Recall of Fear Extinction in Humans Activates the Ventromedial Prefrontal Cortex and Hippocampus in Concert

Mohammed R. Milad, Christopher I. Wright, Scott P. Orr, Roger K. Pitman, Gregory J. Quirk, and Scott L. Rauch

Background: Extinction of conditioned fear is thought to form a new safety memory that is expressed in the context in which the extinction learning took place. Rodent studies implicate the ventromedial prefrontal cortex (vmPFC) and hippocampus in extinction recall and its modulation by context, respectively. The aim of the present study is to investigate the mediating anatomy of extinction recall in healthy humans.

Methods: We used event-related functional magnetic resonance imaging (fMRI) and a 2-day fear conditioning and extinction protocol with skin conductance response as the index of conditioned responses.

Results: During extinction recall, we found significant activations in vmPFC and hippocampus in response to the extinguished versus an unextinguished stimulus. Activation in these brain regions was positively correlated with the magnitude of extinction memory. Functional connectivity analysis revealed significant positive correlation between vmPFC and hippocampal activation during extinction recall.

Conclusions: These results support the involvement of the human hippocampus as well as vmPFC in the recall of extinction memory. Furthermore, this provides a paradigm for future investigations of fronto-temporal function during extinction recall in psychiatric disorders such as posttraumatic stress disorder.

Key Words: anxiety disorders, fear conditioning, fMRI, learning and memory, PTSD, skin conductance response

Extinction deficits might play a role in anxiety disorders, including posttraumatic stress disorder (PTSD) (Bremner 2004; Milad *et al.* 2006; Rauch *et al.* 2006; Rothbaum and Davis 2003). Conversely, behavioral therapies are purported to rely on extinction (Rothbaum and Davis 2003). Therefore, the delineation of the neural underpinnings of fear extinction in humans promises to advance our understanding of anxiety disorders and their treatments.

In Pavlovian conditioning (Charney 2004; Davis *et al.* 2006), a neutral conditioned stimulus (CS) is paired with an aversive unconditioned stimulus (US) (e.g., a shock). After several trials, the CS induces conditioned fear responses (CRs), such as freezing in rodents (Ledoux 2000) and skin conductance responses (SCRs) in humans (Orr *et al.* 2000). Repeatedly presenting the CS without the US extinguishes the CR. Thereafter, two memories exist in the brain: a CS/US association (conditioning memory) and a CS/no-US association (extinction memory) (Quirk 2002; Rescorla 2001). During re-exposure to the extinguished CS in the extinction context, "recall" of the extinction memory is manifest in a small CR. However, re-exposure in a different context produces "renewal," as manifest in a large CR (Bouton 2004).

From the Department of Psychiatry (MRM, CIW, SPO, RKP, SLR), Massachusetts General Hospital and Harvard Medical School, Charlestown; Division of Cognitive and Behavioral Neurology (CIW), Department of Neurology, Brigham and Women's Hospital, Boston, Massachusetts; Department of Veterans Affairs Medical Center (SPO), Manchester, New Hampshire; and the Department of Physiology (GJQ), Ponce School of Medicine, Ponce, Puerto Rico.

Address reprint requests to Scott L. Rauch, M.D., Psychiatric Neuroscience Division, Massachusetts General Hospital, Bldg 149, 13th Street, 2nd Floor, Charlestown, MA 02129; E-mail: rauch@psych.mgh.harvard.edu.

Received June 6, 2006; revised October 6, 2006; accepted October10, 2006.

Animal studies implicate ventromedial prefrontal cortex (vmPFC) and hippocampus in recall of fear extinction and its contextual modulation, respectively (Milad *et al.* 2006; Rauch *et al.* 2006; Sotres-Bayon *et al.* 2006). Rats with vmPFC lesions extinguish but show impaired extinction recall 24 hours later (Lebron *et al.* 2004; Quirk *et al.* 2000). Hippocampal inactivation 24 hours after extinction prevents contextual shifts from modulating extinction recall (Corcoran and Maren 2001). Human functional magnetic resonance imaging (fMRI) studies have shown mPFC activation during extinction of aversive olfactory (Gottfried and Dolan 2004) and eye-blink (Molchan *et al.* 1994) conditioning. In an fMRI experiment by Phelps *et al.* (2004), vmPFC activation during extinction recall in healthy humans significantly correlated with extinction training on the previous day. However, assessment of extinction recall in that study was compromised by the high levels of fear at test, necessitating the removal of the first three recall trials. We found that human vmPFC thickness positively correlated with extinction recall (Milad *et al.* 2005b). However, it is unknown whether extinction recall is associated with activation of the human vmPFC. Furthermore, to date, involvement of the hippocampus in extinction recall in humans has yet to be thoroughly investigated, in that paradigms applied have not explicitly manipulated context.

Here we used a novel design that allowed dissociation of extinction recall from conditioning recall. On Day 1, subjects received conditioning followed by extinction, with photographs of colored lights as CSs (Figure 1). We supplemented the standard design that includes a CS+ (followed by shock) and a CS− (not followed by shock) with a second, distinct CS+ (also followed by shock). Subsequently, one CS+ was extinguished (CS+E) but the other CS+ was not (CS+U). Day 2 tested extinction recall by contrasting responses to the two conditioned stimuli. This within-subjects design can detect neural and behavioral responses associated with extinction memory per se (Quirk *et al.* 2000; Barrett *et al.* 2003). To manipulate context, we displayed visual CSs within photographs of two distinct rooms

0006-3223/07/$32.00
doi:10.1016/j.biopsych.2006.10.011

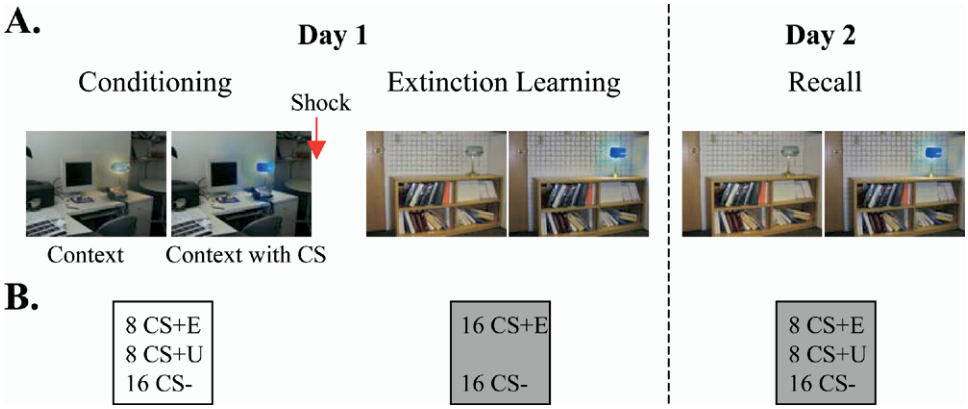

Figure 1. Schematic of experimental protocol. **(A)** Pictures showing the visual contexts used in the experiment, within which conditioned stimuli were presented. In this example, pictures of an office and a conference room represent conditioning and extinction contexts, respectively, whereas the blue light represents the conditioned stimulus (CS)+ that was paired with the shock and later extinguished. **(B)** Schematic representation of the different phases of the experiment. The numbers of each stimulus type presented during the conditioning, extinction learning, and extinction recall are indicated. Gray shading represents the extinction context. Habituation phase is not shown for simplicity. CS+E, extinguished stimulus; CS+U, un-extinguished stimulus.

