Behavioral and neural predictors of upcoming decisions

M. X COHEN
University of Bonn, Bonn, Germany
and University of California, Davis, California

and

C. RANGANATH
University of California, Davis, California

Although it is widely known that brain regions such as the prefrontal cortex, the amygdala, and the ventral striatum play large roles in decision making, their precise contributions remain unclear. Here, we used functional magnetic resonance imaging and principles of reinforcement learning theory to investigate the relationship between current reinforcements and future decisions. In the experiment, subjects chose between high-risk (i.e., low probability of a large monetary reward) and low-risk (high probability of a small reward) decisions. For each subject, we estimated value functions that represented the degree to which reinforcements affected the value of decision options on the subsequent trial. Individual differences in value functions predicted not only trial-to-trial behavioral strategies, such as choosing high-risk decisions following high-risk rewards, but also the relationship between activity in prefrontal and subcortical regions during one trial and the decision made in the subsequent trial. These findings provide a novel link between behavior and neural activity by demonstrating that value functions are manifested both in adjustments in behavioral strategies and in the neural activity that accompanies those adjustments.

Recent research in cognitive neuroscience has focused on understanding the neural mechanisms of decision making under uncertainty (Kahn et al., 2002; Krawczyk, 2002; Manes et al., 2002; Sanfey, Hastie, Colvin, & Grafman, 2003; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). Although existing neuropsychological and neuroimaging research has helped to identify the brain regions involved in decision making, questions remain regarding the precise functions of these regions. For example, substantial evidence from economic and cognitive decision making research has demonstrated that the outcome of one decision strongly influences the following decision (Erev & Roth, 1998; Mookherjee & Sopher, 1994; Sarin & Vahid, 2001), but how brain regions might mediate these trial-to-trial adjustments in behavioral strategies is unknown. If decision making is a dynamic process of choosing a decision, evaluating the outcome, and adjusting future behavior accordingly, brain regions involved in decision making should exhibit activity that reflects trial-to-trial adjustments in behavior that occur as a result of the evaluation process. Thus, a more complete understanding of the neurobiological mechanisms of decision making must take into account how outcomes and neural processes during one trial affect decisions on subsequent trials.

Reinforcement learning theory provides a framework for assessing the predictive value of current outcomes for future decisions (Erev & Roth, 1998; Sarin & Vahid, 2001). Models based on reinforcement learning theory have been used copiously in the realms of machine learning and economic decision making (Barto, 1995; Sutton, 1992), and recently, researchers have utilized it when investigating the neural processes underlying reward-based learning (Barraclough, Conroy, & Lee, 2004; Egelman, Person, & Montague, 1998; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; Paulus, Feinstein, Tapert, & Liu, 2004; Schultz, 2004). These researchers have demonstrated that principles of reinforcement learning theory can be used to predict activity in dopaminergic brain regions, such as the orbitofrontal cortex and the striatum, that correlate both with the predictive power of the cues associated with rewards (Friston, Tononi, Reeke, Sporns, & Edelman, 1994; Montague & Berns, 2002; O'Doherty et al., 2003; Schultz, 2004) and with changes in performance during learning (O'Doherty et al., 2003; Paulus et al., 2004). However, it remains unknown whether reinforcement learning theory can help explain the neural activity that underlies trial-to-trial adjustments in behavioral strategies.

A key prediction of reinforcement learning theory is that individuals use the outcomes of their decisions as reinforcements to update or adjust their decision strategies.

Importantly, the extent to which current outcomes affect the value of future decisions varies for different individuals and can be influenced by experimental variables, such as the probabilities and magnitudes of potential rewards. For example, after receiving a risky reward, some individuals might place a higher value on a high-risk decision option, whereas others might place a lower value on this decision option. The degree to which the outcomes of different types of decisions affect the value of future decisions can be described by parameters termed *value functions*, which can be mathematically estimated for different kinds of decisions for different subjects.

In the present study, we used a novel application of reinforcement learning theory to assess how individual differences in value functions predict the relationship between neural activity during current trials and decisions chosen in subsequent trials. In our functional magnetic resonance imaging (fMRI) study, subjects chose between two economically equivalent (i.e., equal in expected value) decisions—a high-risk (i.e., low probability of a large monetary reward) and a low-risk (high probability of a small reward) decision—and were probabilistically rewarded. Our analyses focused on how these value functions predicted the relationship between neural activity during one trial and the decision chosen in the following trial (e.g., activity during high-risk rewards that were followed either by high-risk or low-risk decisions).

METHOD

Subject

Seventeen subjects (8 male, 22–27 years of age) were recruited from the University of California, Davis, student community. The subjects gave informed consent prior to the experiment and were paid for their participation by the hour.

Procedure

The task involved making rapid behavioral responses in order to win money. On each trial, the subjects first saw a visual cue for 400 msec that indicated the condition and response requirements (see below). When they saw this cue, the subjects had to rapidly indicate their decision by pressing a button or withholding a response (see below). Three hundred milliseconds after the offset of the cue, the amount of money the subjects were rewarded in that trial appeared on the screen (for example, "+$2.50"). A variable intertrial interval (range: 2–8 sec) separated the trials. The subjects had to rapidly decide which of two economically equivalent (i.e., equal in expected value) decisions they wished to make: a low-risk decision, in which they were very likely to win a small reward (80% chance of $1.25 and 20% chance of $0.00), or a high-risk decision, in which they were less likely to win a large reward (40% chance of $2.50 and 60% chance of $0.00). The subjects indicated each decision by either responding or withholding a response, depending on the shape of the cue. For example, a white circle meant that subjects would press the button to indicate a low-risk decision or withhold a response to indicate a high-risk decision, whereas a white square would indicate the opposite. The purpose of having the subjects press a button during half of the trials and withhold responses during the other half of the trials was to counterbalance any effects of response inhibition across trial types. In the analyses reported, the results did not differ according to this variable, so we collapsed across press and withhold response conditions. Additional control trials were included in which the subjects were 100% likely to receive a reward ($2.50) if they responded appropriately to the cue (i.e., press or

withhold, according to the cue). These trials are not discussed in the present article. The subjects were trained extensively on a behavioral version of the task prior to the experiment and were told the probabilities and outcomes associated with each cue type and response. This training minimized initial cue–reward learning effects. There were a total of 300 decision trials. In addition to this task, each subject completed a visuomotor response task that was used to empirically derive a subject-specific hemodynamic response function (HRF; see below; Aguirre, Zarahn, & D'Esposito, 1998).

