

**Neural correlates of conceptual-level fear generalization in posttraumatic stress disorder**

**Cite this article as:** Rajendra A. Morey, Courtney C. Haswell, Daniel Stjepanović, Joseph E. Dunsmoor and Kevin S. LaBar, Neural correlates of conceptual-level fear generalization in posttraumatic stress disorder, *Neuropsychopharmacology* doi:10.1038/s41386-020-0661-8

This Author Accepted Manuscript is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

Terms of use and reuse: academic research for non-commercial purposes, see here for full terms. <https://www.nature.com/authors/policies/license.html#AAMtermsV1>

Author accepted manuscript

Conceptual fear generalization in PTSD

## Neural correlates of conceptual-level fear generalization in posttraumatic stress disorder

Rajendra A. Morey, M.D.<sup>a,b,c,d</sup>, Courtney C. Haswell M.S.<sup>a,b</sup>, Daniel Stjepanović PhD<sup>b</sup>, †Mid-Atlantic MIRECC Workgroup<sup>a</sup>, Joseph E. Dunsmoor Ph.D.<sup>e</sup>, Kevin S. LaBar Ph.D.<sup>c,d</sup>

<sup>a</sup> Mid-Atlantic Mental Illness Research Education and Clinical Center, Durham NC

<sup>b</sup> Duke-UNC Brain Imaging and Analysis Center, Durham NC

<sup>c</sup> Department of Psychiatry, Duke University Medical Center, Durham NC

<sup>d</sup> Center for Cognitive Neuroscience, Duke University, Durham NC

<sup>e</sup> Department of Psychiatry, University of Texas at Austin, Austin TX

### Corresponding author:

Rajendra A. Morey, M.D.  
Duke-UNC Brain Imaging and Analysis Center  
40 Duke Medicine Circle, Room 414  
Durham, NC 27710 USA  
Phone: 1-919-384-8582 ext. 4006  
Facsimile: 1-919-416-5912  
E-mail: [rajendra.morey@duke.edu](mailto:rajendra.morey@duke.edu)

### Key Words:

posttraumatic stress disorder, functional MRI, fear generalization, semantic concepts, fear learning, amygdala

## Conceptual fear generalization in PTSD

**ABSTRACT**

Posttraumatic stress disorder (PTSD) may develop when mechanisms for making accurate distinctions about threat relevance have gone awry. Generalization across conceptually related objects has been hypothesized based on clinical observation in PTSD, but the neural mechanisms remain unexplored. Recent trauma-exposed military veterans (n=46) were grouped into PTSD (n=23) and non-PTSD (n=23). Participants learned to generalize fear across conceptual categories (animals or tools) of semantically-related items that were partially reinforced by shock during functional magnetic resonance imaging (fMRI). Conditioned fear learning was quantified by shock expectancy and skin conductance response (SCR). Relative to Veteran controls, PTSD subjects exhibited a stronger neural response associated with fear generalization to the reinforced object category in the striatum, anterior cingulate cortex, amygdala, occipitotemporal cortex, and insula ( $Z > 2.3$ ;  $p < 0.05$ ; whole-brain corrected). Based on SCR, both groups generalized the shock contingency to the reinforced conceptual category, but learning was not significantly different between groups. We found that PTSD was associated with an enhanced neural response in fronto-limbic, midline, and occipitotemporal regions to a learned representation of threat that is based on previously established conceptual knowledge of the relationship between basic-level exemplars within a semantic category. Behaviorally, veterans with PTSD were somewhat slower to differentiate threat and safety categories as compared to trauma-exposed Veteran controls due in part to an initial overgeneralized behavioral response to the safe category. These results have implications for understanding how fear spreads across semantically-related concepts in PTSD.

## Conceptual fear generalization in PTSD

**INTRODUCTION**

Fear learning can involve making inferences about threat associations, which are influenced by prior knowledge and experience [1]. Evolutionarily, such threat-related assessments facilitated survival such as predator-prey interactions, but have been co-opted along with other related brain circuits by modern-day psychological stressors [2]. Generalization of conditioned fear is a basic, cross-species associative-learning process whereby fear acquired to a conditioned stimulus (CS+), which reliably predicts an aversive unconditioned stimulus (US), transfers to safe stimuli that resemble the CS+ in some way. Although such threat processing is typically adaptive, posttraumatic stress disorder (PTSD) may develop when mechanisms for making accurate distinctions about threat relevance have gone awry.

In PTSD, experience with objects and events associated with a traumatic event may lead to an overgeneralized spread of fear. Neural circuits that process conditioned fears in PTSD, relative to trauma-exposed Veteran controls, overgeneralize to related stimuli based on their overlap with perceptual attributes of the CS+, such as the emotional intensity expressed on a fearful face[3] or the size of a geometric shape[4]. Generalizing fear along a gradient of intensity in facial expression, for instance, yielded greater neural responses in PTSD than trauma-exposed Veteran controls in several brain areas, including the thalamus, locus coeruleus, insula, fusiform gyrus, and primary visual cortex, as well as enhanced functional connectivity between the amygdala and calcarine cortex[3]. When generalizing fear to images of circles that increase parametrically in size relative to the fear-conditioned stimulus, PTSD patients showed a stronger response than Veteran controls in the caudate, anterior insula, left ventral hippocampus, dorsomedial and dorsolateral prefrontal cortex (PFC)[4].

In clinical presentation, patients with PTSD may report symptom triggers to stimuli or situations that share a conceptual association, but little or no perceptual overlap with details of the trauma. For instance, symptom triggers in military PTSD might include idiosyncratic reminders of the event including ethnic food from the region they were deployed, anniversary dates, or military holidays, and even the topic of war in general. Concept-based symptom provocation provides a seemingly vast source of potential triggers that could lead to overgeneralization of

## Conceptual fear generalization in PTSD

threat association in PTSD that extend beyond cues that bear physical resemblance to trauma details[5]. Arguably, the spread of fear that is based on semantic relationships and conceptual knowledge may apply more directly to aspects of real-world generalization mechanisms in PTSD than low-level perceptual attributes. While some prior work has examined perceptual generalization in PTSD[3, 4], and conceptual generalization in non-clinical individuals[6, 7], the neural substrates mediating conceptually-based fear generalization in PTSD remain unknown. Modeling this aspect of threat generalization in PTSD is important because in some cases, learning trauma associations may take place with exposure to only a single traumatic event that later generalizes to other stimuli along abstract conceptual factors.

The present study sought to determine the neural mechanisms by which fear learning generalizes through category-level representations of object concepts in PTSD. Our behavioral protocol[8] exploits established organizational separations in cortical representations of animate and inanimate objects. One group of participants learned with experience, through partial reinforcement, that images of animals predicted an aversive electrical shock but that tools were safe, whereas a separate group learned the opposite contingency. This innovative paradigm presents trial-unique conditioned stimuli to achieve fear learning based on participants' prior knowledge about the conceptual relationships among objects (animals from tools). Each participant viewed the same images chosen to vary widely in perceptual attributes, but different reinforcement contingencies determined which category attained threat-relevance.

Our prior work in non-trauma-exposed young adults showed that conceptual fear generalization engaged distributed brain regions (including amygdala, insula, midbrain, cingulate gyrus, portions of the frontal lobe, and posterior category-selective cortex), which showed greater activity to exemplars from the aversively-reinforced object category (CS+) than the unreinforced category (CS-)[9]. Moreover, the multivariate activity patterns of the items from the reinforced category were more correlated to each other (relative to those from the unreinforced category) in the amygdala and occipitotemporal cortex, suggesting an aversive learning-induced change in their representation. We showed that the object typicality of

## Conceptual fear generalization in PTSD

exemplars influenced left hippocampal activation whereby early trials were modulated by typicality more than late trials. Lastly, there was a temporally graded pattern of task-based functional connectivity between the hippocampus and amygdala that was stronger during early training trials but then dissipated in late trials[9].

