

# The Role of Ventral Frontostriatal Circuitry in Reward-Based Learning in Humans

Adriana Galvan,<sup>1</sup> Todd A. Hare,<sup>1</sup> Matthew Davidson,<sup>1</sup> Julie Spicer,<sup>1</sup> Gary Glover,<sup>2</sup> and B. J. Casey<sup>1</sup>

<sup>1</sup>Sackler Institute, Weill Medical College of Cornell University, New York, New York 10021, and <sup>2</sup>Department of Radiology and Neurosciences Program, Center for Advanced Magnetic Resonance Technology at Stanford, Stanford University, Stanford, California 94305

This study examined changes in behavior and neural activity with reward learning. Using an event-related functional magnetic resonance imaging paradigm, we show that the nucleus accumbens, thalamus, and orbital frontal cortex are each sensitive to reward magnitude, with the accumbens showing the greatest discrimination between reward values. Mean reaction times were significantly faster to cues predicting the greatest reward and slower to cues predicting the smallest reward. This behavioral change over the course of the experiment was paralleled by a shift in peak in accumbens activity from anticipation of the reward (immediately after the response), to the cue predicting the reward. The orbitofrontal and thalamic regions peaked in anticipation of the reward throughout the experiment. Our findings suggest discrete functions of regions within basal ganglia thalamocortical circuitry in adjusting behavior to maximize reward.

**Key words:** reward; learning; basal ganglia; nucleus accumbens; orbital frontal cortex; thalamocortical circuitry; fMRI

## Introduction

The ability to alter behavior based on expected reward outcomes is critical for goal-directed behavior. Animals alter behaviors depending on anticipated reward outcome (Pavlov, 1903, 1927), and electrophysiological studies show that striatal neurons play an important role in rewarded behavior, given their sensitivity to changes in reward magnitude and frequency (Ito et al., 2002; Cromwell and Schultz, 2003). For example, some striatal neurons discriminate between reward magnitudes and some show increased firing for stimuli that predict larger rewards (Cromwell and Schultz, 2003). This is paralleled with behavioral changes of greater anticipatory licking and faster responses to cues predicting large versus small amounts of juice. Dopamine neurons that project to the ventral striatum fire progressively less to predicted reward and increasingly to the cue predicting the reward (Mirenowicz and Schultz, 1994; Hollerman and Schultz, 1998; Fiorillo et al., 2003). These data are consistent with reinforcement learning theory of the role of dopamine in reward-related behavior. According to this theory, dopamine does not signal reward per se, but rather mediates a learning signal that allows better predictions of upcoming rewards, and appropriate behavioral adjustments (Dayan and Balleine, 2002; Montague et al., 2004). One goal of this study was to examine neural circuitry involved in this type of learning using functional magnetic resonance imaging (fMRI) and a behavioral paradigm previously used with nonhuman primates (Cromwell and Schultz, 2003).

The neural circuitry underlying reward has been defined

largely by dopamine projections, which originate in the substantia nigra and ventral tegmental area and project to the prefrontal cortex (PFC) and basal ganglia (Haber, 2003). Thalamic relay nuclei transmit basal ganglia output to the PFC, forming “closed loops” often referred to as the basal ganglia thalamocortical circuits (BGTC) (Alexander et al., 1986; McFarland and Haber, 2002) [although these circuits are described as parallel closed loops, there is cross talk among circuits at the subcortical level that helps drive motivations to executable actions (Haber, 2003)]. Here, we focus on this circuitry and its implication in motivated and reward-related behavior.

Neuroimaging studies show involvement of BGTC in reward-related behavior. For example, the ventral striatum is sensitive to increasing values of reward valence (Delgado et al., 2000, 2003; Knutson et al., 2001) and predictability (Berns et al., 2001; O’Doherty et al., 2003). Thalamocortical regions play a role in linking reward with specific goal-directed actions (MacDonald et al., 2000), because the orbital frontal cortex (OFC) is sensitive to reward valence and outcome (Elliott et al., 2003; Knutson et al., 2003; Rolls et al., 2003; O’Doherty, 2004), and thalamic lesions show impairments in stimulus–response representations (Dagenbach et al., 2001; Kubat-Silman et al., 2002) and learning (Ridley et al., 2004).

We used a spatial delay reward task similar to one in nonhuman primate studies (Cromwell and Schultz, 2003). This paradigm consisted of a cue, response, and reward that were temporally separated in a slow event-related design. We predicted that, behaviorally, subjects would respond fastest to the large reward once cue–reward associations were learned. Here, we define learning as the point in time at which behavioral performance (response latencies) significantly diverge for each reward value. We hypothesized that behavioral changes would be paralleled by neural changes in the ventral striatum, based on previous work (Mirenowicz and Schultz, 1994; Hollerman and Schultz, 1998;

