function. Thus, these regressors maximized the contrast between the early and late phase of each block (20

trials) and were convolved with a prototypical hemodynamic response (Boynton et al., 1996) before inclusion in the regression model. Three regressors were used to model residual motion (in the roll, pitch, and yaw directions). Regressors for baseline and linear trends were used to eliminate slow signal drifts. The AFNI program 3dDeconvolve (Ward, 2002) was used to calculate the estimated voxel-wise response amplitude. A Gaussian filter with FWHM 4 mm was applied to the voxel-wise percent signal change data to account for individual variations in anatomical landmarks. Data of each subject were normalized to Talairach coordinates. The voxel-wise percent signal change data were entered into a mixed model ANOVA with task condition (early versus late) as a fixed factor and subjects as a random factor. A threshold adjustment method based on Monte-Carlo simulations was used to guard against identifying false positive areas of activation (Forman et al., 1995). Functional regions of interest, which were defined via activation clusters that showed a cluster-wise a posteriori $P < 0.05$, in the inferior prefrontal cortex consisting of the ventromedial (BA 10) and ventrolateral (BA 11, 47) prefrontal cortex as well as the pallidum, were used as masks to extract BOLD fMRI time series data from each subject.

Second, three subject response-specific regressors were defined for winning, tying, and losing trials to determine which neural substrates differentially activate with varying outcomes. Analogous to the analysis above, these regressors were convolved with a prototypical hemodynamic response (Boynton et al., 1996) before inclusion in the regression model. Similarly, three motion regressors and regressors for baseline and linear trends were used. The AFNI program 3dDeconvolve was used to calculate the estimated voxel-wise response amplitude.

Third, individual subjects' regressors that were obtained from the trend detection model as derived from the modified temporal difference model, which are described in detail below, were convolved with a prototypical hemodynamic response and entered into a multiple regression model which included the motion regressors, baseline, and linear drift. These individual regressors incorporate both the inter-subject variability of the acquisition of the preferred response as well as the temporal variability of the acquisition process. Spatial filtering and volume-thresholding was applied as specified above.

Statistical analyses

The voxel-wise percent signal change data were entered into a mixed model ANOVA with response type as a fixed factor and subjects as a random factor. Moreover, a within-subjects contrast was computed between winning–losing trials to determine the effects of outcome. Finally, the average percent signal difference was extracted from regions of activation that were found to survive this threshold/cluster method.

Temporal difference model and trend detection

For each response option, two value functions were created, V^+ and V^-, which represent values associated with "winning" and "losing". Temporal discount and learning parameters were used to model the effect of "winning" and "losing" on the subsequent selection of a response. The "discount parameters" (Suri and Schultz, 2001), α^+ and α^-, quantify the degree to which the previous value function is discounted over time to influence the decision-making on the current trial. The learning parameters, β^+ and β^- (Pearce and Bouton, 2001; Wasserman and Miller, 1997), model the degree to which "winning" or "losing" during the current trial affects the value function for the next trial. The value functions for the *Paper* response are thus modeled as:

$$V^+_{\text{Paper}}(i+1) = \alpha^+ V^+_{\text{Paper}}(i) + \beta^+(\delta(i) - V^+_{\text{Paper}}(i)) \quad (1)$$

$$V^-_{\text{Paper}}(i+1) = \alpha^- V^-_{\text{Paper}}(i) + \beta^-(\delta(i) - V^-_{\text{Paper}}(i)) \quad (2)$$

where $\delta(i)$ equals whether the subject experienced a win or a loss at time i, respectively. The value function models for the other responses have similar forms.

The value functions were used to compute the probability of selecting a response using an approach that was previously described in Egelman et al. (1998) and Montague and Berns (2002). For example, the probability for selecting *paper* on trial i was defined as:

$$P_{\text{Paper}}(i) = \exp(V^+_{\text{Paper}} - V^-_{\text{Paper}}) \Big/ \sum_{j\,=\,\{\text{Paper, Rock, Scissors}\}} \exp(V^+_j - V^-_j) \quad (3)$$

For each subject, the discounting and learning parameters were estimated using an optimization algorithm implemented in the statistical package *R* (Ihaka and Gentleman, 1996). Specifically, the per trial model-prediction error was computed as:

$$\text{Error}(\alpha^+, \alpha^-, \beta^+, \beta^-) = \sum_{i\,=\,2}^{n} \sum_{j\,=\,\{\text{Paper, Rock, Scissors}\}} P_j(i) - R_j(i) \Big/ n - 1 \quad (4)$$

Here, $R_{\text{Paper}}(i)$ is 1 if the subject selected paper at the ith trial and zero otherwise. The model-prediction error term corresponds to the average difference between the observed response sequence and the model-predicted probability of the response at any given trial. Using a limited-memory modification of the variable metric algorithm (Byrd, 1995), the model parameters were allowed to fluctuate between {0,1} for each subject, which allows for testing the hypothesis that random response selection (all parameters are 0) resulted in optimal prediction.

To test the hypotheses that (1) subjects used a procedure consistent with the temporal difference model, a multivariate analysis of variance was performed for all model parameters ($\{\alpha^+, \alpha^-, \beta^+, \beta^-\} > 0$). A paired t test was used to examine (1) whether subjects based their selection more sensitively on losses than on wins ($\beta^+ < \beta^-$) and (2) whether subjects showed equal temporal discounting for wins and losses ($\alpha^+ = \alpha^-$).

