ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 24 (2005) 862–873

An fMRI study of reward-related probability learning

M.R. Delgado,[a,*] M.M. Miller,[a] S. Inati,[a,b] and E.A. Phelps[a,b]

[a]*Department of Psychology, New York University, New York, NY 10003, USA*
[b]*Center for Neural Science, New York University, New York, NY 10003, USA*

Received 17 June 2004; revised 27 August 2004; accepted 1 October 2004
Available online 18 November 2004

The human striatum has been implicated in processing reward-related information. More recently, activity in the striatum, particularly the caudate nucleus, has been observed when a contingency between behavior and reward exists, suggesting a role for the caudate in reinforcement-based learning. Using a gambling paradigm, in which affective feedback (reward and punishment) followed simple, random guesses on a trial by trial basis, we sought to investigate the role of the caudate nucleus as reward-related learning progressed. Participants were instructed to make a guess regarding the value of a presented card (if the value of the card was higher or lower than 5). They were told that five different cues would be presented prior to making a guess, and that each cue indicated the probability that the card would be high or low. The goal was to learn the contingencies and maximize the reward attained. Accuracy, as measured by participant's choices, improved throughout the experiment for cues that strongly predicted reward, while no change was observed for unpredictable cues. Event-related fMRI revealed that activity in the caudate nucleus was more robust during the early phases of learning, irrespective of contingencies, suggesting involvement of this region during the initial stages of trial and error learning. Further, the reward feedback signal in the caudate nucleus for well-learned cues decreased as learning progressed, suggesting an evolving adaptation of reward feedback expectancy as a behavior–outcome contingency becomes more predictable.

Keywords: Reward; Punishment; Basal Ganglia; Striatum; Caudate; Incentive; Valence; Feedback; Emotion

Introduction

Learning what choice is best comes with experience. In order to maximize rewards, an organism will strive to make better choices based on trial and error. Thus, it is imperative that brain mechanisms exist to support early learning of contingencies that will lead to a rewarding outcome. The goal of the present study is to investigate how the human brain behaves during a reward-learning paradigm, specifically the acquisition and progression of reward learning. One structure that has been implicated in processing of reward-related information is the striatum, the input unit of the basal ganglia, and specifically the caudate nucleus, part of the dorsal striatum. The striatum is a component of multiple cortico-striatal loops that are modulated by dopaminergic neurons in the midbrain, which have been shown to increase firing to unexpected rewards and conditioned stimuli that predict a reward. Due to its heterogeneity in terms of function and connectivity, the striatum is in a prime position to integrate cognitive and motivational information and influence goal-directed behavior. The human striatum is therefore a possible key structure in the acquisition of contingencies that lead to a reward.

Previous research has suggested a role for the striatum in processing reward-related information across species. Significant increases in dopamine release in the striatum, for example, have been observed during cocaine self-administration in rats (Di Chiara and Imperato, 1988; Ito et al., 2002). Neurons in the monkey striatum have been shown to respond to the anticipation (Apicella et al., 1992; Kawagoe et al., 1998) and delivery (Apicella et al., 1991; Hikosaka et al., 1989) of rewards. In accordance with animal studies, brain imaging studies of the human striatum have observed activity during the processing of both primary and secondary rewards (Aharon et al., 2001; Berns et al., 2001; Breiter et al., 2001; Delgado et al., 2000, 2003; Elliott et al., 2004; Kirsch et al., 2003; Knutson et al., 2000, 2001a; O'Doherty et al., 2002, 2004; Pagnoni et al., 2002). The striatum's response to the anticipation and delivery of rewards and punishments suggests that it may be a key structure in affective learning. Indeed, as argued by Schultz (2003), learning can be viewed as a change in outcome predictions and the acquisition of discriminatory responses to different stimuli may reflect the learning of appropriate behavioral actions.

Although the striatum responds to anticipation and delivery of rewards, the caudate nucleus, a component of the dorsal portion of the striatum, does not seem to respond to the reward per se. Rather, it seems to be more vigorously recruited when an outcome is contingent on an action (Tricomi et al., 2004), suggesting a larger role for reinforcement-based processing, where predictions and feedback help adjust behavior. The plasticity of the striatum allows

* Corresponding author. Department of Psychology, New York University, 6 Washington Place-Room 873, New York, NY 10003. Fax: +1 212 995 4349.
E-mail address: m.delgado@nyu.edu (M.R. Delgado).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.10.002

for such rapid reinforcement of actions as shown in dynamic and efficacious synaptic changes in the rat throughout learning of a procedural task (Jog et al., 1999) and during self-stimulation (Reynolds et al., 2001; Wickens et al., 2003). Thus, the caudate nucleus' unique role in reward processing may be to contribute to the brain's ability to learn though reinforcement.

The caudate nucleus is one of the main regions affected in degenerative disorders such as Parkinson's and Huntington's disease. In accordance with the idea that the caudate is important during feedback-based learning, patients with Parkinson's disease are slower during initial learning of an associative learning task (Myers et al., 2003), as compared to control subjects, and show deficits during a feedback-based learning task, as opposed to intact learning during a nonfeedback version of the same paradigm (Shohamy et al., 2004). Similarly, patients with Huntington's disease have poor performance on a trial and error incidental learning task, a type of learning thought to be dependent on the integrity of the caudate nucleus (Brown et al., 2001).

The striatum, particularly the caudate nucleus, is therefore a structure involved in processing reward-related information and various aspects of learning. Research suggests that the caudate may be an essential component of a brain circuit that allows us to improve our choices through trial and error learning. However, it is unclear whether this observed pattern of results extends from cognitive to more affective learning, where feedback properties are both informative and incentive-laden (representing possible gain or losses). The goal of this experiment was to investigate the role of the human striatum during reward-related contingency learning.

The present study investigated how activity in the caudate nucleus is modulated as reward learning progresses, specifically looking at the early stages of learning, when associations between action and outcome are being formed, and during latter stages, when the well-learned stimulus–responses are performed. We used a gambling paradigm where wins and losses were determined on the basis of guessing, but learning of stimulus–response contingencies could influence future performance. Participants were instructed that different cues, presented prior to making a guess, predicted what type of choice was more likely to lead to a reward. The introduction of a learning component insured that participants had a chance to maximize their rewards based on actual performance, allowing us to investigate how activity in the striatum, particularly the caudate nucleus, is modulated as learning of affective contingencies progresses.

Methods

Participants

Seventeen right-handed volunteers participated in this study (9 male, 8 female). Participants responded to posted advertisement (average age: M = 23.29, SD = 3.31), and all participants gave informed consent.

Procedure

The paradigm involved a series of 120 interleaved trials, divided into 10 runs of 12 trials each. Participants were instructed that they would see a card and were asked to guess if the value of such card was higher or lower than the number 5. Each individual trial represented one specific card or value where a guess could lead to a reward or a punishment. A white border on a black background served as the card and was constantly displayed throughout the experiment. Inside the card, events such as the cue and the feedback were displayed.

Each trial was divided into two periods: the probabilistic cue period and the feedback period. During the probabilistic cue period, participants saw one of five different cues (star, square, circle, triangle or diamond), which was presented for 1.5 s, at the onset of the 12-s probabilistic cue period (Fig. 1). These cues