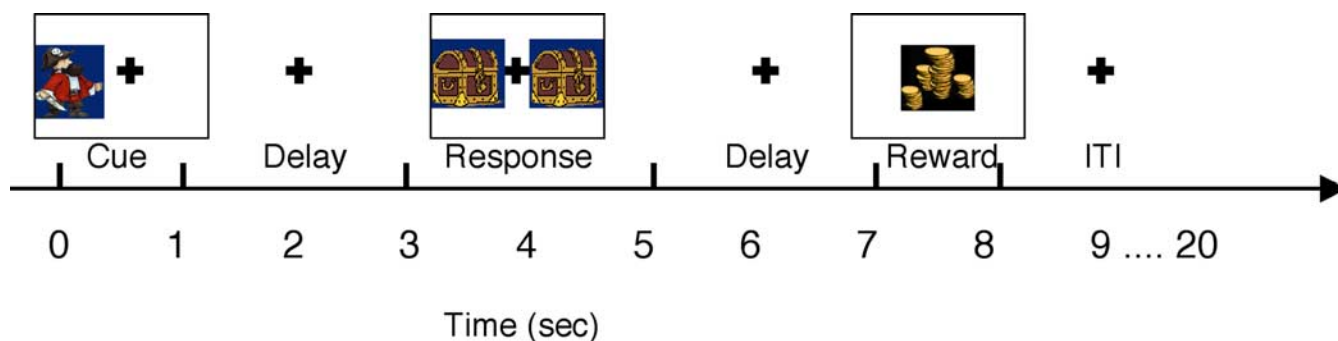
Received Feb. 22, 2005; revised Aug. 5, 2005; accepted Aug. 9, 2005.

This research was supported in part by National Institutes of Health Grants R21 DA15882-02 and R01 DA18879-01 (B.J.C.) and by a training grant from the National Eye Institute (A.G.).

Correspondence should be addressed to Adriana Galvan, 1300 York Avenue, Box 140, New York, NY 10021. E-mail: adg2006@med.cornell.edu.

DOI:10.1523/JNEUROSCI.2431-05.2005

Copyright © 2005 Society for Neuroscience 0270-6474/05/258650-07\$15.00/0



**Figure 1.** Task design. One of three cues (cartoon pirate) appeared on the left or right side of a fixation for 1 s. After a 2 s delay, a response prompt (2 treasure chests) appeared for 2 s, and subjects were instructed to press with their pointer finger if the cue had been on the left and with their middle finger if the cue had been on the right. After another 2 s delay, a reward outcome (small, medium, or large pile of coins) was presented and was followed by a 12 s ITI. The total trial length was 20 s.

Cromwell and Schultz, 2003). Specifically, we predicted that accumbens activity would be sensitive to reward values and would shift from the rewarded response to the reward-predicting cue as a function of time on task, as shown with midbrain neural recordings (Mirenowicz and Schultz, 1994; Hollerman and Schultz, 1998). We predicted that maximum changes in thalamocortical regions would occur in anticipation of reward after an outcome-related response, as previously suggested (Wallis et al., 2001; Ramnani and Miall, 2003).

## Materials and Methods

**Participants.** Twelve right-handed healthy adults (six females), aged 23–29 (mean age, 25.3 years), were included in the fMRI experiment. Subjects had no history of neurological or psychiatric disorder, and each gave informed consent for a protocol approved by the Institutional Review Board.

**Experimental task.** Participants were tested using an adapted version of a delayed-response two-choice task previously used in nonhuman primates (Cromwell and Schultz, 2003) in an event-related fMRI study (Fig. 1). In this task, three cues (counterbalanced) were each associated with a distinct reward magnitude. Subjects were instructed to press either their index or middle finger to indicate the side on which a cue appeared when prompted, and to respond as quickly as possible without making mistakes. One of three pirate cartoon images was presented in random order on either the left or right side of a centered fixation for 1000 ms (Fig. 1). After a 2000 ms delay, subjects were presented with a response prompt of two treasure chests on both sides of the fixation (2000 ms) and instructed to press a button with their right index finger if the pirate was on the left side of the fixation or their right middle finger if the pirate was on the right side of the fixation. After another 2000 ms delay, a reward feedback of a small, medium, or large amount of coins was presented in the center of the screen (1000 ms). Each pirate was associated with a distinct reward amount. There was a 12 s intertrial interval (ITI) before the start of the next trial. Subjects were guaranteed \$50 for participation in the study and were told they could earn up to \$25 more, depending on performance [as indexed by reaction time (RT) and accuracy] on the task. Although the reward amounts were distinctly different from one another, the exact value of each reward was not disclosed to the subject, because during pilot studies, subjects reported counting the money after each trial, and we wanted to avoid this possible distraction. Stimuli were presented with the integrated functional imaging system (PST, Pittsburgh, PA) using a liquid crystal display video display in the bore of the magnetic resonance (MR) scanner and a fiber optic response collection device.

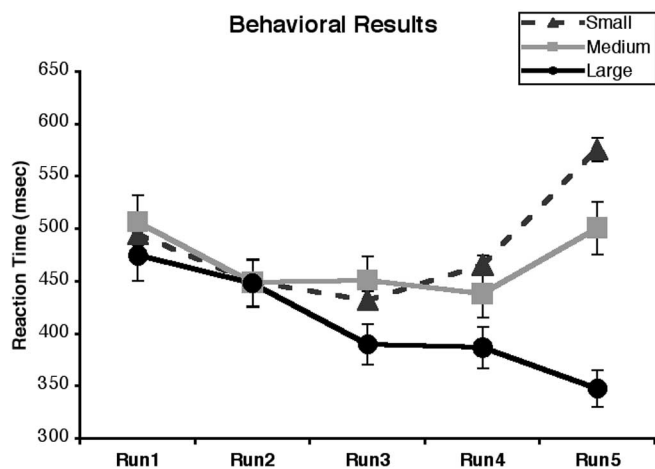
The experiment consisted of five runs of 18 trials (6 each of small, medium, and large reward trials), which lasted 6 min and 8 s each. Each run had 6 trials of each reward value presented in random order. At the end of each run, subjects were updated on how much money they had earned during that run. Before beginning the experiment, subjects were shown the actual money they could earn to ensure motivation. They received detailed instructions that included familiarization with the