MRI Acquisition and Data Processing

MRI data were collected on a 1.5T GE Signa scanner at the University of California, Davis, Research Imaging Center. Functional imaging was done with a gradient echo EPI sequence (TR = 2,000, TE = 40, FOV = 220, 64 × 64 matrix, voxel size = 3.475 × 3.475 × 5 mm), with each volume consisting of 22 oblique axial slices (tilt angle: ~−15° from ac–pc line). In pilot studies, we determined that this set of parameters maximized signal-to-noise ratios and blood oxygenation level dependent (BOLD) contrast in regions such as the orbitofrontal cortex and the amygdala. Coplanar and high-resolution T1-weighted images were also acquired from each subject. EPI data were realigned to the first volume, coregistered with the anatomical scan, spatially normalized to MNI space (Brett, Johnsrude, & Owen, 2002) resampled to 3.5-mm isotropic voxels, and spatially smoothed with an 8-mm FWHM kernel, using SPM99 software.

Data Analyses

Event-related BOLD responses associated with each condition were analyzed using a general linear model (Worsley & Friston, 1995) in VoxBo software (www.voxbo.org). All the models incorporated empirically derived estimates of intrinsic temporal autocorrelation (Zarahn, Aguirre, & D'Esposito, 1997) and filters to attenuate frequencies above 0.24 Hz and below 0.01 Hz. HRFs were empirically derived for each subject, using BOLD responses in the central sulcus during the visuomotor response task (Aguirre et al., 1998). These HRFs were used to model BOLD responses to each event type in all the analyses. In addition, the mean of each scanning run, the global signal (orthogonalized with respect to the design matrix; Desjardins, Kiehl, & Liddle, 2001), and an intercept were included as nuisance covariates. Group analyses were performed by entering statistical parametric maps of beta values from each subject for each contrast into a second-level, random-effects, one-sample t test, in which the mean estimate across subjects at each voxel was tested against zero. For the value function correlation maps, individual differences in value functions were regressed upon each voxel in the brain during contrasts of interest. Significant regions of activation were identified using an uncorrected two-tailed threshold of $p < .001$ and a cluster threshold of 6 contiguous voxels. In the figures, activations are overlaid on a single subject's normalized T1-weighted image, using MRIcro software (Rorden & Brett, 2000).

Reinforcement Learning Theory Parameters

Under reinforcement learning theory, the probability of making the high-risk decision on a particular trial, t, is the logit transform of the difference in the values of the high-risk and the low-risk decisions on that trial (Barraclough et al., 2004; Luce, 1999). In other words:

$$p_t(\text{highrisk}) = \frac{\exp[v_t(\text{highrisk})]}{\exp[v_t(\text{lowrisk})] + \exp[v_t(\text{highrisk})]},$$

where v_t is the value of each decision option on trial t. This value is updated following each trial:

$$v_{t+1}(\text{highrisk}) = \alpha v_t(\text{highrisk}) + \partial(\text{highrisk})$$
$$v_{t+1}(\text{lowrisk}) = \alpha v_t(\text{lowrisk}) + \partial(\text{lowrisk}),$$

where α is a discount parameter that describes the decrease in the effect of previous outcomes (i.e., a *forgetting* parameter) and ∂ is the parameter (termed *value function*) that describes the update in

Table 1
Intercorrelations of Value Functions and Behavioral Response Patterns

	Risk	HR	HN	LR	LN	Discount	High After Reward	High After Nonreward	Low After Reward
HR	.49*	1.00							
HN	.54*	.82*	1.00						
LR	−.15	.06	−.06	1.00					
LN	.19	.40	.36	.78*	1.00				
Discount	−.28	−.12	−.03	−.38	−.36	1.00			
High after reward	.81*	.82*	.74*	.02	.47	−.25	1.00		
High after nonreward	.85*	.59*	.73*	.10	.49*	−.24	.87*	1.00	
Low after reward	−.89*	−.22	−.31	.42	.15	.20	−.59*	−.65*	1.00
Low after nonreward	−.73*	−.03	−.01	.30	.26	.24	−.40	−.47	.90*

Notes—Risk, proportion of high-risk decision options chosen; HR, high-risk reward; HN, high-risk nonreward; LR, low-risk reward; LN, low-risk nonreward; high after reward, high-risk decisions chosen following high-risk rewards; High after nonreward, high-risk decisions chosen following high-risk rewards; Low after nonreward, low-risk decisions chosen following low-risk rewards. Boldface numbers with asterisks indicate that the correlation is significant at $p < .05$.

value according to whether trial t was rewarded or not rewarded. During trials in which the subjects made the high-risk decision, this parameter could take on one of three values: high-risk reward (HR), the change in value during a high-risk trial that was rewarded; high-risk nonreward (HN), the change in value during a high-risk trial that was not rewarded; and 0 during trials in which the subject made the low-risk decision. For low-risk trials, the parameters are LR and LN. Thus, if an individual has a large estimate (regardless of the sign) on, for example, HR, it implies that that individual values high-risk rewards relatively highly and those events have a large impact on the subsequent change in value of that decision, and if the individual has an estimate close to 0, winning a high-risk reward has little impact on that individual's future value of that decision. The sign of the parameter indicates whether the value of that decision option (high-risk for HR and HN and low-risk for LR and LN) increases or decreases in the following trial. Note that these parameters are estimates of the change in the subjective value of a decision option according to current reinforcements and are not estimates of the preference for or use of particular decision strategies. A maximum likelihood minimization procedure (Barraclough et al., 2004; Lagarias, Reeds, Wright, & Wright, 1998) was used to estimate these five parameters for each subject (HR, HN, LR, LN, and the discount factor) in Matlab (www.mathworks.com). The estimation procedure converged for all the subjects. No significant correlations were observed between the discount factor and brain activity or behavior in any conditions and, thus, are not discussed further.

