Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning

Mauricio R. Delgado1*, Rita L. Jou2, Joseph E. LeDoux3 and Elizabeth A. Phelps2,3

1 Department of Psychology, Rutgers University, Newark, NJ, USA
2 Department of Psychology, New York University, New York, NY, USA
3 Center for Neural Science, New York University, New York, NY, USA

Edited by:
Daeyeol Lee, Yale University School of Medicine, USA

Reviewed by:
Jeansok J. Kim, University of Washington, USA
Hackjin Kim, Korea University, Korea

*Correspondence:
Mauricio R. Delgado, Department of Psychology, Rutgers University, 340 Smith Hall, 101 Warren Street, Newark, NJ 07102, USA.
E-mail: delgado@psychology.rutgers.edu

INTRODUCTION

The ability to modify and control our emotional responses is critical for adaptive function and goal-directed behavior. Although learning to fear a potentially dangerous situation is important, it is equally important to be able to modify this fear when new information is available, or use this fear to motivate adaptive action that diminishes the potential threat. Recent research examining the neural systems of regulating fear in humans has highlighted passive extinction techniques (Milad and Quirk, 2002; Knight et al., 2004; Phelps et al., 2004) and the use of cognitive strategies (Kalisch et al., 2005; Ochsner and Gross, 2005; Delgado et al., 2008b). These techniques focus on modifying the fear response in the presence of the fear-eliciting event. Another common response used to regulate fear, however, is to take an action to avoid the potential danger and diminish the fear response. Given how frequently action is used to cope with potential threat and fear outside the laboratory, surprisingly little research conducted in humans has examined the neural system mediating the active coping of fear. Research in non-human animals has suggested that active coping of fear may involve amygdala and striatal interactions (Killcross et al., 1997; Everitt et al., 1999; LeDoux and Gorman, 2001; Cardinal et al., 2002). The goal of the present study is to investigate if an amygdala–striatal circuitry underlies active coping of fear in humans.

Previous research across species has shown that the amygdala is critical for learning about aversive outcomes, while the striatum is involved in reward-related processing. Less is known, however, about the role of the amygdala and the striatum in learning how to exert control over emotions and avoid negative outcomes. One potential mechanism for active avoidance of stressful situations is postulated to involve amygdala–striatal interactions. The goal of this study was to investigate the physiological and neural correlates underlying avoidance learning in humans. Specifically, we used a classical conditioning paradigm where three different conditioned stimuli (CS) were presented. One stimulus predicted the delivery of a shock upon stimulus offset (CS+), while another predicted no negative consequences (CS–). A third conditioned cue also predicted delivery of a shock, but participants were instructed that upon seeing this stimulus, they could avoid the shock if they chose the correct action (AV+). After successful learning, participants could then easily terminate the shock during subsequent stimulus presentations (AV–). Physiological responses (as measured by skin conductance responses) confirmed a main effect of conditioning, particularly showing higher arousal responses during pre (AV+) compared to post (AV–) learning of an avoidance response. Consistent with animal models, amygdala–striatal interactions were observed to underlie the acquisition of an avoidance response. These results support a mechanism of active coping with conditioned fear that allows for the control over emotional responses such as fears that can become maladaptive and influence our decision-making.

Keywords: amygdala, striatum, negative reinforcement, instrumental conditioning, fMRI, reward, punishment
partially independent neural circuits mediate active and passive
means of fear expression and that the amygdala’s connectivity
with the striatum allows for active coping strategies to develop
demand fear induced by a conditioned stimulus (Everitt
et al., 1991; Amorapanth et al., 2000; Cain and LeDoux, 2008).
In support of this hypothesis, both dorsal and ventral striatum in
rats have been implicated in various types of avoidance learning
(Winocur and Mills, 1969; Allen and Davison, 1973; Neill et al.,
1974; McCullough et al., 1993; Li et al., 2004).

In humans, neuroimaging experiments suggest that the
striatum is involved in the expectation of an aversive stimulus,
whether an opportunity to avoid exists or not (Jensen et al., 2003;
Delgado et al., 2008a). However, less is known about the poten-
tial striatal–amygdala interactions that may underlie avoidance
learning in the human brain. In this experiment, we used a modi-
ied aversive conditioning paradigm in conjunction with blood
oxygenated level dependent (BOLD) and autonomic measures
to explore the acquisition of an avoidance learning response.
Participants were instructed that they could: (a) avoid a potential
shock by learning a behavioral response (i.e., a button press), and
(b) express the behavioral response after successful learning to
prevent future shock delivery. As suggested by animal models
(e.g., LeDoux and Gorman, 2001), we hypothesized that interac-
tions of the amygdala and striatum would underlie a measure
of successful avoidance learning, comparing BOLD responses
pre- and post-learning.