stimuli used. For instance, subjects were shown the three cues and three reward amounts they would be seeing during the experiment. They were not told how the cues related to the rewards. We explicitly emphasized that there were three amounts of reward, one being small, another medium, and another large. These amounts are visually obvious in the experiment, because the number of coins in the stimuli increases with increasing reward. Only one of the 12 subjects could articulate the association between specific stimuli and reward amounts, when asked explicitly about this association during debriefing at the end of the experiment.

**Image acquisition.** Imaging was performed using a 3T General Electric (General Electric Medical Systems, Milwaukee, WI) MRI scanner using a quadrature head coil. Functional scans were acquired using a spiral-in and -out sequence (Glover and Thomason, 2004). The parameters included a repetition time (TR) of 2000 ms, echo time (TE) of 30 ms,  $64 \times 64$  matrix, 29.5 mm coronal slices,  $3.125 \times 3.125$  mm in-plane resolution, and flip of  $90^\circ$ , for 184 repetitions, including four discarded acquisitions at the beginning of each run. Anatomical T1-weighted in-plane scans were collected (TR, 500; TE, min;  $256 \times 256$ ; field of view, 200 mm; slice thickness, 5 mm) in the same locations as the functional images in addition to a three-dimensional data set of high-resolution spoiled gradient-recalled acquisition in a steady state images (TR, 25; TE, 5; slice thickness, 1.5 mm; 124 slices).

**Image analysis.** The Brainvoyager QX (Brain Innovations, Maastricht, The Netherlands) software package was used to perform a random effects analysis of the imaging data. Before analysis, the following preprocessing procedures were performed on the raw images: slice scan time correction (using sinc interpolation), linear trend removal, high-pass temporal filtering to remove nonlinear drifts of three or fewer cycles per time course, spatial data smoothing using a Gaussian kernel with a 4 mm full width at half-maximum, and three-dimensional motion correction to detect and correct for small head movements by spatial alignment of all volumes to the first volume by rigid body transformation. Estimated rotation and translation movements never exceeded 2 mm for subjects included in this analysis. Functional data were coregistered to the anatomical volume by alignment of corresponding points and manual adjustments to obtain optimal fit by visual inspection and were then transformed into Talairach space. During Talairach transformation, functional voxels were interpolated to a resolution of  $1 \text{ mm}^3$ . Regions of interest were defined by Talairach coordinates in conjunction with reference to the Duvernoy brain atlas (Talairach and Tournoux, 1988; Duvernoy, 1991).

Statistical analysis of the imaging data was performed on the whole brain using a general linear model (GLM) comprised of 60 (five runs by 12 subjects) z-normalized functional time courses. A GLM analysis of the functional data was conducted with reward magnitudes as the primary predictor. Changes in MR signal were examined at time points immediately after the cue and immediately preceding the reward (i.e., anticipation of reward). The predictors were obtained by convolution of an ideal boxcar response (assuming a value of 1 for the volume of task presentation and a volume of 0 for the remaining time points) with a linear model of the hemodynamic response (Boynton et al., 1996) and used to build



**Figure 2.** Behavioral results. There was a main effect of reward and a significant interaction of time on task (learning) by reward. After learning, subjects were faster when responding to cues associated with large ( $348 \pm 74$  ms) rewards (●) relative to the medium ( $501 \pm 160$  ms) rewards (■) and slower to small ( $576 \pm 176$  ms) rewards (▲). Runs (time on task) are on the x-axis, and reaction time (in seconds) is on the y-axis.

the design matrix of each time course in the experiment. Only correct trials were included, and separate predictors were created for error trials. *Post hoc* contrast analyses were then performed based on *t* tests on the  $\beta$  weights of predictors to identify regions that showed distinct patterns of activity for the different reward types. Contrasts were conducted with a random effects analysis, and a contiguity threshold of 50 transformed voxels (interpolated resolution of  $1 \text{ mm}^3$ ) was used to correct for multiple comparisons (Forman et al., 1995). Percentage changes in the MR signal, for the entire trial relative to 6 s of fixation preceding trial onset, were calculated for selected regions of interest (ROIs) using event-related averaging over significantly active voxels obtained from the contrast analyses. Given our hypothesis regarding the accumbens, percentage changes in the MR signal from this region were calculated using event-related averaging obtained from an ROI-GLM [regions of interest defined by Talairach coordinates in conjunction with reference to the Duvernoy brain atlas (Talairach and Tournoux, 1988; Duvernoy, 1991)].