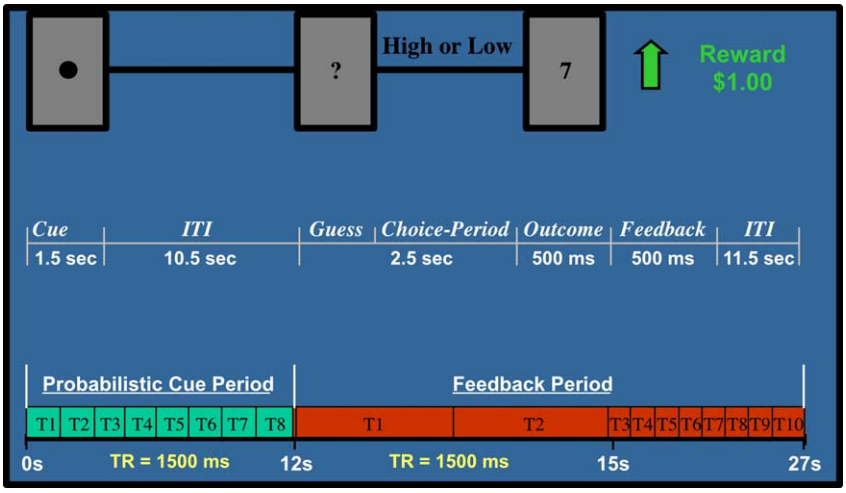

Fig. 1. Description of events in the probabilistic learning card-guessing paradigm. The task involved the presentation of multiple trials each lasting 27 s. Participants were asked to guess if the value of a "card" was higher or lower than the number 5. Each trial represented a different card or value. A trial was divided into a probabilistic cue period and a feedback period. During the probabilistic cue period (the initial 12 s, time-points T1–T8), participants were presented with one of five cues (e.g., the circle). Each cue represented the probability that the card was high or low. Participants' goal was to maximize their attained rewards by learning the probabilities over time. The probabilities were 100% (high or low), 67% (high or low) and 50% (random). At the onset of the feedback period (the last 15 s of the trial, time points T1–T10), a question mark is presented prompting the participant to make their guess. After the choice period (2.5 s), the value of the card (the outcome) was revealed (500 ms) and followed by the appropriate feedback (an up or down arrow presented for 500 ms). There was an 11.5-s delay before the onset of the next trial. There were 120 total trials, with each cue being presented 24 times throughout the experiment.

represented the probable value of the card on each trial. Participants were instructed that during each trial, the card had an unknown value ranging from 1 to 9, and the participants' goal was to make a guess about the value of the card during the feedback period. A question mark appeared in the center of the card at the onset of the feedback period, indicating that the participant had 2.5 s to guess if the card value was higher or lower than the number 5 (potential low outcomes: 1, 2, 3, 4; potential high outcomes: 6, 7, 8, 9; outcome was never 5). Participants pressed the left or right button of a button box to indicate their selection. After the choice period, a number appeared in the center of the card for 500 ms, followed by an arrow that was also displayed for another 500 ms. Each correct guess led to the presentation of a positive feedback, depicted by an upward green arrow, which indicated a monetary reward of $1.00. An incorrect guess was followed by negative feedback, depicted by a downward red arrow, indicating a monetary punishment of $0.50. Finally, there was a rest period of 11.5 s between trials, resulting in a total of 27 s per trial. Thus, one trial was comprised of a probabilistic cue period (where participants were presented with one of 5 possible cues) followed by a feedback period (where participants made a guess regarding the value of the card and received feedback).

Unlike previous paradigms (Delgado et al., 2000), participants were asked to maximize their rewards by doing more than just pure guessing. They were instructed to pay attention to the five different cues that were presented during the probabilistic cue period, and were told that each of these cues represented a different probability regarding the card value that followed (the probability of a high or low number). There were five different types of trials (the five cues) presented 24 times each during the experiment for a total of 120 trials. Participants were not told the contingencies before playing, just that each cue represented a different probability. The participants' goal was therefore to learn the probabilities as the experiment progressed, thus improving their guesses and maximizing their monetary gain.

Participants were compensated according to their performance, although the minimum of $40.00 was guaranteed for volunteering. Unbeknownst to the participants, trial order was predetermined and two versions of the paradigm were administered to counterbalance for order effects. There were 24 trials per cue presented in an interleaved fashion, and learning was measured over the entire duration of the scanning session. Although trial order was predetermined, outcome and feedback were contingent on performance and varied per participant. Trials where a response was not made in time were depicted by a pound sign (#) carried a monetary penalty of $1.00 and were excluded from further analysis. Stimulus presentation and behavioral data acquisition were controlled by a Macintosh computer with PsyScope software (Macwhinney et al., 1997).

fMRI acquisition

A 3-T Siemens Allegra head-only scanner and a Siemens standard head coil were used for data acquisition at NYU's Center for Brain Imaging. Twenty-four contiguous oblique-axial slices (3.3 × 3.3 × 3 mm voxels) parallel to the AC–PC line were obtained. Structural images were acquired in the same location as the functional images, using a standard T1-weighted pulse sequence. Functional images were acquired using a single-shot gradient echo EPI sequence (TR = 1500 ms, TE = 30 ms, FOV = 210 cm, flip angle = 90°, bandwidth = 4340 Hz/px, echo spacing = 0.29 ms). Images were then corrected for motion with Automated Image Registration (AIR 3.08; Woods et al., 1992), and detrended with a simple linear regression to adjust for drift within runs. The structural images of each participant were stripped to remove the skull and coregistered to a common reference brain, chosen from among all participants (Woods et al., 1993). Statistical analyses were performed using each participant's functional images after being transformed into the same common space. The images were also normalized by a mean scaling of each image to match global mean image intensities across participants, and smoothed using a three-dimensional Gaussian filter (8-mm FWHM) to account for between-subject anatomic differences. For visualization purposes, both the statistical maps created in analysis and the reference brain were transformed to standard Talairach stereotaxic space (Talairach and Tournoux, 1988) using AFNI software (Cox, 1996).

Behavioral and imaging analysis

Preliminary analysis of behavioral data included an investigation of possible differences in performance according to gender (male, female) or task version (trial order 1, trial order 2). For these analyses, total monetary score was used as a measure of performance (total amount earned, taking into account rewards, punishments and missed trials). Once established that there were no differences in performance based on gender or task version (described in "Behavioral results" section), we then conducted the principal behavior analysis of overall learning. Accuracy, or learning, was measured using participant's choices (high or low). A trial where the contingency was 100% chance the card value was high, for example, was scored as correct if participants pressed the button corresponding to a high choice. Similarly, a trial where the contingency was 67% chance the card value was low was scored as correct if participants pressed the button corresponding to a low choice. For 50% trials, it did not matter if participants chose high or low, thus we used the high choice as a measure of "accurate" response, expecting variability to be large in this condition. Accuracy scores were obtained for each cue in each participant. The two 100% cues were combined to form the 100% probability condition. Similarly, the 67% cues were combined to form the 67% probability condition. Data for all probability conditions for all participants was subdivided into three learning phases: the early learning phase (consisting of the first eight presentations of a cue), the middle learning phase (consisting of the next eight presentations of a cue), and the late learning phase (consisting of the last eight presentations of a cue). A repeated measures two-way ANOVA was then performed using participants (N = 12) as a random factor, and probability condition (100%, 67% and 50%) and learning phase (early, middle, late) as within-subject factors.

Analysis of neuroimaging data was conducted in a similar manner. Repeated-measures three-way ANOVAs were performed on the entire set of co-registered data for both trial periods (probabilistic cue and feedback period) separately. These whole-brain analyses involved investigation of any regions of interest (ROIs) that showed an interaction with time (i.e., learning phase and time), defined by strength of effect ($P < 0.0001$) and size (5 or more voxels). During the probabilistic cue period, within-subjects factors included learning phase (early, middle, late), probability condition (100%, 67% and 50%), and time (the eight sequential 1.5-s scans in the 12-s period, referred to as T1–T8). This ANOVA allowed us to look at two contrasts of interest: (a) brain regions

showing an interaction of learning phase by time, and (b) regions showing an interaction between probability condition and time. During the feedback period, within-subjects factors included learning phase (early, middle and late), feedback (reward, punishment) and time (the 10 sequential 1.5-s scans in the 15-s period, referred to as T1–T10). This ANOVA allowed us to look at two contrasts of interest: (a) brain regions activated showing an interaction of learning phase by time, and (b) regions showing an interaction between feedback and time. Regions of interest comprised of five or more contiguous voxels were selected, as a precaution against type-1 errors (Forman et al., 1995). Event-related time-series graphs, which represent functional magnetic resonance imaging (fMRI) mean intensity value for each factor for time periods T1–T8 (during probabilistic cue period) and T1–T10 (during feedback period), were displayed for selected ROIs.

