

Two sides of the same coin: Learning via positive and negative reinforcers in the human striatum

Michael A. Niznikiewicz, Mauricio R. Delgado*

Department of Psychology, Rutgers University, Newark, NJ 07102, United States

ARTICLE INFO

Article history:

Received 25 March 2011

Received in revised form 2 July 2011

Accepted 13 July 2011

Keywords:

Approach

Avoidance

Reinforcement learning

Ventral striatum

Nucleus accumbens

Caudate

Dopamine

Cingulate gyrus

ABSTRACT

The human striatum has been previously implicated in the processing of positive reinforcement, but less is known about its role in processing negative reinforcement. In this experiment, participants learn specific approach or avoidance responses, mediated by positive and negative reinforcers respectively to investigate how affective learning and associated neural activity are influenced by the motivational context in which learning occurs. The paradigm was divided into two discrete sessions, where participants could either earn monetary rewards (approach sessions) or avoid monetary losses (avoid sessions) based on successful learning. Specifically, a conditioned cue predicted the chance to win or avoid losing money contingent on a correct button press (pre-learning trials), which upon learning led to the delivery of rewards or termination of losses (post-learning trials). Skin conductance responses (SCRs) and subjective ratings confirmed a learning effect (greater SCRs pre vs. post-learning) irrespective of reinforcer valence. Concurrently, activity in the ventral striatum was characterized by a similar learning effect, with greater responses during pre-learning. Interestingly, such learning effect was enhanced in the presence of a negative reinforcer, as suggested by an interaction between learning phase and session, highlighting the influence negative reinforcers can have on striatal circuits involved in learning and motivated behavior.

Published by Elsevier Ltd.

1. Introduction

Across human development, learning is often motivated by a variety of reinforcers ranging from stimuli necessary for survival (e.g., food) to more abstract stimuli in the environment (e.g., social approval). A common goal of these reinforcers is to increase the frequency of a behavior, although the context in which this occurs can be either positive or negative (Skinner, 1938). For instance, an increase in a student's study habits could be due to the desire to earn a good grade and praise from parents, both which

serve as examples of positive reinforcers. Alternatively, the boosted study time could be attributed to a desire to avoid the negative feelings associated with a failing grade and parental disapproval, serving as negative reinforcers in this context. In both cases, the behavioral output is similar; however, the context in which learning occurs is different and could lead to long-term consequences in future goal-directed behaviors (e.g., excessive approach or avoidance responses). Thus, it is important to understand the influence positive and negative reinforcers can have on behaviors and in turn how they can modulate associated neural mechanisms that are typically involved in reinforcement learning.

One brain region that has been repeatedly implicated in reward processing is the striatum, the input unit of the basal ganglia, and a region that makes important connections with various cortical inputs to influence motor,

* Corresponding author at: Department of Psychology, Rutgers University, Smith Hall, Room 340, 101 Warren Street, Newark, NJ 07102, United States. Tel.: +1 973 353 3949; fax: +1 973 353 1171.

E-mail address: delgado@psychology.rutgers.edu (M.R. Delgado).

cognitive and motivated behavior (Alexander et al., 1986; Haber and Knutson, 2010; Middleton and Strick, 2000). This is illustrated by an elegant animal literature which highlights, for example, that lesions in the ventral striatum of rodents lead to deficits in approach behaviors (for review see Robbins and Everitt, 1996), while neuronal recordings from this same region in non-human primates respond to conditioned stimuli that predict potential rewards (Cromwell and Schultz, 2003; Hassani et al., 2001). In humans, the striatum has been associated with reward-related learning in a variety of paradigms, with positively valenced conditioned cues eliciting approach-like behavior during instrumental learning (for review see Delgado, 2007; Montague and Berns, 2002; O'Doherty, 2004; Rangel et al., 2008).

The striatum has also been posited as a key component of models of motivated behavior throughout development (Casey et al., 2008; Ernst et al., 2006). Such models suggest that prefrontal cortical control centers known to be involved in regulating emotional responses (Ochsner and Gross, 2005) are slower to develop through adolescence in comparison to more subcortical structures such as the striatum (Casey et al., 2008; Ernst et al., 2006). As a result, activation of the striatum in response to rewarding stimuli tends to be exaggerated in adolescence (e.g., Ernst et al., 2005; Galvan et al., 2005; May et al., 2004; but see Bjork et al., 2004) and linked to increased propensity for risky decision-making often observed during this period of development (e.g., Reyna et al., 2011; Van Leijenhorst et al., 2010).

More recently, the human striatum has also been linked with learning in a negative context, such as learning to avoid a mild shock (Delgado et al., 2009; Jensen et al., 2003). However, activation of the striatum during anticipation of aversive events is not always observed (e.g., Breiter et al., 2001; Gottfried et al., 2002; Yacubian et al., 2006), with some reports suggesting that the ventral striatum is primarily involved in reward-related processing, and not responsive when the context of action-learning is more negative, such as the avoidance of monetary loss (e.g., Knutson et al., 2001). Further, the amygdala is the structure most often associated with aversive learning, as evidenced by animal models of fear conditioning (for review see Phelps and LeDoux, 2005) and neuropsychological investigations of fear learning in patients with amygdala lesions (Bechara et al., 1995; LaBar et al., 1995). The amygdala has also been hypothesized in some models to mediate avoidance behaviors during development (Ernst et al., 2006). Thus, more research is necessary to clarify the influence of negative reinforcers on motivated learning in striatal circuits that are more typically associated with reward processing.

In this experiment, we take advantage of the fact that monetary incentives represent a common reinforcer that can be either positive (gains) or negative (losses). Specifically, we adapted the paradigm from Delgado et al. (2009) to include both approach and avoidance learning sessions and allow for within subjects comparisons of learning by positive and negative reinforcers respectively. Before each learning session, participants played a simple gambling game (adapted from Delgado et al., 2000) in order to endow

them with an experimental monetary bank. Their goal for the approach learning sessions was to build upon this bank by learning, through trial and error, the appropriate action that led to a positive outcome. The goal of the avoidance learning sessions was to keep from losing the money they just earned in the gambling session by learning the appropriate action that avoided a monetary loss. Aside from visual characteristics of the conditioned stimuli (i.e., color) and the valence of the reinforcement (i.e., positive or negative), the approach and avoidance learning sessions were comparable, in turn allowing for a direct comparison of striatal learning systems under positive or negative contexts.

Given the previously described role of the human striatum in affective learning, we hypothesized that the striatum would be involved in learning with both positive and negative reinforcers. Furthermore, we predicted that despite overlapping neural circuitry when approaching or avoiding an affective stimulus, negative reinforcers would lead to greater influences on striatum responses involved in mediating reinforcing effects on behavior. This prediction is consistent with the observation that losses loom larger than gains (Kahneman and Tversky, 1979) and decisions made under negative, compared to positive, contexts can have a greater impact on behavior and striatal BOLD signals (Delgado et al., 2008). These results would suggest that neural mechanisms underlying the acquisition of adaptive behavioral responses to attain goals can be shaped by the motivational context in which they are learned.

2. Methods

2.1. Participants

Twenty-five participants were recruited from the population of students at Rutgers University – Newark. From this subset, four participants were excluded due to failure to comply with experimental requirements (e.g., poor understanding of instructions), while two more were excluded because of scanner malfunction. Thus, nineteen participants comprised the final analysis (10 females, mean age 22 ± 4.5 years). Participants were prescreened for contraindications to MRI, for right-handedness, and had normal or corrected vision. All participants gave informed consent according to the Internal Review Board of Rutgers University Newark and the University of Medicine and Dentistry of New Jersey approved the experiment.