## Results

### Behavioral results

The effects of time on task (learning) and reward magnitude were tested with a 5 (runs)  $\times$  3 (small, medium, and large reward) ANOVA for the dependent variables of mean RT and mean accuracy. There was a main effect of reward ( $F_{(2,22)} = 6.64$ ;  $p = 0.006$ ) and a significant interaction of learning by reward ( $F_{(8,88)} = 5.575$ ;  $p = 0.001$ ) (Fig. 2) but no main effect of learning ( $F_{(4,44)} = 0.75$ ;  $p = 0.56$ ) on mean reaction time. The main effect of reward showed that subjects were faster when responding to cues associated with the large reward ( $348 \pm 74$  ms) relative to the medium ( $501 \pm 160$  ms) and small ( $576 \pm 176$  ms) rewards. Of more significance is the interaction of runs by reward depicted in Figure 2, showing that mean reaction times did not differ across all three reward values until the latter trials of the experiment (last run) at which point the subjects appeared to have learned the associations between cues and reward values. There were no significant main effects or interactions with mean accuracy across reward values because subjects were near ceiling performance (small, 98.6%; medium, 99.4%; large, 99.4%).

### Imaging results

A GLM analysis of the functional data was conducted with reward magnitude as the primary predictor. Changes in MR signal were examined across all runs of the experiment, at time points immediately

after the cue and immediately preceding the reward outcome (i.e., anticipation of reward). The right and left nucleus accumbens (NAcc)/ventral striatum ( $x = 8, y = 6, z = -2$ , and  $x = -8, y = 6, z = -5$ , respectively), medial thalamus ( $x = 7, y = -21, z = 9$ ), left inferior frontal ( $x = -46, y = 37, z = 13$ ), anterior cingulate ( $x = -6, y = 19, z = 25$ ), and superior parietal ( $x = -24, y = -57, z = 43$ ) regions were significantly activated during the cue and response intervals (based on a small vs large cue contrast). The orbitofrontal cortex ( $x = 39, y = 45, z = 0$ ) was activated during the anticipatory reward interval but not the cue interval (relative to baseline). *t* tests were performed on the  $\beta$  weights of the reward predictors to identify regions that showed distinct patterns of activity for the different reward values during early, middle, and late trials.

Only the region of the nucleus accumbens/ventral striatum (see Fig. 4, left) showed discrete responses to each reward value (small, medium, and large) by the later trials of the experiment. In parallel, there was a shift in peak of the hemodynamic response in this region from the anticipatory reward interval to the cue predicting the reward (Fig. 3*a–c*). *t* tests confirmed significant differences in MR signal between small and medium rewards ( $t_{(11)} = 3.7$ ;  $p < 0.05$ ), and medium and large rewards ( $t_{(11)} = 4.8$ ;  $p < 0.05$ ) during late trials (Fig. 3*c*). *Post hoc t* tests showed significant differences ( $t_{(11)} = 5.2$ ;  $p < 0.05$ ) in the  $\beta$  weights for the large reward magnitude for early versus late trials in the accumbens region during cue presentation, confirming a shift from reward to cue interval (Fig. 4, right).

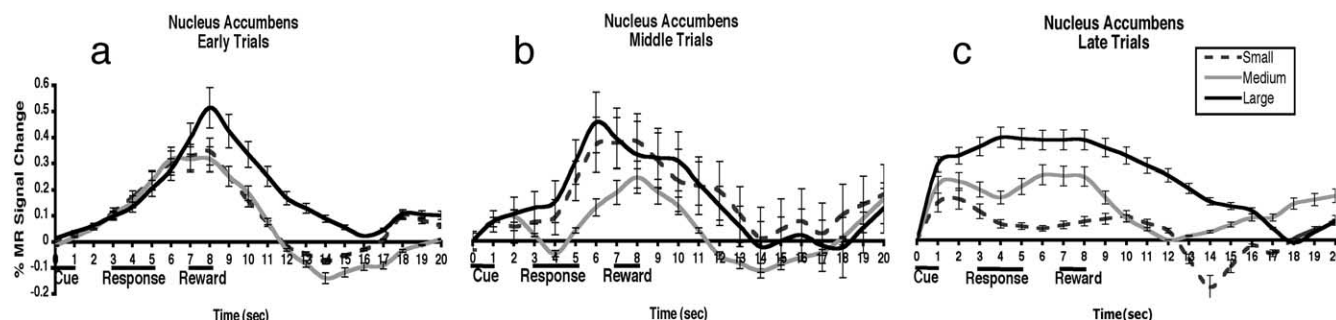
The thalamus and orbitofrontal regions (Fig. 5*a,b*) both showed the greatest activity in anticipation of the reward, but neither region discriminated between the medium and large rewards (thalamus,  $t_{(11)} = 0.86$ ,  $p = 0.45$ ; orbitofrontal region,  $t_{(11)} = 1.3$ ,  $p = 0.21$ ) even during the later trials of the study (Fig. 5*c,d*) and neither region showed an analogous shift in peak in the hemodynamic response as seen in the nucleus accumbens from the reward, to the cue predicting the reward. For both regions, the MR signal peaked in anticipation of the reward after the accumbens response (Fig. 6).

Whereas the change in MR signal in the thalamus and orbitofrontal regions occurred in anticipation of the reward outcome, just after the response, MR signal change in the inferior frontal, anterior cingulate, and superior parietal regions preceded the response. *t* tests confirmed significant differences ( $t_{(11)} = 3.4$ ;  $p < 0.05$ ) between the  $\beta$  weights of the small and large rewards during these time points with greater signal for the larger reward value (see supplemental Fig. 1*a–c* and text, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). A *post hoc* comparison of  $\beta$  weights during the early and late trials for these time points showed a significant decrease ( $t_{(11)} = -5.3$ ;  $p < 0.05$ ) in activity for these regions regardless of reward value (see supplemental Fig. 2*a–c*, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). A discussion of these findings is found in the supplemental material (available at [www.jneurosci.org](http://www.jneurosci.org)).