such that on Day 1, fear conditioning and extinction training were preformed in contexts A and B, respectively, as previously described (Milad *et al.* 2005a). On Day 2, extinction memory recall was tested in the extinction context (B). Such contextual manipulation was done to increase the likelihood of robust extinction recall, because recall of extinction memory becomes less ambiguous when tested in the context where extinction training took place (Bouton 2002). We hypothesized that during extinction recall there would be significant vmPFC and hippocampal activation to the CS+E relative to the CS+U in the extinction context.

Methods and Materials

Subjects

A total of 17 psychiatrically healthy subjects (9 men and 8 women, mean age 25 years old, range 19–39) were recruited from the local community. Written informed consent was obtained in accordance with the requirements of the Partners Healthcare System Human Research Committee. Data from three subjects were excluded owing to technical problems, resulting in a final sample size of 14.

Fear Conditioning, Extinction, and Testing Procedures

The stimuli and experimental protocol used in the present study were similar to those used in a previous psychophysiological study (Milad *et al.* 2005a). The experimental protocol was administered over 2 separate days (see Figure 1). On Day 1, the Habituation phase consisted of 12 trials, in which the to-be CS+s and to-be CS−s (4 of each) were presented in a counterbalanced manner within either the to-be conditioning context or the to-be extinction context. In the Conditioning phase, there were two CS+s (e.g., red and blue lights) that were paired with the US at a partial reinforcement rate of 60%. One of these was extinguished during the subsequent extinction phase on Day 1 (CS+E), whereas the other was not (CS+U). A third CS (e.g., a yellow light) presented during the Conditioning phase was never paired with the US (CS−). The Conditioning phase consisted of 8 CS+E, 8 CS+U, and 16 CS− trials, all presented within the conditioning context. The shock US occurred immediately after CS+ offset with no delay between the CS offset and the US onset. The shock electrodes remained attached to the subject's fingers during all subsequent phases of the experiment, and subjects were instructed throughout the experiment (except during the Habituation phase) that "they may or may not receive the electric shock." However, shocks were actually delivered only during the Conditioning phase. After an approximate 1-min break, the Extinction Learning phase began. During Extinction Learning, 16 CS+E and 16 CS− trials were presented within the extinction context. On Day 2 (the test day), during the Recall phase, 8 CS+E, 8 CS+U, and 16 CS− were presented in the extinction context. No shocks were delivered during or before the onset of the Recall phase.

Presentations of the CS+E and CS+U across the different phases of the experiment were sequential. To elaborate, in any given phase requiring the presentation of both types of CS+s, all 8 trials of one CS+ type (e.g., CS+E) were presented first, followed by 8 trials of the other CS+ type (e.g., CS+U). The CS− trials were intermixed throughout all phases of the experiment. This stimulus presentation method was selected on the basis of pilot work that preceded the formal study, which showed that sequential presentations of the two types of CS+ produced the most effective conditioning. To control for order effects, the order of CS+ type (CS+E or CS+U) presentation was counterbalanced across all phases of the experiment. For each trial during the experiment, the context picture was presented for 9 sec: 3 sec alone, followed by 6 sec in combination with the CS+E, CS+U, or CS−. The mean inter-trial interval was 15 sec (range: 12–18 sec). All the experimental phases were conducted while fMRI data were being acquired.

Psychophysiological Measures

The SCR score was calculated as previously described (Milad *et al.* 2005a; Orr *et al.* 2000; Pitman and Orr 1986). Briefly, SCR for each CS trial was calculated by subtracting the mean skin conductance level during the 2 sec immediately before CS onset (during which the context alone was being presented) from the highest skin conductance level recorded during the 6-sec CS duration. Thus, all skin conductance responses to the CS+E, CS+U, and CS− reported reflect changes in skin conductance level above and beyond any changes in skin conductance level produced by the context. Although there are other equally valid conventions for scoring SCR (e.g., with the first interval response or second interval response; Vansteenwegen *et al.* 2005), limiting our skin conductance scoring to a pre-specified early interval

would not make optimal use of the data gathered and would be more vulnerable to type II error, by virtue of missing peak effects. The magnitude of extinction retention was measured as follows: each subject's SCR to the first two CS+ trial of the extinction Recall phase was divided by their largest SCR to a CS+ trial during the Conditioning phase and then multiplied by 100, yielding a percentage of maximal conditioned responding. This in turn was subtracted from 100% to yield an extinction retention index. Unless otherwise specified, all data are presented as means ± SEM. Student *t* tests were performed to test for statistically significant differences between means, with appropriate Bonferroni corrections when required.

Image Acquisition

A Trio 3.0 Tesla whole body high-speed imaging device equipped for echo planer imaging (EPI) (Siemens Medical Systems, Iselin, New Jersey) with an 8-channel gradient head coil was used. Head movement was restricted with foam cushions. After an automated scout image was obtained and shimming procedures performed, high-resolution three-dimensional magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequences (repetition time [TR]/echo time [TE]/Flip angle = 7.25 msec/3 msec/7°; 1 × 1 mm in plane × 1.3 mm) were collected for spatial normalization and for positioning the subsequent scans. Scans with T1 (TR/TE/Flip angle = 8 sec/39 msec/90°) and T2 (TR/TE/Flip angle = 10 sec/48 msec/120°) sequences were used for registration of individual functional data. Functional MRI images (i.e., blood oxygenation level dependent [BOLD]) were acquired with gradient echo T2*-weighted sequence (TR/TE/Flip angle = 3 sec/30 msec/90°) (Kwong *et al.* 1992). The T1, T2, and gradient-echo functional images were collected in the same plane (45 coronal oblique slices parallel to the anterior-posterior commissure line, tilted 30° anterior) with the same slice thickness (3 mm × 3 mm × 3 mm). The same scanning procedure was conducted on Day 2.

