Neural reactivity tracks fear generalization gradients

Tsafrir Greenberg [a,b], Joshua M. Carlson [b], Jiook Cha [b,c], Greg Hajcak [a], Lilianne R. Mujica-Parodi [b,*]

[a] Department of Psychology, State University of New York at Stony Brook, United States
[b] Department of Biomedical Engineering, State University of New York at Stony Brook, United States
[c] Graduate Program in Neuroscience and Department of Neurobiology and Behavior, State University of New York at Stony Brook, United States

ABSTRACT

Recent studies on fear generalization have demonstrated that fear-potentiated startle and skin conductance responses to a conditioned stimulus (CS) generalize to similar stimuli, with the strength of the fear response linked to perceptual similarity to the CS. The aim of the present study was to extend this work by examining neural correlates of fear generalization. An initial experiment ($N = 8$) revealed that insula reactivity tracks the conditioned fear gradient. We then replicated this effect in a larger independent sample ($N = 25$). Activation in the insula, anterior cingulate, right supplementary motor cortex and caudate increased reactivity as generalization stimuli (GS) were more similar to the CS, consistent with participants' overall ratings of perceived shock likelihood and pupillary response to each stimulus.

© 2012 Published by Elsevier B.V.

1. Introduction

Paradigms that assess fear learning have provided valuable translational tools for understanding the etiology, maintenance and treatment of anxiety disorders (Milad et al., 2006; Mineka and Oehlberg, 2008). The acquisition and extinction of conditioned fear responses involve a common neurocircuitry across species that includes the amygdala, insula, anterior cingulate cortex, hippocampus, sensory areas, and ventromedial prefrontal cortex (Büchel and Dolan, 2000; LeDoux, 2000; Phelps et al., 2004). In addition to acquisition and extinction, there is increasing interest in fear generalization, which describes the transfer of a conditioned fear response to stimuli that are perceptually similar to the conditioned stimulus (CS). Insofar as the transfer of fear responses from threat-related stimuli to potentially innocuous cues is a common feature in anxiety disorders (Lissek et al., 2008), fear generalization may be a key learning process in the development and maintenance of pathological anxiety.

Recent studies have validated laboratory-based procedures for testing fear generalization, which involves the assessment of fear responses to a CS and to generalization stimuli (GS) that vary in perceptual similarity to the CS (Hajcak et al., 2009; Lissek et al., 2008). In these paradigms, fear responses were quantified with the fear-potentiated startle reflex, which followed a generalization

gradient: the strongest startle reflex was elicited during the CS, with a steep decline corresponding to the relative decrease in similarity of the GS to the CS[1] (Hajcak et al., 2009; Lissek et al., 2008). Lissek and colleagues assessed fear generalization in a paradigm in which participants had to learn which stimulus was the CS and which were the GS. On the other hand, Hajcak and colleagues found comparable results even when participants were explicitly instructed regarding the identity of the CS and the reinforcement contingencies to the CS and GS. Despite being told explicitly which stimulus was the CS, and never being shocked following a GS, participants in the Hajcak et al. study had larger startle responses and reported greater shock likelihood as GS were more perceptually similar to the CS.

Fear generalization paradigms could be useful for assessing pathological fear and risk for anxious psychopathology. For instance, patients with panic disorder exhibit a flatter fear gradient with more gradual decreases in fear response to the GS (Lissek et al., 2010). Hajcak et al. (2009) reported fear generalization deficits in a generalization paradigm as a function of variation in the brain-derived neurotrophic factor (*BDNF*) genotype, which has been related to both learning and anxiety-related behaviors.

In the current study, we sought to extend this work by examining neural activity using fMRI in a fear generalization paradigm that we previously employed (Hajcak et al., 2009). The aim was to elucidate the brain regions associated with generalization, which have

* Corresponding author at: Laboratory for Computational Neurodiagnostics, Department of Biomedical Engineering, Bioengineering Building Room 119, State University of New York at Stony Brook, Stony Brook, NY 11794, United States. Tel.: +1 917 669 3934.
E-mail address: lmujicaparodi@gmail.com (L.R. Mujica-Parodi).

[1] Comparable results have also been obtained using skin conductance (Dunsmoor et al., 2009; Vervliet et al., 2010), which is a more general measure of arousal that is not specific to defensive motivation.

received little attention in the literature, and to examine whether reactivity in these regions exhibit a similar generalization gradient to that reported with peripheral measures of fear. These neural gradients may be useful in identifying deficits in the generalization process and may be relevant to future work on pathological anxiety (e.g., Lissek et al., 2010). In the current study, the CS was a middle-sized rectangle and the GS were six additional rectangles varying in width from the CS by ±20%, ±40% or ±60%.

In an initial experiment (N = 8), we examined regions of interest (ROIs) based on neuroimaging studies of fear learning that have implicated key areas in the expression and inhibition of autonomic and behavioral fear responses (Dunsmoor et al., 2011; Sehlmeyer et al., 2009). These ROIs included the *amygdala, insula, thalamus, caudate, anterior cingulate cortex (ACC) and ventromedial prefrontal cortex (vmPFC)*. We hypothesized that reactivity in one or more of these regions would demonstrate a similar gradient response to the pattern reported in previous laboratory-based studies. In a second experiment[2] (N = 25), we conducted a whole-brain analysis and obtained additional self-report ratings and physiological measures.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Eight individuals (6 females and 2 males) participated in the study (Mean age = 23.2; SD = 4.7). All reported being right handed. Potential participants were screened for prescription and recreational drug usage, as well as neurological and psychological histories. The study was approved by the Stony Brook University Institutional Review Board; all participants provided informed consent.

2.1.2. Procedure

Prior to the scan, an electric shock, delivered to the left wrist (Constant Voltage Stimulator STM 200; Biopack Systems Inc.), was individually set for each participant to a level that was "uncomfortable but not painful". Instructions for the task were then provided. Participants were told that the middle-sized rectangle (CS) indicated a 50% probability that they would receive a subsequent electric shock, but that shocks would never follow rectangles of greater or lesser size. A conditioning phase was administered next, which included five presentations of the CS with electric shock (i.e., CS_{paired}) and one presentation of each of the other six rectangles. The task immediately followed. Thus, the current study examined generalization within the context of a paradigm that combined instructed and associative fear learning.