MATERIALS AND METHODS

PARTICIPANTS

Thirty-two participants were initially recruited for this study
(19F/13M, M = 19.8, SD = 2.2). Nine participants were excluded
from further analysis due to excessive motion during scanning
(N = 4, more than 2 mm of movement), failure to learn the task
(N = 3) or equipment malfunction during session (N = 2, shocks
not delivered). The final behavioral and neuroimaging analyses were
conducted on 23 participants (15F/8M, mean age = 19.9, SD = 2.6).
Participants responded to posted advertisement and all partici-
pants gave informed consent. The experiment was approved by the
University Committee on Activities Involving Human Subjects at
New York University.

PROCEDURE

The experiment consisted of an aversive conditioning paradigm
with instruction. Participants were presented with three types of
colored squares (e.g., blue, yellow, green) that served as conditioned
stimuli (CS). Two of the CSs were fully predictable and led to
the delivery of a mild shock to the wrist (the unconditioned stimulus,
US) with either 100% (CS+ trials) or 0% (CS− trials) probabil-
ity (certain trials). The third CS also predicted delivery of a mild
shock, however participants were afforded a chance to avoid the
shock if they learned the appropriate response (avoidable or AV
trials). The AV trials were further subdivided into two types of tri-
als according to learning success. During pre-learning trials (AV+
trials) participants attempted to learn how to avoid the negative
outcome. Post-learning trials included subsequent presentations
of the CS where successful avoidance of the US was maintained
by the previously learned response (AV− trials). Thus, the four
conditions (CS+, CS−, AV+, AV−) comprised two 2 classes of CS
(certain and avoidable) that varied with respect to conditioned
response (aversive and safe; Figure 1).

Each CS presentation lasted 10 s and was broken down into a
CS and a response phase. During the CS phase (4–6 s), partici-
pants were presented with the type of CS and instructed to just
observe and wait for the response phase. The CS phase was the
task period of interest and measures of physiological and BOLD
responses reflect activation at this time point, unaltered by
any motor responses or shock delivery. The response phase was
begun by a question-mark in the middle of the colored square (4–6 s).
At this time, participants were instructed to make a behavioral
response (i.e., a button press). A mild shock was delivered during
CS+ and AV+ trials for 200 ms that co-terminated with the end of
the response phase. The trial ended with a 14-s inter-trial inter-
val, for a total trial time of 24 s. Each session contained 24 trials,
evenly divided into 6 trials for each type of condition (CS+, CS−,
AV+, AV−).

For AV trials, participants had a chance to avoid the shock with
the correct response during the response phase. Specifically, they
were told that one of eight button presses could terminate the
shock delivery. Participants were given an MRI compatible but-
ton box with four buttons and used their right hand to make one
response per trial. They were further instructed that the “correct
button” could be the first or second time a button was pressed, thus
creating eight possible correct buttons and diminishing excessive
motor coordination issues associated with the use of multiple but-
ton boxes. Participants were also asked to make a non-contingent
button press during the response phase for certain trials (CS+, CS−)
to control for motor requirements.

Prior to scanning, participants were instructed what each CS pre-
dicted (certain or avoidable outcome). Unbeknownst to participants,
however, the correct button press during AV trials was always the

FIGURE 1 | Human avoidance paradigm. Participants were presented with
three types of CS. Both CS+ and CS− predicted a certain outcome (an aversive
shock or no shock respectively). The AV+ condition predicted a potential shock
but afforded the participant an opportunity to avoid a shock with the correct
behavioral response. An AV− trial referred to trials post-learning of an
avoidance response. Colors were counterbalanced across scanning sessions.
response made in the sixth AV+ trial, irrespective of which button was pressed. The correct button press was then repeated post-learning, during the remaining six AV− trials. This ensured that all participants experienced the same schedule of reinforcement, with each session containing 24 trials, evenly divided into six trials for each type of condition (CS+, CS−, AV+, AV−). Participants who failed to learn the contingency (i.e., failed to repeat the correct button, and thus never experiencing AV− trials), typically reported not paying attention, and were excluded from all further analysis (N = 3).