To obtain a measure of trend detection, that is, the change of relative frequency of response patterns across trials, the trial-dependent root mean squared, RMS, was computed for each option as

$$\text{RMS}(P_{\text{Paper}}(i)) = \frac{\sum_{k\,=\,i}^{i+n} P_{\text{Paper}}(k)^2}{n} - \left(\sum_{k\,=\,i}^{i+n} P_{\text{Paper}}(k)\right)^2 \quad (5)$$

To obtain a general estimator of change, an average RMS was obtained by averaging across the three options (paper, rock, and

scissors). The average RMS provides a measure of variability of the prediction based on the value functions defined above. Similar to the early versus late regressor analyses, the computation of the RMS was based on eight consecutive trials ($n = 8$). The RMS is maximized when the value function is either monotonically increasing or decreasing, which corresponds precisely to the situation when the optimum response is changing. Finally, the subject-specific average RMS was convolved when a prototypical hemodynamic response was used as a regressor as described above.

Results

Subjects acquired the selection of advantageous actions (Fig. 3) and selected the preferred response 35 ± 2% of the time during the first eight trials and 44 ± 2% during the second eight trials ($F(1,18) = 12.02$, $P < 0.01$). In comparison, subjects selected the worst response 28 ± 1% during the first eight trials and 25 ± 2% during the last eight trials of each block ($F(1,18) = 7.04$, $P < 0.01$). In comparison, there was no significant change in response latency ($F(1,18) = 0.20$, NS). Subjects' reports obtained via structured questionnaires from 13 of the 19 subjects after scanning showed that 7/13 thought the computer chose at random, 4/13 reported that they noticed a pattern but were unable to specify the nature of the pattern, one subject thought the computer was "minimizing losses", and one subject thought the computer was using the subject's previous response.

As shown in Table 1 and Fig. 4a, activations in the medial frontal gyrus (BA 10), left ventrolateral frontal gyrus (BA 11/47), and left pallidum were significantly higher during the first eight trials relative to the last eight trials of each block. Individual differences between the acquisition of the advantageous responses during the first eight and last eight trials of each block correlated significantly with the activation difference between the early and late trial block regressor in the left pallidum ($r = -0.52$, $P < 0.05$) and with the activation in the medial frontal gyrus ($r = -0.52$, $P < 0.05$) but not with left ventrolateral frontal gyrus ($r = -0.17$, NS). In the pallidum and medial frontal gyrus, the smaller the difference between the late and early phase activation the less likely the subject were to acquire the preferred response. Fig. 4b shows the time series signal averaged over functional ROIs identified by the a priori defined early versus late phase regressors. The correlations between the medial frontal gyrus (BA 10), left ventrolateral frontal gyrus (BA 11/47), and left pallidum with the average RMS function obtained from the temporal difference model were 0.63, 0.64, and 0.50, respectively. Thus, although the magnitude of the BOLD-fMRI response according to the a priori defined regressors did not correlate with the acquisition measures of advantageous versus disadvantageous action selection in all areas, the inter-

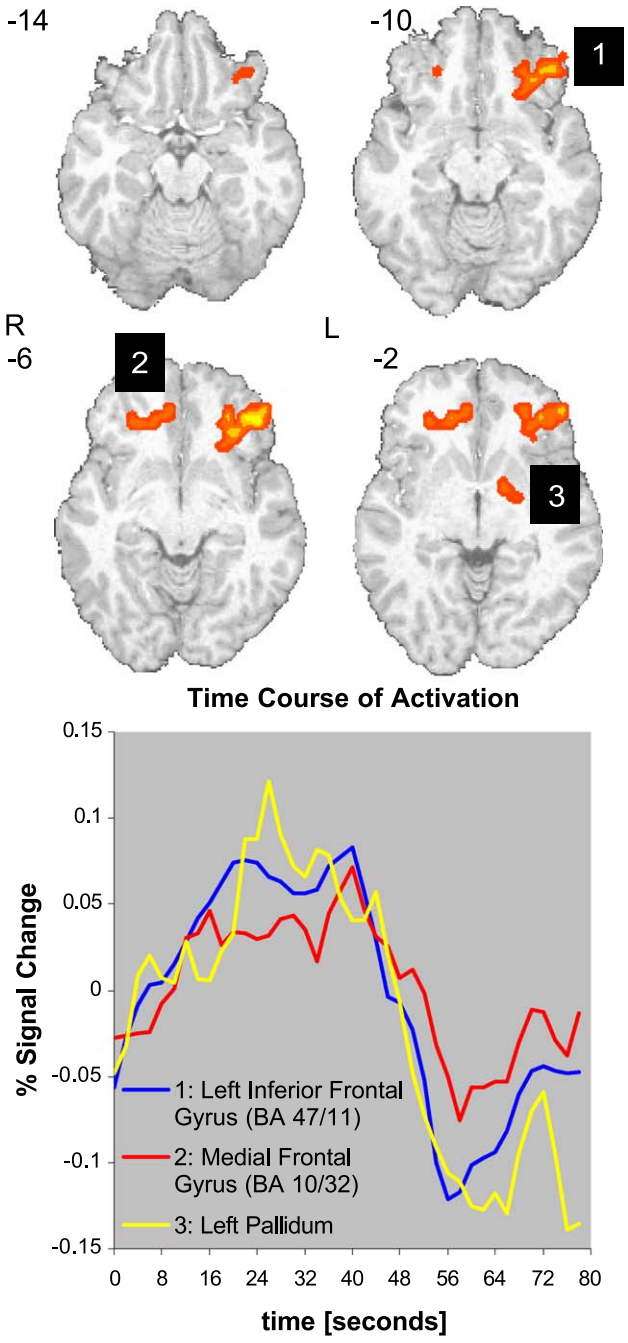

Fig. 4. (upper panel) Three areas showed a significant interaction between the early and late phases of the block. Bilateral medial frontal gyrus (BA 10), left middle frontal gyrus (BA 11/47), and left pallidum show more activation during the early phase relative to the late phase. (lower panel) The time course of these activations.