Through the interaction of learning phase and time contrast, we hoped to identify whether the brain's response, particularly in the striatum, changed over time as learning progressed. Larger blood oxygenation level-dependent (BOLD) responses during the early phases of learning would support our hypothesis that the striatum is an integral structure during the initial acquisition of outcome contingencies. To further investigate how striatum activity was influenced as learning progressed, two post hoc ANOVAs were conducted using only positive feedback (reward) trials during the feedback period, comparing the easily learned probability condition (100%) versus the random probability condition (50%). These ANOVAs allowed us to investigate if the striatum response to reward was modulated as a function of learning, or increased predictability. During the early phase of learning, an ANOVA investigating ROIs defined by a probability condition (100%, 50%) by time (T1–T10) interaction should reveal no striatal activity as reward feedback provides an important learning signal to all conditions. During the late phase of learning, however, a similar analysis should reveal a significant effect in the striatum, as the fully expected, learned reward signal provided in the 100% condition is not as valuable as the unlearned 50% condition, and therefore is hypothesized to be of smaller magnitude. These post hoc ANOVAs were analyzed using a less strict threshold ($P < 0.001$) and cluster size (four or more voxels).

Results

Behavioral results

Analysis was conducted on all 17 participants to investigate behavioral effects of gender, trial order and overall monetary gain. Participants monetary score was calculated at the end of the session and took into account correct ($1.00 gain), incorrect (−$0.50 loss) and missed trials (−$1.00 loss). Participants scores ranged from $34.50 to $78.00 ($M = 58.15$, SD = 12.28). Using the monetary score for each participant, we then looked at any effects of trial order (version 1 and version 2) and gender (9 male and 8 female). Independent two-tailed t tests showed that there were no differences in performance according to different trial orders [version 1, $M = 56.67$, SD = 11.66; version 2, $M = 59.81$, SD = 13.55; t (15) = −0.52, $P = 0.61$] or gender [male, $M = 61.33$, SD = 9.2; version 2, $M = 54.56$, SD = 14.85; t (15) = 1.15, $P = 0.27$].

Out of 17 participants, one was removed due to excessive motion in the scanner (more than 5 mm shift in movement throughout session), while four more were excluded from further analysis due to poor performance, based on monetary score (participants with monetary scores one standard deviation below the mean were excluded-overall group: $N = 17$, $M = 58.15$, SD = 12.28; good group: $N = 12$, $M = 64$, SD = 6.87; poor group: $N = 4$, $M = 39.38$, SD = 4.13) and learning rate (failure to show pattern of increased learning, described below). Therefore, behavioral and neuroimaging data from the 12 remaining participants was analyzed and described further (Fig. 2). A repeated measures two-way ANOVA was performed to measure accuracy and learning rate according participant's choice. The ANOVA used participants ($N = 12$) as a random factor, and probability condition (100%, 67% and 50%) and learning phase (early, middle, late) as within-subject factors. As expected, participants showed progressive learning of the contingencies as displayed by a main effect of phase [$F(2, 22) = 7.31$, $P < 0.004$]. Participants were also more accurate during the 100% condition [$F(2, 22) = 8.54$, $P < 0.002$]. In contrast, performance of the same ANOVA using the four excluded participants revealed no learning of the contingencies over time

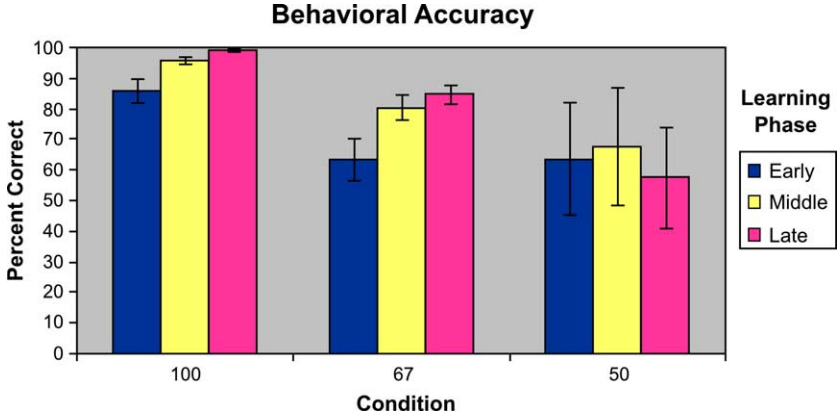

Fig. 2. Accuracy results during learning for all participants. The results are displayed broken down by condition (100, 67 and 50) and learning phase (early, middle and late). Participants were more accurate during trials involving 100% cues, showing that as learning progressed so did accuracy. A similar result was observed during trials involving 67% cues, while learning during the 50% probabilistic cue did not differ as task progressed. Overall, participants showed progressive learning of the contingencies as displayed by a main effect of phase [$F(2, 22) = 7.31$, $P < 0.004$], while being more accurate during the 100% condition [main effect of condition: $F(2, 22) = 8.54$, $P < 0.002$]. In all graphs, early learning is depicted by a dark blue color, middle learning is represented by the color yellow and late learning is illustrated by the color pink.

[$F(2, 6) = 0.41$, $P = 0.68$] and no differences between probability conditions [$F(2, 6) = 1.98$, $P = 0.22$].

Neuroimaging results

Probabilistic cue period

A repeated measures three-way ANOVA was performed encompassing all trials in the probabilistic cue period, which were followed by a behavioral response, using learning phase (early, middle, late), probability condition (100%, 67%, 50%) and time (T1–T8) as within-subject factors. Regions showing an interaction of learning phase (early, middle, late) and time (T1–T8) are listed in Table 1 [$F(14, 154) = 3.36$, $P < 0.0001$]. Activity was observed in the right striatum (caudate nucleus, $x = 13$, $y = 8$, $z = 1$) and right inferior prefrontal cortex [Brodmann area (BA) 45, $x = 30$, $y = 23$, $z = 9$], with both areas showing more activity during earlier phases of learning for all conditions. This is reflected in the time series graphs for the right caudate nucleus (Fig. 3, only trials where feedback was reward are graphed) showing larger responses to the onset of the cue during early phases of learning for 100% trials (easy to learn), 67% (harder to learn) and 50% (where no learning takes place).

An interaction of probability condition (50%, 67%, 100%) and time (T1–T8) during the probability cue period was also observed, yielding activation in two regions [$F(14, 154) = 3.36$, $P < 0.0001$, Table 1]. Activity was observed in the left medial prefrontal cortex/anterior cingulate (BA 32/9, $x = -9$, $y = 42$, $z = 23$) showing a pattern of decreasing activation from the onset of the cue. This decrease was larger for the 50% condition. Activity was also observed in the right inferior parietal cortex (BA 40/7, $x = 30$, $y = -67$, $z = 43$) showing increases in activation that were larger for the 50% condition. Finally, one ROI showed a three-way interaction [learning phase, probability condition and time: $F(28, 308) = 2.45$, $P < 0.0001$, Table 1]. Localized in the left medial prefrontal cortex (BA 10/9, $x = -5$, $y = 52$, $z = 23$), this ROI showed a decreasing pattern of activation from the onset of the cue, which was larger during the 50% condition trials.

Table 1
Probabilistic cue period ($P < 0.0001$)

Region of activation	Brodmann areas	Laterality	Talairach coordinates		
			x	*y*	*z*
Learning phase and time interaction					
Inferior frontal gyrus	45	R	30	23	9
Caudate nucleus		R	13	8	1
Condition and time interaction					
Inferior parietal cortex	40/7	R	30	−67	43
Medial prefrontal/ cingulate	32/9	L	−9	42	23
Learning phase, condition and time interaction					
Medial prefrontal cortex	10/9	L	−5	52	23

Areas showing interactions of learning phase (early, middle, late) and time (T1–T8) and condition (50%, 67%, 100%) and time, during the probabilistic cue period. Activations reflect peaks found in ROIs of five or more contiguous voxels defined by this contrast [$F(14, 154) = 3.36$, $P < 0.0001$]. Brodmann areas are provided when applicable. The laterality of the activated ROIs is also provided where regions were located either in the right (R) or left (L) hemisphere. The stereotaxic coordinates of the peak of the activation are given according to Talairach space (Talairach and Tournoux, 1988).