RESULTS

Behavioral Results

Consistent with findings from behavioral studies demonstrating that people tend to be risk averse (i.e., prefer decision options with higher probability outcomes; Tversky & Kahneman, 1981), the subjects made the low-risk decision more often than the high-risk decision [61% ± 3 vs. 39% ± 3 ($M ± SEM$), respectively; one-sample t test differs from chance: $t(16) = -3.17, p < .01$], but response times were equivalent for both decision options [low-risk, 359 msec ± 10; high-risk, 366 msec ± 8; $t(16) = 0.94$, n.s.]. In addition to this overall decision bias, the subjects used trial-to-trial decision-making strategies: They were more likely to make low-risk than high-risk decisions after low-risk rewards [63%, where

50% is no strategy use; $t(16) = 3.68, p < .01$] and low-risk nonrewards [64%; $t(16) = 3.97, p < .01$] and were less likely to make high-risk than low-risk decisions after high-risk nonrewards [42%; $t(16) = -2.11, p = .05$], but not after high-risk rewards [44%; $t(16) = -1.73$, n.s.]. We calculated these strategies using conditional probabilities (e.g., the number of high-risk decisions chosen, given that the previous trial was a high-risk reward, divided by the number of high-risk rewards).

The overall proportion of high-risk decisions correlated significantly with both HR and HN value functions (HR, $r = .49, p = .047$; HN, $r = .54, p = .025$), but not with LR and LN value functions ($r = -.15$ and .19, respectively; $p > .4$]. The use of trial-to-trial strategies, such as the probability of making a high-risk decision after a high-risk reward was also correlated with both HR and HN parameters (HR, $r = .82, p < .001$; HN, $r = .74, p < .001$], and the probability of making a low-risk decision after a low-risk reward was correlated marginally with the LR value function ($r = .42, p = .095$). These and other correlations among behavioral strategies and value functions are presented in Table 1.

fMRI Results

Although the focus of this article is on reinforcement learning theory and activity during current trials as a function of decisions made in subsequent trials, for completeness we will report the overall effects of brain activity during current trials, regardless of the outcome and decision made in the following trial. The decision to take the high-risk over the low-risk decision was associated with increased activation in the anterior cingulate, the left dorsomedial thalamus, and the bilateral ventral striatum, regions that have previously been identified in cognitive and behavioral control mechanisms that are related to reward processes (Bush et al., 2002; McCullough & Salamone, 1992; Shidara & Richmond, 2002; Stern & Passingham, 1994; Whishaw & Kornelsen, 1993). No regions exhibited increased activation during low-risk, as compared with high-risk, decisions. MNI coordinates for

these and all other activations reported in this article are listed in Table 2.

Our fMRI analyses focus on examining neural activity during different combinations of decision (high risk or low risk) and outcome (rewarded or nonrewarded) as a function of the decision made in the subsequent trial (high risk or low risk). These contrasts therefore reflect the relationship between brain activity during a current decision–outcome combination and the decision made on the subsequent trial. For each of the contrasts, positive activation values indicate that increased activity during the current trial type predicted a high-risk decision in the following trial, and negative activation values indicate that increased activity during the current trial predicted a low-risk decision in the subsequent trial. To examine the role of value functions, individual differences in all five value functions were correlated with activity in each voxel for each contrast across subjects.

Our first set of analyses was focused on adjustments in decision making following high-risk rewards. Ignoring individual differences in value functions, no regions of the brain exhibited significant activations during high-risk rewards that predicted the subjects' next decision. However, when individual differences in value functions were regressed upon these images, several significant findings emerged. We found that HR values, which predicted the making of high-risk decisions following high-risk rewards on the behavioral level (see Table 1), also predicted brain activity during high-risk reward trials as a function of whether, on the next trial, a high- or a low-risk decision was made. This activity was observed in the right amygdala ($r = .84$), the anterior orbital prefrontal cortex (PFC; $r = .83$), the right dorsal cingulate ($r = .80$), the left putamen ($r = .79$), and the right and bilateral middle frontal gyrus ($r = .81$; see Figure 1). Importantly, the magnitudes and signs of activation are consistent with what one would predict from the meaning of the value functions (see the scatterplots in Figure 1): For individuals with a high HR value function, increased activity predicted high-risk decisions; for individuals with a negative HR value function, increased activity predicted low-risk decisions; and for individuals with a value function close to zero, activity during high-risk rewards did not predict subsequent decisions. No significant correlations were observed with HN or LR value functions. The LN value function was significantly correlated with the bilateral lingual gyrus (left, $r = .89$; right, $r = .86$) and the posterior insula ($r = .81$).

We next examined neural activity during adjustments in decision making following high-risk nonrewards. As in the previous analysis, when individual differences were ignored, no significant regions of activation were observed. However, when individual differences in value functions were taken into account, several significant correlations emerged. Consistent with the fact that both HR and HN predicted the making of high-risk decisions following high-risk nonrewards (see Table 1), the HR value function was significantly correlated with activity in the right ventral striatum ($r = .75$; see Figure 2B), and the HN value function was significantly correlated with activity in the left dorsal cingulate ($r = .82$; see Figure 2A), the left visual cortex (BA 18), and the left cerebellum. No overlapping correlations were observed between these two value functions, despite their being correlated with each other and with certain behavioral strategies. No correlations were observed with LR or LN functions.