The US constituted mild shocks delivered to the right wrist through a stimulating bar electrode connected to a Grass Medical Instruments stimulator. The stimulator was shielded for magnetic interference and grounded through an RF filter. Participants used a work-up procedure to set the appropriate shock level prior to the experimental session. Specifically, participants experienced a mild shock (10 V, 200 ms, 50 pulses/s) which was gradually increased up to a fixed maximum (60 V). They were instructed to set a level that was deemed uncomfortable, but not painful (mean shock level = 25.69 V, SD = 8.91).

Task events were programmed using E-PRIME software, v1.0 (PST, Pittsburgh, PA, USA). The color of the CSs was counterbalanced across sessions. Stimuli were presented in a black background and projected onto a screen which was visible inside the scanner through a mirror in the head coil. Right handed responses were made using an MRI compatible button box. At the end of the experimental session, participants were debriefed and compensated.

**PHYSIOLOGICAL SET-UP, ASSESSMENT AND BEHAVIORAL ANALYSIS**

Skin conductance responses (SCRs) were acquired from the participant’s middle phalanges of the second and third fingers in the left hand and amplified by BIOPAC Systems skin conductance module. Shielded Ag–AgCl electrodes were grounded through an RF filter panel and served to acquire data. AcqKnowledge software was used to analyze the analog skin conductance waveforms. The level of SCR response was assessed for each trial as the base to peak amplitude difference in skin conductance of the largest deflection in the 0.5–4.5 s latency window after onset of the CS (see LaBar et al., 1995). A minimum response criterion of 0.02 µS was used with lower responses scored as 0. Raw scores were square-root transformed prior to statistical analysis to normalize the distributions (LaBar et al., 1998). Acquired SCRs during the CS phase were then averaged per participant and per type of trial (CS+, CS−, AV+, AV−). A 2 × 2 repeated measures ANOVA with participants as a random factor was used to test for a main effect of type of CS (certain, avoidable) and conditioned response (aversive, safe). Post hoc paired t-tests were then conducted to probe differences between the contrast of interest, AV+ and AV− trials.

Additional behavioral data was acquired in the form of reaction time during the response phase, using an MRI compatible button box with four buttons. The primary analysis of the reaction time data was a paired t-test comparison of certain (CS+, CS−) and AV trials (AV+, AV−), hypothesized to differ with respect to motivation. Since the schedule of reinforcement was predetermined to reflect learning after six trials, accuracy differences were not expected, and participants who did not learn were excluded as previously described.

**FMRI ACQUISITION AND ANALYSIS**

A 3T Siemens Allegra head-only scanner and a Siemens standard head coil were used for data acquisition at NYU’s Center for Brain Imaging. Anatomical images were acquired using a T1-weighted protocol (256 × 256 matrix, 176 1-mm sagittal slices) Functional images were acquired using a single-shot gradient echo EPI sequence (TR = 2000 ms, TE = 20 ms, FOV = 192 cm, flip angle = 75°, bandwidth = 4340 Hz/px, echo spacing = 0.29 ms). Thirty-five contiguous oblique-axial slices (3 mm × 3 mm × 3 mm voxels) parallel to the AC–PC line were obtained. Analysis of imaging data was conducted using Brain Voyager software (Brain Innovation, Maastricht, The Netherlands). The data was initially corrected for motion (using a threshold of 2 mm or less), and slice scan time using sinc interpolation was applied. 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 (three 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 analysis was performed on the functional data using a general linear model (GLM) on 23 participants. There were four regressors of interest in the CS phase (CS+, CS−, AV+, AV−). There were also six regressors of no interest that modeled the response phase (separated into four types of trials according to condition) and the shock delivery (CS+, US and AV+, US). The principal contrast served to identify regions of interest (ROIs) involved in processing anticipated aversive outcomes during the CS phase, using a conservative threshold of FDR < 0.01 along with a cluster threshold of 10 contiguous voxels. Specifically, trials where an aversive outcome was expected (AV+, CS+) were compared to the most non-aversive control condition (CS−), as some residue conditioned fear could exist in AV− trials. Given the a priori hypothesis with respect to amygdala–striatal interactions, an amygdala ROI was functionally defined with this contrast using a more lenient threshold of p < 0.025 (uncorrected) along with a cluster threshold of four contiguous voxels (Buchel et al., 1998; LaBar et al., 1998). Mean parameter estimates reflecting effect size of a particular condition were extracted from ROIs in the striatum and amygdala for further analysis. A correlation analysis was then conducted comparing learning changes differences (i.e., mean parameter estimates for AV+ minus mean parameter estimates for AV−) between the functionally defined amygdala and striatum ROIs. Additionally, the time course of activation across the entire functional run for each individual participant was extracted from the amygdala ROI and used in an exploratory connectivity analysis. The time-course data was z-transformed and used as a single predictor in a GLM. The resulting activation map was thresholded at FDR < 0.001 and identified regions which hemodynamic patterns correlated with the seed amygdala ROI. Finally, an exploratory analysis was performed comparing AV+ and AV− trials, investigating brain regions associated with avoidance learning changes, and identified ROIs at p < 0.005 with four or more contiguous voxels.