Functional MRI Data Analysis

Functional MRI data were analyzed with the standard processing stream of the Martinos Center for Biomedical Imaging (software and documentation is available at http://www.nmr.mgh.harvard.edu/P41/resourcesSoftDescription.html). Each functional run was motion corrected, spatially smoothed (full width at half maximal [FWHM] = 5 mm) with a three-dimensional Gaussian filter and intensity normalized to the low level baseline. The analysis included a linear correction to account for low-frequency drift. To estimate the stimulus effects at each voxel, an event-related design with gamma fit model was used (Dale 1999). A trial averaging time window of 21 sec (7 TRs) was used that began 6 sec (2 TRs) before trial onset. Group statistical maps were created by calculating a *t* statistic at each voxel on the selectively averaged functional images for the contrasts of interest across the time window (Dale *et al.* 1999). Coordinates for the peak voxels in each region were specified in terms of the Talairach atlas for comparison with previous studies (Talairach and Tournoux 1998).

Main effects of condition for each experimental phase were first assessed. The contrasts used were: 1) first 4 trials of each CS+ (collapsed across CS+E and CS+U) versus CS− for the Conditioning phase; 2) last 12 CS+E versus CS− trials for the Extinction Learning phase; and 3) first 4 trials of CS+E versus CS+U for the Extinction Recall phase. Note that because the shock did not overlap in time with any of the CS+s during the Conditioning phase, the fMRI model for analyzing the conditioning data excluded all shock-induced activity. Once main effects of condition were identified within our *a priori* search territories (vmPFC, hippocampus, and amygdala), functional regions of interest (ROIs) were defined (i.e., clusters of contiguous voxels exceeding the *a priori* statistical threshold) within each of these search territories. Then, for each subject, signal values were extracted from these ROIs to calculate percent signal change for each condition (in reference to the baseline fixation condition).

Given that the main focus of the present study is to examine the neural correlates of extinction recall, we conducted additional analyses to examine the relationship between the fMRI BOLD signal changes in the vmPFC and hippocampus during extinction recall and the magnitude of the psychophysiological index of extinction memory recall. To this end, the percent signal changes obtained from the ROIs within vmPFC and hippocampus were each tested for correlation with extinction retention index. Given that the values were extracted from functional ROIs, we used a threshold of $p < .05$, two-tailed, uncorrected. Furthermore, we also conducted functional connectivity analyses with SPM2 (http://www.fil.ion.ucl.ac.uk/spm/software/spm2) to

Table 1. Significant Activations (+ sign) and Deactivations (− sign) During the Different Phases of the Experiment in *A Priori* Regions of Interest (threshold for peak voxel $p < 10^{-4}$ corrected) and in Other Brain Regions (threshold for peak voxel $p < 10^{-6}$ corrected)

Area of Activation, Brodmann Area	Talairach Coordinates	Peak *p* Value
Fear Conditioning, Contrast: CS+ vs. CS−		
A priori regions		
Amygdala (+)	25, −7, −10	6.9×10^{-5}
vmPFC (−), A25	8, 15, −12.5	5.9×10^{-5}
vmPFC (−), A32	8, 31, −14	7.2×10^{-6}
Regions outside the *a priori* areas		
None		
Late Extinction Learning, Contrast: CS+ vs. CS−		
A priori regions		
Amygdala (+)	−20, −3, −16	1.6×10^{-5}
vmPFC (+), A32	7, 23, −15.5	1.8×10^{-5}
vmPFC (+), A25	−10, 15, −13	4.5×10^{-5}
Regions outside the *a priori* areas		
None		
Extinction Recall, Contrast: CS+E vs. CS+U		
A priori regions		
vmPFC (+), A32	6, 25, −11	2.0×10^{-6}
vmPFC (+), A32	2, 35, −8	6.3×10^{-10}
Hippocampus (+)	−30, −22, −15	2.5×10^{-11}
Hippocampus (+)	29, −20, −14	5.0×10^{-9}
Regions outside the *a priori* areas		
Superior Frontal G. (+), A8	−23, 18, 43	1.9×10^{-9}
Medial Frontal G. (+), A6	23, 1, 20	1.0×10^{-7}
Medial Temporal G. (+), A21	44, 6, −30	6.3×10^{-13}
Superior Prietal G. (+), A40	29, −36, 36	1.0×10^{-13}
Fusiform G. (+), A35	32, −38, −7	1.0×10^{-10}
Globus Pallidus (+)	26, −17.5, −1	1.6×10^{-9}
Cerebellum (+)	7, −74.5, −32	1.0×10^{-12}
Cerebellum (+)	13, −75, −36	1.0×10^{-11}
Cerebellum (+)	10, 68, −13	1.3×10^{-9}
Cerebellum (+)	−10, −75, −36	1.0×10^{-11}
Cerebellum (+)	−21, 75, −42	1.0×10^{-11}

CS, conditioned stimulus; vmPFC, ventromedial prefrontal cortex; E, extinguished; U, not extinguished; G, gyrus.