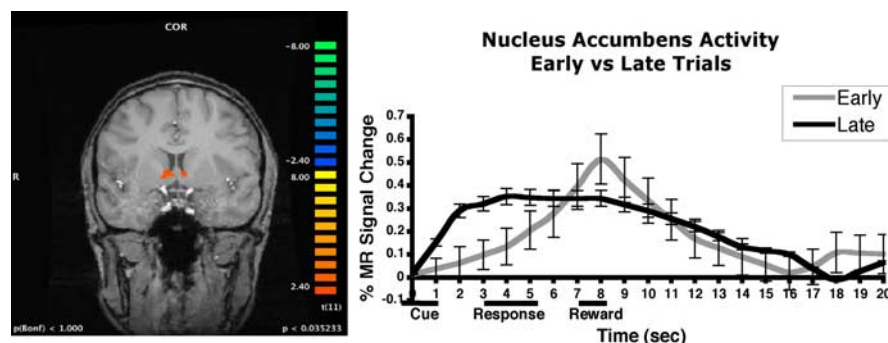
## Discussion

This study examined the effects of reward magnitude on behavior and neural activity in dopamine-rich basal ganglia thalamocortical circuitry. We show that the NAcc, thalamus, and OFC are each sensitive to reward magnitude in anticipation of the reward outcome, with the accumbens showing the greatest discrimination between discrete reward values. This pattern of activity was paralleled by faster RTs to cues predicting increasing reward.





**Figure 3.** Nucleus accumbens percentage MR signal change during early (*a*), middle (*b*), and late (*c*) trials for small, medium, and large rewards after cue presentation. Time is on the x-axis, and the percentage of MR signal change is on the y-axis. Plots are not adjusted for the hemodynamic response.



**Figure 4.** Left, Activation in the accumbens ( $x = 8, y = 6, z = -2$ , and  $x = -8, y = 6, z = -5$ ) was greater for the large relative to the small cue (based on *t* tests between the  $\beta$  weights of the small vs large cue predictors). Right, The percentage of the MR signal change for the large reward magnitude for early versus late trials in the accumbens region during cue presentation. Error bars indicate SE.

### Shifts in accumbens activity with learning

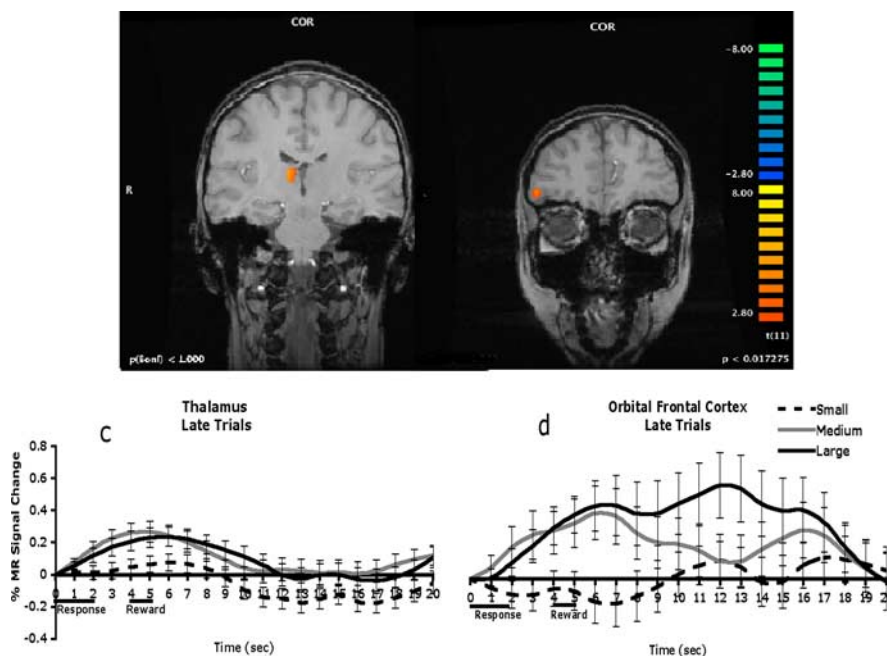
During early trials of the experiment, all three regions showed MR signal changes to reward anticipation. After learning (defined as changes in RT for each reward value), only accumbens activity shifted from the rewarded response to the reward-predicting cue, consistent with electrophysiological studies of reward in dopamine brain regions (Mirenowicz and Schultz, 1994; Hollerman and Schultz, 1998). Shifts in striatal activity with learning occur in the putamen (O'Doherty et al., 2003) and caudate (Delgado et al., 2005), but such changes have not been shown with fMRI specifically in the accumbens, a primary target site for dopamine projections. Our findings suggest that, as cue-reward contingencies are learned, the actual reward gradually elicits less accumbens activation, whereas the reward-predicting cue gradually elicits greater accumbens activity. These findings parallel nonhuman primate studies of learning-induced shifts in neural activity in dopamine-rich circuitry with primary reinforcers (Hollerman and Schultz, 1998).