Our second set of analyses was focused on adjustments in decision making following low-risk decisions. We first examined activity during low-risk rewards followed by high- versus low-risk decisions. When individual differences in value functions were ignored, no activations were observed. Furthermore, no regions exhibited significant correlations with HR, HN, or LN value functions. Differences in the LR value function, however, which marginally predicted the making of low-risk decisions following low-risk rewards ($r = .42, p = .095$), predicted activations in the bilateral inferior temporal gyrus (left, $r = .78$; right, $r = .88$; Figure 3A).
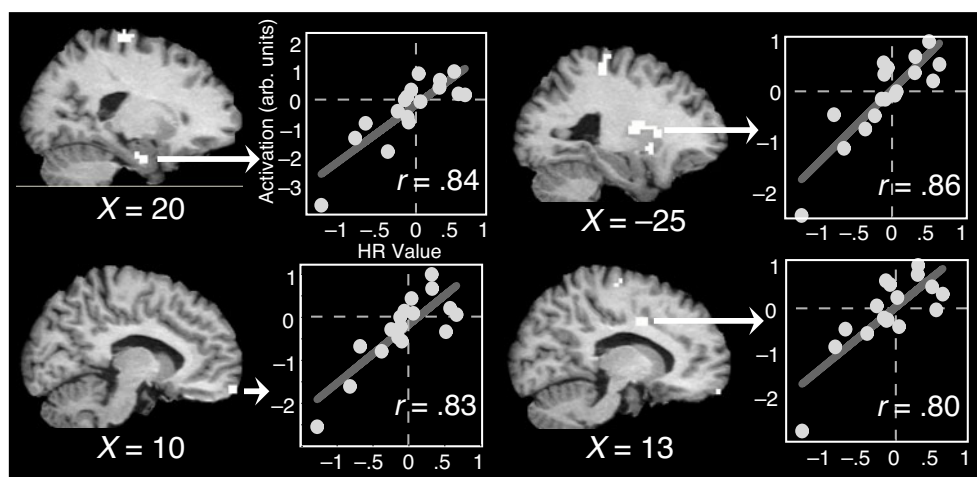
Table 2
Peak Activation Coordinates (MNI) and *t* Values
of Significant Activations

Region	BA	*t*	*x*	*y*	*z*
High- vs. Low-Risk Decision					
Anterior Cingulate	32	6.78	10	43	28
L. Ventral striatum		5.43	10	10	0
R. Ventral striatum		5.18	−13	7	−5
L. Thalamus		4.86	−7	−11	−3
High-Risk Rewards Followed by High- vs. Low-Risk Decisions					
Correlation with HR					
R. Amygdala	36	5.65	20	−8	−28
R. Cingulate	24	5.15	13	1	38
L. Putamen		6.56	−28	−3	14
R. Anterior orbital gyrus	11	4.3	10	64	−18
L. Middle frontal gyrus	6	5.05	−32	−3	38
R. Middle frontal gyrus	6	4.61	27	2	46
Correlation with LN					
L. Lingual	37	5.99	−24	−49	−4
R. Lingual	37	7.72	26	−51	0
L. Insula		5.47	−38	−26	18
R. Insula		5.26	41	−33	18
High-Risk Nonrewards Followed by High- vs. Low-Risk					
Correlation with HR					
R. Ventral striatum		4.61	7	4	−7
Correlation with HN					
L. Cingulate	24/32	6.01	−17	8	35
L. Lingual gyrus	18	6.19	−21	−91	−2
L. Cerebellum		5.12	−34	-65	−31
Low-Risk Rewards Followed by High- vs. Low-Risk Decisions					
Correlation with LR					
R. Inf. Temporal Gyrus	20	4.47	−27	−2	−42
L. Inf. Temporal Gyrus	20	4.22	34	2	−46
Low-Risk Nonrewards Followed by High- vs. Low-Risk Decisions					
Correlation with LN					
Cingulate	32	5.04	−4	40	32

Note—Brodmann areas (BAs), MNI coordinates (*x, y, z*), and *t* values of regions active in each contrast are shown. L, left; R, right; HR, high-risk reward; LN, low-risk nonreward; HN, high-risk nonreward; LR, low-risk reward.

Figure 1. Regions in the prefrontal cortex, the amygdala, and the putamen exhibited significant correlations with the high-risk reward (HR) value function during high-risk rewards that were followed by high-versus low-risk decisions. Scatterplots depict the relationships between individual differences in HR value functions and activation magnitudes from identified regions. The activation units are maximum-value normalized estimates of the difference in activity between choosing a high- versus a low-risk decision on the following trial. Thus, positive numbers indicate that greater activity during the current trial predicted a high-risk decision on the subsequent trial, negative numbers indicate that greater activity during the current trial predicted a low-risk decision on the next trial, and zero indicates that activity during the current trial did not predict the subsequent decision. For ease of comparison, dotted lines are drawn on the zero lines of the value function estimates and predictive neural activity. Consistent with behavioral results (see Table 1), the HR value function predicted the relationship between neural activity during high-risk reward trials and the decision made on the following trial.

Finally, we examined neural activity during adjustments in decision-making strategies following low-risk nonrewards. As in the previous analyses, no activations were observed when individual differences were ignored, and no correlations were observed with HR, HN, or LR value functions. The LN value function, however, significantly predicted activity in the anterior cingulate cortex ($r = .88$; see Figure 3B).

DISCUSSION

In this study, we investigated how brain activity and principles of reinforcement learning theory can be combined to better understand the roles of the PFC and subcortical structures in decision making. We found that both overall behavioral patterns and trial-to-trial decision-making strategies were predicted by individual differences in value functions. Although this was not surprising, given that value functions were estimated using behavioral data, it demonstrates that these are meaningful individual differences that are related to task performance. Critically, individual differences in value functions were correlated with activity in relatively specific areas of the brain that have been previously identified as being involved in reward-based decision making, such as the amygdala, the anterior cingulate, and the orbital PFC.

Models of Reinforcement Learning Theory
Reinforcement learning theory is a broad theory that describes ways for organisms to learn optimal behaviors

in order to maximize rewards. To date, the most common application of reinforcement learning theory in neuroscience has been the temporal difference reinforcement learning model. In a typical temporal difference learning experiment, an arbitrary cue is presented (e.g., a red light) and is followed by a reward several seconds later. The temporal difference model is a set of equations that describes how organisms learn to associate the arbitrary cues (or actions) with the future reward. Researchers have successfully used this model to predict both learning curves and responses of dopamine neurons during learning in several species, including bees, rats, monkeys, and humans (Egelman et al., 1998; Montague & Berns, 2002; O'Doherty et al., 2003; Paulus et al., 2004).