Table 1
Volume-thresholded clusters of areas that activated differently early versus late during trial block

Region	Volume	x	y	z	L/R	$F(2,18)$	Description	BA
1	3072	−34	36	−5	L	25.4	Inferior frontal gyrus	44/11
2	704	16	38	−4	R	13.7	Medial frontal gyrus	10/32
3	576	−15	−1	2	L	10.3	Pallidum	

subject variability of the trend detection process correlated significantly with all three areas.

The time course of individually derived trend detection functions and the BOLD hemodynamic response for the left inferior frontal gyrus, which is based on volume-thresholded activation as defined by the functional ROI derived from the trend detection model, is shown in Fig. 5. As expected, the trend detection model increases early during the trial block and is time-locked to the

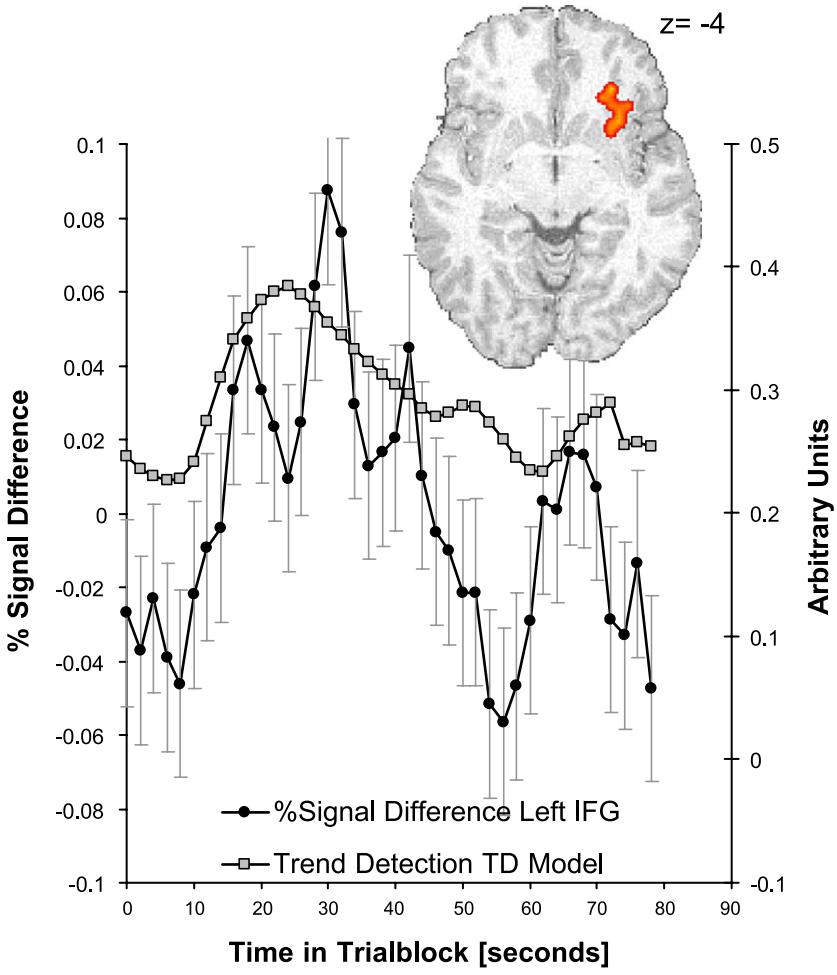

Fig. 5. Average time course and SEM of activation in left inferior frontal gyrus (BA 44) as defined by the functional ROI derived from the trend detection model. Inset shows functional ROI, *z*-coordinate according to Talariarch Atlas.

hemodynamic changes in the inferior frontal gyrus. In addition to this area, trend detection-related activation was also found in the posterior parietal cortex and the superior temporal gyrus (Table 2) but not in the medial frontal gyrus or the left pallidum.

A multivariate analysis of variance revealed that all model parameters differed significantly from zero ($F(4,15) = 532.22$, $P < 0.01$), supporting the hypothesis that all parameters were required in the temporal difference model to successfully predict the behavior observed during the RPS task. Moreover, the model-prediction error differed significantly from chance predictions ($t(18) = 58.0$, $P < 0.01$), indicating that the model predicted the subject's subsequent selection of paper, rock, or scissors significantly better than chance. The estimation of the learning parameters for wins and losses ($\beta^{+} = 0.62 \pm 0.10$; $\beta^{-} = 0.97 \pm 0.02$) showed that subject's selection were more sensitively affected by losses relative to wins ($t_{\text{paired}}(18) = 3.22$, $P < 0.01$). In comparison, the temporal discount parameters ($\alpha^{+} = 0.69 \pm 0.08$; $\alpha^{-} = 0.78 \pm$

Table 2
Volume-thresholded clusters of areas that activated according to the trend detection temporal difference model

Region	Volume	*x*	*y*	*z*	L/R	Description	BA
1	1216	−62	−22	30	L	Inferior parietal lobule	40
2	832	−57	−23	7	L	Superior temporal gyrus	41
3	832	−40	56	56	L	Superior parietal lobule	7
4	768	−28	22	−3	L	Inferior frontal gyrus	11/47

Table 3
Volume-thresholded clusters of areas that activated differently as a function of outcome