**RESULTS**

**BEHAVIORAL AND PHYSIOLOGICAL RESULTS**

SCRs were acquired during the CS phase as a measure of physiological arousal. A main effect of conditioned response was observed [F(1, 21) = 42.34, p < 0.0001; Figure 2] suggesting that participants.

[Figure 2]
were more aroused during presentation of CS that predicted aversive (CS+, AV+) compared to safer outcomes (CS−, AV−) consistent with earlier studies of human aversive conditioning (LaBar et al., 1998; Phelps et al., 2004). Post hoc paired t-tests revealed a differential response between CS+ and CS− trials \( t(22) = 4.71, p < 0.0002 \) as expected based on previous findings. Importantly, a difference was also observed between AV+ and AV− \( t(22) = 4.73, p < 0.0002 \), suggesting that learning an avoidable response in this paradigm decreases previously documented conditioned fear responses. No main effect of type of CS \( F(1, 21) = 0.3, p = 0.59 \) or interaction \( F(2, 21) = 1.3, p = 0.27 \) was observed using SCR data. Instead, differences across type of CS were seen in the reaction time data in the response phase. Specifically, participants were faster to make a response to avoidable \( (M = 593.55, SD = 161) \) compared to certain \( (M = 671.93, SD = 146.47) \) trials \( t(21) = 12.24, p < 0.0002 \), suggesting behavioral reactions were faster when motivation to respond was high (Delgado et al., 2004). These effects were particularly strong when comparing safe trials \( [\text{AV−} \text{ and CS−}; t(22) = 4.24, p < 0.0005] \), but also approaching significance when contrasting the aversive trials \( [\text{AV+} \text{ and CS+}; t(22) = 1.5, p = 0.14] \). No effects of conditioning \( [F(1, 21) = 1.06, p = 0.31] \) or interaction \( [F(1, 21) = 0.31, p = 0.58] \) were observed in the reaction time data.

**NEUROIMAGING RESULTS**

The main analysis used to identify ROIs involved a contrast of aversive trials, where an aversive outcome was expected (AV+, CS+), and the most non-aversive control condition (CS−). This contrast yielded positive activation patterns in an array of cortical regions (Table 1), but central to this investigation, we observed activation in the ventral and dorsal striatum bilaterally. Mean parameter estimates for individual participants were then extracted for the striatum ROIs for further analysis. Within the left ventral striatum ROI identified in this contrast (Figure 3A), which included the putamen, a differential response between AV+ and AV− trials was found using post hoc t-tests \( t(22) = 3.25, p < 0.005 \). This pattern also characterized BOLD responses in the right dorsal striatum ROI (Figure 3B; \( t(22) = 2.37, p < 0.05 \)). Interestingly, a differential response between AV+ and CS+ trials was seen in the left ventral striatum ROI \( t(22) = 2.22, p < 0.05 \), but not within the voxels defined in the right dorsal striatum ROI \( t(22) = 0.94, p = 0.36 \).

An amygdala ROI was functionally defined with the same contrast of aversive and safe trials, but using a more lenient threshold given the a priori hypothesis with respect to amygdala–striatal interactions.