However, the temporal difference model is not well suited for our paradigm, because our main focus was not on how the brain learns to associate an arbitrary cue with a future reward (which, in our study, had already been learned before the scanning session began) but, rather, on how reinforcement information is used to guide future decisions throughout the task. However, the dynamic process of using reinforcements to guide future decisions remains a strong form of learning. Indeed, responses of dopamine neurons continue to reflect changes in reward signals long after initial associations between cues and rewards have been learned (Fiorillo, Tobler, & Schultz, 2003), and in our study, the subjects' use of trial-to-trial strategies demonstrated that outcomes of current decisions had a significant effect on their subsequent decision strategies.
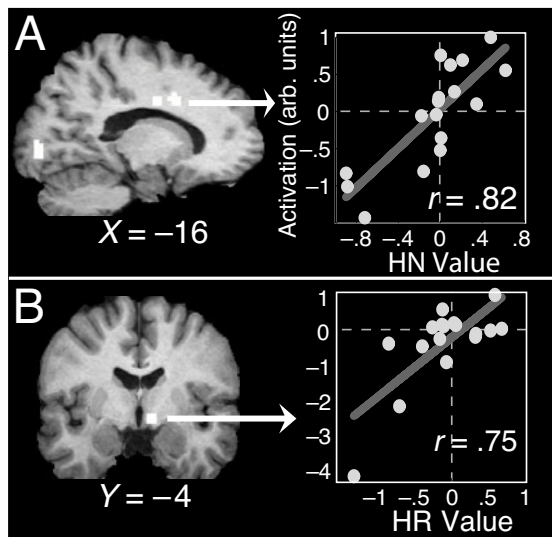
Figure 2. Regions showing significant correlations with (A) high-risk nonreward (HN) and (B) high-risk reward (HR) value functions during HN trials that were followed by high- versus low-risk decisions. Activations show that individual differences in the HN value function (panel A) predicted activation magnitudes in the cingulate gyrus and that the HR parameter (panel B) predicted activation magnitudes in the ventral striatum. Both of these value functions additionally predicted the subjects' decisions on subsequent trials. Conventions are the same as those in Figure 1.

immediately rewarded. Thus, the learning that we modeled in our study was not the initial pairing of cues and their potential rewards but, rather, of how reinforcement signals are used to guide future reward-seeking behavior. Nonetheless, differences between various models derived from reinforcement learning theory are slight, the underlying mathematical formulae are similar, and the differences are more related to the nature of the problem (e.g., long vs. no temporal lag between the cue and the reward) than to the meaning or interpretation of the value functions (see Sutton & Barto, 1998, for an extensive comparison of the similarities and differences between various reinforcement learning models).

Value Functions and the Relationship Between Localized Brain Activity and Future Decisions

Several regions of the brain that have been previously implicated in reward-based decision making showed significant activations in our study. For example, activity within different (and nonoverlapping) subregions of the anterior cingulate cortex predicted subsequent decisions for three of the four value functions we focused on (HR, HL, and LN), and the anterior cingulate was also significantly more active during current trials in which the subjects made high-risk, rather than low-risk, decisions (when future decisions were ignored). The anterior cingulate has been widely implicated in cognitive control and the monitoring of ongoing behavior for errors or conflict (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Devinsky, Morrell, & Vogt, 1995; van Veen & Carter, 2002), as well as in rapidly evaluating the emotional significance of outcomes (Gehring & Willoughby, 2002) and modifying autonomic levels of arousal (Critchley, Corfield, Chandler, Mathias, & Dolan, 2000). In one recent study, Kerns and colleagues have demonstrated that

Our study is thus more amenable to a model derived from reinforcement learning theory that focuses on the "two-armed bandit" problem (Goeree & Holt, 1999; Sutton & Barto, 1998), which assumes that an individual must choose one of two or more options with some knowledge about the reinforcement contingencies and is
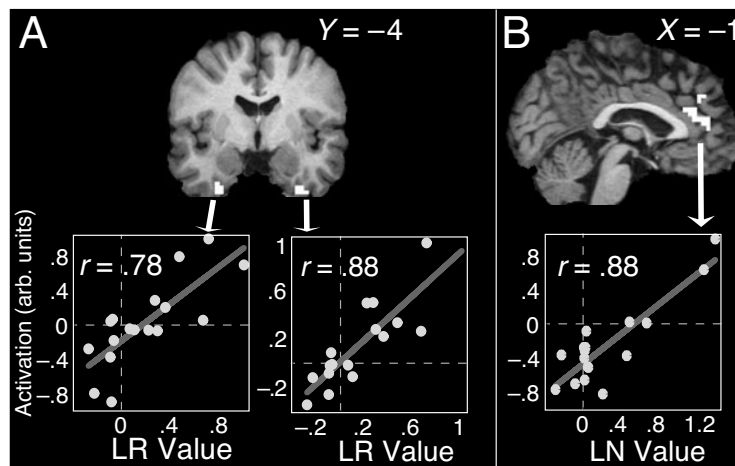

Figure 3. The low-risk reward (LR) and low-risk nonreward (LN) value functions predicted the relationship between activity in the bilateral anterior entorhinal cortex (panel A) during LR trials and the subsequent decision. The LN value function also predicted activity during LN trials (panel B). Conventions are the same as those in Figure 1.

during a Stroop task (where response errors and conflicts are common), activity in the anterior cingulate during high-conflict trials predicted adjustments in performance during subsequent trials (Kerns et al., 2004). Similar findings have been obtained with EEG error-related negativity, which likely originates in the anterior cingulate (Ridderinkhof, Nieuwenhuis, & Bashore, 2003). Our findings are consistent with the idea that the anterior cingulate implements control processes related to the use of performance feedback to guide future behaviors and, in addition, suggest that this relationship might be linked to individual differences in the value of decision options. It is possible that our subjects experienced some degree of conflict during high-risk decisions and that different levels of conflict were related both to anterior cingulate activity and to adjustments in behavioral strategies. This hypothesis could be assessed in future studies by using a paradigm similar to ours but additionally collecting online or postexperimental ratings of levels of response conflict associated with high- versus low-risk choices.