Importantly, these neural circuits have been broadly implicated in functional studies of fear conditioning, symptom provocation, and emotion regulation in PTSD[10-13]. These studies generally report a hyper-reactive salience network in PTSD, including the amygdala, anterior cingulate cortex, and insula, coupled with poor contextual and cognitive regulation of this circuit via the hippocampus and prefrontal cortex (PFC), respectively. Although it remains debated, structural abnormalities in some of these fronto-limbic regions may confer predispositions to develop PTSD following a trauma, in addition to trauma-induced pathological changes[14, 15]. Symptom provocation and fear conditioning studies further implicate interactions between these regions and posterior sectors of the brain that represent sensory features of environmental stimuli and more complex face, scene, and object representations[3, 16, 17].

Thus, we predicted that learning to generalize fears across object category exemplars in PTSD, relative to trauma-exposed Veteran controls, would more extensively modify representations along the occipital temporal cortex, related association cortices, as well as areas implicated in simpler forms of fear conditioning, such as the amygdala. We further predicted that generalization of aversive learning to object categories would more strongly modulate functional connectivity between the amygdala and occipitotemporal cortex in the PTSD compared to trauma-exposed Veteran controls.

## **MATERIALS AND METHODS**

### *Participants*

The participants (N=46) were recruited from February 2014 through August 2016, from a repository of over 3,800 US military veterans[18]. All the participants served since 11-September 2001 (9/11), and most were deployed to the military conflicts in Iraq and/or

## Conceptual fear generalization in PTSD

Afghanistan. All the participants provided written informed consent to participate in procedures approved by the Institutional Review Boards at Duke University Medical Center and the Durham VA Medical Center. Participants underwent screening for inclusion in the study. Exclusions included Axis I diagnosis (other than MDD or PTSD), contraindication to MRI, moderate/severe traumatic brain injury, past or current alcohol dependence, past or current substance dependence, current alcohol abuse, current substance abuse, neurological disorders, and age over 65 years. However, past alcohol or substance abuse was permissible. Abuse criteria met in the past 12 months constitutes current abuse. We administered the Clinician-Administered PTSD Scale[19] based on DSM-5, Alcohol Use Disorders Test (AUDIT)[20], Combat Exposure Scale (CES), Childhood Trauma Questionnaire (CTQ)[21], Trauma Life Events Questionnaire (TLEQ)[22], Beck Depression Inventory (BDI-II)[23], psychotropic medication use based on self-report, and Drug Abuse Screening Test (DAST)[24]. Participants' demographic and clinical features are reported by diagnostic group in **Table 1**. Greater levels of combat exposure, lifetime trauma exposure, depressive symptoms, and serotonergic medication use, in the PTSD group compared to the control group are addressed in follow-up analyses (**Supplemental Tables 1-5**).

### *Stimuli*

Images of the objects used in the study were obtained from the website [www.lifeonwhite.com](http://www.lifeonwhite.com) and from publicly available resources on the internet. The stimuli were the same as in our previous publication that contains a complete list of stimuli and healthy participants' category typicality ratings[8]. We avoided the use of highly threat-relevant images, such as weapons associated with combat trauma and animal exemplars associated with Specific Phobias (snakes/spiders), to mitigate the potential arousal bias evoked by these objects[25]. The presentation of stimuli was controlled using Presentation Software (Neurobehavioral Systems, Albany, CA, USA).

### *Procedure Overview*

The scanning session contained 4 phases: (1) a functional localizer of animal and tool exemplars, (2) pre-conditioning resting state, (3) fear generalization task based on conceptual

## Conceptual fear generalization in PTSD

categories of animals and tools, and (4) post-conditioning resting state. Twenty-four hours later, subjects returned for a recognition memory test and provided typicality ratings of the category exemplars.

### *Animal-Tool Localizer*

The animal, tool, and phase-scrambled localizer contained 12 alternating blocks (separated by 11-s fixation cross) with 10 images/block presented for 750-ms (separated by a 250-ms blank screen). This phase was used to identify category-specific cortex independently of the fear generalization task itself. Localizer images were not reused for fear learning. Analysis of the category localizer included separate regressors for animals, tools, and phase-scrambled blocks. Single-subject contrast images for each event versus baseline were used for group-level random-effects contrasts of animals versus tools and tools versus animals. Activations were masked by (animals + tools) > phase-scrambled images. The fear learning response was interrogated within ROI spheres defined by category localizer. Subjects assigned to animal CS+ were analyzed separately from subjects assigned to tool CS+.

### *Fear Conditioning Task*

Event-related fear conditioning during fMRI was achieved by presenting one category (e.g. animals) designated as CS+ (50% of exemplars co-terminated with the US) and the non-reinforced category (e.g. tools) served as the CS- (**Figure 1**). Importantly, every trial contained a unique basic-level exemplar that was never repeated. Subjects were not instructed on the CS-US contingencies and were not informed that images would be presented only once in the MRI session. A 6-ms electrical shock applied to the right-wrist served as the US. Shock administration was controlled using the STM-100 and STM-200 modules connected to the MP-150 BIOPAC system (BIOPAC systems, Goleta, CA, USA). The level of the shock was calibrated prior to the start of the experiment using an ascending staircase procedure to reach a level deemed “annoying but not painful” by the subject [26].

Each run contained 20 trials (10 animals, 10 tools) presented in a pseudorandomized order, with no more than two objects from the same category occurring in a row. We used four different stimulus presentation orders to counterbalance presentation of animal and tool



## Conceptual fear generalization in PTSD

exemplars across participants (**Figure 1**). Each trial was 6-s in duration, during which time subjects rated expectancy for US delivery on a scale anchored between 0 (certainty the shock will not occur) and 100 (certainty the shock will occur) using an MRI-compatible joystick. Subjects were instructed that the final location of the rating bar at trial offset would be recorded as their response. A fixation cross was displayed for 10–12 s duration (average=11 s) following the offset of each trial. One category (e.g. animals) was designated the CS+, and 50% of exemplars from this category co-terminated with the US. The other object category (e.g. tools) served as the CS-, and none of its exemplars were reinforced with a US. Category assignment was counterbalanced between PTSD and Control groups, such that approximately half of the participants (N=24; 12 PTSD, 12 Veteran controls) were conditioned to animals whereas the remaining subjects (N=22; 11 PTSD, 11 Veteran controls) were conditioned to tools. The selection of CS+ items that were paired with shock was random and counterbalanced between subjects.

### *Recognition Memory Test and Typicality Ratings*

Twenty-four hours later, subjects returned to a different laboratory setting outside the MRI facility where they conducted a follow-up recognition test of memory for the CS+ and CS- items and rated each animal and tool picture for its typicality to the superordinate category. Category typicality ratings were made for each item on a 10-point scale (0=highly atypical, 10=highly typical). Participants provided memory ratings for all of the exemplars that were encountered as CS+ and CS- stimuli during fear learning, as well as a matched set of new exemplar foils belonging to the same superordinate categories (new tool/new animal), by indicating *new* or *old* and the confidence of this response on a continuous visual scale. In addition, participants rated whether the stimulus was followed by a shock to assess source memory. One subject was missing source memory ratings.