2.2. Experimental procedure

2.2.1. Overview

The main goal of the experiment for participants was to learn the appropriate action that either led to a positive reinforcer (approach learning session) or turned off a negative reinforcer (avoidance learning session). To accomplish this, there were two main types of scan sessions. During approach learning sessions, participants were asked to learn associations through positive reinforcement, where the outcomes represented gain or no gain of money. In contrast, avoidance learning sessions required participants to learn by negative reinforcement, where they could only lose or not lose money (adapted from Delgado et al., 2009).

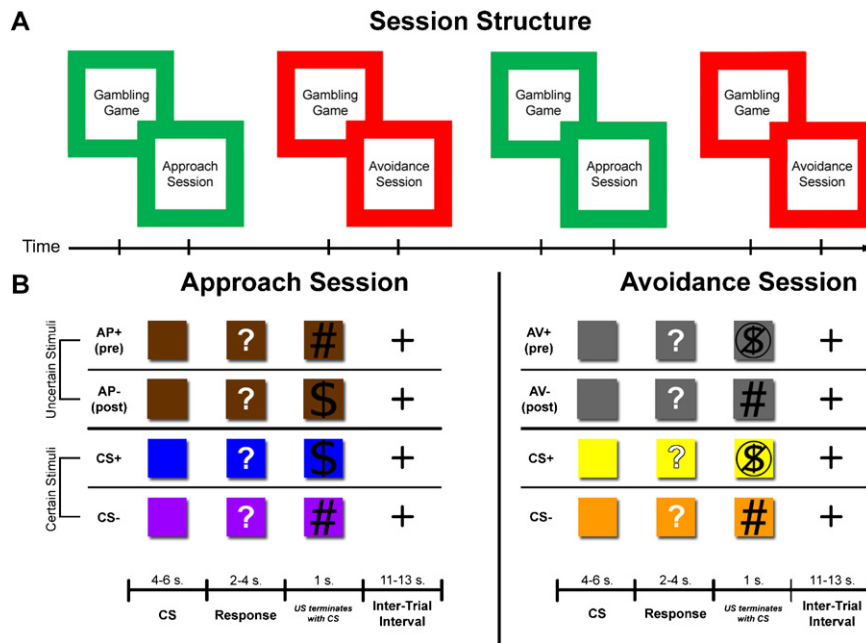


Fig. 1. Experimental procedure. (A) Participants were presented with counterbalanced rounds of approach and avoidance learning twice, each preceded by a gambling game. (B) Approach and avoidance learning sessions contained three types of stimuli, 2 certain and 1 uncertain. In the avoidance session, the CS+ resulted in a monetary loss, the CS- resulted in no loss of money, and the AV resulted in monetary loss (pre) until participants learned the correct response (post). In the approach session, the CS+ resulted in monetary gain, the CS- resulted in no gain of money, and the AP resulted in monetary gain only after subjects had learned the correct response (post).

Each session was preceded by a simple gambling game designed to provide an experimental bank for participants (adapted from Delgado et al., 2000, 2006). Throughout the experiment, participants played four counterbalanced sessions, two approach and two avoidance, each with its own gambling game and own independent pot of money (Fig. 1a). Participants made right hand responses using a 4-button MRI-compatible response unit.

2.2.2. Approach and avoidance learning sessions

During the approach learning session, participants were presented with three colored squares that predicted a monetary outcome and required a motor response. Two squares were fully predictable and led to either a monetary gain (conditioned stimulus or CS+) or no monetary gain (CS-). Such stimuli required a motor response that did not influence the outcome and are henceforth referred to as certain stimuli as they predicted an outcome with 100% certainty. The third square (the approach stimulus or AP) predicted a monetary gain contingent on an appropriate motor response (i.e., a button press), thus referred to as an uncertain stimulus. Once the correct response was learned, repeated actions led to monetary rewards. The structure of the avoidance learning session was identical to the approach session for direct comparison purposes, except for the different color of the stimuli and the negative context of the session. During the avoidance learning session, participants learned through negative instead of positive reinforcement. That is, the uncertain square in this session, known as the avoidable or AV stimulus, resulted in monetary loss until the participant learned the appropriate motor response. After this point, the participant could

avoid losing money. The CS+ within this session predicted monetary loss with 100% certainty, while the CS- predicted no monetary loss. Thus, each session was comprised of 3 colored squares consisting of two fully predictable conditioned stimuli (certain stimuli) and one stimulus whose outcome value was contingent on an appropriate motor response (uncertain stimulus). Overall, there were 12 certain (6 CS+; 6 CS-) and 12 uncertain (AV/AP) trials per session. The color of the squares and the order of the sessions were counterbalanced across participants.

Participants were told that the correct answer to the AV and AP stimuli was one of twelve possible choices. Since the button box used to collect participants' right-hand responses only had 4 buttons, they were informed that the correct choice could be the first, second, or third time they pressed a button – yielding twelve possible choices. Unbeknownst to the participant, he or she determined the correct answer for the AV and AP stimuli with the response to the sixth (out of 12) presentation of that stimulus. Prior to the sixth presentation of the uncertain stimuli, any participant response would lead to a monetary loss (avoid session) or no monetary gain (approach session). This predetermined schedule of reinforcement ensured that all participants experienced the same number of pre and post learning trials over the course of the experiment (e.g., 6 AV+, 6 AV-). Learning, therefore, was operationally defined as the transition between a period of exploration for the appropriate answer (pre-learning trials) to the expression of the learned response (post-learning trials). Any participant that missed the sixth trial, or failed to use the learned response in subsequent trials was excluded from further analysis due to a lack of post-learning phase. Such par-

Participants ($n = 4$) typically reported not paying attention or losing focus on the goal of the experiment.

Every trial of a learning session began with a presentation of a colored square (cue phase; 4–6 s) which predicted a potential outcome (i.e., CS+, CS– or uncertain stimulus). A question mark then appeared serving as an indicator for participants to choose one of four buttons to respond (response phase; 2–4 s). The outcome phase immediately followed (1 s) and co-terminated with the CS. Three potential symbols were presented in the outcome phase: a dollar sign symbolizing a monetary gain, a crossed out dollar sign depicting a monetary loss, and a pound sign representing no outcome. The trial concluded with a jittered inter-trial interval (11–13 s) [see Fig. 1b]. The primary phase of interest was the cue phase as it served as the initial representation of the conditioned stimulus without being affected by motor responses. At the end of each scanning session, participants rated the stimuli they had just seen using a Likert scale from 1 to 7. Results from these ratings questions were used to confirm that participants were paying attention and that they understood the contingencies presented. Specifically, participants rated how much they liked or disliked each conditioned stimulus and also how emotionally arousing each stimulus was.