Individual differences in the HR value function predicted the relationship between activity in the orbital PFC (BA 11) and in the amygdala and decisions made in subsequent trials. Orbital PFC has been heavily implicated in both risky decision making (Bechara, Damasio, & Damasio, 2000, 2003) and behavioral strategy reversals (Cools, Clark, Owen, & Robbins, 2002), suggesting that neurons in the orbital PFC monitor ongoing reinforcements and help guide behavior in the service of gaining more reinforcements (Rolls, 2000). The somatic marker hypothesis proposes that individuals form associations between emotional or physiological states and outcomes of their decisions and that the orbital PFC plays a critical role in using those associations to bias upcoming decision processes (Bechara et al., 2000; Bechara, Damasio, Tranel, & Damasio, 1997). Thus, patients with damage to the orbital PFC are labeled as being insensitive to the future consequences of their current actions (Bechara et al., 2000). Our finding that activity in the anterior orbital PFC during high-risk rewards predicted the subjects' subsequent decisions (see Figure 1, lower left) is consistent with a role for this region in considering or deliberating over upcoming risky decisions and demonstrates that activity in the orbital PFC not only reflects ongoing changes in reinforcements, but additionally predicts trial-to-trial adjustments in risk-taking behavior.

Patients with damage to the amygdala are also impaired in decision-making and reward-based learning tasks (Bechara, Damasio, Damasio, & Lee, 1999; Rolls, 2004). Our findings suggest that the impairments observed in patients with damage to the orbital PFC or the amygdala may be partly related to difficulties in adjusting future behavior in light of current outcomes. The fact that value functions predicted the relationship between activity in these regions and future behavior is consistent with the observation that activity in these regions correlates with relative subjective preference of rewards (Arana et al., 2003; Hikosaka & Watanabe, 2000) and suggests that some reward

preferences can be mathematically estimated and used to predict brain–behavior relationships.

We observed activation in two regions of the striatum—the left putamen and a ventral region that likely corresponded to the nucleus accumbens. Activity in the nucleus accumbens has been shown to reflect reward prediction errors—times when expected rewards were not received (Montague & Berns, 2002)—and has been linked with reward-based motivation, learning processes, and addictive disorders (Berridge & Robinson, 2003; Depue & Collins, 1999; Volkow, Fowler, Wang, & Swanson, 2004). In our study, when high-risk rewards were not received, ventral striatal activity predicted the decision the subjects made in the subsequent trial, demonstrating that activity in this region also reflects future behavior related to the receiving of rewards. In addition, changes in dopamine transmission have been observed in the human putamen during a variable reward schedule in a decision-making task (Zald et al., 2004), suggesting a role for this region in receiving uncertain and probabilistic rewards. The fact that individual differences moderated the relationships between brain activity in the striatum and subsequent decisions suggests that individual differences play an important role in the link between reward processing and reward-seeking behavior.

Regions in the anterior temporal cortex were also sensitive to future decisions. Although most investigations of this brain region have focused on visual associative learning and memory formation (Fell, Klaver, Elger, & Fernandez, 2002; Miyashita, Okuno, Tokuyama, Ihara, & Nakajima, 1996; Ranganath, Cohen, & Brozinsky, in press), Murray and colleagues have shown that the anterior temporal cortex is required for learning visually guided reward associations (Liu, Murray, & Richmond, 2000). The present results suggest that activity in this region may also signal aspects of nonvisual higher level reward associations. Further research will be required to more precisely clarify what kinds of processes this region performs that are related to decision making.

Significant correlations between predictive brain activity and value functions were also observed in the posterior medial occipital and cerebellar areas, regions not typically thought of as being critical for decision making and reward processing. However, in a recent fMRI study, activations in the cerebellum and the medial occipital cortex were observed when trials in which subjects had to make a decision were compared with those in which they did not (Blackwood et al., 2004). It is possible that these activations represent processes, such as response preparation or attentional shifting, that support decision-making reward-seeking behavior (Ramnani & Miall, 2003; Roesch & Olson, 2003).

Discount Factor and Decision-Related Brain Activity

In addition to the value functions, we calculated a discount (i.e., forgetting) parameter that represents the extent to which the history of wins and losses are taken into account when current decision options are chosen. We

observed no significant relationship between individual differences in this parameter and predictive brain activity. It is not likely that this was due to a restricted range of differences on this variable, because the range of values was similar to those observed for the value functions and the average discount factor across subjects was significantly greater than zero [$M = .48$, $SD = .67$; $t(16) = 2.98$, $p < .01$]. Our results might seem to contrast with results recently reported by Tanaka and colleagues (Tanaka et al., 2004), in which they reported that a discount factor predicted brain activity in the striatum, the medial PFC, and the insula during a decision-making task. However, for two reasons, it is difficult to directly compare our results with those in Tanaka et al.

First, in our study, current decisions and outcomes did not influence future outcomes (and the subjects were instructed that this was the case), whereas in Tanaka et al.'s (2004) study, subjects' choices affected both current and future outcomes. Thus, in contrast to Tanaka et al.'s study, reinforcements in our task did not serve as a useful guide for predicting future rewards, and thus, as would be expected, this parameter predicted neither how the subjects used reinforcements to guide future behavior (see Table 1) nor brain activity for trial-to-trial strategies. Second, we estimated discount parameters on the basis of behavioral data for each subject, whereas in Tanaka et al.'s study, various discount parameters were chosen for all the subjects a priori, and these values were used in regressions to determine whether and where activity in the brain corresponded to discounting at various time scales. Thus, it appears that although representations of discount factors exist in the brain and can be empirically observed, these representations might not significantly influence behavior and brain activity during all decision-making situations, especially when they need not be relevant, as in our task.

Caveats and Conclusions

A few caveats should be noted when the present results are interpreted. First, because our subjects performed this experiment only in one setting, it is unknown whether these value functions represent state or trait measures of behavioral strategy usage. To our knowledge, longitudinal studies in which these variables have been assessed over time have not been conducted. Thus, it is possible that an individual's value functions are not stable over time, even though they predicted both behavior and neural activity in our experiment. However, a study has shown that value functions in monkeys are highly stable and reliable over many testing sessions (Barraclough et al., 2004), suggesting that these parameters tap into stable traits. Second, our rapid event-related design did not allow us to separate neural activity related to different stages of decision making, such as choosing the decision, anticipating the outcome, and evaluating the outcome (Rogers et al., 2004). Thus, it is possible that brain activity elicited during different phases of the decision-making process might reveal additional relationships between brain activity and value functions.