Neural activity to the small reward dipped below baseline in later trials of the experiment. The small reward condition is presented in the context of larger rewards and thus may seem a relatively disappointing outcome, occurring 33% of the time, whereas better rewards occurred 66% of the time. Lack of a sufficient reward at an expected time is similar to the lack of an expected event at a given time, which both result in a dip below baseline (decreased striatal activity) in human imaging (Davidson et al., 2004) and nonhuman studies (Fiorillo et al., 2003). Slower RT to the small condition may reflect a violation in outcome expectation to which subjects adjusted behavior as shown previously (Amso et al., 2005).

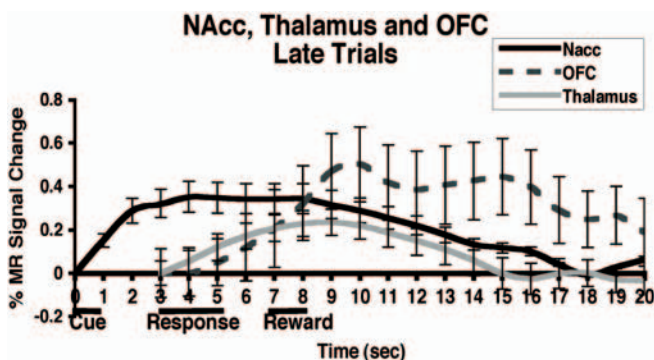
### Reward value influences behavioral and neural responses

The NAcc showed the most robust response to reward value, but the OFC and thalamic regions were also sensitive to reward magnitude differences. Accumbens activity increased monotonically as a function of reward magnitude unlike in the study by Elliott et al. (2003), in which striatal regions responded to reward regardless of value. In contrast, thalamocortical regions showed differences only between the smallest and largest reward values, consistent with the findings of others (Elliott et al., 2003). Activation in thalamocortical regions, particularly in the OFC, to the different rewards may reflect a biasing signal from the NAcc (Montague et al., 2004), given the distinct activation time courses. Whereas the NAcc was most responsive to the reward-predicting cue, the OFC and thalamus were the most active in the anticipation of reward outcome. These data are similar to nonhuman primate findings of the distinct roles of the striatum and PFC during associative learning (Pasupathy and Miller, 2005). First, the striatum identifies rewarded associations and the PFC is subsequently active during behavioral executions to obtain the identified reward (Pasupathy and Miller, 2005). This bias from the accumbens may subsequently influence a change in motivation as evidenced by faster behavioral responses to increasing reward. Behavioral and prefrontal cortical preference for relatively large rewards has been shown in nonhuman primates (Leon and Shadlen, 1999; Tremblay and Schultz, 1999) and humans (O'Doherty et al., 2000, 2001; Krangelbach et al., 2003).

In contrast to changes in accumbens activity, OFC and thalamocortical activity did not shift to the cue but remained active after the response, in anticipation of the reward. These regions changed only in the magnitude of activation, with a decrease in thalamic activity and an increase in OFC activity over time, consistent with literature showing involvement of thalamocortical loops in goal-directed actions (Dagenbach et al., 2001; Wallis and Miller, 2003). Electrophysiological studies of PFC neurons show the greatest firing preceding a prompted response (Fuster, 2001). Our data suggest that, as expectations about upcoming reward values are learned, the accumbens biases thalamocortical activity that helps adjust behavior to maximize outcomes (Montague et al., 2004; Pasupathy and Miller, 2005), suggesting a frontostriatal shift in reward sensitivity. Initially, the



**Figure 5.** Activation in the medial thalamus (*a*) ( $x = 7, y = -21, z = 9$ ) and orbital frontal cortex (*b*) ( $x = 45, y = 48, z = -2$ ) was greatest in anticipation of the reward (after the rewarded response) during late trials, but neither the thalamus (*c*) nor the orbital frontal cortex (*d*) discriminated between medium and large rewards. Time is on the x-axis, and the percentage of MR signal change is on the y-axis.



**Figure 6.** Temporal dissociation of basal ganglia thalamocortical regions. The accumbens MR signal onset occurs first, followed by the thalamus and orbital frontal cortex. Time is on the x-axis, and the percentage of MR signal change is on the y-axis.

accumbens showed greater activation than the OFC to the reward, but with learning, the accumbens showed greatest activation to the cue while the OFC activation increased to the rewarded response.

#### Contrasts between current findings and previous studies

Previous work shows ventral striatal activity during reward anticipation and OFC activation to reward outcome (Knutson et al., 2001; Bjork et al., 2004), but without examining regions before and after learning (both when the reward is unpredicted and later when the cue becomes a predictor), these studies fail to address the possibility that neural responses initially occur to the reward itself when unpredicted but shift to the conditioned stimulus during learning (O'Doherty, 2004). We address this question and provide evidence for the stimulus substitution theory (Pavlov, 1927), which suggests that a conditioned stimulus acquires value by eliciting the same responses that otherwise would have occurred to the unconditioned stimulus, in effect acting as a substi-

tute for the unconditioned stimulus itself (O'Doherty, 2004). By simultaneously examining BGTC as a function of time, we extend previous work (Delgado et al., 2000, 2003; Knutson et al., 2001; Elliott et al., 2003, 2004) but note a number of differences.