---

**Table 1 | Contrast of aversive (CS+ and AV+) and Safe (CS−) trials at FDR <0.01 and contiguity threshold of 10 voxels.**

<table>
<thead>
<tr>
<th>Region of activation</th>
<th>Brodmann areas</th>
<th>Laterality</th>
<th>Talairach coordinates</th>
<th>#Voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medial frontal gyrus</td>
<td>6</td>
<td>Left</td>
<td>0</td>
<td>−1</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>6</td>
<td>Right</td>
<td>28</td>
<td>−5</td>
</tr>
<tr>
<td>Somatosensory cortex</td>
<td>1, 2, 3</td>
<td>Left</td>
<td>−39</td>
<td>−34</td>
</tr>
<tr>
<td>Superior parietal lobe</td>
<td>7</td>
<td>Left</td>
<td>−33</td>
<td>−51</td>
</tr>
<tr>
<td>Precentral gyrus</td>
<td>4</td>
<td>Left</td>
<td>−35</td>
<td>−23</td>
</tr>
<tr>
<td>Precuneus</td>
<td>7</td>
<td>Left</td>
<td>−13</td>
<td>−70</td>
</tr>
<tr>
<td>Medial frontal gyrus</td>
<td>6, 8</td>
<td>Right</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>Dorsolateral PFC</td>
<td>9</td>
<td>Left</td>
<td>−31</td>
<td>52</td>
</tr>
<tr>
<td>Inferior parietal lobe</td>
<td>40</td>
<td>Right</td>
<td>54</td>
<td>−47</td>
</tr>
<tr>
<td>Medial occipital gyrus</td>
<td>19</td>
<td>Left</td>
<td>−23</td>
<td>−86</td>
</tr>
<tr>
<td>Dorsal striatum</td>
<td></td>
<td>Right</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Dorsal striatum</td>
<td></td>
<td>Left</td>
<td>−13</td>
<td>−3</td>
</tr>
<tr>
<td>Ventral striatum</td>
<td></td>
<td>Left</td>
<td>−18</td>
<td>12</td>
</tr>
<tr>
<td>Ventral striatum</td>
<td></td>
<td>Right</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>45</td>
<td>Left</td>
<td>−30</td>
<td>22</td>
</tr>
<tr>
<td>Lingual gyrus</td>
<td>18</td>
<td>Right</td>
<td>12</td>
<td>−62</td>
</tr>
<tr>
<td>Occipital lobe</td>
<td>12, 18</td>
<td></td>
<td>1</td>
<td>−80</td>
</tr>
<tr>
<td>Cerebellum</td>
<td></td>
<td>Left</td>
<td>−33</td>
<td>−52</td>
</tr>
<tr>
<td>Cerebellum</td>
<td></td>
<td>Right</td>
<td>31</td>
<td>−43</td>
</tr>
</tbody>
</table>

---
Activity within the amygdala ROI in the left hemisphere showed differential responses between CS+ and CS− trials \([t(22) = 2.28, p < 0.05; \text{Figure 4}]\), consistent with previous literature and contrast used to define this ROI, while differential response between AV+ and AV− trials approached significance \([t(22) = 1.53, p = 0.14]\).

A measure of the magnitude of avoidance learning was calculated from mean parameter estimates (i.e., mean beta weights) with the goal of contrasting the a priori ROIs (i.e., amygdala and striatum) during task performance. Specifically, we used the difference between mean parameter estimates during AV+ and AV− trials, reflecting the difference between pre- and post-learning of an avoidance response. This measure of the magnitude of avoidance learning for the amygdala ROI was then correlated with the same measure for both left ventral striatum \((r = 0.51, p < 0.05)\) and right dorsal striatum \((r = 0.54, p < 0.05)\) ROIs previously described. To better quantify a potential interaction between the amygdala and striatum during avoidance learning, however, an exploratory connectivity analysis was performed where the time course of amygdala activation for each individual subject was used as a single predictor in a GLM. As hypothesized by animal models (LeDoux and Gorman, 2001), it was expected that the seed ROI, the amygdala activation pattern, would correlate with striatum activity during task performance. With the caveat that this analysis included the entire task, and not selected types of trials (e.g., avoidance trials), activation in the striatum bilaterally was observed to correlate with the seed amygdala ROI. Specifically, ROIs in the left \((x, y, z = -7, 15, 5; 1615 \text{ voxels})\) and right \((x, y, z = 7, 9, 3; 903 \text{ voxels})\) ventral caudate nucleus were observed in this analysis (Figure 5A), with some degree of overlap with the striatum ROI previously defined by the general analysis (Figure 5B).