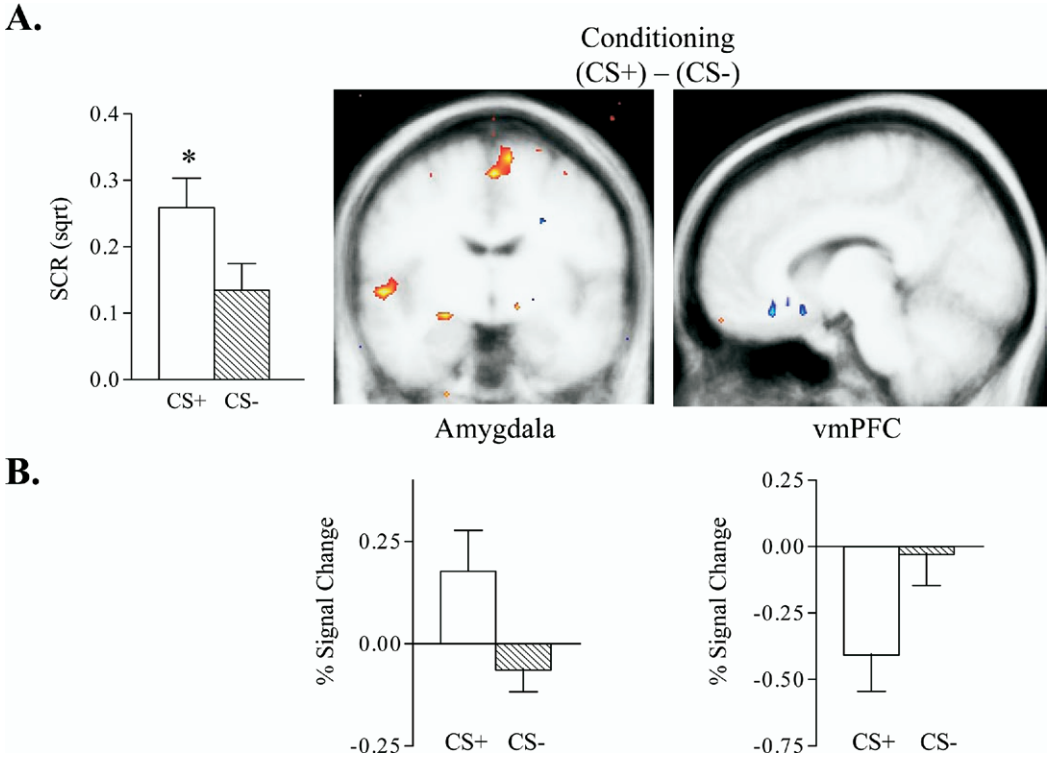

Figure 2. Functional activations and deactivations during early conditioning. **(A)** Mean skin conductance response (SCR) to the first four conditioned stimulus (CS)+ trials was significantly higher than mean SCR to the first four CS− (not followed by shock) trials, indicating successful differential conditioning (left panel) (*$p < .05$). This contrast yielded significant amygdala activation and ventromedial prefrontal cortex (vmPFC) deactivation (right panels). Imaging data are shown in neurological convention (left-right reversed). **(B)** Percent signal change for the amygdala (left) and vmPFC (right) in each condition (relative to fixation).

examine the functional relationship between the vmPFC and the hippocampus as well as the amygdala during extinction recall. Once again, after obtaining a main-effect of condition in the vmPFC (CS+E vs. CS+U), functional ROIs were defined and signal values were extracted for each subject with the MarsBar tool in SPM2. Hence, with the vmPFC main effect locus as the "seed," we performed a voxel-wise correlational analysis within SPM2. The significance threshold for the functional connectivity analysis was set at $p < .001$, two-tailed, uncorrected, reflecting the *a priori* hypothesis that significant associations would be found within hippocampus and/or the amygdala.

Results

Statistical Significance, ROIs

The statistical significance threshold for the contrasts described was $p < 10^{-4}$ for the peak voxel, accompanied by at least four adjacent voxels with $p < 10^{-3}$. These calculations incorporated small volume correction on the basis of the volumes of *a priori* ROIs, viz. vmPFC, hippocampus, and amygdala. A separate Bonferroni correction based on whole brain volume was used to set a significance threshold of $p < 10^{-6}$ for loci outside the *a priori* areas (Kennedy *et al.* 1998). Table 1 lists all brain regions (both within and outside the *a priori* ROIs) that showed activations or deactivations meeting these thresholds, for each phase of the experiment.

Brain Correlates of Fear Conditioning (Day 1)

The average shock intensity selected by the subjects was 2.4 mA ± .24 SE (range: 1.4–4.0 mA). The SCRs to the CS+ and CS− during early conditioning are shown in Figure 2 [SCRs to the CS+E and CS+U were averaged, because they did not differ from one another: $t(13) = .61$, $p > .5$]. The SCRs to the CS+ were significantly larger than SCRs to the CS− [$t(13) = 2.2$, $p < .05$], indicating that differential conditioning was achieved. During early fear conditioning, we found significant right dorsal amygdala activation to the CS+ relative to the CS− (Figure 2A) (*p* value and Talairach coordinates for this locus and all others to be described in the Results section can be found in Table 1). A more medial region of the right amygdala also showed a trend toward activation that was just below our statistical threshold ($p = 1.1 \times 10^{-3}$, x = 23, y = −6, z = −20). In the vmPFC ROI, two discrete loci of significant deactivation to the CS+ relative to the CS− were observed (Figure 2A). Percent signal change for each condition (CS+ and CS−) in the amygdala and the vmPFC is shown in Figure 2B. As can be observed, amygdala activity increased in response to the CS+ presentation, whereas vmPFC activity decreased. Contrary to the findings from the first four trials alone, analysis of all eight conditioning trials did not reveal significant amygdala activation in response to the CS+ relative to the CS−. This is consistent with habituation of amygdala activity in the later conditioning trials. Decreased vmPFC activity to the CS+ relative to the CS− remained significant across the eight conditioning trials.

Brain Correlates of Extinction Learning (Day 1)

Early extinction learning trials typically reflect persistence of the recently acquired, short-term conditioning memory, whereas later trials better reflect extinction learning. Therefore, analysis of extinction learning was limited to the last 12 of the 16 extinction

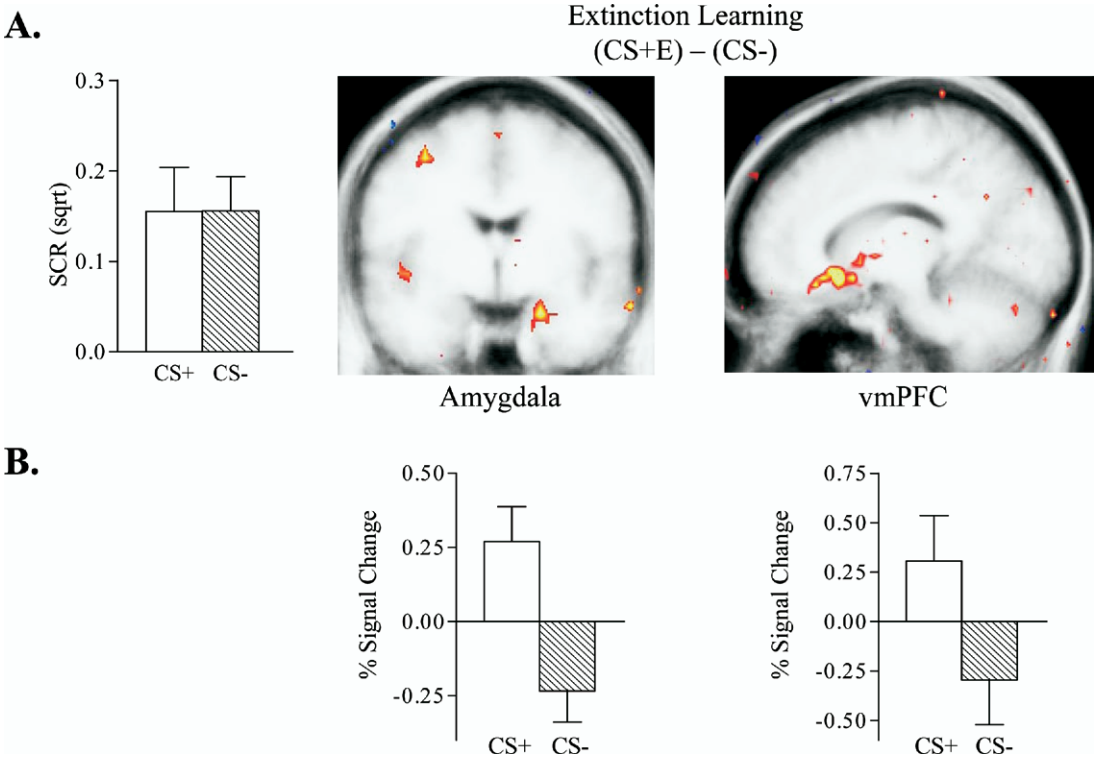

Figure 3. Functional activations and deactivations during late extinction training. **(A)** The absence of a significant difference in mean SCR between the CS+ and CS− during the last 12 extinction learning trials, along with lower levels of both, indicates that CRs were extinguished (left). This contrast yielded significant vmPFC and amygdala activations (right). **(B)** Percent signal change for the amygdala (left) and vmPFC (right) in each condition (relative to fixation). Abbreviations as in Figure 2.