2.2.3. Gambling sessions

Prior to every learning session, participants engaged in a simple gambling task (adapted from Delgado et al., 2000, 2006). There were two important goals for the gambling session. First, it served to provide participants with an experimental bank which could either be added to (approach session) or subtracted from (avoid session). Second, the gambling session served as an independent way to define reward circuitry regions of interest (ROIs) as previously shown in experiments using this task (see Delgado, 2007 for review). In the gambling session, participants were told to guess whether the presented card had a value higher (e.g., 6, 7, 8, 9) or lower (e.g., 1, 2, 3, 4) than 5. At the onset of the trial, participants were presented with a question mark prompting them to enter their response (i.e., high or low) within 2 s. The question mark was then replaced by the actual value of the card and an outcome symbol for 2 s that indicated whether the participant was correct (a monetary reward depicted by a green check mark) or incorrect (a monetary loss depicted by a red "X"). A jittered 10–12 s inter-trial interval followed each trial for a total of 14–16 s per trial. The amount of times they won and lost was predetermined (9 reward trials, 6 loss trials, randomized) so that they would end each gambling session with a net gain of money. Over the course of the four gambling sessions, the participant experienced 60 trials (36 reward, 24 loss) trials.

2.2.4. Method of reinforcement

In this experiment, the actual monetary value of a single trial in any of the sessions was ambiguous until the end of the experiment. The goal of this procedure was to ensure that the only thing that mattered for participants was the occurrence (or non-occurrence) of a reinforcer. That is, participants considered the ultimate valence of the outcome of a trial (i.e., positive or negative) rather than its magnitude

or absolute value. Specifically, participants were instructed that they would spin a wheel with 8 values ranging from \$1.50 to \$5.00 increasing in \$0.50 increments at the end of the experiment. The value the wheel landed on would be applied to every gain and loss in the experiment. Thus, participants are made aware that the monetary incentives are real, but that they need only focus on the affective components of the outcomes during the task. Participants' final compensation was comprised of an experimental rate (\$25/h) and any monetary incentives earned during the experiment (total of \$55).

2.2.5. Materials

The experimental paradigm was programmed using E-PRIME software, v2.0 (PST, Pittsburgh, PA). Stimuli were set against a black background and projected onto a screen, which was visible inside the scanner using a mirror attached to the head coil. At the end of the experimental session, participants were debriefed and compensated.

2.3. Data acquisition and analysis

2.3.1. Behavioral data acquisition and analysis

The schedule of reinforcement was programmed in such a way that all participants learn the correct answer to AV/AP stimuli after 6 trials. Therefore, differences in accuracy were not expected, and failure to stick with the correct answer resulted in exclusion. The main behavioral measure was reaction time (RT), which has been previously used to show differences in motivation between stimuli in similar paradigms (e.g., Delgado et al., 2009). In particular, stimuli that allowed participants the opportunity to avoid a punishment elicited faster reaction times and as such were considered more motivating than those that were uncontrollable. Reaction time was tested using a 2×2 repeated measures analysis of variance (ANOVA) with session (approach and avoidance learning session) and stimulus type (certain and uncertain) as within-subjects factors. Specifically, this analysis tested the prediction that uncertain stimuli (e.g., AV+/AV–) would be more motivating and thus elicit faster responses than responses recorded during certain trials (CS+/CS–). Finally, subjective ratings were acquired at the end of each of the four learning sessions and served as manipulation checks. These ratings probed both the valence ("How much did you like this stimulus?") and the intensity ("How much emotion of any kind did you feel when you saw this stimulus?") associated with each conditioned stimulus using a Likert scale from 1 to 7. A 2×3 repeated measures ANOVA with session (approach and avoidance learning session) by stimulus type (AV/AP, CS+, CS–) as within-subjects factors was conducted to probe subjective ratings of valence and intensity.

2.3.2. Physiological data acquisition and analysis

Throughout the experiment, skin conductance responses (SCRs) were gathered from the first and second fingers of the participant's left hand using BIOPAC systems skin conductance module. Data acquired through shielded Ag–AgCl electrodes, which were grounded through an RF filter panel, was transmitted to a data collection station within the control room at the scanning facility.

ACQKNOWLEDGE software facilitated the analysis of SCR waveforms. Each response was scored using a 0.5–4.5 s window after the onset of the stimulus. A minimum base to peak difference of 0.02 mS (micro Siemens) for each response was used as a criterion, with lower responses scored as 0. Given this criteria, data for 5 participants were not included in the final analysis due to low amount of responses. The square root of the SCR was taken prior to statistical analysis to reduce skewness (LaBar et al., 1998). SCRs across the four learning sessions were averaged per participant and per type of trial focusing on the cue phase. The main analysis of interest consisted of a 2×2 session (approach and avoidance learning sessions) by learning phase (AV+/AP+ i.e., pre-learning, and AV-/AP- i.e., post-learning) repeated measures ANOVA, comparing affective measures of arousal during approach and avoidance learning. A second 2×2 repeated measures ANOVA was also conducted using type of session (approach and avoidance) and certain stimulus type (CS+, CS-).

2.4. fMRI acquisition and analysis

2.4.1. Acquisition

Functional magnetic resonance imaging (fMRI) data was acquired using a 3 T Siemens Allegra head-only scanner and a Siemens standard head coil at the University of Medicine and Dentistry of New Jersey's Advanced Imaging Center. A T1-weighted protocol (256×256 matrix, 176 1-mm sagittal slices) was used to gather high-resolution anatomical images. Functional images were acquired using a single-shot gradient echo EPI sequence (TR = 2000 ms, TE = 25 ms, FOV = 192 cm, flip angle = 80° , bandwidth = 2604 Hz/px, echo spacing = 0.29 ms). Thirty-five contiguous oblique-axial slices ($3 \text{ mm} \times 3 \text{ mm} \times 3 \text{ mm}$ voxels) parallel to the anterior commissure–posterior commissure (AC–PC) line were obtained.

2.4.2. Analysis

The Brain Voyager statistical analysis package (Brain Innovation, Maastricht, The Netherlands; v2.2) was used to analyze the imaging data. Motion correction (using a threshold of 3 mm or less), and slice scan time correction using Trilinear/sinc interpolation was applied to the data to correct for movement and to align data to a single time point. Further, spatial smoothing was performed using a three-dimensional Gaussian filter (4-mm FWHM), along with voxel-wise linear detrending and high-pass filtering of frequencies (3 cycles per time course). Structural and functional data of each participant was then transformed to standard Talairach stereotaxic space (Talairach and Tournoux, 1988).

A random-effects general linear model (GLM) was used to analyze the data of the 19 final participants. The gambling sessions were modeled using two regressors representing reward and loss trials. The learning sessions contained four regressors at the cue phases of the uncertain AV and AP stimuli across pre and post learning (AV+, AV-, AP+, AP-) and four regressors modeled for the certain stimuli (CS+, CS-, in each session respectively). There were also eight regressors of no interest used to model the

response and outcome phase, along with six regressors for motion and one for missed trials.

There were three analyses performed to investigate the neural correlates underlying approach and avoidance learning in this paradigm. The primary analysis involved functionally defining independent ROIs implicated in reward processing (see Delgado, 2007 for review). This was done by contrasting reward and loss trials within the gambling task. These functionally defined ROIs then served as task-independent regions to compare approach and avoidance learning. We then extracted mean parameter estimates (i.e., beta weights) from these ROIs during the learning sessions. The resulting data was input into a 2×2 repeated measures ANOVA to investigate the effect of session and learning phase (pre-learning, post-learning). The Statistical Parametric Map (SPM) for the gambling session analysis was corrected at a False Discovery Rate (FDR) < 0.005 .