Finally, an exploratory analysis was performed comparing AV+ and AV− trials, which investigated brain regions distinctly associated with learning changes during avoidance trials. Corticostriatal circuits typically involved in reinforcement learning (for review see O’Doherty, 2004; Daw and Doya, 2006; Balleine et al., 2007) were identified in this contrast (Table 2). These included ROIs in the dorsal and ventral striatum, along with dorsal (BA 6) and ventromedial (BA 10/32) prefrontal regions. Interestingly, both striatum ROIs showed a pattern of response resembling learning signals, as only the AV+ trials, where learning could occur, elicited a strong BOLD signal as represented by higher mean parameter estimate values.

**DISCUSSION**

The goal of this study was to explore the neural circuitry underlying active coping of fear in humans using a variant of an aversive conditioning paradigm where conditioned fear could be
Participants first acquired a behavioral response to terminate delivery of a mild shock, and then continued to use this response to refrain from further aversive outcomes. Physiological arousal, as index by SCRs, supported the behavioral evidence of learning as arousal levels were decreased post-learning of an avoidance response. Additionally, instrumental behavior was faster during avoidance trials, compared to trials where a certain outcome was expected (i.e., non-contingent response), potentially indicating higher motivational levels when an opportunity to exert control over an emotional event is present. A contrast of aversive and safe trials identified *a priori* ROIs in both dorsal and ventral striatum along with amygdala, with BOLD signals within the striatum differing between pre- and post-learning of an avoidance response, a measure that correlated with BOLD signals in the amygdala. This was supported by a connectivity analysis using the amygdala as a seed ROI which found correlations with the striatum. As postulated by non-human animal models (Killcross et al., 1997; Everitt et al., 1999; LeDoux and Gorman, 2001; Cardinal et al., 2002), active coping of fear in humans may involve amygdala and striatal interactions as a means of diminishing conditioned fear and exerting control over emotional responses.

The involvement of the striatum in active avoidance has been previously observed in animal studies (Winocur and Mills, 1969; Allen and Davison, 1973; Neill et al., 1974; McCullough et al., 1993; Li et al., 2004). In the context of this human paradigm, the striatum was particularly involved in learning a behavioral action that allowed for control over conditioned fear, highlighting the role
of the striatum in action-contingency during learning (O’Doherty et al., 2004; Tricomi et al., 2004), while extending it to aversive states. Despite its functional heterogeneity and connectivity to regions such as the amygdala, the human striatum is typically discussed in the context of reward processing (for review see Rangel et al., 2008), although evidence continues to suggest the striatum’s involvement during affective learning irrespective of type of reinforcer (positive or negative). This paradigm provides additional support for the involvement of the human striatum in processing negative reinforcement.

The amygdala is a structure linked to aversive processes, particularly the acquisition of fears (for review see Phelps and LeDoux, 2005). Studies in animals also link the amygdala with certain forms of avoidance learning (e.g., Killcross et al., 1997; Machado et al., 2009), or escaping from fear (Amorapanth et al., 2000), leading to the hypothesis that amygdala–striatum interaction could underlie one’s ability to actively cope with conditioned fear (LeDoux and Gorman, 2001). Given this a priori hypothesis, we used a lenient threshold previously used by other human fear conditioning studies (e.g., LaBar et al., 1998) to investigate the role of the amygdala in human avoidance learning. Although the results have to be carefully considered given the lenient threshold, we observed a correlation between the time course of amygdala activation during task performance and the striatum, supportive of a potential interaction between the two structures during avoidance learning.

Previous research investigating the regulation of fear in humans has examined passive extinction techniques (Milad and Quirk, 2002; Knight et al., 2004; Phelps et al., 2004) and the use of cognitive strategies (for review see Ochsner and Gross, 2005). In the current paper, we examine the role of active coping strategies, particularly taking an action to avoid a potential threat, to adaptively control fear. A common finding across the three types of techniques is that conditioned fear is diminished, as evidenced by a physiological correlate of fear (SCRs) and decreases in BOLD response in the amygdala. Interestingly, the left amygdala ROI identified in a previous cognitive regulation study of conditioned fear (Delgado et al., 2008b; x, y, z = −20, 0, −20) was quite similar to the left amygdala ROI identified in the current study using an avoidance paradigm (x, y, z = −20, 2, −19). One potential difference across the techniques, however, is the role of the striatum in the control of fears. Striatum activation has been reported in previous papers examining the control of fear through passive extinction and cognitive strategies techniques (Phelps et al., 2004; Delgado et al., 2008b), although the particular contrast was a general effect of conditioning. The current findings suggest that the motivation to avoid a negative outcome with an instrumental response engages the striatum even more than just simple conditioning as evidenced by increased activation in the left striatum during AV+ trials, when subjects were learning the avoidance response, compared to CS+ trials when they were simply anticipating an aversive outcome, further highlighting the involvement of the striatum in learning via negative reinforcement.