trials to specifically examine the neural correlates of extinction learning per se. During the last 12 extinction trials, no difference was observed in SCRs between the CS+E and CS− [$t(13) = .01$, $p > .9$, Figure 3A]. By fMRI, the same CS+E versus CS− contrast revealed significant bilateral activations in the vmPFC (Figure 3A). Significant activation within the left amygdala (Figure 3A) was also observed. Figure 3b depicts the corresponding percent signal change to the CS+ and CS− in both the amygdala and the vmPFC. These data show that increased fMRI BOLD signal observed in these two brain regions resulted from both an increase in activity to the CS+ as well as a decrease in activity to the CS−.

Brain Correlates of Extinction Recall (Day 2)

The contrast of interest for the Recall phase was CS+E versus CS+U (both within the extinction context), corresponding to recall of extinction, while controlling for recall of conditioning. Mean SCRs to these two stimuli during this phase are shown in Figure 4. The SCRs to the CS+E were significantly lower than to the CS+U [$t(13) = 3.12$, $p < .01$], suggesting maintenance of conditioned fear responses across the 2 days and recall of extinction memory. The CS+E versus CS+U contrast revealed significant activations at two distinct loci within vmPFC (Figure 4A). Moreover, significant activation was also observed in bilateral hippocampi (Figure 4A). Analysis of the percent signal change revealed that the source of this difference between these two conditions is decreased activity to the CS+U in both the vmPFC and hippocampus (Figure 4B).

SCR Correlations and Functional Connectivity Analyses During Extinction Recall

To assess the relationship between extinction memory recall and activations in both the vmPFC and hippocampus, we conducted a regression analysis between the difference in percent signal change (CS+E − CS+U) extracted from our functional ROIs previously described and the extinction retention index. Figure 4C shows that activations in the vmPFC (x = 2, y = 35, z = −8) and right hippocampus (x = 29, y = −20, z = −14) were positively correlated with extinction memory; the higher the activation in these brain regions to the CS+E, the higher the magnitude of extinction memory (vmPFC: $r = .66$, $p = .01$; hippocampus: $r = .72$, $p = .004$; Figure 4C). Activation at the other vmPFC locus (x = 6, y = 25, z = −11) and the left hippocampus (x = −30, y = −22, z = −15) did not significantly correlate with extinction retention index values (vmPFC: $r = -.13$, $p = .65$; hippocampus: $r = .46$, $p = .11$). In addition, no significant correlation was observed between the success of extinction learning on Day 1 (measured as the difference between early and late extinction learning) and the percent signal change in vmPFC extracted from the functional ROI on Day 2 ($r = -.12$, $p = .67$).

To assess the interregional correlations between the vmPFC, hippocampus, and amygdala during extinction recall, we conducted functional connectivity analyses with SPM2 tools as described in the Methods and Materials section. To first identify the vmPFC seed for such correlations, we reanalyzed the extinction recall data with SPM2 and used the same contrast previously described (CS+E vs. CS+U). The findings