2.1.3. Task

The task consisted of three blocks presented consecutively. Each block included 40 trials (5 trials × 8 conditions) for a total of 120 trials. The stimuli were seven red rectangles with identical height (56 pixels) and varying width (112–448 pixels) presented against a black background. The middle-sized rectangle (280 pixels) was the conditioned stimulus (CS). Half of the time the CS co-terminated with a 500 ms electric shock (CS_{paired}), while half of the time it did not ($CS_{unpaired}$). The six remaining rectangles differed by ±20%, ±40% or ±60% in width from the CS and served as the generalization stimuli (GS). Stimuli were presented pseudorandomly for 2 s with a jittered interstimulus interval ranging from 4 to 10 s, during which a white fixation cross was shown on a black background. The task was programmed with E-prime 1.2 (Psychology Software Tools, Inc. Pittsburg, PA) and presented with an MRI compatible 60 Hz projector with 1024 × 768 resolution. The duration of the task was 15 min and 12 s.

2.1.4. Image acquisition

Participants were scanned with a 3 tesla Siemens Trio scanner at the Stony Brook Social, Cognitive and Affective Neuroscience (SCAN) center. A total of 456 T2*-weighted echoplanar images were acquired with an oblique coronal angle and TR = 2000 ms, TE = 22 ms, flip angle = 83°, matrix = 96 × 96, FOV = 224 mm × 224 mm, slices = 36 and slice thickness = 3.5 mm. In addition, we obtained T1-weighted structural scans with TR = 1900 ms, TE = 2.53, flip angle = 9°, FOV = 176 mm × 250 mm × 250 mm, and matrix = 176 mm × 256 mm × 256 mm.

2.1.5. Image analysis

Preprocessing procedures were performed in SPM8 and included slice time correction, motion correction, normalization and smoothing with a 6-mm full width at half maximum Gaussian kernel. Preprocessed images were entered into a general

linear model in which each rectangle was modeled as an event with no duration; CS_{paired} and $CS_{unpaired}$ were modeled separately. The six motion parameters estimated during realignment were included as regressors of no interest and serial autocorrelations were modeled using an AR (1) process. First-level single subject statistical parameter maps were created for the 'CS_{paired} − Baseline' (i.e., fixation), '$CS_{unpaired}$ − Baseline' and each of the 'GS − Baseline' contrasts. These contrasts, except for 'CS_{paired} − Baseline', were used in a second-level random effects repeated measures analysis.

2.1.6. Gradients of neural reactivity

Individual bilateral masks were created for the amygdala, insula, thalamus, caudate nucleus, anterior cingulate cortex (ACC) and ventromedial prefrontal cortex (vmPFC) using the Masks for Regions of Interest Analysis software (Walter et al., 2003). A region of interest (ROI) analysis for the F-contrast (main group effect) was performed using an initial threshold of $\alpha = .01_{(uncorrected)}$ and extent threshold = 20 contiguous voxels, and a small volume familywise error rate corrected $\alpha = .05$, for each mask.

Neural gradients were generated for the right and left insula (which were the only regions that showed significant activation with these thresholds) by extracting the first eigenvariate (i.e., the principal component) from a 6 mm sphere centered on the local maxima within each region, for each of the '$CS_{unpaired}$ − Baseline' and 'GS − Baseline' contrasts, across all participants. Mean values for $CS_{unpaired}$, as well as GS ± 20%, GS ± 40% and GS ± 60%, were plotted as a four-point gradient.

2.2. Results

2.2.1. Gradients of neural reactivity

Generalization gradients for the right and left insula are shown in Fig. 1b and c, respectively. Reactivity in the right ($F(3,21) = 18$, $p < .001$) and left ($F(3,21) = 13.3$, $p < .001$) insula varied as a function of stimulus type. For the right insula, pairwise comparisons revealed higher reactivity for the $CS_{unpaired}$ versus GS ± 40% ($p = .004$) and GS ± 60% ($p = .01$), and for the GS ± 20% versus GS ± 40% ($p = .02$). A comparison of the GS ± 20% to GS ± 60% was marginally significant ($p = .053$). For the left insula, reactivity was higher for the $CS_{unpaired}$ versus GS ± 40% ($p = .007$) and GS ± 60% ($p = .03$), and for the GS ± 20% versus both GS ± 40% ($p = .03$) and GS ± 60% ($p = .04$).

3. Experiment 2

3.1. Methods

3.1.1. Participants

Twenty-five women participated in the study (Mean age = 21.6; SD = 5.1). All reported being right-handed except for one participant, who reported being ambidextrous. Participants were screened for psychiatric illness with the Structured Clinical Interview for DSM-IV Axis I Disorders – Patient Edition, Version 2 (SCID-I/P; First et al., 2002). All other screening procedures were identical to Experiment 1. The study was approved by the Stony Brook University Institutional Review Board; all participants provided informed consent.

3.1.2. Experimental paradigm

The experimental paradigm was identical to Experiment 1 except for the addition of post-task ratings of shock likelihood for each rectangle, obtained on a Likert scale of 1 ("certainly not shocked") to 5 ("certainly shocked"), acquisition of pupillary response with the Eyelink-1000 (SR Research Ltd., Ontario) as a measure of activation of the sympathetic nervous system, as well as a 12 s increase in task length to accommodate a change in TR and TE due to scanner requirements.

3.1.3. Image acquisition and analysis

A total of 440 T2*-weighted echoplanar images were acquired with an oblique coronal angle and TR = 2100 ms, TE = 23 ms, flip angle = 83°, matrix = 96 × 96, FOV = 224 mm × 224 mm, slices = 37 and slice thickness = 3.5 mm. Parameters for acquisition of structural images, as well as preprocessing procedures and statistical analysis were identical to Experiment 1.