The other two analyses involved separate whole-brain 2×2 ANOVAS to examine the main effects and interactions of (a) session and learning phase (pre vs. post) and (b) session and certain stimuli (CS+ and CS-). Statistical Parametric Maps (SPMs) were thresholded at $p < 0.001$ and then corrected using a voxel cluster method. Specifically, a voxel cluster threshold of 4 contiguous voxels (mm^3) yielded a corrected alpha < 0.05 for this analysis according to Brain Voyager's Cluster Thresholding Plugin (Forman et al., 1995; Goebel et al., 2006). This level of statistical correction was used for each main effect and interaction SPM defined by these two ANOVAs. Mean parameter estimates were then extracted from regions surviving these criteria for further analysis, and *post hoc* two-tailed paired sample *t*-tests were used to investigate the differences between regressors.

3. Results

3.1. Behavioral and physiological results

3.1.1. Subjective ratings

Ratings were acquired after each session (gambling and learning) as a manipulation check to ensure participants were engaged in the task. In the gambling sessions, participants showed a preference for stimuli associated with winning money over those associated with losing money [$t(19) = 14.9$, $p < 0.001$]. In the learning sessions, two subjective rating measures were acquired: ratings of perceived valence and intensity of the stimulus. A 2×3 session by stimulus repeated measures ANOVA probing ratings of intensity revealed a main effect of session ($F(1, 18) = 10.3$, $p < 0.005$), stimulus type ($F(2,36) = 22.12$, $p < 0.001$), and a session by condition interaction ($F(2, 36) = 14.675$, $p < 0.001$) [Fig. 2a]. In the approach learning sessions, the AP stimulus was ranked as more emotionally arousing than both the CS+ and the CS- stimulus [$t(18) = 2.67$, $p < 0.015$ and $t(18) = 5.93$, $p < 0.001$ respectively]. The CS+ was also ranked more emotionally arousing than the CS- [$t(18) = 3.54$, $p < 0.002$]. In the negative reinforcement sessions, the AV stimulus was ranked more emotionally arousing than the CS+ and CS- [$t(18) = 6.02$, $p < 0.001$ and $t(18) = 2.87$, $p < 0.01$ respectively], and the CS- was ranked as more emotionally arousing than the

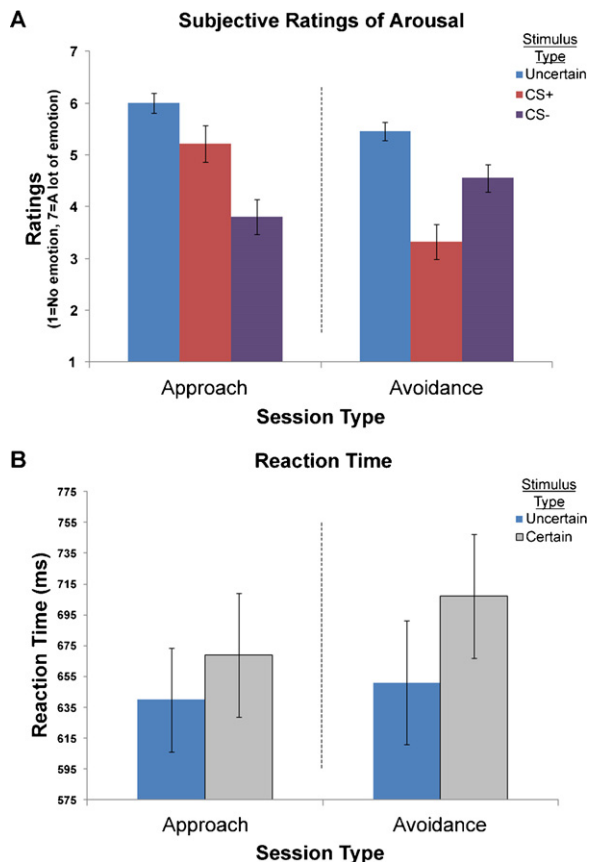


Fig. 2. Behavioral results. (A) Subjective ratings of arousal showing greater responses for uncertain (AP/AV) compared to certain (CS+, CS–) stimuli (\pm s.e.m.). Uncertain trials refer to AP trials during approach sessions and AV trials during avoidance sessions. (B) Reaction time measures indexing potential motivational differences between the uncertain and certain stimuli as expressed by faster responses during uncertain stimuli irrespective of type of reinforcer (positive or negative; \pm s.e.m.).

CS+ [$t(18)=3.15, p<0.006$]. When comparing the AP and AV stimuli, pair-samples two-tailed t -tests showed the AP stimuli being ranked as more emotionally arousing than the AV stimuli [$t(18)=3.07, p<0.007$]. Similar results were observed for the ANOVA investigating subjective ratings of valence, with the exception of the difference between AP and AV stimuli, which was merely a trend [$t(18)=1.94, p=0.068$].

3.1.2. Reaction time

A 2×2 repeated measures ANOVA revealed no effect of session or interaction, but did show a main effect of stimulus type (certain and uncertain stimuli; $F(1, 18)=14.09, p<0.001$) as hypothesized (Fig. 2b). *Post hoc* paired samples two-tailed t -tests revealed that uncertain stimuli elicited faster reaction times than certain stimuli [$t(18)=3.917, p=0.001$]. Because of a priori hypotheses that negative reinforcers would lead to greater influences on striatum signals and behavioral responses, we conducted an exploratory *post hoc* paired samples t -test probing the differences between uncertain and certain stimuli in both avoidance and approach session separately. We found a sig-

nificant difference between uncertain and certain stimuli during avoidance [$t(18)=3.07, p<0.01$], but not approach [$t(18)=1.44, p<0.17$] learning sessions. However, it should be noted that this result is exploratory since no interaction was observed.

3.1.3. Skin conductance responses

A 2×2 repeated measures ANOVA on the SCR cue phase data investigating session by learning phase revealed no effect of session [$F(1, 13)=0.417, p>0.05$], but a trending main effect of learning phase [$F(1, 13)=4.64, p=0.051$]. Within this factor, each pre-learning phase elicited a slightly higher SCR response than the post-learning phase. Importantly, no interaction between session and stimulus type was observed [$F(1, 13)=0.909, p>0.05$], suggesting that participants' physiological level of arousal was not different across approach and avoid learning sessions. Within the certain stimuli, a 2×2 session by stimulus repeated measures ANOVA also revealed a significant main effect of stimulus [$F(1, 13)=5.79, p<0.03$], driven by greater responses in the CS– compared to CS+ trials, but no effect of session [$F(1, 13)=0.395, p>0.05$] or interaction [$F(1, 13)=0.001, p>0.05$].

3.2. Neuroimaging results

3.2.1. Region of interest analysis

Our primary analysis used the functionally defined ventral striatum ROIs generated by the contrast of reward and loss trials during the gambling session (Table 1; Fig. 3a). Mean parameter estimates were extracted from these independent ROIs for each subject using the model of the learning sessions for further analysis. In the right ventral striatum ROI ($x, y, z=17, 7, -6$; Fig. 3b), a 2×2 repeated measures ANOVA with session (approach and avoidance learning) by learning phase (pre and post-learning of uncertain stimuli) as factors revealed a main effect of session [$F(1, 18)=9.01, p<0.008$] and a main effect of learning phase [$F(1, 18)=11.82, p<0.003$], characterized by greater blood oxygen level dependent (BOLD) responses to pre-compared to post-learning trials [$t(18)=3.44, p<0.003$]. An interaction between session and learning phase [$F(1, 18)=4.426, p<0.05$] suggested that striatal signals during motivated learning were modulated by the context in which learning occurred (i.e., negative reinforcer).