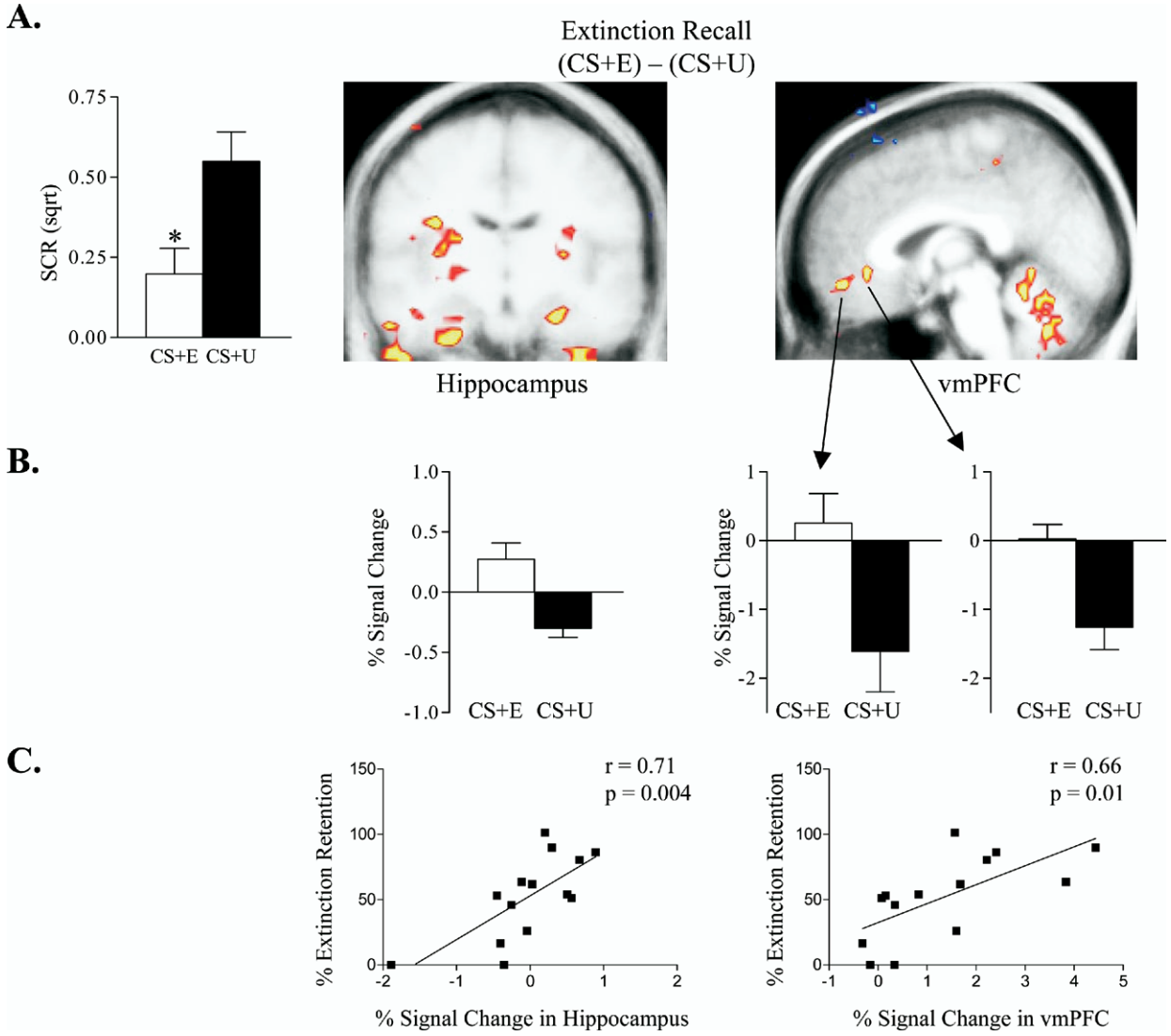

Figure 4. Functional activations and correlations during extinction recall. **(A)** Mean SCR to the extinguished stimulus (CS+E) was significantly lower than mean SCR to the un-extinguished stimulus (CS+U) during the first 4 extinction recall trials, indicating expression of extinction memory in the extinction context (left) (*$p < .05$). This contrast yielded significant hippocampal and vmPFC activations (right). **(B)** Percent signal change for the hippocampus (left) and vmPFC (right) in each condition (relative to fixation). **(C)** Regression plot showing positive correlations between the percent signal change in the vmPFC and hippocampus and extinction retention index. Other abbreviations as in Figure 2.

of these new analyses were largely confirmatory of those previously obtained, specifically in terms of vmPFC loci and hippocampal activations; the locations of peak voxels were similar but not identical, most likely owing to differences in normalization methods between the two applications. Functional connectivity analysis, in which vmPFC activation during extinction recall was taken as the seed (Figure 5A), revealed a significant correlation with activation at a locus within the hippocampus (x = 20, y = −11, z = −20, $r = .94$, $p < .0001$, Figure 5B). In addition, we observed a positive correlation between vmPFC and amygdala activation (x = −17, y = −7, z = −18. $r = .83$, $p < .0005$).

Discussion

We observed vmPFC activation to an extinguished CS+ relative to an unextinguished CS+ during extinction recall. This was accompanied by hippocampal activation. Activations in both of these brain regions were positively correlated with an index of extinction memory recall. In addition, activations in these two brain regions during extinction recall were positively correlated with one another. We also observed vmPFC activation during the late stage of extinction learning. Furthermore, we replicated previous neuroimaging findings showing amygdala activation during fear acquisition as well as during extinction learning (Phelps *et al.* 2004). Together, these findings bridge animal and human data by illustrating the homologous involvement of vmPFC and hippocampus in mediating extinction recall.

There is a notable similarity between the location of the vmPFC regions activated during extinction recall in the present study and locations previously reported in structural (Milad *et al.* 2005b) and functional (Phelps *et al.* 2004) neuroimaging studies linking vmPFC to extinction recall in humans. Taking into consideration the magnitude of spatial smoothing employed in

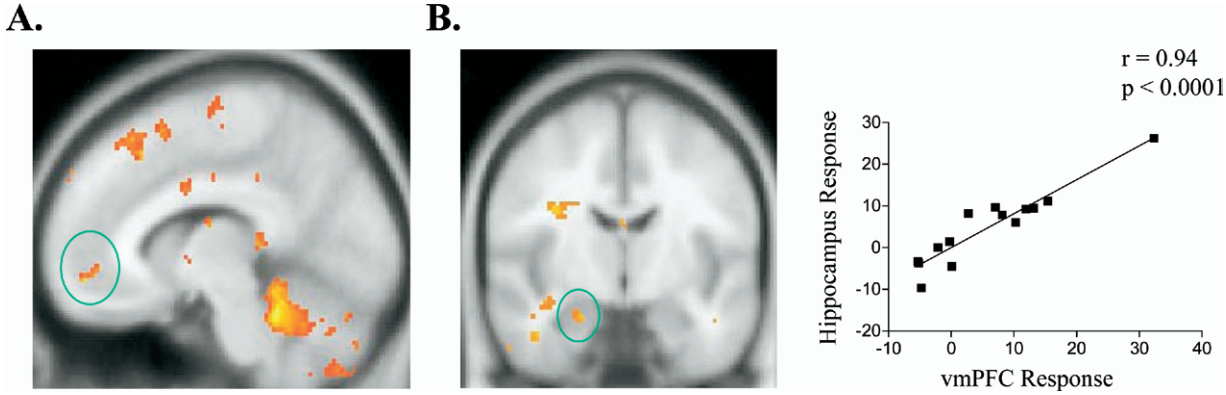

Figure 5. Positive correlations between vmPFC and hippocampal activations during extinction recall. **(A)** Location of vmPFC seed used for the functional connectivity analysis. The seed of activation was obtained from a main-effect of condition with the contrast of CS+E versus CS+U. **(B)** Functional magnetic resonance imaging (fMRI) maps and regression plots for hippocampal region found to be correlated with vmPFC activation. Other abbreviations as in Figure 2.

each of these studies and the fact that these coordinates represent the location of a peak voxel within a larger cluster, the results from these three studies converge to implicate a relatively circumscribed region within the vmPFC in extinction recall.

An apparent homologue to this region in rodents, the infralimbic cortex, is also activated when rats recall extinction (Barrett *et al.* 2003; Herry and Garcia 2002; Milad and Quirk 2002). Indeed, on the basis of data in rodents and in patients with anxiety disorders, several investigators have hypothesized that the human vmPFC plays a specific role in extinction recall (Maren and Quirk 2004; Milad *et al.* 2006; Rauch *et al.* 1998, 2006; Sotres-Bayon *et al.* 2004). The data presented herein not only provide additional support for this hypothesis but also extend previous findings in several important respects. First, the physiological responses measured in the present study validated the occurrence of extinction recall. Second, vmPFC activation was observed during the first few extinction recall trials, when the meaning of the CS is most ambiguous, and was positively correlated with extinction memory recall. Third, from a methodological standpoint, the vmPFC activation was uniquely obtained by contrasting two CS+s, one extinguished and the other not, thereby providing a within-subject control for recall of conditioning.