3.1.4. Gradients of neural reactivity

Gradients of neural reactivity were generated for all brain regions for which we found significant clusters for the main effect group F-contrast using a whole brain threshold of $\alpha = .001_{(uncorrected)}$ and extent threshold of 20 contiguous voxels.

3.1.5. Preprocessing of pupil data

Pupil data was processed using custom MATLAB codes (MathWorks). First, we excluded periods of eye blinks detected by an on-line parsing system (Eyelink; SR Research Ltd., Ontario). We used a window of 100 ms prior to onsets of eye blinks and 300 ms following their offset in order to minimize after-blink constriction effects. Missing values were linearly interpolated. We adopted pre-processing procedures from Hupé et al. (2009). Specifically, a baseline for each trial was calculated by averaging data points from 500 ms immediately preceding the onset of the stimulus and then subtracting this mean from each trial. The baseline corrected values were
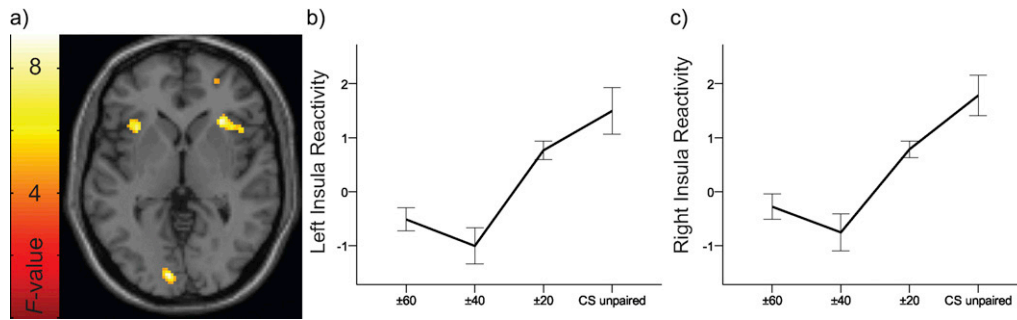
Fig. 1. Activation map and neural gradients for the right and left insula (Experiment 1). (a) An axial slice shows activation in the right and left insula for the F-contrast main group effect ($P < .001_{uncorrected}$). Left (b) and right (c) insula reactivity as a function of stimulus type. Error bars indicate standard errors of means.

z-scored to allow comparison across participants and filtered using a low-pass filter (4 Hz cutoff frequency) to reduce measurement noise. After preprocessing, trials were examined and excluded if they had outliers, exceeding two standard deviations and no pupillary light reflex in response to luminosity changes of presented stimuli. The average number of trials excluded for each participant was 19.1% (SD = 9.1). For statistical analysis, the pupillary response was defined as the overall pupil diameter change (i.e., area under curve) within a 1000 ms window, beginning 1 s after stimulus onset. Data for one participant was excluded due to technical problems.

3.2. Results

Self-report and neural measures were evaluated with repeated measures analysis of variance. Pupillary response was evaluated with a mixed linear model in order to account for missing trials across conditions and to increase sensitivity of the analysis. Pairwise comparisons for all measures were made using the Bonferroni test for multiple comparisons.

3.2.1. Self-reported shock likelihood

Shock likelihood ratings varied as a function of stimulus type ($F_{(3,72)} = 77.9$, $p < .001$; Fig. 2a). Pairwise comparisons showed that shocks were rated as more likely following the CS ($M = 3.9$, SD = 1) compared to GS ± 20% ($M = 3.02$, SD = .88; $p = .007$), GS ± 40% ($M = 1.66$, SD = .67; $p < .001$) or GS ± 60% ($M = 1$, SD = .14; $p < .001$). In addition, shocks were rated as more likely following the GS ± 20% compared to both GS ± 40% ($p < .001$) and GS ± 60% ($p < .001$), and the GS ± 40% compared to GS ± 60% ($p < .001$).

3.2.2. Pupillary response

A generalization gradient of pupillary response is presented in Fig. 2b. Pupillary response varied as a function of stimulus type ($F_{(4,44)} = 30.4$, $p < .001$). Pairwise comparisons showed that pupillary response was larger for the CS$_{unpaired}$ versus GS ± 40% ($p < .001$) and GS ± 60% ($p < .001$). The difference between the CS$_{unpaired}$ and the GS ± 20% was marginally significant ($p = .078$). Comparisons for GS ± 20% versus GS ± 40% and GS ± 60%, and for GS ± 40% versus GS ± 60% were not significant (all ps > .10).

3.2.3. Gradient of neural reactivity

Generalization gradients for the right and left insula are presented in Fig. 3b and c, respectively. Reactivity in both regions varied as a function of stimulus type (right insula: $F_{(3,72)} = 21.8$, $p < .001$; left insula: $F_{(3,72)} = 9.2$, $p < .001$). For the right insula, there was higher reactivity for the CS$_{unpaired}$ versus GS ± 40% ($p < .001$) and GS ± 60% ($p < .001$), and for the GS ± 20% versus GS ± 40% ($p < .001$). A comparison of the GS ± 20% to GS ± 60% was marginally significant ($p = .055$). For the left insula, reactivity was higher for the CS$_{unpaired}$ versus GS ± 40% ($p = .002$) and GS ± 60% ($p = .04$), and for the GS ± 20% versus GS ± 40% ($p < .001$) and GS ± 60% ($p < .03$). In addition to the insula, the F-contrast revealed significant clusters for the *anterior cingulate cortex (ACC), right supplementary motor area (SMA), caudate, ventromedial prefrontal cortex (vmPFC), somatosensory cortex and primary visual cortex.*. We therefore, generated gradients for these brain regions in order to compare their response patterns to that of the insula. Reactivity in the ACC (Fig. 4a; $F_{(3,72)} = 9.7$, $p < .003$), right SMA (Fig. 4b; $F_{(3,72)} = 15.7$, $p < .001$) and the right and left caudate (Fig. 4c; right caudate: $F_{(3,72)} = 11.52$, $p < .001$; left caudate: ($F_{(3,72)} = 9.39$, $p < .001$) showed a similar response pattern with higher reactivity associated with increased similarity of GS to CS. For the ACC, reactivity was higher for the CS$_{unpaired}$ and for the GS ± 20 relative to GS ± 40% ($p < .001$ and $p = .001$, respectively). For the right SMA, reactivity was higher for the CS$_{unpaired}$ versus GS ± 40 ($p < .001$) and GS ± 60% ($p = .001$), and for the GS ± 20% versus GS ± 40 ($p = .003$) and GS ± 60% ($p = .03$). For the right caudate, reactivity was higher for the CS$_{unpaired}$ versus GS ± 40 ($p < .001$) and GS ± 60% ($p = .04$), and for the GS ± 20% versus GS ± 40 ($p = .001$). For the left caudate, reactivity was higher for the CS$_{unpaired}$ versus GS ± 40 ($p < .001$) and GS ± 60% ($p = .03$), and for the GS ± 20% versus GS ± 40 ($p = .01$). In contrast, reactivity in the vmPFC (Fig. 5a; $F_{(3,72)} = 13$, $p < .001$) and somatosensory cortex (Fig. 5b; $F_{(3,72)} = 13$, $p < .001$) showed a reverse response pattern. For the vmPFC, reactivity was higher for the