In the left ventral striatal ROI ($x, y, z=-19, 4, -6$), a 2×2 session by learning phase repeated measures ANOVA revealed a main effect of session [$F(1, 18)=6.474, p<0.02$], a main effect of learning phase [$F(1, 18)=16.17, p<0.001$], and no significant interaction [$F(1, 18)=2.71, p=0.12$]. This ROI was large enough to contain two distinct peaks of activation (Fig. 3a), thus as an exploratory analysis, we performed the same ANOVA in both the medial ($x, y, z=-10, 4, -6$) and lateral ($x, y, z=-19, 4, -6$) peaks to examine potential interactions between session and learning phase. No interaction was observed in the more medial peak of the left ventral striatum [$F(1, 18)=1.94, p=0.18$]. In contrast, the more lateral peak of the ventral striatum ROI resembled the right striatal ROI both in terms of location and pattern of activity, showing an interaction between session and learning phase [$F(1, 18)=4.55, p<0.05$].

Table 1
Gambling session contrast (Reward > Loss).

Region of activation	Laterality	Talairach coordinates			Voxels (1 mm ³)	T-stat
		x	y	z		
Ventral striatum	Right	17	7	-6	591	7.27
Ventral striatum	Left	-19	4	-6	789	7.45
Medial peak	Left	-10	4	-6	228	7.55
Lateral peak	Left	-19	4	-6	302	7.52

An additional analysis was conducted in each ROI to investigate any effects of certain stimuli across sessions. In the right ventral striatum ROI, a 2×2 repeated measures ANOVA between session and type of certain stimulus (CS+, CS-) revealed a main effect of type of stimulus [$F(1, 18) = 8.15, p < 0.01$], driven by greater BOLD responses to CS- compared to CS+ trials [$t(18) = 2.09, p < 0.05$], but no effect of session or interaction. Within the left ventral striatum ROI, there were no effects with respect to the certain stimulus observed.

3.2.2. Whole-brain analysis

In order to explore other regions involved in approach and avoid learning in this specific paradigm, two whole-brain analyses were performed within the learning sessions. First, a 2×2 session by learning phase repeated measures ANOVA was conducted (Table 2; Fig. 4a). A main effect of learning phase was revealed in regions such as the striatum bilaterally, cingulate gyrus, and insula each showing greater responses during the pre, compared to post, learning phase (e.g., Fig. 4b and c for cingulate gyrus and right striatum respectively). Within this analysis, no voxels were identified showing a greater response for post- compared to pre-learning stimuli. Voxels showing a main effect of session were identified in a different region within the cingulate gyrus ($x, y, z = -22, -23, 36$). This region displayed a greater BOLD response during avoidance compared to the approach learning sessions [$t(18) = 3.99, p < 0.001$]. No voxels corresponding to an interaction of session and learning phase were identified.

Second, a 2×2 session by CS type repeated measures ANOVA was conducted to investigate the effect of the CS+ and CS- stimuli. A main effect of CS type revealed two dis-

tinct regions in the middle frontal gyrus (BA 6 and BA 10) and one region in the right amygdala (Table 3). Akin to the SCR analysis, *post hoc* paired samples *t*-tests showed the CS- eliciting higher BOLD responses than the CS+ in the amygdala [$t(18) = 5.38, p < 0.001$] and the middle frontal gyrus [BA 6; $t(18) = 6.34, p < 0.001$]. Conversely, the more anterior ROI in the middle frontal gyrus (BA 10) showed a higher BOLD response in CS+ compared to CS- trials [BA 10; $t(18) = 5.79, p < 0.001$]. A main effect of session revealed an ROI in the post-central gyrus, where paired sample two-tailed *t*-tests showed the approach learning session eliciting higher BOLD activity than the avoidance learning session [$t(18) = 7.37, p < 0.001$]. No voxels corresponding to an interaction of CS type and session were identified.

4. Discussion

The goal of this study was to use fMRI to investigate neural circuits involved in learning via positive and negative reinforcers. Specifically, this experiment probed how the human striatum, a structure typically implicated in reward-related processes, was modulated during learning when the motivational context is driven by the presence of a negative reinforcer. Participants acquired an adaptive behavioral response (i.e., a correct button press) via positive (approach learning) or negative (avoidance learning) reinforcers separately, in a within-subjects design that allowed direct comparisons when learning occurred under each motivational context. Participants showed greater subjective and physiological responses across learning (pre vs. post-learning), particularly when presented with trials that afforded the opportunity to either attain a monetary reward or avoid a monetary loss, compared to trials

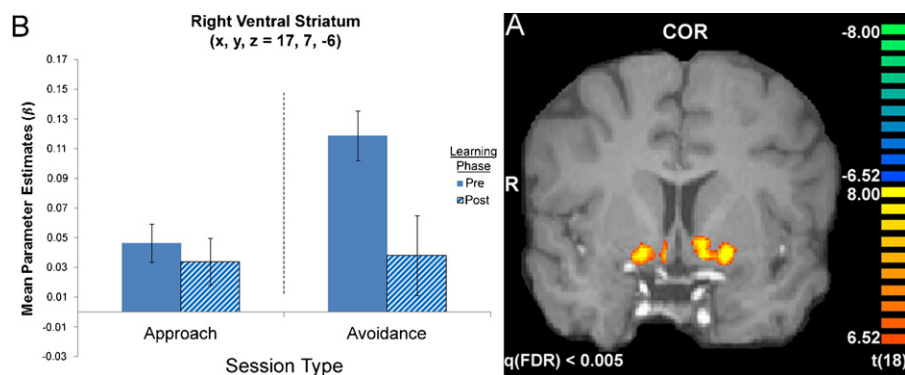
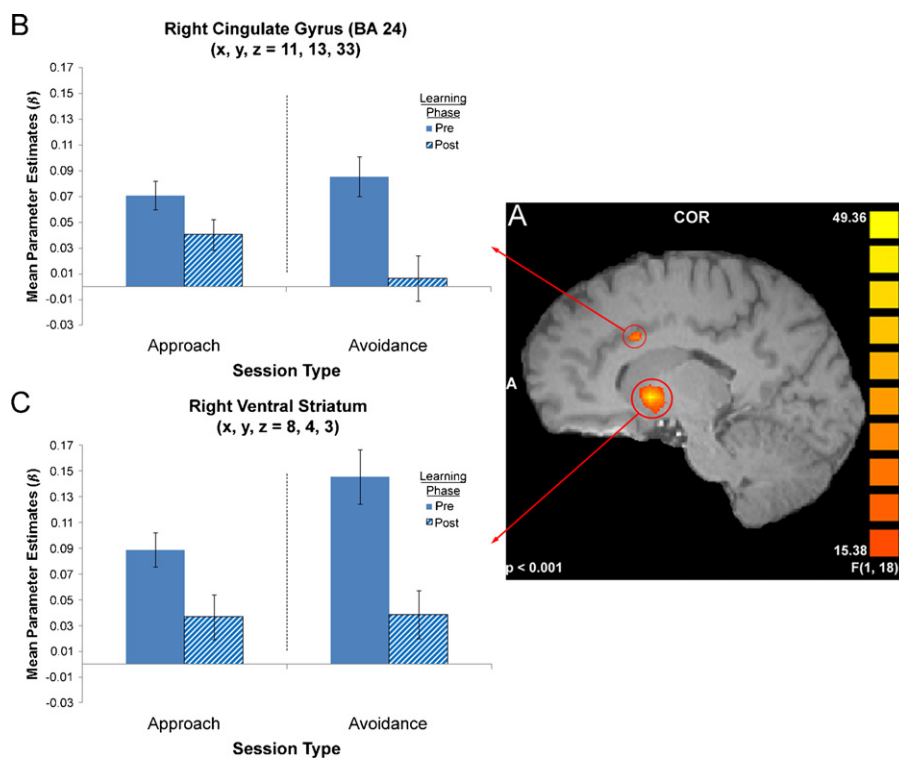