### *Skin Conductance Response Methods and Analysis*

Psychophysiological recording electrodes were placed on palmar surface of the left hand, and the data were analyzed as previously described [9] using procedures in the **Supplemental Materials**. We collected these data as an additional index of fear conditioning. We obtained

## Conceptual fear generalization in PTSD

usable data from 36 participants (19 Veteran controls, 17 PTSD) who elicited non-zero shock responses. Ten additional participants were excluded from analysis: six participants had zero shock responses, and four participants lacked measurable data because of technical issues with the equipment, possibly due to excessive noise associated with the scanner environment. The CS+ and CS- responses were baseline corrected, normalized to the shock response within subjects to account for possible group differences in overall arousal, and square-root transformed to reduce skewness.

### *fMRI Parameters*

High-resolution T1-weighted structural MRI scans were acquired with a 3D-FSPGR sequence on a General Electric 3T MR750 scanner (1-mm<sup>3</sup> voxels, TR/TE/flip angle=7.84 ms/2.98 ms/12°, FOV=256 mm<sup>2</sup>, 166 slices, and 1 excitation). Functional images were acquired parallel to the anterior commissure-posterior commissure line using a SENSE™ spiral-in sequence (matrix=64<sup>2</sup>; FOV=256 mm<sup>2</sup>; 34 interleaved slices with no gap; 3.75 mm<sup>3</sup> voxels; TR/TE/flip angle=2000 ms/27 ms/60°). Functional imaging data were preprocessed using FSL version 5.0.9[FMRIB Software Library; Oxford University, UK][27] including brain extraction, motion correction, slice-timing correction, affine registration, 5-mm FWHM smoothing, and time series filtering < 0.01 Hz to minimize low-frequency drift.

### *fMRI Univariate and Task-Based Functional Connectivity Analysis*

Only fMRI data from the task period is presented; resting-state data were not analyzed. The main analysis for interrogating the generalization response used whole-brain voxelwise analysis that contrasted responses to all exemplars from the reinforced category (CS+) versus those to the unreinforced category (CS-) between the PTSD and trauma-exposed control groups. The contrasts used whole-brain voxelwise regression with the FSL FMRI Expert Analysis Tool (FEAT). Shock delivery was modeled as a regressor of no interest (50% of CS+ trials) along with the customary six parameters specifying head motion. Statistical contrasts for CS+ > CS- trials (and the reverse contrast of CS- > CS+ trials) were corrected for multiple comparisons, which was performed on whole-brain data with the FSL two-step cluster-threshold procedure to identify clusters of contiguous voxels with  $Z > 2.3$  and then test the significance of resulting

## Conceptual fear generalization in PTSD

cluster(s) at a corrected threshold of  $p < 0.05$  according to Gaussian random field theory[28, 29]. The results of the first-level contrast (CS+ > CS-) were subjected to between-group contrasts for PTSD > Control and Control > PTSD. For our second-level analysis, we used FMRIB's Local Analysis of Mixed Effects (FLAME 1 + 2). Eklund and colleagues reported revealed the false positive rate of FLAME1 is similar to using a cluster-forming p-value threshold of 0.001[29]. All voxels from FLAME1, which are close to threshold according to the selected contrasts and thresholds are processed with FLAME 2. This process involves implicit estimation of the mixed-effects variance using the Metropolis-Hastings Markov Chain Monte Carlo simulation. In response to peer-review, we compared the results of *randomise* and TFCE to FLAME 1+2. *Randomise* and TFCE yielded more voxels than FLAME 1+2, which was consistent with a recent report by Han and colleagues, who compared TFCE to AFNI ClustSim (method most similar to FLAME1 + 2) [30].

Task-modulated functional connectivity measures were generated for each exemplar in the reinforced category using an interaction term between trial-specific onset times as the *task regressor* and the activation time course of the seed region as the *physiological regressor*. We adopted the generalized psychophysiological interaction (gPPI) approach, which provides improved model fit compared with standard PPI[31]. Mean and maximum connectivity strengths within each target region from the single subject level were input into a model contrasting groups (control, PTSD).

*Representational Similarity Analysis (RSA)*

RSA is one type of multi-voxel pattern analysis approach to fMRI in which a similarity structure is derived and compared across spatially-distributed patterns of fMRI activity. Visual responses were compared between groups by considering the similarity structure of neural responses of exemplars from each object category[32]. See Deen et al.[33] and **Supplemental Material** for analysis details. Briefly, we constructed a general linear model with each regressor containing a single trial. Voxels identified with the animal localizer were omitted for subjects who learned with an animal CS+ for the purpose of generating an ROI to test generalization to animal exemplars. Similarly, voxels identified with the tool localizer were omitted for subjects who

## Conceptual fear generalization in PTSD

learned with a tool CS+ for the purpose of generating an ROI to test fear generalization of tool exemplars. Voxel patterns were computed in each subject for the entire occipitotemporal cortex, which is where distributed representations of these object categories reside, excluding voxels that showed a main effect of the CS+ category in the localizer task. This exclusive mask was created to avoid biasing the RSA results with voxels that reflect a main univariate category effect. Mean activation from each ROI provided trial-wise ( $80 \times 80$ ) representational similarity matrices (RSM) of correlations between all the possible pairwise combinations of the animal and tool exemplars for each subject. Kendall-Tau rank-order correlations were used to determine whether orderings were more consistent within diagnostic groups than between-groups separately for animal-animal, tool-tool, and animal-tool exemplar correlations. Differences between PTSD-PTSD, Control-Control, and PTSD-Control rank-order correlations and zero were calculated with exact permutation testing.

## RESULTS

### *Clinical and demographic information*

Detailed clinical and demographic information is reported by group in **Table 1**.

### *Skin conductance responses (SCRs)*

A 2x2 (Group X CS type) ANOVA was computed on the SCR responses in 36 (78.3%) participants (19 Veteran controls, 17 PTSD) with usable data (see *Methods*). A significant main effect (**Figure 2A**) of CS type [ $F(1,34)=14.701$ ,  $p=0.001$ , partial- $\eta^2=0.302$ ] showed higher magnitude responses to the CS+ (mean=0.380; SD =0.248) relative to the CS- (mean=0.282; SD=0.222). The interaction of group and CS Type was non-significant [ $F(1,34)=0.001$ ,  $p=0.991$ , partial- $\eta^2=0.001$ ].

### *Shock expectancy*

A 2x2 (Group X CS Type) mixed ANOVA was computed on the shock expectancy ratings averaged over all four trial runs (**Figure 2C**). Greater shock expectancy ratings to the CS+, relative to the CS-, indicated successful fear conditioning [ $F(1,44)=61.233$ ,  $p < 0.001$ , partial- $\eta^2=0.582$ ]. The main effect of Group ( $p=0.415$ ), and interaction of Group and CS Type ( $p=0.103$ )

## Conceptual fear generalization in PTSD

were non-significant (**Figure 2B**). Importantly, the shock expectancy results were not confounded by significant group differences in shock voltage levels set by the participants themselves (PTSD subjects: mean=51.204, SE=4.173; Veteran controls: mean=43.936, SE=3.329;  $t(44)=1.383$ ,  $p=0.17$ ).

Because most of the learning took place within Run 1, we further interrogated the temporal pattern of responding across early (first-five trials) and late (last-five trials) phases of Run 1 to acquire a fine-grained picture of initial learning rates. Linear trend analysis of Run 1 showed an increasing linear trend to the CS+ shock expectancy ratings in the PTSD group ( $r=0.868$ ,  $p=0.0006$ ) but not the Control group ( $r=0.152$ ,  $p=0.675$ ). By contrast, both groups showed decreasing linear trends in shock expectancy for the CS- (Control:  $r=0.768$ ,  $p=0.001$ ; PTSD:  $r=0.843$ ,  $p=0.002$ ). The data were also analyzed in a 2x2x2 (Group x CS Type x Phase) mixed ANOVA (**Figure 2B**). These results showed that learning progressed from early-phase trials to late-phase trials [ $F(1,44)=8.467$ ,  $p=0.006$ , partial- $\eta^2=0.161$ ] and interacted at the non-significant trend-level with Group [ $F(1,44)=3.776$ ,  $p=0.058$ , partial- $\eta^2=0.079$ ]. Expectancy ratings differed between groups in early [ $t(44)=2.770$ ,  $p=0.016$  partial- $\eta^2=0.149$ ; Bonferroni corrected], but not late [ $t(44)=0.092$ ,  $p>0.9$  partial- $\eta^2<0.001$ ; Bonferroni corrected], phases of Run 1. On average, PTSD patients were six trials slower to initially differentiate the CS+ from CS- based on shock expectancy ratings out of a total of 40 trials per trial-type (80 total trials).