The paradigm used for this experiment was adapted from previous animal (Amorapanth et al., 2000) and human (Phelps et al., 2004; Delgado et al., 2008b) studies of aversive conditioning. It is a simple task that has distinct advantages for studying a complex process such as avoidance learning. The inclusion of separate CS and response phase, for instance, allows probing of neural responses to the initial representation of the CS without the potential motor and motor preparation confounds associated with this instrumental procedure. This paradigm can also be adapted to study avoidance learning with secondary reinforcers (e.g., money; see Kim et al., 2006), comparisons between reinforcers of different valence (positive and negative reinforcement), and varying levels of probability or complexity of avoidance response (e.g., manipulation of effort required to successfully avoid negative outcome). This avoidance paradigm also has its limitations, however, such as the minimal amount of trials experienced by a participant per condition (6), which required a fixed reinforcement schedule. The task length was designed to limit the amount of shocks administered and, due to piloting, provide an ideal time window where participants felt that they could indeed be successful. It is also possible that some type of habituation can occur in this design as AV+ and AV− trials are separated in time within a scanning session. An argument against habituation being an explanation for the observed results, however, is that individuals who failed to learn the task did not show the differential responses in the striatum when comparing AV+ and AV− trials, which was characteristic of successful task learning (see Supplementary Materials).

In summary, we used a variant of a fear conditioning study where participants had a chance to avoid a negative outcome with an instrumental behavior. Consistent with animal models (e.g., LeDoux and Gorman, 2001), we found amygdala–striatal interactions in humans potentially underlying avoidance learning and exerting control over conditioned fears. Future studies will probe how human mechanisms of affective learning through negative

---

### Table 2 | Contrast of AV+ and AV− trials at p < 0.005 and contiguity threshold of four voxels.

<table>
<thead>
<tr>
<th>Region of activation</th>
<th>Brodmann areas</th>
<th>Laterality</th>
<th>Talairach coordinates</th>
<th># Voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dorsomedial frontal gyrus</td>
<td>6</td>
<td>Right</td>
<td>x = 33, y = 1, z = 55</td>
<td>377</td>
</tr>
<tr>
<td>Dorsal striatum</td>
<td>Right</td>
<td>16, 12, 19</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>Ventral striatum</td>
<td>Left</td>
<td>−20, 11, 1</td>
<td>126</td>
<td></td>
</tr>
<tr>
<td>Frontal medial gyrus</td>
<td>Right</td>
<td>39, 46, 1</td>
<td>199</td>
<td></td>
</tr>
<tr>
<td>Ventromedial PFC</td>
<td>32/10</td>
<td>Right</td>
<td>x = 25, y = 45, z = −1</td>
<td>248</td>
</tr>
<tr>
<td>Ventromedial PFC</td>
<td>32/10</td>
<td>Right</td>
<td>x = 11, y = 37, z = −6</td>
<td>578</td>
</tr>
</tbody>
</table>
reinforcers (i.e., avoidance learning) compare to learning through positive reinforcers (i.e., approach learning) to further understand the range of involvement of regions such as the striatum in affective learning, behavioral control and decision-making.

ACKNOWLEDGMENTS

This work was supported by James S. McDonnell foundation grant to Elizabeth A. Phelps, the Beatrice and Samuel A. Seaver Foundation and a National Institute on Drug Abuse grant to Mauricio R. Delgado (DA022998). We would like to acknowledge Mike Niznikiewicz for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/behavioralneuroscience/paper/10.3389/neuro.08.033.2009/

REFERENCES


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2009 Delgado, Jou, LeDoux and Phelps. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.


Received: 10 June 2009; paper pending published: 28 July 2009; accepted: 12 September 2009; published online: 01 October 2009.


Copyright © 2009 Delgado, Jou, LeDoux and Phelps. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.