Fig. 3. (A) Ventral striatum ROIs defined by the linear contrast of Reward > Loss in the gambling sessions at an FDR correction < 0.005. (B) Mean parameter estimates in the right ventral striatum ROI ($x, y, z = 17, 7, -6$) showing an interaction of session (approach and avoidance) and learning phase (pre and post-learning; \pm s.e.m.).

Table 2
Session (approach and avoidance) \times learning phase (pre and post-learning) ANOVA.

Region of activation	Brodmann area (BA)	Laterality	Talairach coordinates			Voxels (1 mm ³)	F-stat
			x	y	z		
<i>Main effect of learning phase</i>							
Medial frontal gyrus	BA 6	Left	-7	7	51	290	18.31
Middle frontal gyrus	BA 6	Right	26	-2	49	251	16.99
Supramarginal gyrus	BA 40	Left	-40	-44	36	139	18.51
Cingulate	BA 24	Right	11	13	33	128	18.86
Middle frontal gyrus	BA 9	Right	32	25	30	183	19.74
Middle frontal gyrus	BA 10	Left	-34	49	21	223	17.35
Middle frontal gyrus	BA 10	Left	-40	49	12	248	18.4
Insula	BA 13	Left	-31	22	6	141	18.34
Insula		Right	29	25	3	126	18.41
Striatum		Right	8	4	3	989	24.92
Striatum		Left	-10	4	0	572	25.75
<i>Main effect of session</i>							
Cingulate	BA 24, 31	Left	-22	-23	36	206	19.31

**Fig. 4.** (A) Whole brain analysis exploring a main effect of learning phase from a 2 \times 2 session by learning phase ANOVA. Shown here are regions of the ventral caudate nucleus (x, y, z = 8, 4, 3) and cingulate gyrus (x, y, z = 11, 13, 33; BA 24) at a threshold of $p < 0.001$. Mean parameter estimates from the (B) cingulate gyrus and the (C) right ventral caudate nucleus are displayed showing greater responses during pre compared to post-learning trials (\pm s.e.m.).**Table 3**
Session (approach and avoidance) \times certain stimuli (CS+ and CS-) ANOVA.

Region of activation	Brodmann area (BA)	Laterality	Talairach coordinates			Voxels (1 mm ³)	F-stat
			x	y	z		
<i>Main effect of certain stimulus type</i>							
Middle frontal gyrus	BA 6	Left	-28	4	45	365	18.39
Middle frontal gyrus	BA 10	Right	26	64	9	167	22.41
Amygdala		Right	17	1	-15	108	18.19
<i>Main effect of session</i>							
Postcentral gyrus	BA 3	Left	-40	-29	51	216	19.3

where the positive or negative outcome was fully predictable. Increased motivated behavior was also observed during approach and avoidance learning trials overall, as indexed by faster responses than those recorded during trials with certain outcomes. Activity within an independently defined ROI in the ventral striatum revealed an interaction between type of session (approach and avoidance) and type of learning phase (pre and post), highlighted by greater responses during the acquisition of a behavior aimed at avoiding a negative outcome. These results suggest that despite overlapping neural circuitry when approaching or avoiding a conditioned stimulus, negative reinforcers can lead to greater influences on ventral striatum signals involved in mediating reinforcing effects on behavior.

The striatum is a multi-faceted structure with several anatomical connections that facilitate goal-directed behavior (for review see Haber and Knutson, 2010). Across species, the striatum has been found to be important for affective learning, particularly in the context of predicting potential rewards (for review see Delgado, 2007; Montague and Berns, 2002; O'Doherty, 2004; Rangel et al., 2008; Robbins and Everitt, 1996). For instance, signals corresponding to prediction errors, or the mismatch between expected and experienced rewards, are often correlated with BOLD signals in dorsal and ventral striatum (O'Doherty et al., 2003; O'Doherty, 2004; van den Bos et al., 2009) with greater correlations suggestive of increased behavioral performance during reward-learning tasks (Schonberg et al., 2007). Further, striatum signals are found to be important particularly during the acquisition of reward contingencies, showing a decrement as associations become fully predictable (Delgado et al., 2005; Haruno et al., 2004; Pasupathy and Miller, 2005). Our findings are consistent with this literature, as striatum BOLD responses show main effect of learning phase during approach learning sessions, with greater responses during the initial acquisition of a behavioral action to attain a reward.

More recently, neuroimaging experiments have also implicated the human striatum in aversive learning. For instance, aversive prediction errors have been found to correlate with striatum BOLD signals during classical conditioning paradigms (Delgado et al., 2008; Seymour et al., 2004, 2007), with striatum activity correlating with predictions of a potentially negative outcome regardless if an opportunity to avoid it existed or not (Jensen et al., 2003). Furthermore, studies using active avoidance of negative outcomes have found striatal activation during the initial acquisition of avoidance contingencies (Delgado et al., 2009) and expression of learned avoidance (Schlund and Cataldo, 2010; Schlund et al., 2010). Taken together, these studies support a role for the striatum in learning with negative reinforcers, which is also echoed in the current study.

Our study has two distinct features that helps advance the understanding of the role of the striatum in affective learning and processing of monetary incentives. First, it is one of the few studies where learning can take place in both a positive and a negative context using the same reinforcer (money), thus ensuring a within-subject comparison

of the contribution of the striatum across affective learning with both reinforcers. Second, it presents a new way of comparing positive with negative contexts using monetary reinforcers that attempt to control for issues typically associated with this type of comparison. With respect to the first feature, it was observed that BOLD signals within an independent functionally defined ventral striatum ROI showed an interaction between type of session and learning phase, which suggested that learning signals within the striatum were greater when learning via negative, compared to positive reinforcers. One plausible explanation for this finding is the idea that the saliency of a stimulus can drive activity in the striatum (Zink et al., 2004), which can be exaggerated in a negative context using primary reinforcers such as shock (Jensen et al., 2007). However, increases in striatum activity are not always modulated by the occurrence of salient events such as monetary loss (Delgado et al., 2000), a gamble signifying loss (Tom et al., 2007) or even shock itself (Seymour et al., 2004). In the current study, the certain stimuli are examples of potentially salient stimuli as they fully predict positive (approach CS+) or negative (avoidance CS+) outcomes. Previous studies have used CS+ stimuli to signal an outcome (e.g., Delgado et al., 2009; Jensen et al., 2003, 2007), and have seen robust neural responding to such stimuli, but many of these studies either had participants learn the nature of the CS (for a review see Phelps and LeDoux, 2005), or used primary reinforcers (Delgado et al., 2009; Jensen et al., 2003, 2007). In our experiment, little to no activity was observed in the striatum in response to these stimuli, potentially because they were fully predictable, which has shown to be less dependent on striatal responses (Berns et al., 2001; Delgado et al., 2005) and participants had no control over their outcome (Tricomi et al., 2004).