### *Recognition Memory*

Recognition memory for the CS items was measured by  $d'$  and analyzed with a 2x2 (Group X CS type) mixed ANOVA. A significant influence of reinforcement during fear learning was found on subsequent memory discriminability [ $F(1,44)=7.956$ ,  $p=.007$ , partial- $\eta^2=.156$ ], with greater recognition memory for CS+ (mean=1.444, SE=0.120) than CS- items (mean=1.264, SE=0.097). The specific object category assigned as the CS+ did not influence results [ $F(1,44)=2.377$ ,  $p=0.130$ , partial- $\eta^2=0.051$ ]. Diagnosis did not influence responding on this measure as either a main effect or in interaction with any of the other factors ( $p$ -values > 0.05).

### *Source Memory*

## Conceptual fear generalization in PTSD

Source memory for which items were paired with shock was analyzed with a 2x2x2 (Group X CS type X memoranda type) mixed ANOVA. A significant main effect of CS type (CS+ vs CS-; mean=65.826, SE=2.085 vs. 90.306, SE=2.167) [ $F(1,43)=102.572, p < 0.001, \text{partial-}\eta^2=.705$ ], a significant main effect of memoranda type (old vs new; mean=71.688, SE=1.799 vs mean=84.444, SE=2.033) [ $F(1,43)=65.354, p < 0.001, \text{partial-}\eta^2=0.603$ ], and a significant interaction between these variables [ $F(1,43)=65.797, p < 0.001, \text{partial-}\eta^2=0.605$ ]. Examining this interaction, we found that the CS+ items were significantly more likely to be rated as paired with shock than the CS- items [ $t(43)=10.823, p < 0.001, \text{partial-}\eta^2=.731$ ].

Additionally, when examining only items presented as CS+ stimuli during fear learning (2x2 mixed ANOVA; Reinforcement x Group), we observed significantly higher endorsement of shock association for the items that were actually paired with a shock than those that were not, (paired with shock vs not paired with shock; mean=46.093, SE=4.014 vs mean=62.625, SE=3.737), [ $F(1,43)=35.859, p < 0.001, \text{partial-}\eta^2=0.449$ ]. There was no main effect of diagnosis [ $F(1,43)=0.455, p=0.504, \text{partial-}\eta^2=0.010$ ] and no interaction of shock pairing with diagnosis [ $F(1,43)=0.149, p=0.702, \text{partial-}\eta^2=0.003$ ].

### *Typicality Results*

Subjects rated each exemplar (**Supplemental Figure 1**) based on how representative the exemplar was of the superordinate category (animals or tools). There was a significant main effect of typicality [ $F(1,43)=10.615, p=0.002, \text{partial } \eta^2=0.19$ ]. Ratings for typicality were not influenced by shock pairing or diagnosis [ $F(1,43)=0.731, p=0.397, \text{partial } \eta^2=0.017$ ].

### *fMRI results from the animal-tool functional localizer*

Whole-brain animal-selective activation in lateral FFG and inferior occipital gyrus, extending into the posterior superior temporal sulcus, and tool-selective activation in medial FFG/parahippocampal gyrus and middle occipital gyrus (**Supplemental Figure 3** and **Supplemental Material**) formed an occipitotemporal ROI.

## Conceptual fear generalization in PTSD

There were no between-group differences among subjects assigned to animal CS+ in either the animal- or tool-selective localizer-defined ROIs [ $F(1,22)=0.012$ ;  $p=0.912$ ] and no interaction of group and ROI [ $F(1,22)=0.735$ ;  $p=0.401$ ]. Similarly, there were no between-group differences among subjects assigned to tool CS+ in either the animal-selective or tool-selective localizer-defined ROIs [ $F(1,22)=0.203$ ;  $p=0.657$ ] and no interaction of group and ROI [ $F(1,22)=0.458$ ;  $p=0.506$ ]. Thus, PTSD was not associated with an altered response to broad stimulus categories with learned fear associations in visual perceptual brain regions.

### *Univariate fMRI results from the conditioning task*

Activation for the effect of conditioned learning (CS+ > CS-) irrespective of CS identity (animal CS+ or tool CS+) was significantly greater in the PTSD than the trauma-exposed control group (**Figure 3** and **Table 2**) in the amygdala, ventral striatum particularly in the putamen, insula, anterior cingulate cortex (ACC), subgenual ACC (sgACC), and visual association areas including lingual gyrus and precuneus. All results were significant after whole-brain correction (cluster threshold  $Z > 2.3$ ;  $p < 0.05$  corrected). The use of unique (non-repeated) exemplars for each CS+ presentation demonstrates that the neural response to the reinforced object category was the result of a generalized threat association to the entire conceptual category. Greater activation in the PTSD than Veteran controls to the CS+ > CS- contrast, when each exemplar was modulated by its typicality rating, recapitulated the main fMRI results.

### *Effects of Trauma Exposure and Clinical Variables on Neural Activity*

The pre-existing group differences in combat trauma, lifetime trauma, serotonergic medication use, and depression symptoms between the PTSD and Control groups (**Table 1**) meant that group differences in fMRI activation might be attributable to either PTSD or to one or more of these variables. Their inclusion as covariates offered inconclusive information because the severity of trauma exposure and diagnostic grouping are collinear and therefore difficult to interpret, according to Miller & Chapman[34]. The role of trauma/clinical variables on the univariate generalization response was further scrutinized vis-à-vis PTSD by assessing the strength of correlations between each covariate and activation separately in PTSD and Control

## Conceptual fear generalization in PTSD

groups in each of five key brain regions. Correlation of activation with lifetime trauma exposure (**Supplemental Table 1**), combat exposure (**Supplemental Table 2**), serotonergic medication (**Supplemental Table 3**), and depression symptoms (**Supplemental Table 4**), had non-significant interactions with diagnosis as assessed by the Fisher  $r$ -to- $z$  test[35]. Re-analysis of the five brain regions with a covariate, generally revealed significant between group differences in activation, as did the results without this covariate (**Supplemental Table 5**). Two exceptions were role of serotonergic medication on the insula and paracingulate gyrus, which produced a non-significant effect of diagnosis and a trend-level diagnosis effect on the amygdala. Finally, removing all subjects who were prescribed serotonergic medication from the analysis, yielded significant between-group differences in vmPFC ( $p=0.002$ ), putamen ( $p=0.01$ ), and trend level differences in the paracingulate gyrus ( $p=0.095$ ) and insula ( $p=0.101$ ).