The positive correlation between vmPFC activation during extinction recall and the magnitude of extinction retention supports the involvement of this brain region in the expression of extinction memory. Furthermore, the functional correlation between the vmPFC and amygdala during extinction recall is also consistent with the idea that vmPFC mediates extinction memory by suppressing the output of the amygdala (Maren and Quirk 2004; Milad *et al.* 2006; Rauch *et al.* 2006). Superficially one might have expected an inverse correlation between these brain regions rather than a positive one. However, the positive correlation is consistent with the top-down control of the amygdala by the vmPFC, if as previously suggested, the vmPFC dampens the output of the amygdala by activating the intercalated clusters of γ-aminobutyric acid [GABA]ergic neurons (ITC) within the amygdala (Berretta *et al.* 2005; Quirk *et al.* 2003; Rosenkranz *et al.* 2003). Considering that BOLD responses can reflect either excitatory or inhibitory processes (Heeger and Ress 2002), the positive correlation observed herein is consistent with activation of inhibitory networks within this structure.

The current paradigm was designed with the express purpose of examining the role of the hippocampus in context-modulated extinction recall. Our experiment was conducted with an ABB design: conditioning in context A, extinction learning in context B, and extinction recall in context B. Indeed, extinction recall with the present design revealed activation in the hippocampus in addition to the vmPFC. Phelps *et al.*'s 2004 pioneering fMRI study of extinction recall in humans was performed within a single context (AAA design) and showed vmPFC activation without accompanying hippocampal activation. Given the documented role of the hippocampus in context-modulated extinction recall in rodents (Corcoran *et al.* 2005), we speculate that the observed hippocampal activation during extinction recall in the present study is related to signaling the extinguished context. The significant functional correlation between the vmPFC and hippocampus as well as their positive correlation with extinction memory recall suggests that these two structures comprise a network that mediates the expression of extinction memory in the appropriate context. A recent animal study showed that low frequency stimulation of the hippocampus immediately after extinction training suppressed long-term potentiation (LTP) induction in the vmPFC–hippocampal pathways and disrupted extinction recall (Farinelli *et al.* 2006), providing further support for this circuit in extinction recall. Such a result is also broadly convergent with findings of hippocampal differences in mood and anxiety disorders, such as PTSD, for which deficiencies in the capacity to appreciate safe contexts have been hypothesized (e.g., Bremner *et al.* 2004; Gilbertson *et al.* 2002; Shin *et al.* 2004b). Further studies are needed to specifically examine the role of the hippocampus in context-specific extinction recall, for example by adding a renewal test phase.

Consistent with previous neuroimaging studies (Knight *et al.* 2004; Phelps *et al.* 2004; Tabbert *et al.* 2005), we observed amygdala activation not only during fear conditioning but also during extinction learning. Amygdala involvement in extinction learning is also suggested by rodent studies. For example, Davis *et al.* (2006) have shown that pharmacological agents that facilitate or block the activity of the basolateral complex of the amygdala can enhance or retard fear extinction, respectively. Moreover, it has been recently shown that extinction training activates protein kinases in the amygdala that are necessary for extinction behavior (Herry *et al.* 2006). As for the vmPFC, we observed deactivation of this brain region during fear acquisition and activation during late extinction learning. The significant vmPFC deactivation during fear conditioning observed in the present study suggests that activity in this brain region is

suppressed to allow for the expression of conditioned fear responses. The significant vmPFC activation during extinction learning was somewhat unexpected, given that previous human and animal studies found no correlations between vmPFC function and behavioral outcomes during extinction learning (e.g., Milad and Quirk 2002; Phelps *et al.* 2004; Santini *et al.* 2004). However, our findings are consistent with the possibility that interactions between the amygdala and vmPFC during extinction training might be necessary for long-term encoding and subsequent retrieval of extinction.

Studies of PTSD that have employed matched comparison subjects have consistently found between-group differences in the vmPFC and hippocampus, with respect to both structure and function (for review see Rauch *et al.* 2006). Considering the mounting evidence of vmPFC deficiency in PTSD patients, it has been hypothesized that PTSD patients might have deficits in fear extinction (Liberzon *et al.* 2003; Milad *et al.* 2006; Rauch *et al.* 1998). Indeed, physiological extinction deficits have been reported in PTSD patients (Orr *et al.* 2000). As previously suggested, activation of the vmPFC and hippocampus is thought to inhibit the amygdala and suppress fear (Milad *et al.* 2006; Rauch *et al.* 2006). Such a model is consistent with PTSD neuroimaging findings (Shin *et al.* 2004a). However, neuroimaging studies investigating the pathophysiology of PTSD have not yet specifically tested the function of the vmPFC and hippocampus during extinction recall (Rauch *et al.* 2006). The fMRI protocol used in the present study could provide the means to accomplish this directly in matched samples of subjects with and without PTSD.

This work was supported by a grant from the National Institute of Mental Health (1R21MH072156-1) to SLR.

We thank Michelle Wedig and Lisa Shin for technical assistance.

Barrett D, Shumake J, Jones D, Gonzalez-Lima F (2003): Metabolic mapping of mouse brain activity after extinction of a conditioned emotional response. *J Neurosci* 23:5740–5749.

Berretta S, Pantazopoulos H, Caldera M, Pantazopoulos P, Pare D (2005): Infralimbic cortex activation increases c-Fos expression in intercalated neurons of the amygdala. *Neuroscience* 132:943–953.

Bouton ME (2002): Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biol Psychiatry* 52:976–986.

Bouton ME (2004): Context and behavioral processes in extinction. *Learn Mem* 11:485–494.

Bremner JD (2004): Brain imaging in anxiety disorders. *Expert Rev Neurother* 4:275–284.

Bremner JD, Vythilingam M, Vermetten E, Vaccarino V, Charney DS (2004): Deficits in hippocampal and anterior cingulate functioning during verbal declarative memory encoding in midlife major depression. *Am J Psychiatry* 161:637–645.

Charney DS (2004): Discovering the neural basis of human social anxiety: A diagnostic and therapeutic imperative. *Am J Psychiatry* 161:1–2.

Corcoran KA, Desmond TJ, Frey KA, Maren S (2005): Hippocampal inactivation disrupts the acquisition and contextual encoding of fear extinction. *J Neurosci* 25:8978–8987.

Corcoran KA, Maren S (2001): Hippocampal inactivation disrupts contextual retrieval of fear memory after extinction. *J Neurosci* 21:1720–1726.

Dale AM (1999): Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8:109–114.

Dale AM, Fischl B, Sereno MI (1999): Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9:179–194.

Davis M, Ressler K, Rothbaum BO, Richardson R (2006): Effects of d-cycloserine on extinction: Translation from preclinical to clinical work. *Biol Psychiatry* 60:369–375.

Farinelli M, Deschaux O, Hugues S, Thevenet A, Garcia R (2006): Hippocampal train stimulation modulates recall of fear extinction independently of prefrontal cortex synaptic plasticity and lesions. *Learn Mem* 13:329–334.