Another potential explanation for differences in striatum signals between avoidance and approach learning could be due to our choice of reinforcer (money). Specifically, when participants are presented with the avoidance learning sessions, they may be displaying behavioral tendencies akin to loss aversion, or a preference for avoiding losses rather than acquiring gains (Kahneman and Tversky, 1979). Within this idea, neural signals in the ventral striatum have been found to correlate with individual differences in loss aversion (Tom et al., 2007) and value computations related to changes with respect to a reference point (Breiter et al., 2001; De Martino et al., 2009).

In the current paradigm, participants also acquire an experimental bank via a gambling task before each approach and avoidance learning session. This bank is essential for participants to feel like they are actually losing something that has been earned and thus creates an endowment that may enhance the subjective value of accrued losses during the avoidance learning sessions (Delgado et al., 2006; Tom et al., 2007). In this experiment, the experimental banks are equated across approach and avoidance to allow for a direct comparison during learning sessions, but one could conjure up a scenario where gambling sessions are created to present a context in which avoidance sessions start with either more or less than what was earned in the approach sessions. This context-

tual manipulation with respect to endowment size is an interesting manipulation for future studies.

A second distinct feature of our paradigm is the use of secondary reinforcers, such as monetary incentives, as a common reinforcer that can be either positive (reward) or negative (loss), unlike primary reinforcers such as shock or food which are more difficult to equate. To adopt this type of incentive, we used a spinner procedure, described in detail in the methods, which kept the actual monetary value of a single trial ambiguous until the end of the experiment. The goal of this procedure was to ensure that the only thing that mattered for participants was the occurrence (or non-occurrence) of a reinforcer. Indeed, this was important, as the concept of marginal utility (value of gains decreases based on individual's asset) is known to influence reward-related circuitry, particularly the striatum (Tobler et al., 2007). While others have elegantly tried to take absolute value out of the equation and primarily examine questions related to the magnitude of the incentive (Galvan et al., 2005), our procedure allowed participants to treat positive and negative outcomes as just that, without any influence of actual value or magnitude. This procedure is promising for studies across development that use monetary incentives as a potential tool for isolating the affective meaning, rather than value of the presented incentives.

In this paradigm, the absolute value gained or lost is unknown, thus participants presumably calculate the value of their actions based on internal tendencies associated with positive and negative reinforcers. For instance, people are more likely to avoid social situations where they can be evaluated than approach them, despite the possibility of forming rewarding relationships (Beck and Clark, 2009), while striatum responses to losses, but not monetary rewards, correlate with increased behavioral choices in some contexts such as social competitions (Delgado et al., 2008). The current study was limited by simple choices (i.e., find appropriate response), thus investigating the influence of negative contexts on complex behavioral choices therefore becomes another interesting future investigation.

Within the striatum, we observed greater influences of negative reinforcers on more lateral regions of the ventral striatum. In contrast, more ventromedial striatum regions including ventral caudate nucleus showed a main effect of learning phase, irrespective of type of reinforcer. Further studies are necessary to fully understand this potential dissociation within the striatum, although given the vast connectivity in this structure (see Haber and Knutson, 2010 for review) it is not surprising that different regions within the striatum would express sensitivity to different task factors. Interestingly, no amygdala activation was observed during either approach or avoidance learning cues. Amygdala activity was apparent in the certain stimuli contrast, but not during the learning trials. The lack of amygdala activity is in contrast with animal studies implicating this structure in avoidance learning (see Cain and LeDoux, 2008), and human neuroimaging studies of avoidance learning using primary reinforcers (Delgado et al., 2009) or in contexts in which participants acquired stable avoidance responding prior

to scanning (Schlund and Cataldo, 2010; Schlund et al., 2010). Our design, on the other hand, used secondary reinforcers and had participants acquire the avoidance response during scanning, potentially creating a quick response coping mechanism which can be driven primarily by the striatum (for review see LeDoux and Gorman, 2001). Importantly, it is difficult to interpret a null result in neuroimaging, so the lack of amygdala activity during learning trials in this paradigm should be treated with caution.

Our paradigm and findings have implications for developmental studies of affective processing. First, as already discussed, the paradigm presents an opportunity to compare the influence of positive and negative reinforcers across development while attempting to control for valuation of monetary reinforcers (also see Galvan et al., 2005). Second, our results present an interesting complement to the influential triadic model of motivated behavior during adolescence (Ernst et al., 2006; Ernst and Fudge, 2009). Briefly, this model suggests that increased reward responses (ventral striatum), decreased avoidance responses (amygdala) and poor regulation (prefrontal cortex) contribute to aberrant behavior seen in adolescents. In the current experiment, young adults show a propensity to learn from both positive and negative reinforcers, engaging the striatum irrespective of motivational context, but not the amygdala. Interestingly, behaviorally inhibited adolescents show an augmented response to both positive and negative conditioned cues of increasing value in both the striatum and amygdala (Guyer et al., 2006). While our discussion of the amygdala is limited due to it being a null finding, our study does raise questions about a role, if any, of the striatum during negative motivational contexts across development.

In conclusion, this study extends the growing literature implicating the striatum in learning from both positive and negative reinforcers. Our results further suggest that specific regions in the lateral ventral striatum are modulated in particular by learning from negative reinforcers. The results provide a direct comparison between the influence of positive and negative reinforcers on acquisition of behaviors and the human striatum, setting up future studies that further probe similarities and differences across development which can translate to clinical studies focusing on acquisition and extinction of maladaptive behaviors (e.g., drug use) reinforced by positive or negative outcomes.

Acknowledgment

This study was funded by a National Institute on Drug Abuse grant to M.R.D. (DA027764).

References

- Alexander, G.E., DeLong, M.R., Strick, P.L., 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357–381.
- Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., Damasio, A.R., 1995. Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* 269, 1115–1118.
- Beck, L.A., Clark, M.S., 2009. Choosing to enter or avoid diagnostic social situations. *Psychol. Sci.* 20, 1175–1181.