Region	Volume	*x*	*y*	*z*	L/R	*F*(2,18)	Description	BA
1	2112	−7	−71	48	L	15.21	Left precuneus	7
2	1792	34	0	46	R	33.03	Right middle frontal gyrus	6/9
3	1344	−15	6	8	L	15.82	Left caudate	
4	1216	31	49	6	R	11.18	Right middle frontal gyrus	10/46
5	1216	−40	45	6	L	14.01	Left middle frontal gyrus	10/46
6	1216	−1	−14	13	L	12.54	Left thalamus	
7	1216	15	−1	24	R	22.23	Right caudate	
8	1216	20	12	58	R	20.4	Right middle frontal gyrus	8

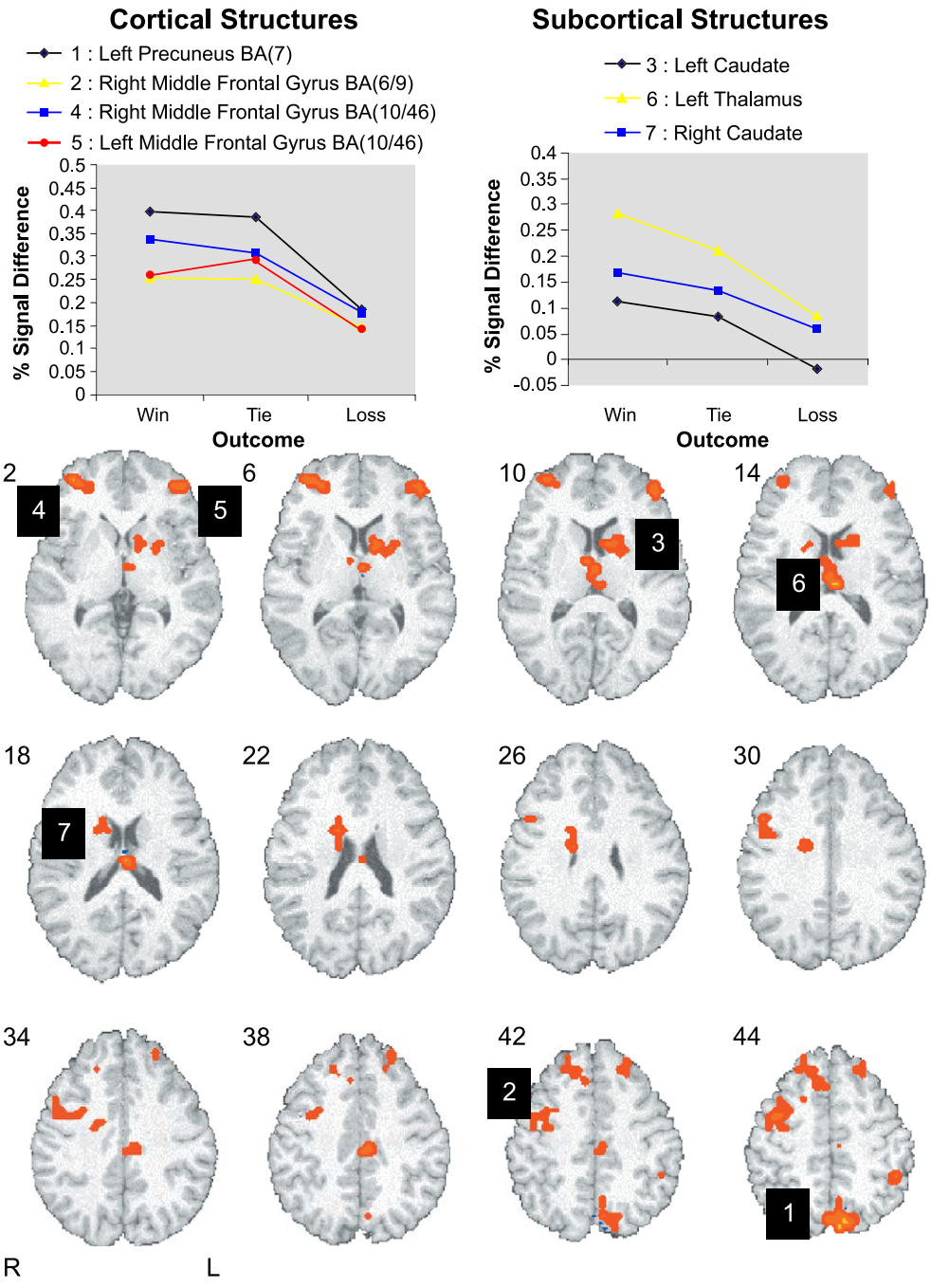

Fig. 6. Volume thresholded cluster of activation due to outcome differences, black numbers indicate *z*-coordinate, white numbers indicate area.

0.04) did not differ significantly for wins or losses ($t_{paired}(18)$ = 1.07, NS).

The behavioral results of selecting the preferred, even, or worst response during the first eight trials versus the last eight trials in a trial block were correlated with the model parameters obtained from each subject to determine whether the model parameter variations predicted variations in observed behavior. A larger value of learning parameter for wins, β^{+}, was associated with more frequent selection of the preferred response during the first eight ($r = 0.62$, $P < 0.01$) and the last eight trials ($r = 0.65$, $P < 0.01$) of a trial block and with a reduced frequency of the worst response during both parts of the trial block ($r = -0.62$, $P < 0.01$ and $r = -0.50$, $P < 0.01$). No corresponding correlations were found for the variability of the learning parameter for losses. In comparison, slower temporal discounting for losses (α^{-}) was associated with fewer selection of the worst response during both parts of the trial block ($r = -0.49$, $P < 0.05$ and $r = 0.59$, $P < 0.01$). In combination, whereas variability of the learning parameter for wins appears to be critical for the acquisition of the preferred response, the variability of the temporal discount parameter for losses explained variability in avoiding the worst response.