Gilbertson MW, Shenton ME, Ciszewski A, Kasai K, Lasko NB, Orr SP, Pitman RK (2002): Smaller hippocampal volume predicts pathologic vulnerability to psychological trauma. *Nat Neurosci* 5:1242–1247.

Gottfried JA, Dolan RJ (2004): Human orbitofrontal cortex mediates extinction learning while accessing conditioned representations of value. *Nat Neurosci* 7:1144–1152.

Heeger DJ, Ress D (2002): What does fMRI tell us about neuronal activity? *Nat Rev Neurosci* 3:142–151.

Herry C, Garcia R (2002): Prefrontal cortex long-term potentiation, but not long-term depression, is associated with the maintenance of extinction of learned fear in mice. *J Neurosci* 22:577–583.

Herry C, Trifilieff P, Micheau J, Luthi A, Mons N (2006): Extinction of auditory fear conditioning requires MAPK/ERK activation in the basolateral amygdala. *Eur J Neurosci* 24:261–269.

Kennedy DN, Lange N, Makris N, Bates J, Meyer J, Caviness VS Jr (1998): Gyri of the human neocortex: An MRI-based analysis of volume and variance. *Cereb Cortex* 8:372–384.

Knight DC, Smith CN, Cheng DT, Stein EA, Helmstetter FJ (2004): Amygdala and hippocampal activity during acquisition and extinction of human fear conditioning. *Cogn Affect Behav Neurosci* 4:317–325.

Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, *et al.* (1992): Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci U S A* 89:5675–5679.

Lebron K, Milad MR, Quirk GJ (2004): Delayed recall of fear extinction in rats with lesions of ventral medial prefrontal cortex. *Learn Mem* 11:544–548.

Ledoux JE (2000): Emotion circuits in the brain. *Annu Rev Neurosci* 23:155–184.

Liberzon I, Britton JC, Phan KL (2003): Neural correlates of traumatic recall in posttraumatic stress disorder. *Stress* 6:151–156.

Maren S, Quirk GJ (2004): Neuronal signalling of fear memory. *Nat Rev Neurosci* 5:844–852.

Milad MR, Orr SP, Pitman RK, Rauch SL (2005a): Context modulation of memory for fear extinction in humans. *Psychophysiology* 42:456–464.

Milad MR, Quinn BT, Pitman RK, Orr SP, Fischl B, Rauch SL (2005b): Thickness of ventromedial prefrontal cortex in humans is correlated with extinction memory.*Proc Natl Acad Sci U S A* 102:10706–10711.

Milad MR, Quirk GJ (2002): Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* 420:70–74.

Milad MR, Rauch SL, Pitman RK, Quirk GJ (2006): Fear extinction in rats: Implications for human brain imaging and anxiety disorders. *Biol Psychol* 73:61–71.

Molchan SE, Sunderland T, McIntosh AR, Herscovitch P, Schreurs BG (1994): A functional anatomical study of associative learning in humans. *Proc Natl Acad Sci U S A* 91:8122–8126.

Orr SP, Metzger LJ, Lasko NB, Macklin ML, Peri T, Pitman RK (2000): De novo conditioning in trauma-exposed individuals with and without posttraumatic stress disorder. *J Abnorm Psychol* 109:290–298.

Phelps EA, Delgado MR, Nearing KI, Ledoux JE (2004): Extinction learning in humans: Role of the amygdala and vmPFC. *Neuron* 43:897–905.

Pitman RK, Orr SP (1986): Test of the conditioning model of neurosis: Differential aversive conditioning of angry and neutral facial expressions in anxiety disorder patients. *J Abnorm Psychol* 95:208–213.

Quirk GJ (2002): Memory for extinction of conditioned fear is long-lasting and persists following spontaneous recovery. *Learn Mem* 9:402–407.

Quirk GJ, Likhtik E, Pelletier JG, Pare D (2003): Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *J Neurosci* 23:8800–8807.

Quirk GJ, Russo GK, Barron JL, Lebron K (2000): The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *J Neurosci* 20:6225–6231.

Rauch SL, Shin LM, Phelps EA (2006): Neurocircuitry models of posttraumatic stress disorder and extinction: Human neuroimaging research-past, present, and future. *Biol Psychiatry* 60:376–382.

Rauch SL, Shin LM, Whalen PJ, Pitman RK (1998): Neuroimaging and the neuroanatomy of PTSD. *CNS Spectr* 3:30–41.

Rescorla RA (2001): Retraining of extinguished Pavlovian stimuli. *J Exp Psychol Anim Behav Process* 27:115–124.

Rosenkranz JA, Moore H, Grace AA (2003): The prefrontal cortex regulates lateral amygdala neuronal plasticity and responses to previously conditioned stimuli. *J Neurosci* 23:11054–11064.

Rothbaum BO, Davis M (2003): Applying learning principles to the treatment of post-trauma reactions. *Ann N Y Acad Sci* 1008:112–121.

Santini E, Ge H, Ren K, Pena dO, Quirk GJ (2004): Consolidation of fear extinction requires protein synthesis in the medial prefrontal cortex. *J Neurosci* 24:5704–5710.

Shin LM, Orr SP, Carson MA, Rauch SL, Macklin ML, Lasko NB, *et al.* (2004a): Regional cerebral blood flow in the amygdala and medial prefrontal cortex during traumatic imagery in male and female Vietnam veterans with PTSD. *Arch Gen Psychiatry* 61:168–176.

Shin LM, Shin PS, Heckers S, Krangel TS, Macklin ML, Orr SP, *et al.* (2004b): Hippocampal function in posttraumatic stress disorder. *Hippocampus* 14:292–300.

Sotres-Bayon F, Bush DE, Ledoux JE (2004): Emotional perseveration: An update on prefrontal-amygdala interactions in fear extinction. *Learn Mem* 11:525–535.

Sotres-Bayon F, Cain CK, Ledoux JE (2006): Brain mechanisms of fear extinction: Historical perspectives on the contribution of prefrontal cortex. *Biol Psychiatry* 60:329–336.

Tabbert K, Stark R, Kirsch P, Vaitl D (2005): Hemodynamic responses of the amygdala, the orbitofrontal cortex and the visual cortex during a fear conditioning paradigm. *Int J Psychophysiol* 57:15–23.

Talairach J, Tournoux P (1998): *Co-Planar Stereotaxic Atlas of the Human Brain. 3-D Proportional System: An Approach to Cerebral Imaging.* New York: Thieme Publishers.

Vansteenwegen D, Hermans D, Vervliet B, Francken G, Beckers T, Baeyens F, Eelen P (2005): Return of fear in a human differential conditioning paradigm caused by a return to the original acquisition context. *Behav Res Ther* 43:323–336.