### *Functional Connectivity Results from the Conditioning Task*

There was a significant temporal influence (across four runs) on task-related amygdala-hippocampal functional connectivity modulated by exemplar typicality [ $F(1,44)=6.954$ ,  $p=0.012$ ], but no main effect of group [ $F(1,44)=0.350$ ,  $p=0.557$ ] or no interaction of group and time [ $F(1,44)=0.025$ ,  $p=0.874$ ]. Similarly, there was no interaction of diagnostic group and time for tool CS+ [ $F(1,21)=0.184$ ,  $p=0.906$ ], or animal CS+ [ $F(1,21)=0.076$ ,  $p=0.972$ ] exemplars. Functional connectivity between amygdala seed and target regions obtained from the GLM results revealed stronger category generalization (CS+ > CS-) in the PTSD group than trauma-exposed control group (**Supplemental Figure 2** and **Supplemental Material**) for left hippocampus [ $t(44)=2.552$ ,  $p=0.018$ ], and trend-level for sgACC [ $t(44)=1.835$ ,  $p=0.080$ ].

### *Representational Similarity in Occipitotemporal Cortex*

Visual responses were compared between PTSD and trauma-exposed control groups by considering the similarity structure of spatial patterns of neural response to category exemplars for both the animal and tool CS+ categories[32, 33]. The hypothesized higher dissimilarity in Control-PTSD as compared to PTSD-PTSD and Control-Control was not demonstrated for animal



Conceptual fear generalization in PTSD

CS+ (**Supplemental Figure 4a**) and tool CS+ exemplars (**Supplemental Figure 4b**) in the relevant perceptual cortex. Detailed results in **Supplemental Material**.

## DISCUSSION

The present study extends knowledge about fear learning mechanisms in PTSD by demonstrating how acquired threat associations generalize across exemplars that belong to a conceptual category. Unlike traditional fear conditioning paradigms that examine repeated reinforcement to a single item, here a novel stimulus was introduced on each trial. In order to successfully predict reinforcement contingencies, participants had to generalize threat associations across the subset of reinforced exemplars to the superordinate category level. Modeling this aspect of threat generalization in PTSD is important because in some cases learning trauma associations may take place with exposure to only a single traumatic event that later generalizes to other stimuli along abstract conceptual features.

Patients with PTSD had a stronger generalization of the neural response than trauma-exposed control veterans in a fronto-limbic network including the amygdala, insula, striatum, ventromedial PFC, and rostral ACC. Although this network has generally been implicated in fear learning in PTSD[3, 4], here we show that its function extends to the generalization of fear to conceptual level object categories. Activity in these regions was especially tuned to the more typical exemplars of the reinforced category in PTSD. In addition, we reveal enhanced functional coupling of the amygdala with the hippocampus to support category-based fear generalization in PTSD. Contrary to our hypothesis, univariate analysis showed no PTSD-associated differences in activation of the category-specific perceptual cortex during generalized fear learning of semantic categories. The multivariate RSA conducted on ventral occipitotemporal cortex also showed negative results associated with PTSD that were consistent with the univariate analysis. Nevertheless, whole brain analysis successfully showed that PTSD patients more strongly engaged specific posterior areas to generalize their fear learning in the left temporo-occipital fusiform cortex, left lateral occipital cortex (superior division), cuneus, and right intracalcarine cortex (**Table 2**). Thus, generalization learning functions, which rely on visual and other constructive processes in these posterior areas, are

## Conceptual fear generalization in PTSD

biased by PTSD only in these circumscribed regions. This PTSD bias extends to visual association areas such as precuneus and lingual gyrus, which have established functions in the processing and early interpretation of complex visual objects and scenes[36-38]. The present results extend our earlier work demonstrating greater engagement in patients with PTSD of the right fusiform gyrus and right intracalcarine cortex in generalizing fear from shock-reinforced fearful faces[3] to conceptual forms of generalization across object exemplars.

PTSD and trauma-exposed Veterans showed similar enhancement shock expectancy ratings to the reinforced category, although the PTSD group was slower in their ability to behaviorally learn the category discrimination. In contrast to prior work on fear generalization in PTSD, where a single CS+ and single CS- are presented repeatedly during learning, here we present 40 unique stimuli of each trial type that were never repeated. Since shock expectancy ratings reflect, in part, declarative knowledge of the reinforcement parameters, this group difference may simply reflect impaired declarative aspects of early learning in PTSD on this complex paradigm. This is not an *under-generalization* effect in PTSD because inspection of the data by trial-type shows that PTSD subjects gave higher shock expectancy ratings to the safe CS- compared to Veteran controls on early trials, which partly accounts for a lack of early differential learning. Indeed, this observation could be leveraged as additional evidence for initial *overgeneralization* of shock expectancy to an innocuous cue (although such evidence should not be over-interpreted). Accordingly, linear trend analysis confirmed a strong linear pattern of threat learning only in the PTSD group, whereas both groups showed strong linear patterns for safety learning.

Both groups exhibited a similar boost in 24-hr delayed recognition memory for the exemplars from the reinforced category, and they both had intact source memory for the specific items that were followed by shock reinforcement. In combination with the shock expectancy ratings, these behavioral results help to rule out behavioral performance differences between groups as a potential confound when interpreting fMRI differences between groups. However, our results raise questions about the interpretation of group differences in fMRI activation that are not accompanied by more widespread behavioral differences. One possibility is that the magnitude

## Conceptual fear generalization in PTSD

of fMRI differences we observed in the PTSD group are too subtle to manifest in physiological learning indices and memory biases. A second possibility is that distributed brain processes are able to compensate for localized differences in activation to maintain a similar level of behavioral performance. Both hypotheses will require further investigation to confirm or disprove.

The stronger generalization of neural fear response in PTSD to broad conceptual categories such as animals and tools is fundamentally different than previous reports of generalization to perceptual gradients such as affect intensity of fearful faces[3] or the size of geometric shapes[4]. Generalization to semantically-defined categories requires a higher level of abstraction that is based on binding the function and taxonomy of specific exemplars to an object category in contrast to the more basic generalization processes from these prior studies, which involve perceptual level binding of geometric dimensions and intensity of fearful facial expressions to threats. This level of abstraction during fear learning involves the recruitment of cortical association areas that are not required when learning to generalize fear associations to concrete perceptual attributes, as exemplified by broader engagement of posterior and midline cortical regions in the present study as compared to Morey et al.[3] and Kaczurkin et al.[4]. Although there are many possible interpretations of this activity, we speculate that our task placed higher demands in the PTSD veterans on relating threats to existing knowledge structures, which may recruit the vmPFC and related structures as part of a broader semantic/schema network[39, 40]. These demands differ from those in perceptual generalization studies in PTSD, for which the vmPFC exhibits an activation pattern that signals the generalization of the safety cue (CS-)[3, 4].

Generalization across abstract conceptual categories is a more naturalistic model for humans and bears greater clinical relevance to the environmental cues that trigger re-experiencing symptoms of PTSD, which typically possess only a vague connection to the original trauma material[5, 41]. Indeed, generalization based purely on perceptual gradients of intensity or magnitude have limited clinical relevance in which generalization phenomena may span

## Conceptual fear generalization in PTSD

emotional, spatial, semantic, perceptual, and cognitive constructs[6, 7, 42]. The highly developed association cortices in humans, which represent an evolutionary rerouting of cortical information flow, may facilitate generalization across a broad range of stimuli that have conceptual similarities even when they lack a common set of perceptual features[43]. Thus, the ability to process complex conceptual information at an abstract level generally helps humans with more judicious threat assessment than other animals, but comes at a penalty when it elicits a rather broad fear generalization response.