Third, our experiment contained only two decision options to choose from in order to obtain monetary rewards, and it is thus unknown whether our brain–behavior correlations will generalize to other decision-making situations, such as those that involve monetary losses or situations in which optimal decision strategies can be learned (Bechara et al., 2003).

Taken together, these findings indicate that principles of reinforcement learning theory can be used to estimate value functions that describe individual differences both in decision-making strategies and in the relationship between neural activity and decisions made in future trials. These findings provide a novel method of elucidating the neural mechanisms of decision making by examining how neural activity during current decisions and outcomes relates to future decisions and how this relationship is moderated by individual differences in the value of outcomes associated with different decisions. Although we found activity in several localized regions of the brain, substantial evidence suggests that these regions are heavily interconnected, both anatomically (Baleydier & Mauguiere, 1980; Cavada, Company, Tejedor, Cruz-Rizzolo, & Reinoso-Suarez, 2000; Zahm, 1999) and functionally (Cohen, Heller, & Ranganath, 2005). The present results demonstrate that regions involved in decision making do more than perform "on the fly" processes related to the trial in which the subject is currently focused; they additionally help guide future behavior in light of current outcomes. In addition, we suggest that future researchers would profit by taking into account the context in which decisions are made and how individual-differences factors affect both behavioral and neural processes.

REFERENCES

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *NeuroImage*, **8**, 360-369.

Arana, F. S., Parkinson, J. A., Hinton, E., Holland, A. J., Owen, A. M., & Roberts, A. C. (2003). Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. *Journal of Neuroscience*, **23**, 9632-9638.

Baleydier, C., & Mauguiere, F. (1980). The duality of the cingulate gyrus in monkey: Neuroanatomical study and functional hypothesis. *Brain*, **103**, 525-554.

Barraclough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience*, **7**, 404-410.

Barto, A. G. (1995). Reinforcement learning. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 804-809). Cambridge, MA: MIT Press.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, **10**, 295-307.

Bechara, A., Damasio, H., & Damasio, A. R. (2003). Risky business: Emotion, decision-making, and addiction. *Journal of Gambling Studies*, **19**, 23-51.

Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience*, **19**, 5473-5481.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, **275**, 1293-1295.

BERRIDGE, K. C., & ROBINSON, T. E. (2003). Parsing reward. *Trends in Neurosciences*, **26**, 507-513.

BLACKWOOD, N., FFYTCHE, D., SIMMONS, A., BENTALL, R., MURRAY, R., & HOWARD, R. (2004). The cerebellum and decision making under uncertainty. *Cognitive Brain Research*, **20**, 46-53.

BOTVINICK, M. M., BRAVER, T. S., BARCH, D. M., CARTER, C. S., & COHEN, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, **108**, 624-652.

BRETT, M., JOHNSRUDE, I. S., & OWEN, A. M. (2002). The problem of functional localization in the human brain. *Nature Reviews Neuroscience*, **3**, 243-249.

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. *Proceedings of the National Academy of Sciences*, **99**, 523-528.

CAVADA, C., COMPANY, T., TEJEDOR, J., CRUZ-RIZZOLO, R. J., & REINOSO-SUAREZ, F. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex: A review. *Cerebral Cortex*, **10**, 220-242.

COHEN, M. X., HELLER, A. S., & RANGANATH, C. (2005). Functional connectivity with anterior cingulate and orbitofrontal cortices during decision-making. *Cognitive Brain Research*, **23**, 61-70.

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. *Journal of Neuroscience*, **22**, 4563-4567.

CRITCHLEY, H. D., CORFIELD, D. R., CHANDLER, M. P., MATHIAS, C. J., & DOLAN, R. J. (2000). Cerebral correlates of autonomic cardiovascular arousal: A functional neuroimaging investigation in humans. *Journal of Physiology*, **523**, 259-270.

DEPUE, R. A., & COLLINS, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral & Brain Sciences*, **22**, 491-569.

DESJARDINS, A. E., KIEHL, K. A., & LIDDLE, P. F. (2001). Removal of confounding effects of global signal in functional MRI analyses. *NeuroImage*, **13**, 751-758.

DEVINSKY, O., MORRELL, M. J., & VOGT, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain*, **118**, 279-306.

EGELMAN, D. M., PERSON, C., & MONTAGUE, P. R. (1998). A computational role for dopamine delivery in human decision-making. *Journal of Cognitive Neuroscience*, **10**, 623-630.

EREV, I., & ROTH, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equalibria. *American Economic Review*, **88**, 848-881.

FELL, J., KLAVER, P., ELGER, C. E., & FERNANDEZ, G. (2002). The interaction of rhinal cortex and hippocampus in human declarative memory formation. *Review of Neuroscience*, **13**, 299-312.

FIORILLO, C. D., TOBLER, P. N., & SCHULTZ, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, **299**, 1898-1902.

FRISTON, K. J., TONONI, G., REEKE, G. N., SPORNS, O., & EDELMAN, G. M. (1994). Value-dependent selection in the brain: Simulation in a synthetic neural model. *Neuroscience*, **59**, 229-243.

GEHRING, W. J., & WILLOUGHBY, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, **295**, 2279-2282.

GOEREE, J. K., & HOLT, C. A. (1999). Stochastic game theory: For playing games, not just for doing theory. *Proceedings of the National Academy of Sciences*, **96**, 10564-10567.

HIKOSAKA, K., & WATANABE, M. (2000). Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cerebral Cortex*, **10**, 263-271.

KAHN, I., YESHURUN, Y., ROTHSTEIN, P., FRIED, I., BEN-BASHAT, D., & HENDLER, T. (2002). The role of the amygdala in signaling prospective outcome of choice. *Neuron*, **33**, 983-994.

KERNS, J. G., COHEN, J. D., MacDONALD, A. W., III, CHO, R. Y., STENGER, V. A., & CARTER, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, **303**, 1023-1026.

KRAWCZYK, D. C. (2002). Contributions of the prefrontal cortex to the neural basis of human decision making. *Neuroscience and Biobehavioral Reviews*, **26**, 631-664.

LAGARIAS, J. C., REEDS, J. A., WRIGHT, M. H., & WRIGHT, P. E. (1998). Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM Journal of Optimization*, **9**, 112-147.