Several areas showed significant activation differences between winning, tying, and losing trials (Table 3). Specifically, bilateral striatum and thalami showed significantly higher activation during winning trials relative to losing trials (Fig. 6). Similarly, both posterior parietal cortex (BA 7) and inferior and dorsolateral prefrontal cortices (BA 10/46) were more active during winning versus losing trials (Fig. 6).

Discussion

In repeated decision-making situations, the ventrolateral prefrontal cortex was more active after the association between actions and outcomes changed and when subjects acquired the advantageous action. A trend detection model was able to explain the time course of activation in this area. This model was based on a value function derived from the temporal difference model with parameters for both rewarding and punishing events. In comparison, bilateral striatum and thalami, posterior parietal cortex and inferior as well as dorsolateral prefrontal cortices activated differentially as a function of outcome. This dissociation of trend detection and reward-related processing during decision-making is consistent with the notion that decision-making involves separable processes that are implemented across different brain areas.

The current findings support the hypothesis that the inferior prefrontal cortex, which includes both ventromedial and ventrolateral prefrontal cortices, is critical for the acquisition of advantageous actions during the RPS task. Moreover, the degree to which these processes are translated into observable actions may depend on the link between assessment, planning, and motor execution in the pallidum, an area that is critical for conditional learning (Robbins et al., 1989). The current results are consistent with those previously reported for probabilistic reversal learning (Cools et al., 2002) and sequence detection (Huettel et al., 2002). The adjustment process during the RPS task may recruit areas that are involved in reward-related processing, for example, the medial or orbitofrontal cortex (O'Doherty et al., 2001, 2003a), and in acquisition of rules related to successful outcome, which may be represented in the ventrolateral prefrontal cortex (Passingham et al., 2000).

The role of the inferior prefrontal cortex in the RPS task is consistent with findings by others who propose that the inferior prefrontal cortex may develop a moment-to-moment model for patterns of events (Huettel et al., 2002). Computing trends via estimating time-dependent variability is a convenient way of determining a number of different aspects of reward-related processes. In animal experiments, spatial foraging for food has been sensitive to both the mean and variance in reward distributions (Real, 1991). Moreover, avoidance of options during decision-making situations with uncertain outcome has been related to their increased outcome variability (Rode et al., 1999), which has resulted in the proposition that the action selection in decision-making situations is influenced primarily by expected outcome and outcome variability. The current finding provides a detailed view of the potential process that is carried out by the inferior prefrontal cortex. The computations for the trend detection function, which is based on computing the outcome variability, can be performed easily online, that is, the neural structure may keep a representation of changing trends during the task to guide action selection. For example, multiplicative neuron units have been described as part of dendro-dendritic interactions (see for review Schmitt, 2002) or during the computation of visual coordinate systems in the parietal cortex (Salinas and Abbott, 1996). Therefore, implementation of trend detection based on an RMS-type procedure may be carried using local recurrent connection circuits in the inferior prefrontal cortex. This process would also explain the parametrically increased activation in this area to "runs" of similar stimuli in a simple choice response task (Huettel et al., 2002).

The temporal difference model originally provided a neural-based conceptualization of stimulus–reward processing. This model has also been used to account for changes in the stimulus–reward relationships in decision-making situations (Egelman et al., 1998). Moreover, BOLD fMRI activation changes were consistent with predictions by the temporal difference model in ventral striatum and orbitofrontal cortex during acquisition of appetitive conditioned stimuli (O'Doherty et al., 2003b) and in the striatum during the expectancy violation of positive and negative stimuli (McClure et al., 2003). The current model is based on a Mackintosh extension of the Rescorla Wagner rule (for a review, see Pearce and Bouton, 2001; Wasserman and Miller, 1997), which accounts for most cases of cue competition, acquisition, extinction, discrimination, conditioned inhibition, contingency effects, and the US-preexposure effect. The Rescorla Wagner rule treats extinction as unlearning and therefore does not predict external inhibition, spontaneous recovery, or reminder-induced recovery, and falls short in predicting such phenomena as latent inhibition. The acquisition of the advantageous response in this task falls within the category of Thorndikean instrumental conditioning (Wasserman and Miller, 1997), which is fundamentally tied to the notion of temporal contiguity (the delay between stimulus and outcome) and contingency (the characteristics of the joint probability of stimulus and outcome). Whereas the former was kept constant in this experiment, the latter is a combination of the subject's response and the computer-preset outcome relationship, which was examined here by creating a subject-specific implementation of the temporal difference model. All model parameters were found to be necessary to significantly predict the subjects' responses as well as inter-subject's response variability. The trend detection extension of the temporal difference model provided here allows one to quantify the subject-specific ability to detect trend changes when acquiring contingency rules ("what follows what when things change"). This process is extremely important and ubiquitous because stimulus–outcome associations change frequently in daily life.