Feedback period

A repeated measures three-way ANOVA was performed encompassing all trials where a response was recorded in the feedback period, using learning phase (early, middle, late), feedback (reward, punishment) and time (T1–T10) as within-subject factors. Regions showing an interaction of learning phase and time are listed in Table 2 [$F(18, 198) = 2.97$, $P < 0.0001$]. Activity was observed in the right striatum (caudate nucleus, $x = 4$, $y = 7$, $z = 3$) and left thalamus ($x = -5$, $y = -5$, $z = 11$), with both areas showing more activity during earlier phases of learning. The time series graphs for the right striatum once again show larger responses during early phases of learning in response to a positive feedback or reward. However, this larger response to reward feedback during early phases of learning is only observed in the 100% trials (easy to learn) as the magnitude of the feedback response in the 50% trials (where no learning takes place) seems to be the same throughout the experiment (Fig. 4).

During the feedback period, an interaction of type of feedback and time was also observed, yielding activation in 10 ROIs [$F(9, 99) = 4.29$, $P < 0.0001$, Table 2]. Notably, robust activation was observed in the striatum comprising several peaks in both hemispheres of the dorsal and ventral striatum. In order to properly investigate the hemodynamic response in these regions, we increased the threshold to look at these peaks separately [$F(9, 99) = 6.96$, $P < 0.0000001$]. Activity in both the left and right, dorsal and ventral striatum replicated previous findings (Delgado et al., 2000), showing a differential response to reward and punishment feedback. A true analysis of how this differential response differed across probability conditions was not possible due to the small number of punishment trials in the 100% condition. A three-way interaction (learning phase, feedback and time) was also investigated [$F(18, 198) = 2.97$, $P < 0.0001$, Table 2]. Activation was found in the right inferior parietal cortex (BA 40, $x = 19$, $y = -32$, $z = 49$), characterized by higher activity during early punishment trials.

Interaction of probability condition and time during different phases of learning

The time series analysis of the learning period by time interaction during the feedback period suggested that the response to positive feedback during 100% trials decreases as learning progresses, in contrast to the response to positive feedback during the unlearned, 50% condition that did not change over time. To further investigate these results, two post hoc ANOVAs were conducted using only reward trials looking at any brain regions that showed a differential response to positive feedback between 100% and 50% probability conditions (interaction of probability condition and time, Table 3). During the early phases of learning, no ROIs were identified [$F(9, 99) = 3.44$, $P < 0.001$]. During the late phases of learning, however, 3 ROIs showed differential reward responses according to the type of probability condition [$F(9, 99) = 3.44$, $P < 0.001$]. Activation was observed in the right caudate nucleus ($x = 8$, $y = 3$, $z = 4$, Fig. 5), showing a decreased response to the learned contingency (100%) and a larger response during unlearned trials (50%). Two other ROIs comprising four voxels each were identified, including a left inferior parietal cortex region (BA 40, $x = -55$, $y = -35$, $z = 25$)

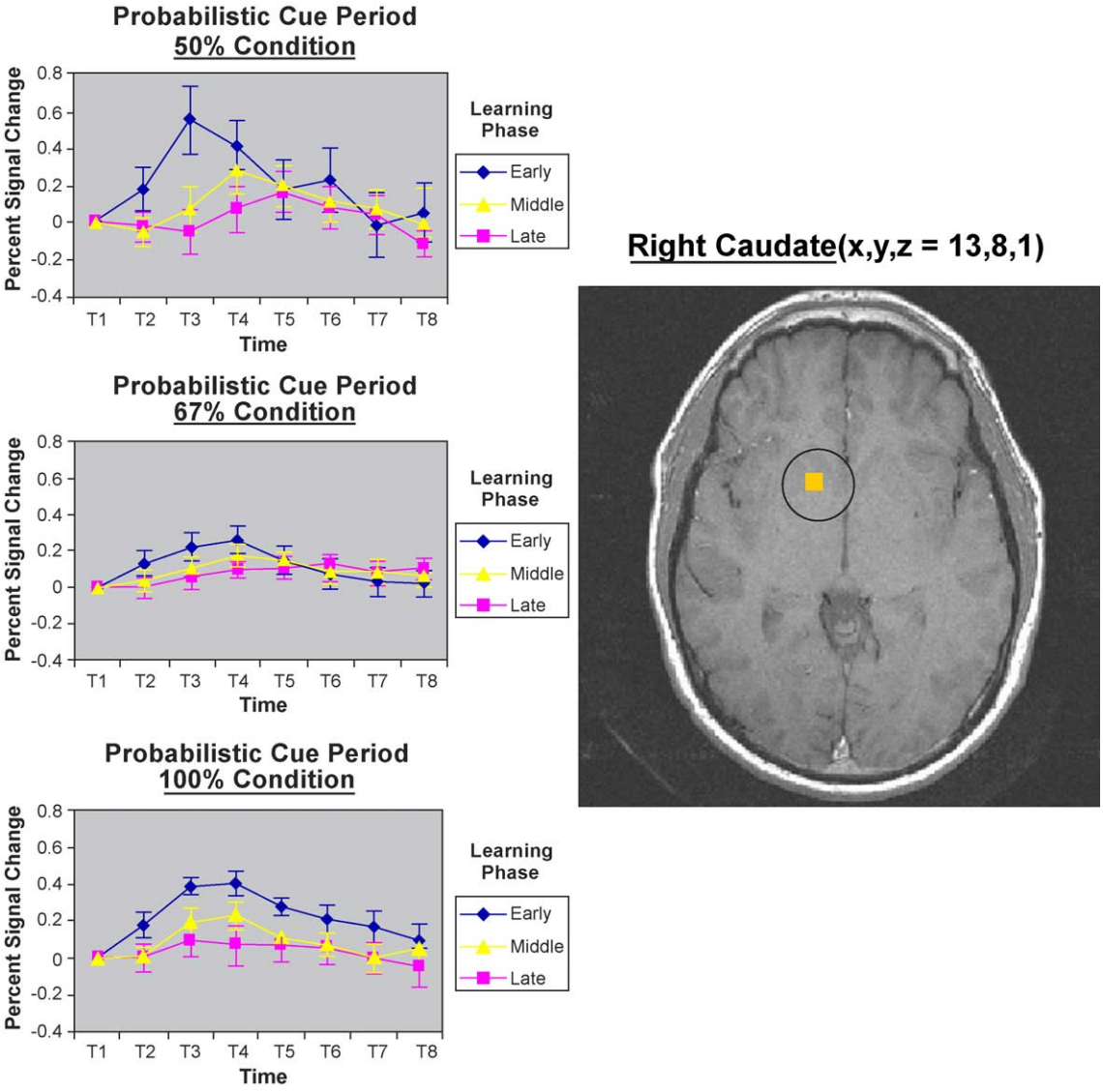

Fig. 3. Activation of the right caudate nucleus (x, y, z = 13, 8, 1) showing an interaction of learning phase (early, middle, late) and time (T1–T8) during the probabilistic cue period [$F(14, 154) = 3.36$, $P < 0.0001$]. Time-series graphs for each condition (50, 67, 100) reflect only trials where feedback was reward. Larger responses to the onset of the cue are observed during early phases of learning for 100% trials (easy to learn), 67% (harder to learn) and 50% (where no learning takes place). Standard error bars for all graphs displayed were calculated on a per participant basis using fMRI mean intensity values. These values were then standardized to percent signal change differences and plotted.

also showing a larger response to 50% condition. The reverse pattern was observed in the right prefrontal cortex ROI (BA 44, x = 32, y = 3, z = 32), where activity was larger during the 100% condition.