First, other imaging studies have not shown behavioral differences for distinct reward values (Knutson et al., 2001; Bjork et al., 2004; May et al., 2004), perhaps because of the continuous reinforcement schedule (rewarding 100% of correct trials) used here that differs from other studies that included conditions of gains, losses, and no reward. Because continuous reward leads to faster learning than intermittent reinforcement, the reward schedule may partially explain our behavioral findings and are consistent with work using a similar paradigm (Cromwell and Schultz, 2003). Also, we instructed subjects to respond as quickly as possible, because their RT could influence the reward outcome. Previous studies do not indicate emphasizing speed.

A second difference was in the changes in accumbens activity as function of time on task (from early to middle to late trials). Others (Knutson et al., 2001a,b) have shown robust accumbens activation in anticipation of increasing monetary reward similar to these data during the initial trials of the experiment. However, our findings begin to diverge during later trials when accumbens activity shifts from anticipation of the reward to the reward-predicting cue, coincident with distinct RTs that differ between reward values. Differences in analyses may explain the divergence, because previous work (Knutson et al. 2001a,b) was conducted with analyses that focused exclusively on the MR signal change during anticipatory delay periods over the entire experiment. Instead, we examined MR signal changes during early and late trials as well as during the periods after the cue and before the reward. O'Doherty et al. (2003) showed a similar shift in the ventral putamen but not in the nucleus accumbens. That study did not have a continuous reinforcement schedule either but omitted rewards on some trials that may explain the differences in regional activity across the two studies.

Finally, Elliott et al. (2003) did not observe striatal sensitivity to differences in reward value. Rather, striatal regions in that report responded to the presence of rewards, regardless of value. The differences between our data may be because the striatal region of interest they report is not in the accumbens but in the putamen, which may not be as sensitive to differences in reward, and also because it was a blocked-design paradigm, with each reward level in a different block.

#### Limitations

A potential confound in the response-related activity of the OFC and thalamus is the differences in RTs for different rewards during later trials of the experiment. Thus, we examined the extent that RTs accounted for variation in activity in these regions. RT accounted for limited activation variance in the OFC ( $r^2 = 0.04$ ) and thalamus ( $r^2 = 0.06$ ).

We interpret our results primarily in terms of implicit biasing

of performance by reward magnitude. However, one subject articulated the specific stimuli–reward associations during debriefing. Thus, the behavior and imaging results, although predominantly driven by implicit learning gained throughout the experiment, could have been influenced by explicit awareness of the cue–reward associations.

## Implications

We show that dopamine-rich regions in ventral frontostriatal circuitry have distinct roles in altering and optimizing behavior for reward. Furthermore, we found that reward magnitude biases learning of stimulus–reward associations in the nucleus accumbens, whereas thalamocortical regions are critical for executing behavioral goals. Together, these findings suggest a mechanism by which reward-predicting cues (e.g., drug-associated stimuli) could elicit behavioral and neural responses that influence reward-seeking behavior. Dopamine-rich regions are implicated in addiction (Volkow et al., 2004) and contextual cues have been linked to relapse-triggered reward-seeking (e.g., drug abuse, gambling) behavior (Robinson and Berridge, 2000; Hyman and Malenka, 2001). In the context of those reports, these data suggest that enhanced responses elicited by cues predicting reward is accompanied by changes in a complex dopamine-rich circuit that may help explain how specific contextual cues can trigger cravings and reward-seeking behavior. Cocaine addicts' self-reported craving is associated with increased activity in the ventral striatum and other dopamine-rich regions (Breiter et al., 1997), similar to rodent work implicating the nucleus accumbens in cocaine-seeking behavior (Ito et al., 2004). Our findings on the learned associations between cues and rewards within the accumbens are consistent with those data and have implications for the underlying mechanism of addictive behaviors and context-driven relapse.