The somatic marker hypothesis (Bechara et al., 1997; Damasio, 1996), the Log-likelihood model (Gold and Shadlen, 2001, 2002), and the prediction error or temporal difference model (Schultz et al., 1997; Suri and Schultz, 1999) may represent neural system implementations to differentially process affective states, compute contingencies and contiguities, and to optimize decision-making in the presence of complex stimulus–outcome relationships. Whereas the somatic marker hypothesis is based on the brain-representation of the condition of the body as it relates to anticipated outcomes of actions (Craig, 2003), the Log-likelihood model and temporal difference model are ways of computing a value associated with a stimulus or the outcomes associated with an action. Although the specific implementations differ, that is, the Log-likelihood model is based on Bayesian adjustments versus the temporal difference model, which is a difference equation, the basic goal accomplished by these models is similar. A value is associated with each action via computational models based on

repeated exposure to advantageous and disadvantageous outcomes. Both heuristic (Wasserman and Miller, 1997) and probabilistic (Dayan et al., 1995; Egelman et al., 1998) rules have been proposed for the selection among competing actions based on the computed relative value. Some investigators have recently extended the temporal difference model to include punishing as well as rewarding events (Daw et al., 2002) and have suggested that dopamine and serotonin may act as opponent processes in this manner. The findings of the temporal difference model and the trend detection process in this investigation support the notion that punishing (losing) events may have a stronger effect on the adjustment of the value function than rewarding (winning) events, which is consistent with the general shape of the value function that has been established in the classical psychology decision-making literature (Tversky and Kahneman, 1981).

Previous investigations have revealed a success/failure-dependent activation pattern in the prefrontal cortex (Elliott et al., 1999, 2000b), anterior cingulate (Elliott and Dolan, 1998), insula (Critchley et al., 2001), amygdala (Kahn et al., 2002), and parietal cortex (Paulus et al., 2001). Some investigators have suggested a medio-lateral gradient of enabling versus suppressing actions associated with favorable outcomes (Elliott et al., 2000a) or positive and negative outcomes (O'Doherty et al., 2001); however, others (Critchley et al., 2001) have suggested that the degree of anticipation and associated arousal may mediate activation differences in these areas. The activation-related differences between winning and losing trials in this task will need further study using an event-related paradigm to differentiate the outcome-specific from the anticipatory modulation of the hemodynamic response.

There are several limitations to the current study. A continuous fixed trial time design was used here to observe trend-related changes during this task. Therefore, it was not possible to obtain activation patterns specific to the anticipation or the evaluation of positive or negative outcomes. Moreover, other processes that are involved in this task such as mentalizing (Gallagher et al., 2002), as defined by the ability to explain and predict the behaviors of others by attributing them to independent mental states, working memory, or strategy selection could not be disambiguated with this experimental design. Lastly, although several subjects reported that the computer responded non-randomly, no significant neural activation differences were observed in this group (data not shown). Therefore, a larger group of subjects may be needed to extract subject-related variability in the activation pattern.

Various aspects of decision-making can be dysfunctional in a number of neuropsychiatric disorders (American Psychiatric Association, 1994; Mogg et al., 1991; Rahman et al., 1999, 2001) including substance-related syndromes (Bechara et al., 2001; Grant et al., 2000; Rogers et al., 1999). Behavior of some substance-dependent subjects is similar to that of patients with bilateral ventromedial prefrontal cortex lesions, and characterized by selecting choices that are associated with immediate benefit, even if these choices yield long-term negative future consequences (Bechara and Damasio, 2002). Dysfunctions of the ventromedial, ventrolateral, and dorsolateral prefrontal cortex have been observed in stimulant-dependent subjects (London et al., 2000; Paulus et al., 2002; Volkow and Fowler, 2000). Others have shown increased activation of the inferior medial and lateral prefrontal cortex in substance-dependent subjects in response to cues that elicit craving responses (Breiter et al., 1997; Childress et al., 1999; Grant et al., 1996; London et al., 2000; Tapert et al., 2003; Volkow and Fowler, 2000; Wang et al., 1999). Moreover, a longer period of abstinence in substance-dependent subjects was associated to reduced activation of the inferior prefrontal cortex (Volkow and Fowler, 2000). This altered functionality of the inferior prefrontal cortex in substance-using subjects may result in an inadequate computation of trend detection with respect to drug-related experiences. Therefore, dysfunctions in the adjustment of one's action when the association between action and outcomes changes could explain subjects' transition from the initiation to the maintenance of using substances despite the development of adverse consequences.

Acknowledgments

We acknowledge the invaluable help of Kelly Winternheimer and Thuy Le. This work was supported by grants from NIMH (R21DA13186, M.P.P.) and support from the Veterans Administration via a Merit Award (M.P.P.).

References

American Psychiatric Association, 1994. Diagnostic and Statistical Manual of Mental Disorders (4th Edition): DSM-IV. The American Psychiatric Association, Washington.

Baxter, M.G., Murray, E.A., 2002. The amygdala and reward. Nat. Rev. Neurosci. 3, 563–573.

Bechara, A., Damasio, H., 2002. Decision-making and addiction (Part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. Neuropsychologia 40, 1675–1689.

Bechara, A., Damasio, H., Tranel, D., Damasio, A.R., 1997. Deciding advantageously before knowing the advantageous strategy [see comments]. Science 275, 1293–1295.

Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S.W., Nathan, P.E., 2001. Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. Neuropsychologia 39, 376–389.

Boynton, G.M., Engel, S.A., Glover, G.H., Heeger, D.J., 1996. Linear systems analysis of functional magnetic resonance imaging in human V1. J. Neurosci. 16, 4207–4221.

Breiter, H.C., Gollub, R.L., Weisskoff, R.M., Kennedy, D.N., Makris, N., Berke, J.D., Goodman, J.M., Kantor, H.L., Gastfriend, D.R., Riorden, J.P., Mathew, R.T., Rosen, B.R., Hyman, S.E., 1997. Acute effects of cocaine on human brain activity and emotion. Neuron 19, 591–611.

Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A., Rosen, B.R., 2002. Dorsal anterior cingulate cortex: a role in reward-based decision making. Proc. Natl. Acad. Sci. U. S. A. 99, 523–528.

Byrd, R.H.L.P., 1995. A limited memory algorithm for bound constrained optimization. SIAM J. Sci. Comput. 16, 1190–1208.

Childress, A.R., Mozley, P.D., McElgin, W., Fitzgerald, J., Reivich, M., O'Brien, C.P., 1999. Limbic activation during cue-induced cocaine craving. Am. J. Psychiatry 156, 11–18.

Cools, R., Clark, L., Owen, A.M., Robbins, T.W., 2002. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. J. Neurosci. 22, 4563–4567.

Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.

Craig, A.D., 2003. A new view of pain as a homeostatic emotion. Trends Neurosci. 26, 303–307.

Critchley, H.D., Mathias, C.J., Dolan, R.J., 2001. Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron 29, 537–545.

Damasio, A.R., 1996. The somatic marker hypothesis and the possible functions of the prefrontal cortex. Philos. Trans. R. Soc. Lond., B Biol. Sci. 351, 1413–1420.

Damasio, A.R., Damasio, H., Christen, Y., 1996. Neurobiology of Decision-Making. Springer, Berlin.

Daw, N.D., Kakade, S., Dayan, P., 2002. Opponent interactions between serotonin and dopamine. Neural Netw. 15, 603–616.

Dayan, P., Hinton, G.E., Neal, R.M., Zemel, R.S., 1995. The Helmholtz machine. Neural Comput. 7, 889–904.

Egelman, D.M., Person, C., Montague, P.R., 1998. A computational role for dopamine delivery in human decision-making. J. Cogn. Neurosci. 10, 623–630.

Elliott, R., Dolan, R.J., 1998. Activation of different anterior cingulate foci in association with hypothesis testing and response selection. NeuroImage 8, 17–29.

Elliott, R., Rees, G., Dolan, R.J., 1999. Ventromedial prefrontal cortex mediates guessing. Neuropsychologia 37, 403–411.

Elliott, R., Dolan, R.J., Frith, C.D., 2000a. Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. Cereb. Cortex 10, 308–317.

Elliott, R., Friston, K.J., Dolan, R.J., 2000b. Dissociable neural responses in human reward systems. J. Neurosci. 20, 6159–6165.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn. Reson. Med. 33, 636–647.

Frith, C.D., Friston, K., Liddle, P.F., Frackowiak, R.S., 1991. Willed action and the prefrontal cortex in man: a study with PET. Proc. R. Soc. Lond., B. Biol. Sci. 244, 241–246.

Gallagher, H.L., Jack, A.I., Roepstorff, A., Frith, C.D., 2002. Imaging the intentional stance in a competitive game. NeuroImage 16, 814–821.

Gold, J.I., Shadlen, M.N., 2001. Neural computations that underlie decisions about sensory stimuli. Trends Cogn. Sci. 5, 10–16.

Gold, J.I., Shadlen, M.N., 2002. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. Neuron 36, 299–308.

Grant, S., London, E.D., Newlin, D.B., Villemagne, V.L., Liu, X., Contoreggi, C., Phillips, R.L., Kimes, A.S., Margolin, A., 1996. Activation of memory circuits during cue-elicited cocaine craving. Proc. Natl. Acad. Sci. U. S. A. 93, 12040–12045.

Grant, S., Contoreggi, C., London, E.D., 2000. Drug abusers show impaired performance in a laboratory test of decision making. Neuropsychologia 38, 1180–1187.

Huettel, S.A., Mack, P.B., McCarthy, G., 2002. Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. Nat. Neurosci. 5, 485–490.

Ihaka, R., Gentleman, R.R., 1996. A language for data analysis and graphics. J. Comput. Graph. Stat. 5 (3), 299–314 ((GENERIC) Ref Type: Journal (Full)) .

Jueptner, M., Stephan, K.M., Frith, C.D., Brooks, D.J., Frackowiak, R.S., Passingham, R.E., 1997. Anatomy of motor learning: I. Frontal cortex and attention to action. J. Neurophysiol. 77, 1313–1324.

Kahn, I., Yeshurun, Y., Rotshtein, P., Fried, I., Ben Bashat, D., Hendler, T., 2002. The role of the amygdala in signaling prospective outcome of choice. Neuron 33, 983–994.

Knutson, B., Adams, C.M., Fong, G.W., Hommer, D., 2001. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. J. Neurosci. 21, 159–164.

Knutson, B., Fong, G.W., Bennett, S.M., Adams, C.M., Hommer, D., 2003. A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. NeuroImage 18, 263–272.

London, E.D., Ernst, M., Grant, S., Bonson, K., Weinstein, A., 2000. Orbitofrontal cortex and human drug abuse: functional imaging. Cereb. Cortex 10, 334–342.

McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal prediction errors in a passive learning task activate human striatum. Neuron 38, 339–346.

Mogg, K., Mathews, A., Eysenck, M., May, J., 1991. Biased cognitive operations in anxiety: artefact, processing priorities or attentional search? Behav. Res. Ther. 29, 459–467.

Montague, P., Berns, G., 2002. Neural economics and the biological substrates of valuation. Neuron 36, 265.

O'Doherty, J., Kringelbach, M.L., Hornak, J., Andrews, C., Rolls, E.T., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat. Neurosci. 4, 95–102.

O'Doherty, J.P., Critchley, H.D., Deichmann, R., Dolan, R.J., 2003a. Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. J. Neurosci. 23, 7931–7939.