Nonlearners group

Four participants were excluded from imaging analysis due to poor behavioral data (and post-questionnaire self-reports), demonstrating that no learning of the contingencies over time took place within this group. Exploratory analyses were performed in this group to further investigate the role of the caudate in learning contingencies over time. Two repeated measures three-way ANOVAs were performed encompassing all trials for all non-learners (n = 4), similarly to the ones previously described. The goal was to identify any activation in the caudate nucleus using the same analysis of interest previously used in the good learners group (interactions of learning phase and time) in the non-learners group (ROIs were identified at a threshold of $P < 0.001$ and contiguity threshold of five or more voxels). During the probabilistic cue period (using learning phase, probability condition and time as within-subject factors), no regions showed an interaction of learning phase (early, middle, late) and time [T1–T8: $F(7, 21) = 5.55$, $P < 0.001$]. During the feedback period (using learning phase, feedback and time as within-subject factors), an interaction of learning phase (early, middle, late) and time (T1–T10) did not yield activation in the caudate nucleus [$F(9, 27) = 4.56$, $P < 0.001$]. Although these exploratory results suggest lack of caudate activation during learning over time in the non-learners group, they must be carefully considered due to the small sample size (n = 4). Yet, these data serve to complement the data from the good learners group and further support the idea

Table 2
Feedback period ($P < 0.0001$)

Region of activation	Brodmann areas	Laterality	Talairach coordinates		
			x	*y*	*z*
Learning phase and time interaction					
Thalamus		L	−5	−5	11
Caudate nucleus		R	4	7	3
Feedback and time interaction					
Precuneus	7	L	−10	−42	46
Inferior parietal cortex	40	L	−54	−63	45
Superior frontal gyrus	8	R	34	49	39
Inferior parietal cortex	40	R	54	−52	38
Caudate nucleus		L	−19	−18	33
Medial frontal gyrus	9/10	L	−3	53	23
Inferior frontal gyrus	44	R	40	10	16
Cingulate gyrus	32	L	−10	38	0
Inferior frontal gyrus	47	R	30	20	−1
Caudate nucleus		R	10	9	5
Caudate nucleus		L	−12	9	4
Ventral striatum		R	10	3	−6
Ventral striatum		L	−14	4	−6
Learning phase, feedback and time interaction					
Inferior parietal cortex	40	R	19	−32	49

Areas showing interactions of learning phase (early, middle, late) and time (T1–T10) and feedback (reward, punishment) and time, during the feedback period. Activations reflect peaks found in ROIs of five or more contiguous voxels defined by this contrast [$F(18, 198) = 2.97$, $P < 0.0001$, for learning phase and time; $F(9, 99) = 4.29$, $P < 0.0001$, for feedback and time]. One large striatal ROI was found in the feedback and time interaction. To investigate both dorsal and ventral striatum, in both hemispheres, an increased threshold was use to separate them [$F(9, 99) = 6.96$, $P < 0.0000001$] and its coordinates are displayed. Brodmann areas are provided when applicable. The laterality of the activated ROIs is also provided where regions were located either in the right (R) or left (L) hemisphere. The stereotaxic coordinates of the peak of the activation are given according to Talairach space (Talairach and Tournoux, 1988).

that the caudate nucleus is an important structure during learning based on feedback.

Discussion

The goal of this study was to investigate how the human brain processes learning of reward contingencies. Specifically, we investigated brain regions thought to be important during the acquisition of reward associations and their modulation as learning progresses. By using a gambling paradigm (where rewards were attained on the basis of guessing) that contained probabilistic cues (which educated the participant in regards to which choice or guess was more likely to lead to a reward), we were able to look at two aspects of reward-related learning. First, what regions were important during acquisition of the cue-outcome relationship? Second, was activity in such regions modulated as a function of increased learning? Activity was observed in the striatum, particularly the caudate nucleus, during both the delay between cue and action (probabilistic cue period) and during the action-outcome, or feedback period. Further, this activity was larger during the early phases of learning, suggesting an integral role for the caudate nucleus during the initial learning of reward contingencies. The observed activity in the striatum also decreased as a function of successful learning, as the response during the feedback period differed depending on how predictable a reward was.

The striatum is a large basal ganglia structure interconnected with frontal regions, collectively forming circuits essential for goal-directed behavior to occur (Haber, 2003; Middleton and Strick, 2000). The striatum can be subdivided into a dorsal (primarily caudate and putamen) and a ventral (primarily nucleus accumbens) components. Although more reward-related neurons are found in the ventral striatum (Apicella et al., 1991), research also suggests that some neurons in the dorsal striatum, specifically the caudate, also respond to reward (Kawagoe et al., 1998, 2004; Lauwereyns et al., 2002; Ravel et al., 2003; Takikawa et al., 2002; Watanabe et al., 2003). This is supported by neuroimaging studies implicating the dorsal striatum in processing an array of rewards and punishments, such as money (Breiter et al., 2001; Delgado et al., 2000, 2003, 2004; Elliott et al., 2000, 2004; Knutson et al., 2001a, 2001b), liquids and odors(Berns et al., 2001; Gottfried et al., 2002; O'Doherty et al., 2002, 2003, 2004). Further, dopamine is released in the rat dorsal striatum during drug-seeking behavior (Ito et al., 2002) and in the human dorsal striatum during food stimulation (Volkow et al., 2002), video-game playing (Koepp et al., 1998) and receipt of monetary rewards (Zald et al., 2004). The activation of the human dorsal striatum in a gambling game where the rewards are contingent on an appropriate choice is therefore consistent with the vast literature linking the striatum with processing reward-related information.

Recently, neuroimaging research has suggested a further link between the dorsal striatum and reward-related processing, by finding evidence for its involvement in paradigms where an action is necessary to attain a reward (Elliott et al., 2004; Haruno et al., 2004; O'Doherty et al., 2004; Tricomi et al., 2004), or when task distractors are behaviorally relevant or salient (Zink et al., 2003, 2004). Of particular interest is the role of the caudate nucleus, which may be processing the properties of the feedback in a reinforcement learning context to improve choice behavior (Tricomi et al., 2004). This is an idea originating from models of reinforcement learning (Barto, 1994), which suggests that stimulus–response associations, perhaps formed in the caudate nucleus (O'Doherty et al., 2004; White and McDonald, 2002), may be strengthened by temporal-prediction signals, potentially mediated by dopamine error signals (Schultz et al., 1997, 2000). In the current study, participants were able to improve their choice behavior in order to maximize potential rewards, a process that required trial and error learning. Thus, we were able to investigate if the caudate nucleus response to attained rewards was modulated as a function of learning and predictions.

The first finding observed in the caudate nucleus in the present study was stronger BOLD responses during the initial stages of learning during both probabilistic cue and feedback period. This is in agreement with the idea that the striatum is important for the initial acquisition of associations that will lead to a reward. Indeed, neuropsychological research on Parkinson's and Huntington's patients (Brown et al., 2001; Knowlton et al., 1996; Myers et al., 2003; Shohamy et al., 2004), as well as lesion studies in animals (Packard and Knowlton, 2002), suggest that formation of stimulus–response associations seems to be dependent on the functional integrity of the dorsal striatum, particularly the caudate nucleus. For example, rats with dorsal striatum lesions show deficits in the initial acquisition of win–stay tasks (a paradigm where rats visit different arms of a serial maze in which some of the arms have a