GS ± 60% versus GS ± 40 ($p = .01$), GS ± 20% ($p < .001$) and CS$_{unpaired}$ ($p < .001$), and for the GS ± 40 versus CS$_{unpaired}$ ($p = .02$). Similarly, for the somatosensory cortex, reactivity was higher for the GS ± 60% versus GS ± 40 ($p = .03$), GS ± 20% ($p = .007$) and CS$_{unpaired}$ ($p < .001$), and for the GS ± 40 versus CS$_{unpaired}$ ($p = .02$). Finally, there was no significant effect of stimulus type for the right ($F_{(3,72)} = 2.5$, $p = .08$) and left ($F_{(3,72)} = .5$, $p = 66$) visual cortex.

3.2.4. Direct comparisons of CS$_{(unpaired)}$ versus GS

Areas of activation for the CS$_{unpaired}$ versus GS ± 40% and GS ± 60% are presented in Table 1. This direct comparison showed enhanced activation in the anterior insula, SMA, cingulate gyrus, caudate, thalamus and frontal areas. Reactivity in these regions is commonly reported in neuroimaging studies of fear learning (Sehlmeyer et al., 2009) and suggests that generalization engages the same circuitry. There was no significant activation for the CS$_{unpaired}$ versus GS ± 20% contrast.

Examination of the reverse contrast (i.e., GS ± 40% and GS ± 60% versus CS$_{unpaired}$; Table 1) showed increased activation in the vmPFC, somatosensory and motor areas, and subcallosal and posterior cingulate. Both the vmPFC and cingulate are associated with modulation of the fear response. There was no significant activation for the GS ± 20% versus CS$_{unpaired}$ contrast.

3.2.5. Amygdala reactivity

The CS$_{unpaired}$ versus GS comparisons did not reveal significant activation in the amygdala. However, since several studies of fear conditioning have demonstrated habituation in the amygdala (Büchel et al., 1998; LaBar et al., 1998), we conducted a linear time modulation analysis to test whether amygdala reactivity showed a decline in reactivity over time. For this analysis, we included linear regressors in a new statistical model to account for time effects in reactivity to the CS$_{unpaired}$, GS ± 20%, GS ± 40% and GS ± 60%. There was a significant decrease in amygdala reactivity over time for all four trial groups (CS$_{unpaired}$: right: $x = 32$, $y = -2$, $z = -12$, $t_{(96)} = 4.9$, $p < .001$, left: $x = -26$, $y = -2$, $z = -12$, $t_{(96)} = 3.82$, $p = .01$; GS ± 20%: right: $x = 32$, $y = -2$, $z = -18$, $t_{(96)} = 4.12$, $p = .005$; GS ± 40%: right: $x = 18$, $y = 4$, $z = -18$, $t_{(96)} = 4.07$, $p = .03$, left: $x = -26$, $y = 0$, $z = -16$, $t_{(96)} = 4.07$, $p = .006$; GS ± 60%: right: $x = 26$, $y = -2$, $z = -22$, $t_{(96)} = 3.56$, $p = .03$; small volume corrected with a bilateral amygdala mask and $\alpha = .05$). There was no interaction of time and stimulus type.

4. General discussion

Across two experiments, the right and left insula showed increased activation to the CS and decreases in response amplitude as a function of increasing dissimilarity between the GS and CS. In addition to the insula, the anterior cingulate cortex (ACC), the right supplementary motor area (SMA) and caudate showed a similar reactivity pattern in the second experiment.[3] Reactivity in the ventromedial prefrontal cortex (vmPFC) and primary somatosensory cortex showed the opposite pattern (i.e., largest response to the GS most dissimilar to the CS), and reactivity in the visual cortex was sensitive to increases in stimuli size. Thus, a pattern consistent with autonomic quantifications of generalization was restricted to the insula, ACC, right SMA and caudate, and was not universally present across the brain. The inverse reactivity pattern in the vmPFC, an area linked to extinction recall and safety learning (Corcoran and Quirk, 2007; Milad et al., 2007) may be associated with inhibition of the