Our understanding of the precise role of the amygdala in fear generalization is evolving. A delicate balance within subregions of the central nucleus of the amygdala appears to regulate fear generalization[44]. Neuronal activity in the medial portion of the central nucleus (the primary output pathway for initiating fear behaviors) is tonically inhibited by the lateral portion of the central nucleus. Accordingly, changes in tonic activity can tip the balance from expression of fear for only the CS+, toward generalization and expression of fear in response to the CS- or other related stimuli. In healthy humans, the amygdala plays a key role in the generalization of fear along perceptual gradients[9, 45] in part due to its functional connectivity to the hippocampus and higher-level sensory cortices. In PTSD, the amygdala has been implicated in perceptual forms of fear generalization across faces expressing different intensities of fear[3] and across circles varying in size[4]. However, these prior studies showed somewhat different patterns of amygdala connectivity in PTSD. In Morey et al.[3], amygdala connectivity with the calcarine (and thalamus at a trend-level) was greater in PTSD than Veteran controls to generalized faces expressing more fear than the CS+, and amygdala connectivity to the vmPFC was greater in PTSD to generalized faces expressing less fear than the CS+. In Kaczurkin et al.[9], amygdala connectivity to the hippocampus was greater in PTSD than Veteran controls to geometric stimuli resembling threats; other connectivity patterns were found in an activation seed that encompassed both the left amygdala and ventral hippocampus, so it is unclear whether the remaining connectivity patterns are in the amygdala proper and are not discussed further here. Our amygdala-hippocampal connectivity results are consistent with Kaczurkin et al.[4], but may reflect different processes. Our mechanistic study in healthy subjects[9] showed

## Conceptual fear generalization in PTSD

that hippocampal activity signaled the more typical category exemplars to provide a foundation of fear learning at the superordinate category level. We hypothesize that the hippocampus is providing information to the amygdala through enhanced functional coupling to the reinforced category. These processes are not engaged during perceptual forms of generalization, but the hippocampus may provide common functions for generalization across studies.

Computational modeling could be applied to yield valuable knowledge and insights about how stimulus value, prediction errors, learning rates, and attentional or associability parameters from different exemplar categories contribute to the clinical phenomenology of PTSD. Several design features of our experimental paradigm, however, limit the kinds of modeling that could be applied and inferences that could be drawn from such an endeavor. First, we intermixed reinforced and generalized stimuli during learning. This design feature precludes modeling learning parameters to a CS+ and then using those parameters to predict generalization during a subsequent test. In addition, we used trial-unique stimuli, which precludes applying classic associative learning models that track parameters based on a stimulus' prior reinforcement history. Finally, unlike prior fear generalization studies that manipulate a single perceptual feature, here the exemplars differ across numerous perceptual dimensions (color, shape, size) as well as conceptual knowledge (threat relevance, typicality) that would have to be jointly considered in terms of their relationship to learning parameters, and these relationships might differ across the two categories tested (animals, tools). Nonetheless, future studies should be designed specifically to test conceptual similarity gradients using Bayesian hierarchical modeling or other approaches that have been applied to mechanistically understand conditioned fear response generalization across a perceptual gradient [46].

Our results have implications for developing new exposure-based treatments to target overgeneralization of fear to conceptual categories[6, 7, 42]. For instance, treatment for spider phobia with extinction therapy successfully generalized to help patients with untreated cockroach phobia to reduce subjective fear, approach distance, and psychophysiological arousal[47]. These treatments might focus on reducing fear reactivity not just to the specific triggers, but also semantically-related stimuli. For example, *in vivo* exposure exercises could use

## Conceptual fear generalization in PTSD

stimulus categories partitioned according to semantic or other abstract feature dimensions and a hierarchy of feared stimuli, with exposures to stimuli resembling the feared stimulus added at each stage. In addition, patients could undergo discrimination training, whereby patients learn to differentiate trauma cues indicative of genuine danger from benign cues with inconsequential resemblance to the trauma.

### *Limitations*

The reinforcement of a single category was accomplished by administering an electrical shock on 50% of trials, which raises the concern that activation was confounded by exposure to shock. However, this is unlikely because the stimulus onset preceded shock onset by 6-s, which is sufficiently long for separating the hemodynamic response between these two stimuli, and because the effect of shock was explicitly controlled in the data analysis. Nonetheless, there remains a possibility that unaccounted variance contributed to the signal contained in the CS+ > CS- fMRI contrasts. The preexisting group difference in lifetime trauma exposure, combat exposure, serotonergic medication, and depression symptoms between the PTSD and Control groups meant that group differences in outcomes could be partially attributable to either PTSD or other clinical variables such as trauma exposure; however, these variables lacked significant effects when explicitly tested for their contributions. While the present results focused on animals and tools as conceptual categories due to their known representation in the human brain, future work should examine other categories that have greater ecological validity for PTSD. Finally, future work should examine generalization at different levels of specificity within the hierarchy of categorical knowledge.

### *Conclusions*

Fear in PTSD is associated with enhanced neural response to a learned representation of threat based on an established conceptual knowledge of the relationship between basic-level exemplars. PTSD may be one outcome of an evolutionary reality that misidentifying a dangerous stimulus as safe is far costlier to survival than incorrectly identifying a safe stimulus

## Conceptual fear generalization in PTSD

as dangerous. Real world situations rarely involve simple unimodal stimuli, but rather involve complex objects and situations that can be represented across a variety of dimensions. One approach to capture the complexity of human fear generalization is to incorporate theoretical knowledge from other psychological disciplines that examine generalization of human knowledge, including Bayesian models of inferential reasoning, categorization, and the organization of conceptual knowledge[48].

Author accepted manuscript

Conceptual fear generalization in PTSD

### **Funding and Disclosures**

Rajendra A. Morey has no financial interests and no potential conflicts of interest. Courtney C. Haswell has no financial interests and no potential conflicts of interest. Daniel Stjepanovic has no financial interests and no potential conflicts of interest. Joseph E. Dunsmoor has no financial interests and no potential conflicts of interest. Kevin S. LaBar has no financial interests and no potential conflicts of interest. The contributing members of the VA Mid-Atlantic MIRECC Workgroup have no financial interests and no potential conflicts of interest.

This research was supported by the U.S. Department of Veterans Affairs (VA) Mid-Atlantic Mental Illness Research, Education, and Clinical Center (MIRECC) core funds of the Department of Veterans Affairs Office of Mental Health Services. Dr. Morey also received financial support from the VA Office of Research and Development (5I01CX000748-01, 5I01CX000120-02). Additional financial support was provided by the National Institute for Neurological Disorders and Stroke (R01NS086885-01A1). Writing of this manuscript was supported by the Department of Veterans Affairs Office of Academic Affiliations Advanced Fellowship Program in Mental Illness Research and Treatment, the Medical Research Service of the Durham VA Health Care System, and the Department of Veterans Affairs Mid-Atlantic MIRECC. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States Government.

### **Acknowledgements**

The VA Mid-Atlantic MIRECC Workgroup contributors for this paper include: Mira Brancu, PhD, Jean C. Beckham, PhD, Patrick S. Calhoun, PhD, Eric Dedert, PhD, Eric B. Elbogen, PhD, John A. Fairbank, PhD, Robin A. Hurley, MD, Jason D. Kilts, PhD, Nathan A. Kimbrel, PhD, Angela Kirby, MS, Christine E. Marx, MD, MA, Scott D. McDonald, PhD, Scott D. Moore, MD, PhD, Jennifer C. Naylor, PhD, Jared Rowland, PhD, Cindy Swinkels, PhD, Steven T. Szabo, MD, PhD, Katherine H. Taber, PhD., Larry A. Tupler, PhD, Elizabeth E. Van Voorhees, PhD, H. Ryan Wagner, Ph.D., Ruth E. Yoash-Gantz, PsyD.



Conceptual fear generalization in PTSD

### **Author Contributions**

RAM wrote the manuscript, contributed to data analysis, contributed to the interpretation of the data, contributed to the acquisition of data, gave final approval for the version to be published and is accountable for all aspects of the work including accuracy and integrity.

CCH made substantial contributions to data analysis, contributed to the interpretation of the data, contributed to the acquisition of data, and gave final approval for the version to be published.