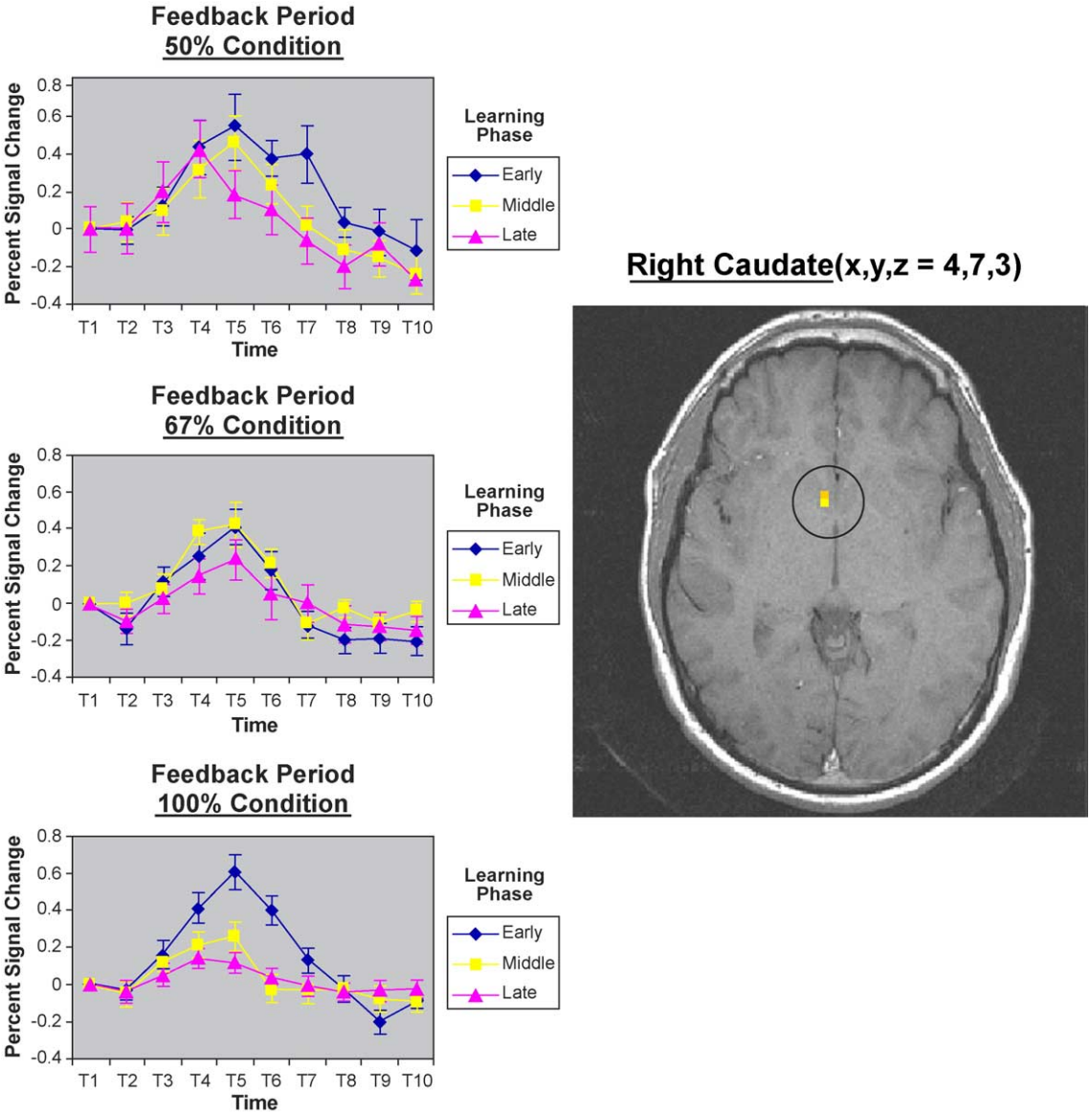

Fig. 4. Activation of the right caudate nucleus (x, y, z = 4, 7, 3) showing an interaction of learning phase (early, middle, late) and time (T1–T10) during the feedback period [$F(18, 198) = 2.97$, $P < 0.0001$]. Time-series graphs for each condition (50, 67, 100) reflect only trials where feedback was reward. Larger responses to the positive feedback or reward are observed during early phases of learning for all trials, when trial and error feedback is important. As learning progresses, however, the magnitude of the response to reward feedback in the 100% trials (easy to learn) seems to decrease, while the magnitude of the feedback response in the unpredictable 50% trials (where no learning takes place) seems to be the same throughout the experiment. Standard error bars for all graphs displayed were calculated on a per participant basis using fMRI mean intensity values. These values were then standardized to percent signal change differences and plotted.

light cue indicating a reward at the end) (Packard and Knowlton, 2002; Packard and McGaugh, 1996). Another example finds that rat striatal forelimb neurons decrease in firing after overtraining in a lever pressing task (Carelli et al., 1997), suggesting striatum function is important during acquisition but not expression of learned motor responses. Our findings that the caudate nucleus is active during the initial stages of learning different contingencies in a probabilistic reward learning task therefore supports the idea that the caudate nucleus is important during early acquisition of associations.

A second finding was how the caudate nucleus activation was modulated as a function of learning associations that varied in their predictability. This is in accordance with physiology (Fiorillo et al., 2003) and neuroimaging (Aron et al., 2004) data showing that dopamine neurons in the midbrain are sensitive to predictability by varying their firing rates according to uncertainty. Aron et al. (2004) also found that this midbrain activity correlated with striatum activity, a target of such dopaminergic projections. In the current study, behavioral data suggested that while some contingencies (100% probability condition) were acquired within the initial trials (early learning phase), other contingencies were acquired only later (67% probability condition) or not at all (50% probability condition). The pattern of activity observed in the caudate nucleus paralleled this behavior, as the response was modulated as a function of increased learning or predictability of the upcoming feedback. Large BOLD signals

Table 3
Feedback period: late phases of learning ($P < 0.001$)

Region of activation	Brodmann areas	Laterality	Talairach coordinates		
			x	*y*	*z*
Inferior parietal cortex	40	L	−55	−35	25
Inferior frontal gyrus	44	R	32	3	32
Caudate nucleus		R	8	3	4

Areas showing interactions of condition (50% and 100%) and time (T1–T10) during the late phases of learning of the feedback period. Activations reflect peaks found in ROIs of four or more contiguous voxels defined by this contrast [$F(9, 99) = 3.44$, $P < 0.001$]. A similar analysis conducted during the early phases of learning revealed no ROIs. Brodmann areas are provided when applicable. The laterality of the activated ROIs is also provided where regions were located either in the right (R) or left (L) hemisphere. The stereotaxic coordinates of the peak of the activation are given according to Talairach space (Talairach and Tournoux, 1988).

were observed during the early phases of learning for reward feedback attained in the 100% condition, which decreased in magnitude as learning progressed. During the late phases of learning, when the feedback was fully expected, reward feedback of well-learned contingencies yielded a significantly lower BOLD response than the one observed during early phases of learning, when trial and error predictions were still of importance. Perhaps after learning, explicit memory strategies not dependent on the striatum are used to perform the task, leading to an eventual decrease in striatum activity (Packard and Knowlton, 2002; Poldrack et al., 1999). In contrast, the 50% probability condition reward feedback (where successful learning could not occur) was characterized by an increase in BOLD response during early phases of learning that did not significantly differ from the signal observed during late phases of learning. The reward response for the 50% condition during late phases of learning was also larger than the 100% condition, suggesting that trial and error prediction-based learning was still occurring, thus still requiring caudate nucleus involvement.

The observed decrease in caudate activation during learning of a probabilistic reward task is in accordance with a recent study by Haruno et al. (2004), which observed modulation of BOLD activation in the caudate nucleus during a stochastic decision task in which difficulty (i.e., probability of success) was manipulated. The findings, however, are in contrast with cognitive studies of probabilistic classification learning, which find increases in caudate activity over task progression (Poldrack et al., 2001). This disparity, however, may be due to several differences between paradigms. First, in the weather prediction task used by Poldrack et al. (2001), it is hard to dissociate the predictive cue from the feedback, and an array of potential strategies could be implemented to solve such a task (Gluck et al., 2002). Our paradigm attempted to separate cue and feedback effects by looking at those periods separately. Second, the type of feedback attained in the weather prediction task (purely cognitive feedback) and the one used in the current manuscript (monetary reward feedback) may be processed differently according to the context of the task (Aron et al., 2004; Delgado et al., 2004). Finally, the locus of activity within the caudate nucleus may also be important as the right caudate nucleus ROI found to be deactivated at initial stages of learning by Poldrack et al. (2001) was more dorsal. (x, y, z = 9, 6, 21) in comparison to the right caudate nucleus ROI observed in the current study that used a reward-related probability learning task (x, y, z = 4, 7, 3, in the feedback period).