- Berns, G.S., McClure, S.M., Pagnoni, G., Montague, P.R., 2001. Predictability modulates human brain response to reward. *J. Neurosci.* 21, 2793–2798.
- Bjork, J.M., Knutson, B., Fong, G.W., Caggiano, D.M., Bennett, S.M., Hommer, D.W., 2004. Incentive-elicited brain activation in adolescents: similarities and differences from young adults. *J. Neurosci.* 24, 1793–1802.
- 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.
- Cain, C., LeDoux, J.E., 2008. Emotional processing and motivation: in search of brain mechanisms. In: Elliot, A. (Ed.), *Handbook of Approach and Avoidance Motivation*. Taylor & Francis Group, LLC.
- Casey, B.J., Jones, R.M., Hare, T.A., 2008. The adolescent brain. *Ann. N.Y. Acad. Sci.* 1124, 111–126.
- Cromwell, H.C., Schultz, W., 2003. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol.* 89, 2823–2838.
- De Martino, B., Kumaran, D., Holt, B., Dolan, R.J., 2009. The neurobiology of reference-dependent value computation. *J. Neurosci.* 29, 3833–3842.
- Delgado, M.R., 2007. Reward-related responses in the human striatum. *Ann. N.Y. Acad. Sci.* 1104, 70–88.
- Delgado, M.R., Jou, R.L., LeDoux, J.E., Phelps, E.A., 2009. Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning. *Front. Behav. Neurosci.* 3, 33.
- Delgado, M.R., Labouliere, C.D., Phelps, E.A., 2006. Fear of losing money? Aversive conditioning with secondary reinforcers. *Soc. Cogn. Affect. Neurosci.* 1, 250–259.
- Delgado, M.R., Li, J., Schiller, D., Phelps, E.A., 2008. The role of the striatum in aversive learning and aversive prediction errors. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 363, 3787–3800.
- Delgado, M.R., Miller, M.M., Inati, S., Phelps, E.A., 2005. An fMRI study of reward-related probability learning. *Neuroimage* 24, 862–873.
- 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.
- Ernst, M., Fudge, J.L., 2009. A developmental neurobiological model of motivated behavior: anatomy, connectivity and ontogeny of the triadic nodes. *Neurosci. Biobehav. Rev.* 33, 367–382.
- Ernst, M., Nelson, E.E., Jazbec, S., McClure, E.B., Monk, C.S., Leibenluft, E., Blair, J., Pine, D.S., 2005. Amygdala and nucleus accumbens in responses to receipt and omission of gains in adults and adolescents. *Neuroimage* 25, 1279–1291.
- Ernst, M., Pine, D.S., Hardin, M., 2006. Triadic model of the neurobiology of motivated behavior in adolescence. *Psychol. Med.* 36, 299–312.
- 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.
- Galvan, A., Hare, T.A., Davidson, M., Spicer, J., Glover, G., Casey, B.J., 2005. The role of ventral frontostriatal circuitry in reward-based learning in humans. *J. Neurosci.* 25, 8650–8656.
- Goebel, R., Esposito, F., Formisano, E., 2006. Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: from single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum Brain Mapp.* 27, 392–401.
- 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.
- Guyer, A.E., Nelson, E.E., Perez-Edgar, K., Hardin, M.G., Roberson-Nay, R., Monk, C.S., Bjork, J.M., Henderson, H.A., Pine, D.S., Fox, N.A., Ernst, M., 2006. Striatal functional alteration in adolescents characterized by early childhood behavioral inhibition. *J. Neurosci.* 26, 6399–6405.
- Haber, S.N., Knutson, B., 2010. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35, 4–26.
- 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.
- Hassani, O.K., Cromwell, H.C., Schultz, W., 2001. Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.* 85, 2477–2489.
- Jensen, J., McIntosh, A.R., Crawley, A.P., Mikulis, D.J., Remington, G., Kapur, S., 2003. Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron* 40, 1251–1257.
- Jensen, J., Smith, A.J., Willeit, M., Crawley, A.P., Mikulis, D.J., Vitcu, I., Kapur, S., 2007. Separate brain regions code for salience vs. valence during reward prediction in humans. *Hum Brain Mapp.* 28, 294–302.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–292.
- Knutson, B., Adams, C.M., Fong, G.W., Hommer, D., 2001. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21, RC159.
- LaBar, K.S., Gatenby, J.C., Gore, J.C., LeDoux, J.E., Phelps, E.A., 1998. Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron* 20, 937–945.
- LaBar, K.S., LeDoux, J.E., Spencer, D.D., Phelps, E.A., 1995. Impaired fear conditioning following unilateral temporal lobectomy in humans. *J. Neurosci.* 15, 6846–6855.
- LeDoux, J.E., Gorman, J.M., 2001. A call to action: overcoming anxiety through active coping. *Am. J. Psychiatry* 158, 1953–1955.
- May, J.C., Delgado, M.R., Dahl, R.E., Stenger, V.A., Ryan, N.D., Fiez, J.A., Carter, C.S., 2004. Event-related functional magnetic resonance imaging of reward-related brain circuitry in children and adolescents. *Biol. Psychiatry* 55, 359–366.
- Middleton, F.A., Strick, P.L., 2000. Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. *Brain Cogn.* 42, 183–200.
- Montague, P.R., Berns, G.S., 2002. Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284.
- O'Doherty, J.P., 2004. Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776.
- 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.
- Ochsner, K.N., Gross, J.J., 2005. The cognitive control of emotion. *Trends Cogn. Sci.* 9, 242–249.
- Pasupathy, A., Miller, E.K., 2005. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433, 873–876.
- Phelps, E.A., LeDoux, J.E., 2005. Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48, 175–187.
- Rangel, A., Camerer, C., Montague, P.R., 2008. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556.
- Reyna, V.F., Estrada, S.M., Demarinis, J.A., Myers, R.M., Stanisz, J.M., Mills, B.A., 2011. Neurobiological and memory models of risky decision making in adolescents versus young adults. *J. Exp. Psychol. Learn. Mem. Cogn.*
- Robbins, T.W., Everitt, B.J., 1996. Neurobehavioural mechanisms of reward and motivation. *Curr. Opin. Neurobiol.* 6, 228–236.
- Schlund, M.W., Cataldo, M.F., 2010. Amygdala involvement in human avoidance, escape and approach behavior. *Neuroimage* 53, 769–776.
- Schlund, M.W., Siegle, G.J., Ladouceur, C.D., Silk, J.S., Cataldo, M.F., Forbes, E.E., Dahl, R.E., Ryan, N.D., 2010. Nothing to fear? Neural systems supporting avoidance behavior in healthy youths. *Neuroimage* 52, 710–719.
- Schonberg, T., Daw, N.D., Joel, D., O'Doherty, J.P., 2007. Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *J. Neurosci.* 27, 12860–12867.
- Seymour, B., Daw, N., Dayan, P., Singer, T., Dolan, R., 2007. Differential encoding of losses and gains in the human striatum. *J. Neurosci.* 27, 4826–4831.
- Seymour, B., O'Doherty, J.P., Dayan, P., Koltzenburg, M., Jones, A.K., Dolan, R.J., Friston, K.J., Frackowiak, R.S., 2004. Temporal difference models describe higher-order learning in humans. *Nature* 429, 664–667.
- Skinner, B.F., 1938. *The Behavior of Organisms: An Experimental Analysis*. Copley Publishing Group, Cambridge.
- Talairach, J., Tournoux, P., 1988. *Co-Planar Stereotaxic Atlas of the Human Brain*. Thieme Medical Publishers, Inc., New York.
- Tobler, P.N., Fletcher, P.C., Bullmore, E.T., Schultz, W., 2007. Learning-related human brain activations reflecting individual finances. *Neuron* 54, 167–175.
- Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518.
- Tricomi, E.M., Delgado, M.R., Fiez, J.A., 2004. Modulation of caudate activity by action contingency. *Neuron* 41, 281–292.
- van den Bos, W., Guroglu, B., van den Bulk, B.G., Rombouts, S.A., Crone, E.A., 2009. Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Front. Hum. Neurosci.* 3, 52.

- Van Leijenhorst, L., Gunther Moor, B., Op de Macks, Z.A., Rombouts, S.A., Westenberg, P.M., Crone, E.A., 2010. Adolescent risky decision-making: neurocognitive development of reward and control regions. *Neuroimage* 51, 345–355.
- Yacubian, J., Glascher, J., Schroeder, K., Sommer, T., Braus, D.F., Buchel, C., 2006. Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J. Neurosci.* 26, 9530–9537.
- 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.