DS contributed to data analysis, contributed to the interpretation of the data, contributed to the acquisition of data, gave final approval for the version to be published.

MAM Workgroup contributed to the acquisition of data, gave final approval for the version to be published.

JED made substantial contributions to the conception and design, contributed to the interpretation of the data, gave final approval for the version to be published.

KSL made substantial contributions to the conception and design, contributed to the interpretation of the data, contributed to writing the manuscript, gave final approval for the version to be published, and is accountable for all aspects of the work including accuracy and integrity.

## Conceptual fear generalization in PTSD

## REFERENCES:

1. Dunsmoor JE, Paz R (2015): Fear generalization and anxiety: Behavioral and neural mechanisms. *Biological Psychiatry*
2. LeDoux J (2012): Rethinking the emotional brain. *Neuron* 73:653-76.
3. Morey R, Dunsmoor J, Haswell C, Brown V, Vora A, Weiner J, et al. (2015): Fear learning circuitry is biased toward generalization of fear associations in posttraumatic stress disorder. *Translational psychiatry* 5:e700.
4. Kaczurkin AN, Burton PC, Chazin SM, Manbeck AB, Espensen-Sturges T, Cooper SE, et al. (2016): Neural substrates of overgeneralized conditioned fear in ptsd. *American journal of psychiatry:appi. ajp.* 2016.15121549.
5. Ehlers A (2015): Understanding and treating unwanted trauma memories in posttraumatic stress disorder. *Zeitschrift für Psychologie/Journal of Psychology*
6. Vervoort E, Vervliet B, Bennett M, Baeyens F (2014): Generalization of human fear acquisition and extinction within a novel arbitrary stimulus category. *PloS one* 9:e96569.
7. Dymond S, Dunsmoor JE, Vervliet B, Roche B, Hermans D (2015): Fear generalization in humans: Systematic review and implications for anxiety disorder research. *Behavior therapy* 46:561-82.
8. Dunsmoor JE, Martin A, LaBar KS (2012): Role of conceptual knowledge in learning and retention of conditioned fear. *Biol Psychol* 89:300-5.
9. Dunsmoor JE, Kragel PA, Martin A, Labar KS (2014): Aversive learning modulates cortical representations of object categories. *Cereb Cortex*
10. Pitman RK, Rasmusson AM, Koenen KC, Shin LM, Orr SP, Gilbertson MW, et al. (2012): Biological studies of post-traumatic stress disorder. *nature Reviews neuroscience* 13:769.
11. Akiki TJ, Averill CL, Wrocklage KM, Schweinsburg B, Scott JC, Martini B, et al. (2017): The association of ptsd symptom severity with localized hippocampus and amygdala abnormalities. *Chronic Stress* 1:2470547017724069.
12. Sheynin J, Liberzon I (2017): Circuit dysregulation and circuit-based treatments in posttraumatic stress disorder. *Neuroscience letters* 649:133-8.
13. Fitzgerald JM, DiGangi JA, Phan KL (2018): Functional neuroanatomy of emotion and its regulation in ptsd. *Harvard review of psychiatry* 26:116-28.
14. Clausen AN, Clarke E, Phillips RD, Haswell C, Morey RA (2019): Combat exposure, posttraumatic stress disorder, and head injuries differentially relate to alterations in cortical thickness in military veterans. *Neuropsychopharmacology*:1-9.
15. Admon R, Milad MR, Hendler T (2013): A causal model of post-traumatic stress disorder: Disentangling predisposed from acquired neural abnormalities. *Trends in Cognitive Sciences* 17:337-47.
16. Lanius R, Bluhm R, Lanius U, Pain C (2006): A review of neuroimaging studies in ptsd: Heterogeneity of response to symptom provocation. *Journal of psychiatric research* 40:709-29.

## Conceptual fear generalization in PTSD

17. Neumeister P, Feldker K, Heitmann CY, Helmich R, Gathmann B, Becker MP, et al. (2017): Interpersonal violence in posttraumatic women: Brain networks triggered by trauma-related pictures. *Social cognitive and affective neuroscience* 12:555-68.
18. Brancu M, Wagner HR, Morey RA, Beckham JC, Calhoun PS, Tupler LA, et al. (2017): The post-deployment mental health (pdmh) study and repository: A multi-site study of us afghanistan and iraq era veterans. *International Journal of Methods in Psychiatric Research* 26
19. Blake DD, Weathers FW, Nagy LM, Kaloupek DG, Gusman FD, Charney DS, et al. (1995): The development of a clinician-administered ptsd scale. *Journal of Traumatic Stress* 8:75-90.
20. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M (1993): Development of the alcohol use disorders identification test (audit): Who collaborative project on early detection of persons with harmful alcohol consumption--ii. *Addiction* 88:791-804.
21. Bernstein DP, Fink L, Handelsman L, Foote J, Lovejoy M, Wenzel K, et al. (1994): Initial reliability and validity of a new retrospective measure of child abuse and neglect. *The American journal of psychiatry* 151:1132.
22. Kubany ES, Haynes SN, Leisen MB, Owens JA, Kaplan AS, Watson SB, et al. (2000): Development and preliminary validation of a brief broad-spectrum measure of trauma exposure: The traumatic life events questionnaire. *Psychological Assessment* 12:210-24.
23. Beck AT, Steer RA, Brown GK (ed)^(eds): Manual for the beck depression inventory-ii. San Antonio, TX: Psychological Corp, 1996
24. Skinner HA (1982): The drug abuse screening test. *Addictive Behaviors* 7:363-71.
25. Ohman A, Mineka S (2001): Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychol Rev* 108:483-522.
26. Dunsmoor JE, Mitroff SR, Labar KS (2009): Generalization of conditioned fear along a dimension of increasing fear intensity. *Learning & Memory* 16:460-9.
27. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, et al. (2004): Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage* 23 Suppl 1:S208-19.
28. Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996): A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping* 4:58-73.
29. Eklund A, Nichols TE, Knutsson H (2016): Cluster failure: Why fmri inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences* 113:7900-5.
30. Han H, Glenn AL, Dawson KJ (2019): Evaluating alternative correction methods for multiple comparison in functional neuroimaging research. *Brain sciences* 9:198.
31. McLaren DG, Ries ML, Xu G, Johnson SC (2012): A generalized form of context-dependent psychophysiological interactions (gppi): A comparison to standard approaches. *Neuroimage* 61:1277-86.