---

[3] Reactivity patterns for the ACC and SMA in Experiment 1 also showed a linear trend but only with a less stringent whole-brain threshold of $\alpha = .01$.
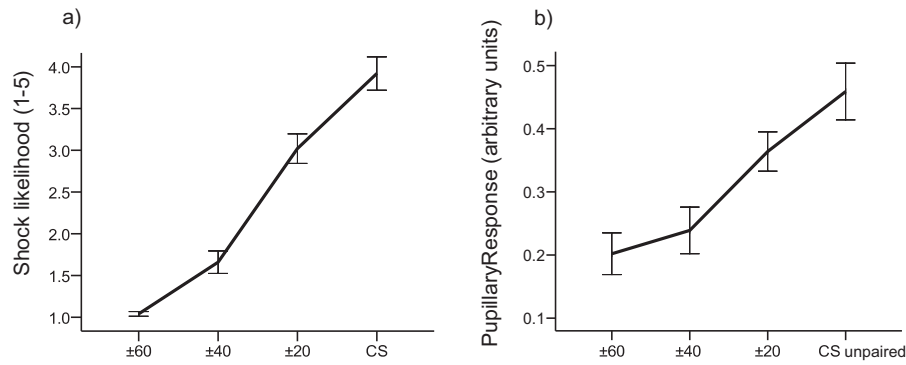
Fig. 2. Post-task ratings of shock likelihood and pupillary response as a function of stimulus type (Experiment 2). (a) Likelihood of shock was rated on a 5-point Likert scale with 1 = "certainly not shocked" and 5 = "certainly shocked". (b) Pupillary response measures (arbitrary units). Error bars indicate standard errors of means.
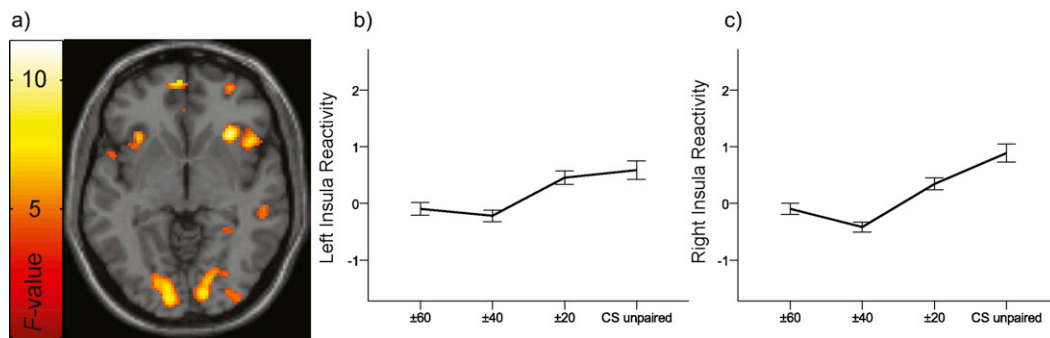

Fig. 3. Activation maps and neural gradients for the right and left insula (Experiment 2). (a) An axial slice shows activation in the right and left insula for the F-contrast main group effect ($P < .001_{uncorrected}$). Left (b) and right (c) insula reactivity as a function of stimulus type. Error bars indicate standard errors of means.

fear response as GS decrease in similarity to the CS. With respect to the somatosensory cortex, the reduced response in this region during presentation of the $CS_{unpaired}$ may be indicative of participants' preparatory response to expected shocks that are not presented. Similarly, a recent study found deactivation in somatosensory areas for a condition in which shocks were expected but not delivered relative to CS_{minus} (Linnman et al., 2011).

The insula is a key structure implicated in anticipatory processing and receives information about the salience, relative value and interoception associated with stimuli (Craig, 2002; Lovero et al., 2009; Menon and Uddin, 2010; Nitschke et al., 2006). The anterior portion, in particular, is commonly activated in paradigms that elicit autonomic arousal, and is thought to be involved in prediction of future aversive body states linked to conditioned

Table 1
Brain activation associated with response to the $CS_{unpaired}$ and generalization stimuli.

Analysis and region	Hemisphere	MNI coordinates			Maximum voxel	
		x	y	z	Voxels	t value
$CS_{unpaired} > GS \pm 40$ and ± 60						
Insula	R	32	22	−4	1087	6.55
	L	−32	20	−12	267	4.64
Anterior cingulate	R	10	20	34	64	4.40
Cingulate gyrus (BA 23)	R	4	−22	32	276	4.77
Caudate	R	10	12	8	190	5.14
	L	−12	0	10	199	4.24
Thalamus	L	−8	−14	14	46	3.96
Superior frontal gyrus (including SMA, BA 8)	R	4	36	46	505	5.61
Middle frontal gyrus	R	32	54	−6	215	4.85
BA 40 (extending to BA 7)	R	44	−60	56	608	4.63
	L	−52	−44	56	79	4.44
Middle temporal gyrus	R	56	−28	−6	201	4.69
$GS \pm 40$ and $\pm 60 > CS_{unpaired}$						
Ventromedial prefrontal cortex	R–L	6	34	−18	591	5.37
Subcallosal cingulate	R–L	−4	12	−10	133	5.15
Medial frontal gyrus	L	−4	60	−4	174	4.91
Precentral gyrus (extends to postcentral gyrus)	R–L	8	−30	70	1964	5.58
Posterior cingulate	R	32	−52	22	313	4.65
	L	−6	−58	20	63	4.07
Angular gyrus	R	46	−76	32	474	4.86
	L	−46	−74	8	184	4.03