## References

- Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 9:357–381.
- Amso D, Davidson MC, Johnson SP, Glover G, Casey BJ (2005) Contributions of the hippocampus and the striatum to simple association and frequency-based learning. *NeuroImage* 27:291–298.
- Berns GS, McClure SM, Pagnoni P, Montague PR (2001) Predictability modulates human brain response to reward. *J Neurosci* 21:2793–2798.
- Bjork JM, Knutson B, Fong GW, Caggiano DM, Bennett SM, Hommer DW (2004) Incentive-elicited brain activation in adolescents: similarities and differences from young adults. *J Neurosci* 24:1793–1802.
- Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16:4207–4221.
- Breiter HC, Gollub RL, Weisskoff RM, Kennedy DN, Makris N, Berke JD, Goodman JM, Kantor HL, Gastfriend DR, Riordan JP, Mathew RT, Rosen BR, Hyman SE (1997) Acute effects of cocaine on human brain activity and emotion. *Neuron* 19:591–611.
- Cromwell HC, Schultz W (2003) Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J Neurophysiol* 89:2823–2838.
- Dagenbach D, Kubat-Silman AK, Absher JR (2001) Human verbal working memory impairments associated with thalamic damage. *Int J Neurosci* 111:67–87.
- Davidson MC, Horvitz JC, Tottenham N, Fossella JA, Watts R, Ulug AM, Casey BJ (2004) Differential cingulate and caudate activation following unexpected nonrewarding stimuli. *NeuroImage* 23:1039–1045.
- Dayan P, Balleine BW (2002) Reward, motivation, and reinforcement learning. *Neuron* 36:285–298.
- Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA (2000) Tracking the hemodynamic responses to reward and punishment in the striatum. *J Neurophysiol* 84:3072–3077.
- Delgado MR, Locke HM, Stenger VA, Fiez JA (2003) Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cogn Affect Behav Neurosci* 3:27–38.
- Delgado MR, Miller MM, Inati S, Phelps EA (2005) An fMRI study of reward-related probability learning. *NeuroImage* 24:862–873.
- Duvernoy HM (1991) The human brain, Ed 2. New York: Springer Wien.
- Elliott R, Newman JL, Longe OA, Deakin JF (2003) Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J Neurosci* 23:303–307.
- Elliott R, Newman JL, Longe OA, William Deakin JF (2004) Instrumental responding for rewards is associated with enhanced neuronal response in subcortical reward systems. *NeuroImage* 21:984–990.
- Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299:1898–1902.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med* 33:636–647.
- Fuster J (2001) The prefrontal cortex—an update: time is of the essence. *Neuron* 30:319–333.
- Glover GH, Thomason ME (2004) Improved combination of spiral-in/out images for BOLD fMRI. *Magn Reson Med* 51:863–868.
- Haber SN (2003) The primate basal ganglia: parallel and integrative networks. *J Chem Neuroanat* 26:317–330.
- Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1:304–309.
- Hyman SE, Malenka RC (2001) Addiction and the brain: the neurobiology of compulsion and its persistence. *Nat Rev Neurosci* 2:695–703.
- Ito R, Dalley JW, Robbins TW, Everitt BJ (2002) Dopamine release in the dorsal striatum during cocaine-seeking behavior under the control of a drug-associated cue. *J Neurosci* 22:6247–6253.
- Ito R, Robbins TW, Everitt BJ (2004) Differential control over cocaine-seeking behavior by nucleus accumbens core and shell. *Nat Neurosci* 7:389–397.
- Knutson B, Adams CM, Fong GW, Hommer D (2001a) Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci* 21:RC159(1–5).
- Knutson B, Fong GW, Adams CM, Varner JL, Hommer D (2001b) Dissociation of reward anticipation and outcome with event-related fMRI. *NeuroReport* 12:3683–3687.
- Knutson B, Fong GW, Bennett SM, Adams CM, Hommer D (2003) A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *NeuroImage* 18:263–272.
- Kringelbach ML, O'Doherty J, Rolls ET, Andrews C (2003) Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cereb Cortex* 13:1064–1071.
- Kubat-Silman AK, Dagenbach D, Absher JR (2002) Patterns of impaired verbal, spatial, and object working memory after thalamic lesions. *Brain Cogn* 50:178–193.
- Leon MI, Shadlen MN (1999) Effect of expected reward magnitude on the response in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24:415–425.
- MacDonald III AW, Cohen JD, Stenger VA, Carter CS (2000) Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288:1835–1838.
- May JC, Delgado MR, Dahl RE, Stenger VA, Ryan ND, Fiez JA, Carter CS (2004) Event-related functional magnetic resonance imaging of reward-related brain circuitry in children and adolescents. *Biol Psychiatry* 55:359–366.
- McFarland NR, Haber SN (2002) Thalamic relay nuclei of the basal ganglia form both reciprocal and nonreciprocal cortical connections, linking multiple frontal cortical areas. *J Neurosci* 22:8117–8132.
- Mirenzowicz J, Schultz W (1994) Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* 72:1024–1027.
- Montague PR, Hyman SE, Cohen JD (2004) Computational roles for dopamine in behavioral control. *Nature* 431:760–767.
- O'Doherty J, Rolls ET, Francis S, Bowtell R, McGlone F, Kobal G, Renner B, Ahne G (2000) Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *NeuroReport* 11:893–897.



- O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 4:95–102.
- O'Doherty JP (2004) Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol* 14:769–776.
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337.
- Pasupathy A, Miller EK (2005) Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433:873–876.
- Pavlov IP (1903) The experimental psychology and psychopathology of animals. Paper presented at 14th International Medical Congress, Madrid, Spain.
- Pavlov IP (1927) Conditioned reflexes. Oxford: Oxford UP.
- Ramnani N, Miall RC (2003) Instructed delay activity in the human prefrontal cortex is modulated by monetary reward expectation. *Cereb Cortex* 13:318–327.
- Ridley RM, Baker HF, Mills DA, Green ME, Cummings RM (2004) Topographical memory impairments after unilateral lesions of the anterior thalamus and contralateral inferotemporal cortex. *Neuropsychologia* 42:1178–1191.
- Robinson TE, Berridge KC (2000) The psychology and neurobiology of addiction: an incentive-sensitization view. *Addiction* 95 [Suppl 2]:S91–S117.
- Rolls ET, Kringelbach ML, De Araujo IE (2003) Different representations of pleasant and unpleasant odours in the human brain. *Eur J Neurosci* 18:695–703.
- Talairach J, Tournoux (1988) Co-planar stereotaxic atlas of the human brain. New York: Thieme.
- Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398:704–708.
- Volkow ND, Fowler JS, Wang GJ, Swanson JM (2004) Dopamine in drug abuse and addiction: results from imaging studies and treatment implications. *Mol Psychiatry* 9:557–569.
- Wallis JD, Miller EK (2003) Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 18:2069–2081.
- Wallis JD, Dias R, Robbins TW, Roberts AC (2001) Dissociable contributions of the orbitofrontal and lateral prefrontal cortex of the marmoset to performance on a detour reaching task. *Eur J Neurosci* 13:1797–1808.