LIU, Z., MURRAY, E. A., & RICHMOND, B. J. (2000). Learning motivational significance of visual cues for reward schedules requires rhinal cortex. *Nature Neuroscience*, **3**, 1307-1315.

LUCE, D. P. (1999). *Individual choice behavior*. New York: Wiley.

MANES, F., SAHAKIAN, B., CLARK, L., ROGERS, R., ANTOUN, N., AITKEN, M., & ROBBINS, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain*, **125**, 624-639.

McCULLOUGH, L. D., & SALAMONE, J. D. (1992). Involvement of nucleus accumbens dopamine in the motor activity induced by periodic food presentation: A microdialysis and behavioral study. *Brain Research*, **592**, 29-36.

MIYASHITA, Y., OKUNO, H., TOKUYAMA, W., IHARA, T., & NAKAJIMA, K. (1996). Feedback signal from medial temporal lobe mediates visual associative mnemonic codes of inferotemporal neurons. *Cognitive Brain Research*, **5**, 81-86.

MONTAGUE, P. R., & BERNS, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, **36**, 265-284.

MOOKHERJEE, D., & SOPHER, B. (1994). Learning behavior in an experimental matching pennies game. *Games & Economic Behavior*, **7**, 62-91.

O'DOHERTY, J. P., DAYAN, P., FRISTON, K., CRITCHLEY, H., & DOLAN, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, **38**, 329-337.

PAULUS, M. P., FEINSTEIN, J. S., TAPERT, S. F., & LIU, T. T. (2004). Trend detection via temporal difference model predicts inferior prefrontal cortex activation during acquisition of advantageous action selection. *NeuroImage*, **21**, 733-743.

RAMNANI, N., & MIALL, R. C. (2003). Instructed delay activity in the human prefrontal cortex is modulated by monetary reward expectation. *Cerebral Cortex*, **13**, 318-327.

RANGANATH, C., COHEN, M. X., & BROZINSKY, C. J. (in press). Working memory maintenance contributes to long-term memory formation: Neural and behavioral evidence. *Journal of Cognitive Neuroscience*.

RIDDERINKHOF, K. R., NIEUWENHUIS, S., & BASHORE, T. R. (2003). Errors are foreshadowed in brain potentials associated with action monitoring in cingulate cortex in humans. *Neuroscience Letters*, **348**, 1-4.

ROESCH, M. R., & OLSON, C. R. (2003). Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *Journal of Neurophysiology*, **90**, 1766-1789.

ROGERS, R., RAMNANI, N., MACKAY, C., WILSON, J. L., JEZZARD, P., CARTER, C. S., & SMITH, S. M. (2004). Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological Psychiatry*, **55**, 594-602.

ROLLS, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex*, **10**, 284-294.

ROLLS, E. T. (2004). The functions of the orbitofrontal cortex. *Brain & Cognition*, **55**, 11-29.

RORDEN, C., & BRETT, M. (2000). Stereotaxic display of brain lesions. *Behavioral Neurology*, **12**, 191-200.

SANFEY, A. G., HASTIE, R., COLVIN, M. K., & GRAFMAN, J. (2003). Phineas gauged: Decision-making and the human prefrontal cortex. *Neuropsychologia*, **41**, 1218-1229.

SANFEY, A. G., RILLING, J. K., ARONSON, J. A., NYSTROM, L. E., & COHEN, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, **300**, 1755-1758.

SARIN, R., & VAHID, F. (2001). Predicting how people play games: A simple dynamic model of choice. *Games & Economic Behavior*, **34**, 104-122.

SCHULTZ, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Current Opinion in Neurobiology*, **14**, 139-147.

SHIDARA, M., & RICHMOND, B. J. (2002). Anterior cingulate: Single neuronal signals related to degree of reward expectancy. *Science*, **296**, 1709-1711.

STERN, C. E., & PASSINGHAM, R. E. (1994). The nucleus accumbens in

monkeys (*Macaca fascicularis*): I. The organization of behaviour. *Behavioural Brain Research*, **61**, 9-21.

SUTTON, R. S. (Ed.) (1992). Special issue on reinforcement learning. *Machine Learning*, **8**.

SUTTON, R. S., & BARTO, A. G. (1998). *Reinforcement learning*. Cambridge, MA: MIT Press.

TANAKA, S. C., DOYA, K., OKADA, G., UEDA, K., OKAMOTO, Y., & YAMAWAKI, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature Neuroscience*, **7**, 887-893.

TVERSKY, A., & KAHNEMAN, D. (1981). The framing of decisions and the psychology of choice. *Science*, **211**, 453-458.

VAN VEEN, V., & CARTER, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & Behavior*, **77**, 477-482.

VOLKOW, N. D., FOWLER, J. S., WANG, G. J., & SWANSON, J. M. (2004). Dopamine in drug abuse and addiction: Results from imaging studies and treatment implications. *Molecular Psychiatry*, **9**, 557-569.

WHISHAW, I. Q., & KORNELSEN, R. A. (1993). Two types of motivation revealed by ibotenic acid nucleus accumbens lesions: Dissociation of food carrying and hoarding and the role of primary and incentive motivation. *Behavioural Brain Research*, **55**, 283-295.

WORSLEY, K. J., & FRISTON, K. J. (1995). Analysis of fMRI time-series revisited—again. *NeuroImage*, **2**, 173-181.

ZAHM, D. S. (1999). Functional-anatomical implications of the nucleus accumbens core and shell subterritories. In J. F. McGinty (Ed.), *Advanciing from the ventral striatum to the extended amygdala: Implications for neuropsychiatry and drug abuse* (Annals of the New York Academy of Sciences, Vol. 877, pp. 113-128). New York: New York Academy of Sciences.

ZALD, D. H., BOILEAU, I., EL-DEAREDY, W., GUNN, R., MCGLONE, F., DICHTER, G. S., & DAGHER, A. (2004). Dopamine transmission in the human striatum during monetary reward tasks. *Journal of Neuroscience*, **24**, 4105-4112.

ZARAHN, E., AGUIRRE, G. K., & D'ESPOSITO, M. (1997). Empirical analyses of BOLD fMRI statistics: I. Spatially unsmoothed data collected under null-hypothesis conditions. *NeuroImage*, **5**, 179-197.