Whole brain threshold $\alpha = .001_{(uncorrected)}$, extent threshold = 20 voxels; BA = Brodmann area.
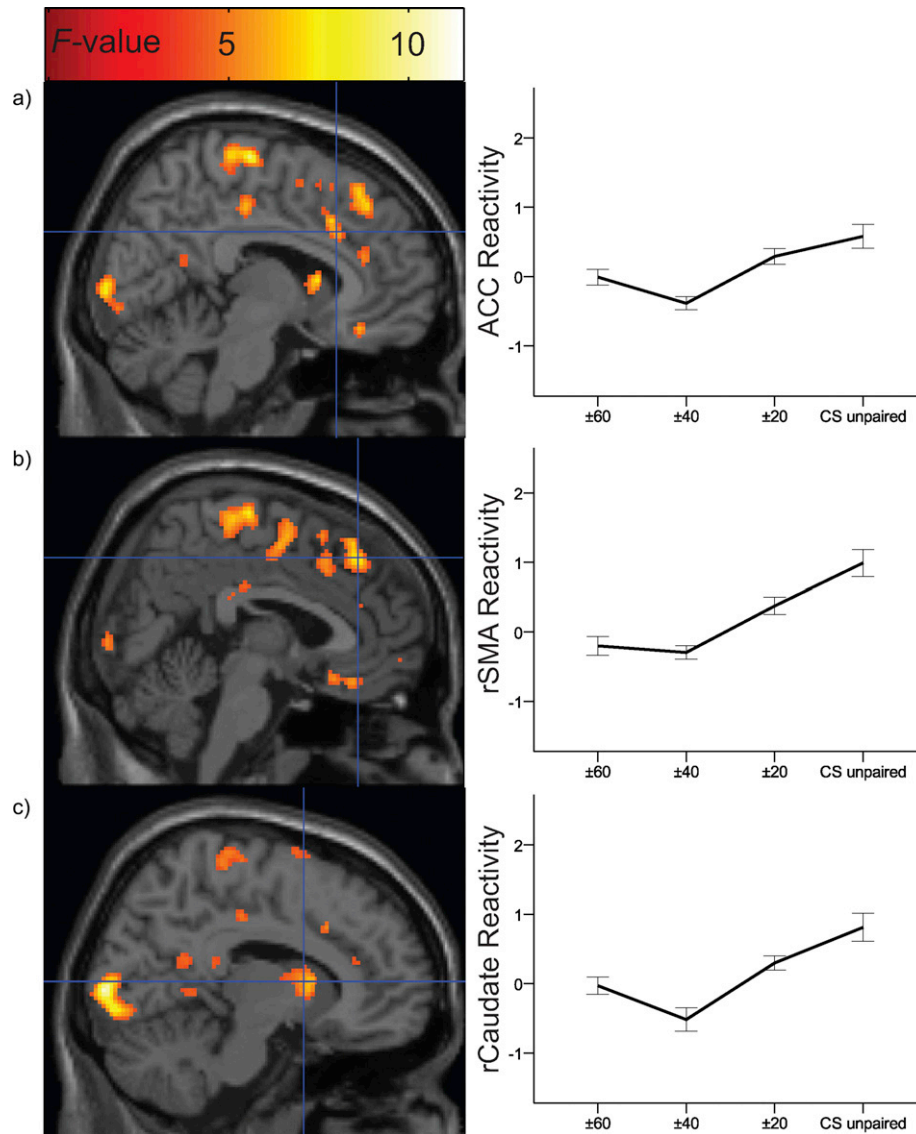
Fig. 4. Activation maps and neural gradients for the anterior cingulate cortex, right supplementary motor area and right caudate (Experiment 2). Sagittal slices showing activation in the anterior cingulate (a), right supplementary motor area (b) and right caudate (c; the gradient for the left caudate was nearly identical and is not presented) for the F-contrast main group effect ($P < .001_{\text{uncorrected}}$) are presented on the left. Cross-hairs indicate the maximum activated voxel within each region used for extraction of the first eigenvariates. Neural gradients for each region are presented on the right. Error bars indicate standard errors of means.

stimuli (Paulus and Stein, 2006) and interoception (Critchley et al., 2004). Information received by the insula is integrated and relayed to other regions that execute attentional, physiological and motor responses. The anterior insula and ACC are reciprocally connected and co-activate in a range of behaviors, including in preparatory responses to painful stimuli (Craig, 2009; Medford and Critchley, 2010). In the context of the current paradigm, we speculate that the pattern of ACC activation observed may be associated with increases in attention and defensive physiological changes linked to perceived similarity of the GS to the CS, and that SMA reactivity may be associated with the initiation of a motor defense response to potential threat of the shock. The insula may modulate these responses, which is consistent with its proposed role in linking anticipatory processing and autonomic arousal with action-planning aimed at avoidance behavior (Simmons et al., 2006).

The striatum (i.e., caudate and putamen) has been implicated in prediction of both appetitive (O'Doherty, 2004) and aversive (Delgado et al., 2008; Schiller et al., 2008) stimuli.

Specifically, studies using shocks (Schiller et al., 2008) and monetary loss (Delgado et al., 2008) as reinforcers report that activity in this region correlates with prediction error (PE) signals that represent the difference between an actual and an expected outcome. For the generalization task, these PE signals would be hypothesized to gradually decrease as anticipation of a shock diminishes. In addition, the striatum is involved in inhibitory motor control (Vink et al., 2005) and may be associated with inhibition of motor responses to anticipated shocks.

The amygdala did not show activation for the CS_{unpaired} versus GS comparisons. However, we found a decline in amygdala reactivity to the CS_{unpaired}, $GS \pm 20\%$, $GS \pm 40\%$ and $GS \pm 60\%$ overtime. Previous studies of fear conditioning reported habituation in amygdala response with increasing number of trials (Büchel et al., 1998), which supports an amygdala role in learning CS-UCS (unconditioned stimulus) associations in the early stages of conditioning (Büchel et al., 1998; LaBar et al., 1998; but see also Bach et al., 2011). Attenuation in amygdala response was also observed when
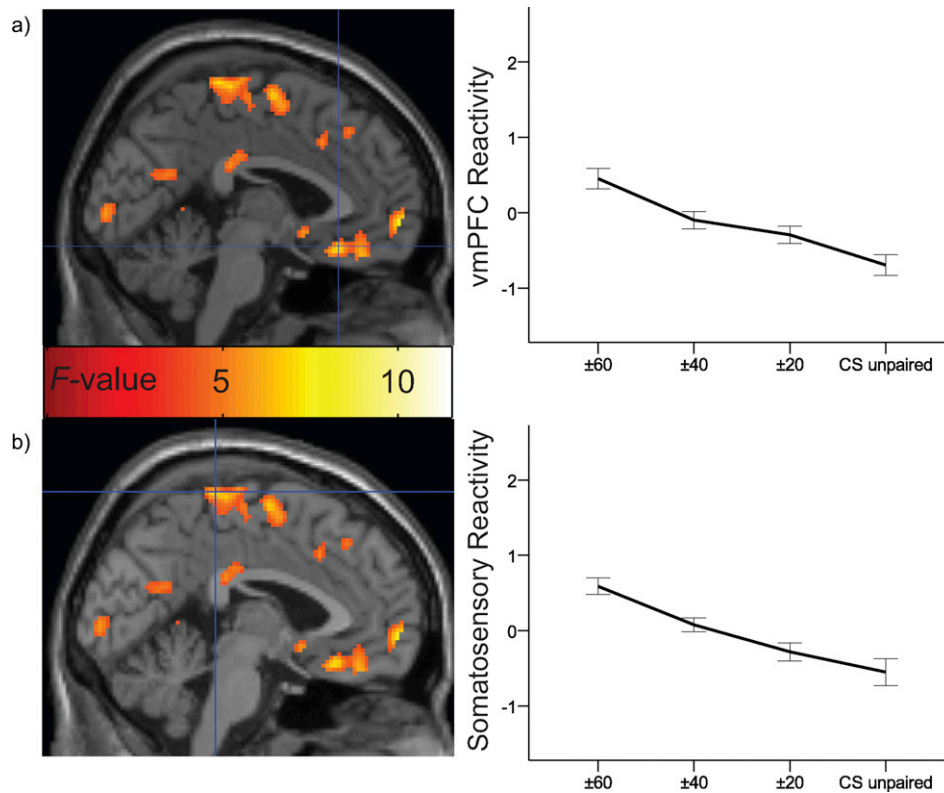
Fig. 5. Activation maps and neural gradients for the ventromedial prefrontal cortex (vmPFC) and somatosensory area as a function of stimulus type (Experiment 2). Sagittal slices showing activation in vmPFC (a), and somatosensory area (b) for the F-contrast main group effect ($P < .001_{\text{uncorrected}}$) are presented on the left. Cross-hairs indicate the maximum activated voxel within each region used for extraction of the first eigenvariates. Neural gradients for each region are presented on the right. Error bars indicate standard errors of means.