The results in the caudate nucleus in our probabilistic reward-related learning task show that this structure is important for the initial acquisition of contingencies, and its activity is modulated as a function of learning and predictability. As a potential reward becomes more predictable, the signal observed in the caudate nucleus is decreased, as opposed to the unpredictable reinforcement signal (elicited by the 50% condition), which does not significantly differ between early and late stages of learning. This is in

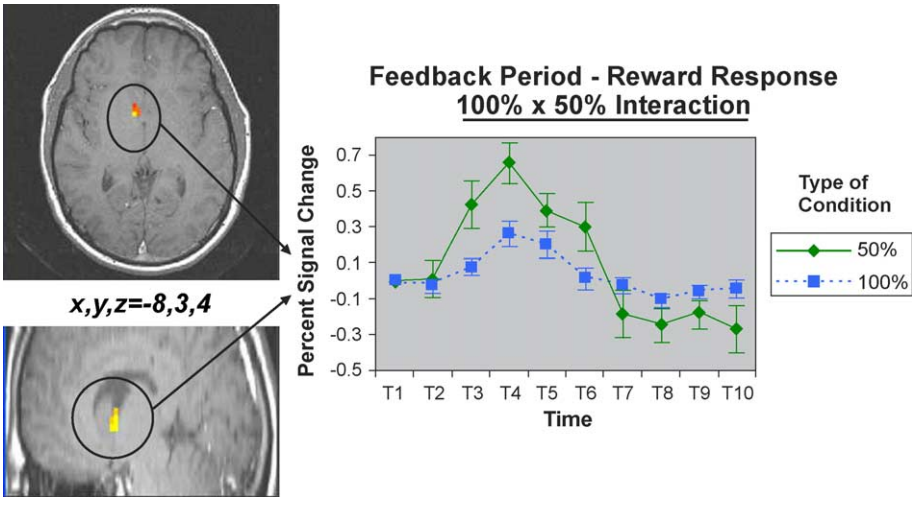

Fig. 5. Axial and sagittal views of right caudate nucleus activation (x, y, z = 8, 3, 4) showing an interaction of probability condition (100% and 50%) and time (T1–T10) during the feedback period of late phases of learning [$F(9, 99) = 3.44$, $P < 0.001$]. Two post hoc ANOVAs investigated differences between the reward feedback response of 100% (easy to learn) and 50% (random) trials during both early and late phases of learning. No ROIs were reported showing a difference between conditions during early learning [F (9, 99) = 3.44, $P < 0.001$], suggesting that the feedback signal is important while the outcome is still unpredictable. As learning progressed, however, the right caudate nucleus showed a decreased response to the learned contingency (100%) and a larger response during unlearned, still unpredictable trials (50%) during the late phases of learning. Green line represents the 50% condition while the blue line represents the 100% condition.

agreement with studies investigating predictability and reward, which find striatum activity to be greater during unpredictable delivery of rewards, such as liquids (Berns et al., 2001). This also concurs with recent neuroimaging literature looking at prediction error coding in the striatum, observing increases in BOLD response with positive prediction errors (when an unexpected reward is delivered) (McClure et al., 2003) and decreases with negative prediction errors (when an expected reward is withdrawn; McClure et al., 2003; O'Doherty et al., 2003).

Interestingly, these recent prediction learning studies have observed that the response to reward shifts during learning. Specifically, once the reward outcome becomes fully predictable, the response to the reward is no longer observed, instead shifting to the earliest predictor of the reward after learning. The results observed in the current experiment show that during presentation of a predictive cue in the probabilistic cue period (the potential earliest predictor of a reward once a contingency is well-learned), caudate activity is more robust during early learning of contingencies. This is in accordance with cognitive studies of probability learning (Myers et al., 2003; Shohamy et al., 2004), but in contrast with reward learning studies involving temporal prediction (McClure et al., 2003; O'Doherty et al., 2003). This discrepancy may be due to distinct features of the paradigms, specifically type of experimental procedure (classical × instrumental), striatum focus of interest (putamen × caudate) and type of reinforcer (primary × secondary). Indeed, in the previously mentioned tasks (McClure et al., 2003; O'Doherty et al., 2003), activity was reported in the dorsal and ventral putamen using paradigms where gustatory rewards (a primary reinforcer) were delivered in a passive manner (classical conditioning). In contrast, the current experiment observed activation in the caudate nucleus using a task where the monetary outcome (a secondary reinforcer) was contingent on an action (instrumental conditioning). However, it is possible that both response profiles discussed, decreasing over time and shifting in time, occur in the striatum, just in distinct subregions. Indeed, exploratory investigations of the cue period response using the ROI acquired in the comparison between 100% and 50% condition reward responses during the late phases of learning suggest that there is some degree of shift in time, as the activity for the 100% condition is slightly larger after cue presentation. Further work, specifically testing for some of the variables mentioned before (such as experimental procedure and type of reinforcer), will be necessary to fully understand shifts over time of conditioned responses in the caudate nucleus.

The potential affinity of the caudate nucleus to processing reinforcement signals from outcomes contingent on an action explains why our activity was stronger during the feedback period rather than the cue period (Tricomi et al., 2004). This does not indicate that the caudate nucleus is responding to choice, but it suggests, in accordance to actor–critic models of learning (Barto, 1994), that the caudate nucleus may be integral for reinforcement learning, and thus reinforcing good choices. Still, the recruitment of the caudate nucleus during the probabilistic cue period may be based on making predictions that are contingent on a latter action. Perhaps, the inclusion of trials where no action is necessary (passive or classical conditioning task) would not recruit the dorsal, but instead the ventral striatum would be apparent. It should also be noted that the caudate nucleus location in this study is slightly more ventral than other caudate nucleus focus of interest found in other reward-related paradigms that did not involve learning (Delgado et al., 2000, 2003), but in accordance with other studies that had a learning component (Haruno et al., 2004; O'Doherty et al., 2004).

Finally, activation of a frontal-striatal-thalamic circuit is not uncommon in learning studies, and perhaps highlights the steady contribution of these areas during learning. The activity observed in the frontal cortex during the late stages of learning, showing higher activity of predictable rather than unpredictable outcomes, may suggest that after the striatum aids in learning an association, prefrontal regions assume a larger role in mediating behavior, although more data is necessary to support this claim. Overall, our results suggest that the dorsal striatum, and in particular the caudate nucleus, has an important role in processing reward-related learning. It is integral during early acquisition of action–outcome contingencies and is modulated as learning progresses, showing less response to reward feedback as action–outcome associations become fully learned and more predictable.

Acknowledgments

This work was supported by NIMH 62104 to EAP and Center for Brain Imaging, NYU. The authors wish to acknowledge Kate Fissell, Brett Sedgewick and Ben Holmes for technical assistance, Susan Ravizza and Elizabeth Tricomi for informative discussion and constructive criticism. The authors would also like to acknowledge the support of the Beatrice and Samuel A. Seaver Foundation.

References

Aharon, I., Etcoff, N., Ariely, D., Chabris, C.F., O'Connor, E., Breiter, H.C., 2001. Beautiful faces have variable reward value: fMRI and behavioral evidence. Neuron 32, 537–551.

Apicella, P., Ljungberg, T., Scarnati, E., Schultz, W., 1991. Responses to reward in monkey dorsal and ventral striatum. Exp. Brain Res. 85, 491–500.

Apicella, P., Scarnati, E., Ljungberg, T., Schultz, W., 1992. Neuronal activity in monkey striatum related to the expectation of predictable environmental events. J. Neurophysiol. 68, 945–960.

Aron, A.R., Shohamy, D., Clark, J., Myers, C., Gluck, M.A., Poldrack, R.A., 2004. Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. J. Neurophysiol. 92, 1144–1152.

Barto, A.G., 1994. Reinforcement learning control. Curr. Opin. Neurobiol. 4, 888–893.

Berns, G.S., McClure, S.M., Pagnoni, G., Montague, P.R., 2001. Predictability modulates human brain response to reward. J. Neurosci. 21, 2793–2798.

Breiter, H.C., Aharon, I., Kahneman, D., Dale, A., Shizgal, P., 2001. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. Neuron 30, 619–639.

Brown, R.G., Redondo-Verge, L., Chacon, J.R., Lucas, M.L., Channon, S., 2001. Dissociation between intentional and incidental sequence learning in Huntington's disease. Brain 124, 2188–2202.

Carelli, R.M., Wolske, M., West, M.O., 1997. Loss of lever press-related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. J. Neurosci. 17, 1804–1814.

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

Delgado, M.R., Nystrom, L.E., Fissell, C., Noll, D.C., Fiez, J.A., 2000. Tracking the hemodynamic responses to reward and punishment in the striatum. J. Neurophysiol. 84, 3072–3077.