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

Pagnoni, G., Zink, C.F., Montague, P.R., Berns, G.S., 2002. Activity in human ventral striatum locked to errors of reward prediction. Nat. Neurosci. 5, 97–98.

Passingham, R.E., Toni, I., Rushworth, M.F., 2000. Specialisation within the prefrontal cortex: the ventral prefrontal cortex and associative learning. Exp. Brain Res. 133, 103–113.

Paulus, M.P., Hozack, N., Zauscher, B., McDowell, J.E., Frank, L., Brown, G.G., Braff, D.L., 2001. Prefrontal, parietal, and temporal cortex networks underlie decision-making in the presence of uncertainty. NeuroImage 13, 91–100.

Paulus, M.P., Hozack, N.E., Zauscher, B.E., Frank, L., Brown, G.G., Braff, D.L., Schuckit, M.A., 2002. Behavioral and functional neuroimaging evidence for prefrontal dysfunction in methamphetamine-dependent subjects. Neuropsychopharmacology 26, 53–63.

Pearce, J.M., Bouton, M.E., 2001. Theories of associative learning in animals. Annu. Rev. Psychol. 52, 111–139.

Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. Nature 400, 233–238.

Pochon, J.B., Levy, R., Fossati, P., Lehericy, S., Poline, J.B., Pillon, B., Le Bihan, D., Dubois, B., 2002. The neural system that bridges reward and cognition in humans: an fMRI study. Proc. Natl. Acad. Sci. U. S. A. 99, 5669–5674.

Pouget, A., Dayan, P., Zemel, R.S., 2003. Inference and computation with population codes. Annu. Rev. Neurosci. 26, 381–410.

Rahman, S., Sahakian, B.J., Hodges, J.R., Rogers, R.D., Robbins, T.W., 1999. Specific cognitive deficits in mild frontal variant frontotemporal dementia. Brain 122, 1469–1493.

Rahman, S., Sahakia, J., Cardinal, N., Rogers, R., Robbins, T., 2001. Decision making and neuropsychiatry. Trends Cogn. Sci. 5, 271–277.

Real, L.A., 1991. Animal choice behavior and the evolution of cognitive architecture. Science 253, 980–986.

Robbins, T.W., Everitt, B.J., Ryan, C.N., Marston, H.M., Jones, G.H., Page, K.J., 1989. Comparative effects of quisqualic and ibotenic acid-induced lesions of the substantia innominata and globus pallidus on the acquisition of a conditional visual discrimination: differential effects on cholinergic mechanisms. Neuroscience 28, 337–352.

Rode, C., Cosmides, L., Hell, W., Tooby, J., 1999. When and why do people avoid unknown probabilities in decisions under uncertainty? Testing some predictions from optimal foraging theory. Cognition 72, 269–304.

Rogers, R.D., Everitt, B.J., Baldacchino, A., Blackshaw, A.J., Swainson, K., Wynne, K., Baker, N.B., Hunter, J., Carthy, T., Booker, E., London, M., Deakin, J.F., Sahakian, B.J., Robbins, T.W., 1999. Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. Neuropsychopharmacology 20, 322–339.

Salinas, E., Abbott, L.F., 1996. A model of multiplicative neural responses in parietal cortex. Proc. Natl. Acad. Sci. U. S. A. 93, 11956–11961.

Schmitt, M., 2002. On the complexity of computing and learning with multiplicative neural networks. Neural Comput. 14, 241–301.

Schultz, W., 2002. Getting formal with dopamine and reward. Neuron 36, 241–263.

Schultz, W., Dickinson, A., 2000. Neuronal coding of prediction errors. Annu. Rev. Neurosci. 23, 473–500.

Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science 275, 1593–1599.

Shadlen, M.N., Newsome, W.T., 1996. Motion perception: seeing and deciding. Proc. Natl. Acad. Sci. U. S. A. 93, 628–633.

Shadlen, M.N., Newsome, W.T., 2001. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiol. 86, 1916–1936.

Suri, R.E., Schultz, W., 1999. A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. Neuroscience 91, 871–890.

Suri, R.E., Schultz, W., 2001. Temporal difference model reproduces anticipatory neural activity. Neural Comput. 13, 841–862.

Tapert, S.F., Cheung, E.H., Brown, G.G., Frank, L.R., Paulus, M.P., Schweinsburg, A.D., Meloy, M.J., Brown, S.A., 2003. Neural response to alcohol stimuli in adolescents with alcohol use disorder. Arch. Gen. Psychiatry 60, 727–735.

Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. Science 211, 453–458.

Volkow, N.D., Fowler, J.S., 2000. Addiction, a disease of compulsion and drive: involvement of the orbitofrontal cortex. Cereb. Cortex 10, 318–325.

Wang, G.J., Volkow, N.D., Fowler, J.S., Cervany, P., Hitzemann, R.J., Pappas, N.R., Wong, C.T., Felder, C., 1999. Regional brain metabolic activation during craving elicited by recall of previous drug experiences. Life Sci. 64, 775–784.

Ward, B.D., Deconvolution analysis of FMRI time series data. 5-1-2002. Medical College of Wisconsin. AFNI 3dDeconvolve Documentation. (GENERIC) Ref Type: Report.

Wasserman, E.A., Miller, R.R., 1997. What's elementary about associative learning? Annu. Rev. Psychol. 48, 573–607.

Zysset, S., Huber, O., Ferstl, E., von Cramon, D.Y., 2002. The anterior frontomedian cortex and evaluative judgment: an fMRI study. NeuroImage 15, 983–991.