participants were informed about the CS-UCS contingency (i.e., instructed fear paradigm) and was correlated with a physiological expression of fear (Phelps et al., 2001). Animal studies have long demonstrated that the amygdala is involved in both acquisition and expression of fear responses (Davis, 1992; LeDoux, 2000). More recently, plasticity in GABAergic synapses in the lateral (Shaban et al., 2006) and central nuclei (Ciocchi et al., 2010) of the amygdala was linked to generalization of conditioned fear responses. The amygdala response pattern in this task was the same for the CS_{unpaired} and GS and may reflect greater engagement of the amygdala by all stimuli, early on, when the relationship between stimuli and shock is established. The absence of a differential response for the CS_{unpaired} and GS may be due to their similarity. The process of learning the stimulus-shock associations likely depends on inputs from mPFC – a region involved in conscious threat appraisal (Mechias et al., 2010). Additionally, the amygdala may be involved in gating the expression of fear responses, as proposed by Ciocchi et al. (2010), possibly via inhibitory inputs from vmPFC. The amygdala and other regions engaged by the generalization task are part of a core neurocircuitry underlying fear conditioning (LeDoux, 2000). Similar regions were implicated in a recent study that examined generalization in response to graded facial expressions (Dunsmoor et al., 2011). Together, these findings provide strong evidence that a shared neurocircuitry supports both forms of fear learning.

Neural reactivity in the insula, ACC, right SMA and caudate paralleled participants' mean post-task ratings of perceived shock likelihood and pupillary response for each stimulus. This reactivity pattern across measures, demonstrating a peak response to the CS and decline in responding associated with greater perceptual dissimilarity of the GS to the CS, was detected despite explicit

pre-task instructions regarding the identity of the CS and the reinforcement contingencies to the CS and GS. These findings are consistent with generalization paradigms using fear-potentiated startle, which also showed that healthy individuals primarily generalize defensive physiological responses to GS that are closest in similarity to the CS (Hajcak et al., 2009; Lissek et al., 2008). The convergent validity of these various measures suggests that, used together, they may facilitate assessment of individual variability in the generalization of fear responses, which may help distinguish healthy and clinical populations. In line with preliminary findings in panic disorder, which show a physiological generalization pattern that extends to GS with greater perceptual difference from the CS (Lissek et al., 2010), we expect that anxious patients, relative to healthy individuals, will demonstrate flatter generalization gradients to more perceptually dissimilar stimuli. Furthermore, based on evidence for vmPFC hypoactivation in PTSD (Etkin and Wager, 2007) it is plausible that the neural gradient in this region may track patients' deficiencies in top-down modulation of the fear response. Together, these gradients may serve as potential markers for diagnosis and/or for prediction of treatment response and therapeutic efficacy. Finally, the inclusion of disorder-specific GS in future studies may help to further discriminate fear responses in different anxiety disorders.

References

Bach, D.R., Weiskopf, N., Dolan, R.J., 2011. A stable sparse fear memory trace in human amygdala. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience 31 (25), 9383–9389.

Büchel, C., Dolan, R.J., 2000. Classical fear conditioning in functional neuroimaging. Current Opinion in Neurobiology 10 (2), 219–223.

Büchel, C., Morris, J., Dolan, R.J., Friston, K.J., 1998. Brain systems mediating aversive conditioning: an event-related fMRI study. Neuron 20 (5), 947–957.

Ciocchi, S., Herry, C., Grenier, F., Wolff, S.B., Letzkus, J.J., Vlachos, I., et al., 2010. Encoding of conditioned fear in central amygdala inhibitory circuits. Nature 468 (7321), 277–282.

Corcoran, K.A., Quirk, G.J., 2007. Recalling safety: cooperative functions of the ventromedial prefrontal cortex and the hippocampus in extinction. CNS Spectrums 12 (3), 200–206.

Craig, A.D., 2002. How do you feel? Interoception: the sense of the physiological condition of the body. Nature Reviews Neuroscience 3 (8), 655–666.

Craig, A.D., 2009. How do you feel now? The anterior insula and human awareness. Nature Reviews Neuroscience 10 (1), 59–70.

Critchley, H.D., Wiens, S., Rotshtein, P., Ohman, A., Dolan, R.J., 2004. Neural systems supporting interoceptive awareness. Nature Neuroscience 7 (2), 189–195.

Davis, M., 1992. The role of the amygdala in conditioned fear. In: Aggleton, J.P. (Ed.), The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction. Wiley-Liss, New York, NY, USA, pp. 255–306.