Delgado, M.R., Locke, H.M., Stenger, V.A., Fiez, J.A., 2003. Dorsal striatum

responses to reward and punishment: effects of valence and magnitude manipulations. Cognit. Affective Behav. Neurosci. 3, 27–38.
Delgado, M.R., Stenger, V.A., Fiez, J.A., 2004. Motivation-dependent responses in the human caudate nucleus. Cereb. Cortex 14, 1022–1030.
Di Chiara, G., Imperato, A., 1988. Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. Proc. Natl. Acad. Sci. U. S. A. 85, 5274–5278.
Elliott, R., Friston, K.J., Dolan, R.J., 2000. Dissociable neural responses in human reward systems. J. Neurosci. 20, 6159–6165.
Elliott, R., Newman, J.L., Longe, O.A., William Deakin, J.F., 2004. Instrumental responding for rewards is associated with enhanced neuronal response in subcortical reward systems. NeuroImage 21, 984–990.
Fiorillo, C.D., Tobler, P.N., Schultz, W., 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. Science 299, 1898–1902.
Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn. Reson. Med. 33, 636–647.
Gluck, M.A., Shohamy, D., Myers, C., 2002. How do people solve the "weather prediction" task?: individual variability in strategies for probabilistic category learning. Learn. Mem. 9, 408–418.
Gottfried, J.A., O'Doherty, J., Dolan, R.J., 2002. Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. J. Neurosci. 22, 10829–10837.
Haber, S.N., 2003. The primate basal ganglia: parallel and integrative networks. J. Chem. Neuroanat. 26, 317–330.
Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., Imamizu, H., Kawato, M., 2004. A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. J. Neurosci. 24, 1660–1665.
Hikosaka, O., Sakamoto, M., Usui, S., 1989. Functional properties of monkey caudate neurons: III. Activities related to expectation of target and reward. J. Neurophysiol. 61, 814–832.
Ito, R., Dalley, J.W., Robbins, T.W., Everitt, B.J., 2002. Dopamine release in the dorsal striatum during cocaine-seeking behavior under the control of a drug-associated cue. J. Neurosci. 22, 6247–6253.
Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V., Graybiel, A.M., 1999. Building neural representations of habits. Science 286, 1745–1749.
Kawagoe, R., Takikawa, Y., Hikosaka, O., 1998. Expectation of reward modulates cognitive signals in the basal ganglia. Nat. Neurosci. 1, 411–416.
Kawagoe, R., Takikawa, Y., Hikosaka, O., 2004. Reward-predicting activity of dopamine and caudate neurons—a possible mechanism of motivational control of saccadic eye movement. J. Neurophysiol. 91, 1013–1024.
Kirsch, P., Schienle, A., Stark, R., Sammer, G., Blecker, C., Walter, B., Ott, U., Burkart, J., Vaitl, D., 2003. Anticipation of reward in a nonaversive differential conditioning paradigm and the brain reward system: an event-related fMRI study. NeuroImage 20, 1086–1095.
Knowlton, B.J., Mangels, J.A., Squire, L.R., 1996. A neostriatal habit learning system in humans. Science 273, 1399–1402.
Knutson, B., Westdorp, A., Kaiser, E., Hommer, D., 2000. fMRI visualization of brain activity during a monetary incentive delay task. NeuroImage 12, 20–27.
Knutson, B., Adams, C.M., Fong, G.W., Hommer, D., 2001a. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. J. Neurosci. 21, RC159.
Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L., Hommer, D., 2001b. Dissociation of reward anticipation and outcome with event-related fMRI. NeuroReport 12, 3683–3687.
Koepp, M.J., Gunn, R.N., Lawrence, A.D., Cunningham, V.J., Dagher, A., Jones, T., Brooks, D.J., Bench, C.J., Grasby, P.M., 1998. Evidence for striatal dopamine release during a video game. Nature 393, 266–268.
Lauwereyns, J., Watanabe, K., Coe, B., Hikosaka, O., 2002. A neural correlate of response bias in monkey caudate nucleus. Nature 418, 413–417.
Macwhinney, B., Cohen, J., Provost, J., 1997. The PsyScope experiment-building system. Spat. Vis. 11, 99–101.
McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal prediction errors in a passive learning task activate human striatum. Neuron 38, 339–346.
Middleton, F.A., Strick, P.L., 2000. Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. Brain Cogn. 42, 183–200.
Myers, C.E., Shohamy, D., Gluck, M.A., Grossman, S., Kluger, A., Ferris, S., Golomb, J., Schnirman, G., Schwartz, R., 2003. Dissociating hippocampal versus basal ganglia contributions to learning and transfer. J. Cogn. Neurosci. 15, 185–193.
O'Doherty, J.P., Deichmann, R., Critchley, H.D., Dolan, R.J., 2002. Neural responses during anticipation of a primary taste reward. Neuron 33, 815–826.
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.
O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., Dolan, R.J., 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304, 452–454.
Packard, M.G., Knowlton, B.J., 2002. Learning and memory functions of the basal ganglia. Annu. Rev. Neurosci. 25, 563–593.
Packard, M.G., McGaugh, J.L., 1996. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. Neurobiol. Learn. Mem. 65, 65–72.
Pagnoni, G., Zink, C.F., Montague, P.R., Berns, G.S., 2002. Activity in human ventral striatum locked to errors of reward prediction. Nat. Neurosci. 5, 97–98.
Poldrack, R.A., Prabhakaran, V., Seger, C.A., Gabrieli, J.D., 1999. Striatal activation during acquisition of a cognitive skill. Neuropsychology 13, 564–574.
Poldrack, R.A., Clark, J., Pare-Blagoev, E.J., Shohamy, D., Creso Moyano, J., Myers, C., Gluck, M.A., 2001. Interactive memory systems in the human brain. Nature 414, 546–550.
Ravel, S., Legallet, E., Apicella, P., 2003. Responses of tonically active neurons in the monkey striatum discriminate between motivationally opposing stimuli. J. Neurosci. 23, 8489–8497.
Reynolds, J.N., Hyland, B.I., Wickens, J.R., 2001. A cellular mechanism of reward-related learning. Nature 413, 67–70.
Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science 275, 1593–1599.
Schultz, W., Tremblay, L., Hollerman, J.R., 2000. Reward processing in primate orbitofrontal cortex and basal ganglia. Cereb. Cortex 10, 272–283.
Schultz, W., Tremblay, L., Hollerman, J.R., 2003. Changes in behavior-related neuronal activity in the striatum during learning. Trends Neurosci. 26, 321–328.
Shohamy, D., Myers, C.E., Grossman, S., Sage, J., Gluck, M.A., Poldrack, R.A., 2004. Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. Brain 127, 851–859.
Takikawa, Y., Kawagoe, R., Hikosaka, O., 2002. Reward-dependent spatial selectivity of anticipatory activity in monkey caudate neurons. J. Neurophysiol. 87, 508–515.
Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain: An Approach to Medical Cerebral Imaging. G. Thieme, Thieme Medical Publishers, Stuttgart, New York.
Tricomi, E.M., Delgado, M.R., Fiez, J.A., 2004. Modulation of caudate activity by action contingency. Neuron 41, 281–292.
Volkow, N.D., Wang, G.J., Fowler, J.S., Logan, J., Jayne, M., Franceschi, D., Wong, C., Gatley, S.J., Gifford, A.N., Ding, Y.S., Pappas, N., 2002. "Nonhedonic" food motivation in humans involves dopamine in the dorsal striatum and methylphenidate amplifies this effect. Synapse 44, 175–180.
Watanabe, K., Lauwereyns, J., Hikosaka, O., 2003. Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. J. Neurosci. 23, 10052–10057.
White, N.M., McDonald, R.J., 2002. Multiple parallel memory systems in the brain of the rat. Neurobiol. Learn. Mem. 77, 125–184.

Wickens, J.R., Reynolds, J.N., Hyland, B.I., 2003. Neural mechanisms of reward-related motor learning. Curr. Opin. Neurobiol. 13, 685–690.

Woods, R.P., Cherry, S.R., Mazziotta, J.C., 1992. Rapid automated algorithm for aligning and reslicing PET images. J. Comput. Assist. Tomogr. 16, 620–633.

Woods, R.P., Mazziotta, J.C., Cherry, S.R., 1993. MRI-PET registration with automated algorithm. J. Comput. Assist. Tomogr. 17, 536–546.

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. J. Neurosci. 24, 4105–4112.

Zink, C.F., Pagnoni, G., Martin, M.E., Dhamala, M., Berns, G.S., 2003. Human striatal response to salient nonrewarding stimuli. J. Neurosci. 23, 8092–8097.

Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., Chappelow, J.C., Berns, G.S., 2004. Human striatal responses to monetary reward depend on saliency. Neuron 42, 509–517.