Delgado, M.R., Li, J., Schiller, D., Phelps, E.A., 2008. The role of the striatum in aversive learning and aversive prediction errors. Philosophical Transactions of the Royal Society of London Series B, Biological sciences 363 (1511), 3787–3800.

Dunsmoor, J.E., Mitroff, S.R., LaBar, K.S., 2009. Generalization of conditioned fear along a dimension of increasing fear intensity. Learning and Memory 16 (7), 460–469.

Dunsmoor, J.E., Prince, S.E., Murty, V.P., Kragel, P.A., LaBar, K.S., 2011. Neurobehavioral mechanisms of human fear generalization. NeuroImage.

Etkin, A., Wager, T.D., 2007. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. The American Journal of Psychiatry 164 (10), 1476–1488.

First, M. B., Spitzer, R. L., Gibbon, M., Williams J.B.W., 2002.Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P) New York: Biometrics Research, New York State Psychiatric Institute.

Hajcak, G., Castille, C., Olvet, D.M., Roohi, J., Hatchwell, E., Dunning, J.P., 2009. Genetic variation in brain-derived neurotrophic factor and human fear conditioning. Genes, Brain and Behavior 8 (1), 80–85.

Hupé, J.-M., Lamirel, C., & Lorenceau, J., 2009. Pupil dynamics during bistable motion perception. Journal of Vision, 9(7):10, 1–19.

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 (5), 937–945.

LeDoux, J.E., 2000. Emotion circuits in the brain. Annual Review of Neuroscience 23, 155–184.

Linnman, C., Rougemont-Bucking, A., Beucke, J.C., Zeffiro, T.A., Milad, M.R., 2011. Unconditioned responses and functional fear networks in human classical conditioning. Behavioural Brain Research 221 (1), 237–245.

Lissek, S., Biggs, A.L., Rabin, S.J., Cornwell, B.R., Alvarez, R.P., Pine, D.S., et al., 2008. Generalization of conditioned fear-potentiated startle in humans: experimental validation and clinical relevance. Behaviour Research and Therapy 46 (5), 678–687.

Lissek, S., Rabin, S., Heller, R.E., Lukenbaugh, D., Geraci, M., Pine, D.S., et al., 2010. Overgeneralization of conditioned fear as a pathogenic marker of panic disorder. The American Journal of Psychiatry 167 (1), 47–55.

Lovero, K.L., Simmons, A.N., Aron, J.L., Paulus, M.P., 2009. Anterior insular cortex anticipates impending stimulus significance. NeuroImage 45 (3), 976–983.

Mechias, M.L., Etkin, A., Kalisch, R., 2010. A meta-analysis of instructed fear studies: implications for conscious appraisal of threat. NeuroImage 49 (2), 1760–1768.

Medford, N., Critchley, H.D., 2010. Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. Brain Structure and Function 214 (5–6), 535–549.

Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. Brain Structure and Function 214 (5–6), 655–667.

Milad, M.R., Rauch, S.L., Pitman, R.K., Quirk, G.J., 2006. Fear extinction in rats: implications for human brain imaging and anxiety disorders. Biological Psychology 73 (1), 61–71.

Milad, M.R., Wright, C.I., Orr, S.P., Pitman, R.K., Quirk, G.J., Rauch, S.L., 2007. Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. Biological Psychiatry 62 (5), 446–454.

Mineka, S., Oehlberg, K., 2008. The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. Acta Psychologica 127 (3), 567–580.

Nitschke, J.B., Sarinopoulos, I., Mackiewicz, K.L., Schaefer, H.S., Davidson, R.J., 2006. Functional neuroanatomy of aversion and its anticipation. NeuroImage 29 (1), 106–116.

O'Doherty, J.P., 2004. Reward representations and reward-related learning in the human brain: insights from neuroimaging. Current Opinion in Neurobiology 14 (6), 769–776 (Review).

Paulus, M.P., Stein, M.B., 2006. An insular view of anxiety. Biological Psychiatry 60 (4), 383–387.

Phelps, E.A., Delgado, M.R., Nearing, K.I., LeDoux, J.E., 2004. Extinction learning in humans: role of the amygdala and vmPFC. Neuron 43 (6), 897–905.

Phelps, E.A., O'Connor, K.J., Gatenby, J.C., Gore, J.C., Grillon, C., Davis, M., 2001. Activation of the left amygdala to a cognitive representation of fear. Nature Neuroscience 4 (4), 437–441.

Schiller, D., Levy, I., Niv, Y., LeDoux, J.E., Phelps, E.A., 2008. From fear to safety and back: reversal of fear in the human brain. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience 28 (45), 11517–11525.

Sehlmeyer, C., Schöning, S., Zwitserlood, P., Pfleiderer, B., Kircher, T., Arolt, V., et al., 2009. Human fear conditioning and extinction in neuroimaging: a systematic review. Plos One 4 (6), e5865–e5865.

Shaban, H., Humeau, Y., Herry, C., Cassasus, G., Shigemoto, R., Ciocchi, S., et al., 2006. Generalization of amygdala LTP and conditioned fear in the absence of presynaptic inhibition. Nature Neuroscience 9 (8), 1028–1035.

Simmons, A., Strigo, I., Matthews, S.C., Paulus, M.P., Stein, M.B., 2006. Anticipation of aversive visual stimuli is associated with increased insula activation in anxiety-prone subjects. Biological Psychiatry 60 (4), 402–409.

Vervliet, B., Kindt, M., Vansteenwegen, D., Hermans, D., 2010. Fear generalization in humans: impact of verbal instructions. Behaviour Research and Therapy 48 (1), 38–43.

Vink, M., Kahn, R.S., Raemaekers, M., van den Heuvel, M., Boersma, M., Ramsey, N.F., 2005. Function of striatum beyond inhibition and execution of motor responses. Human Brain Mapping 25 (3), 336–344.

Walter, B., Blecker, C., Kirsch, P., Sammer, G., Schienle, A., Stark, R., et al., 2003. MARINA: an easy to use tool for the creation of MAsks for Region of INterest Analyses. In: Paper presented at the Ninth International Conference on Functional Mapping of the Human Brain, NY, New York.