## Conceptual fear generalization in PTSD

32. Kriegeskorte N, Mur M, Bandettini P (2008): Representational similarity analysis—connecting the branches of systems neuroscience. *Frontiers in systems neuroscience* 2
33. Deen B, Richardson H, Dilks DD, Takahashi A, Keil B, Wald LL, et al. (2017): Organization of high-level visual cortex in human infants. *Nature communications* 8:13995.
34. Miller GA, Chapman JP (2001): Misunderstanding analysis of covariance. *Journal of Abnormal Psychology* 110:40-8.
35. Diedenhofen B, Musch J (2015): Cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS one* 10:e0121945.
36. Cavanna AE, Trimble MR (2006): The precuneus: A review of its functional anatomy and behavioural correlates. *Brain* 129:564-83.
37. Fink GR, Halligan PW, Marshall JC, Frith CD, Frackowiak R, Dolan RJ (1996): Where in the brain does visual attention select the forest and the trees? *Nature* 382:626.
38. Huth AG, Nishimoto S, Vu AT, Gallant JL (2012): A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron* 76:1210-24.
39. Binder JR, Desai RH, Graves WW, Conant LL (2009): Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex* 19:2767-96.
40. Gilboa A, Marlatte H (2017): Neurobiology of schemas and schema-mediated memory. *Trends in cognitive sciences* 21:618-31.
41. Ehlers A, Hackmann A, Steil R, Clohessy S, Wenninger K, Winter H (2002): The nature of intrusive memories after trauma: The warning signal hypothesis. *Behaviour research and therapy* 40:995-1002.
42. Dymond S, Dunsmoor JE, Vervliet B, Roche B, Hermans D (2014): Fear generalization in humans: Systematic review and implications for anxiety disorder research. *Behavior Therapy*
43. Striedter GF (2016): Evolution of the hippocampus in reptiles and birds. *Journal of Comparative Neurology* 524:496-517.
44. Ciochi S, Herry C, Grenier F, Wolff SB, Letzkus JJ, Vlachos I, et al. (2010): Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature* 468:277-82.
45. Dunsmoor JE, White AJ, LaBar KS (2011): Conceptual similarity promotes generalization of higher order fear learning. *Learn Mem* 18:156-60.
46. Onat S, Büchel C (2015): The neuronal basis of fear generalization in humans. *Nature neuroscience* 18:1811-8.
47. Preusser F, Margraf J, Zlomuzica A (2017): Generalization of extinguished fear to untreated fear stimuli after exposure. *Neuropsychopharmacology* 42:2545.
48. Dunsmoor JE, Murphy GL (2015): Categories, concepts, and conditioning: How humans generalize fear. *Trends in cognitive sciences* 19:73-7.

## Conceptual fear generalization in PTSD

**FIGURES**

**Figure 1.** Subjects were presented with 80 unique exemplars of animals (40) and tools (40) for 6 s each. Exemplars were never repeated during the fear acquisition session. Half the objects from one category (CS+) co-terminated with a 6-ms electrical shock unconditioned stimulus (US), whereas objects from the other category were never reinforced (CS-). Each trial was separated by a 6-8 s jittered inter-trial interval. Subjects were not instructed on the CS-US contingency and were not informed that images were presented only once. In this example, tools served as the CS+ and animals served as the CS- (category assignment was counterbalanced across subjects). Subjects learned through experience to generalize their conditioned fear response from the reinforced exemplars of the CS+ category (e.g., tools as shown here) to the other exemplars from the same object category, thus forming a broad threat response to the entire category.

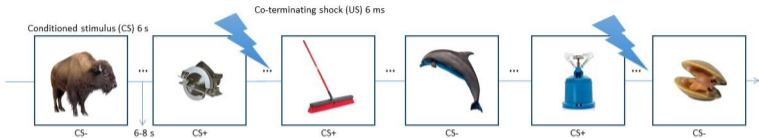
**Figure 2.** Physiological and behavioral indices of conceptual fear generalization. **(a)** Mean skin conductance response (SCR) for CS+ and CS- in Control and PTSD groups averaged across all trials. **(b)** Compared to the trauma-exposed control group, the PTSD group demonstrated delayed contingency learning, as measured by shock expectancy ratings. On average, subjects' differential shock expectancy ratings began to diverge from the 2<sup>nd</sup> trial of each category exemplar in the control group but only from the 7<sup>th</sup> trial onward in the PTSD group. Linear trend analysis confirmed a strong linear pattern of threat learning (CS+) only in the PTSD group, whereas both groups showed strong linear patterns for safety learning (CS-). **(c)** All four runs confirm that learning was achieved in the first run and was maintained for the duration of the experiment. Error bars represent mean  $\pm$  SEM. \* $P < 0.05$ , 2-tailed t-tests. CONT = control.

**Figure 3.** fMRI correlates of conceptual fear generalization. Data are averaged across all exemplars from the reinforced category compared to the unreinforced category (CS+ > CS-). **(A)** Relative to trauma-exposed Veteran controls, PTSD patients exhibited enhanced activation in an extensive corticolimbic network, including the amygdala (amyg) [x y z] = 28, -2, -16, ventromedial prefrontal cortex (vmPFC) [x y z] = 6, 44, -14, striatum, particularly in the putamen (putam) [x y z] = 30, -8, -0, insula [x y z] = 38, -16, 12, rostral anterior cingulate cortex (rACC) [x y z] = 2, 36, -2, subgenual ACC (sgACC) [x y z] = -4, 18, -6, lingual gyrus (lingual) [x y z] = 16, -58, 0,

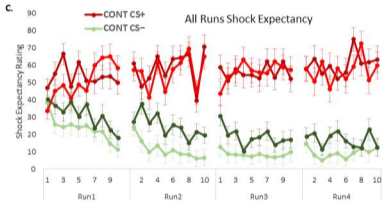
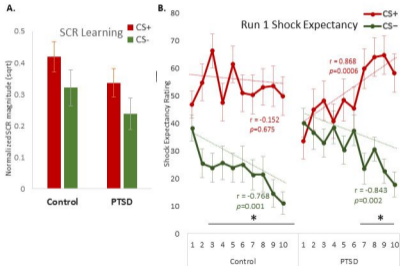
## Conceptual fear generalization in PTSD

and intracalcarine cortex (intracalcarine ) [x y z] = 16, -82, 8. All results were significant with whole brain correction (cluster threshold  $Z > 2.3$ ;  $p < 0.05$  corrected). **(B)** The bar graph offers a visualization of regional activation presented in the activation maps presented as Z-scores. Error bars represent means  $\pm$  SEMs.

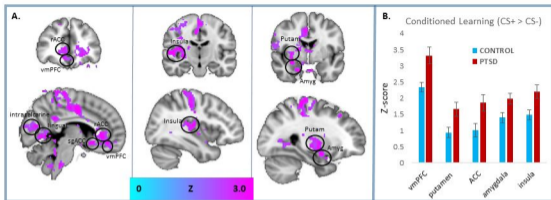
Author accepted manuscript



Under exclusive licence to American College of Neuropsychopharmacology







der exclusive licence to American College of Neuropsychopharm

## Conceptual fear generalization in PTSD

**Table 1. Clinical and Demographic Features of Sample**

Group	Control (n=23) Mean $\pm$ SD	PTSD (n=23) Mean $\pm$ SD	statistic	p-value
Age	38.65 $\pm$ 8.07	41.83 $\pm$ 8.37	t = 1.31	0.20
Race (AA), n [%]	10 [43.5]	12 [56.5]	$X^2 = 2.19$	0.34
White), n [%]	14 [60.8]	9 [39.1]	n/a	n/a
Multiracial [n%]	0 [0]	1 [100]	n/a	n/a
female gender; n [%]	3 [13.0]	7 [30.4]	$X^2 = 2.04$	0.15
CAPS (DSM-5)	1.39 $\pm$ 2.86	46.64 $\pm$ 32.03	t = 6.60	< .0001
*Trauma Chronicity (y)	13.57 $\pm$ 10.14	14.43 $\pm$ 8.78	t = 0.30	0.86
BDI-II	5.25 $\pm$ 7.76	16.14 $\pm$ 12.25	t = 3.63	0.001
TLEQ	9.75 $\pm$ 8.21	19.77 $\pm$ 13.15	t = 3.13	0.003
Combat Exposure	3.00 $\pm$ 5.47	8.23 $\pm$ 8.85	t = 2.43	0.02
Child Trauma (CTQ)	63.83 $\pm$ 10.77	66.00 $\pm$ 10.71	t = .68	0.50
AUDIT	2.25 $\pm$ 1.62	4.09 $\pm$ 4.20	t = 1.93	0.064
DAST	0.25 $\pm$ 0.61	0.55 $\pm$ 1.18	t = 1.08	0.29
Serotonergic n [%]	0 [0]	12 [54.5]	$X^2 = 17.71$	< .0001
Antipsychotic n [%]	0 [0]	0 [0]	$X^2 = 0$	1.0
Mood Stabilizer n [%]	0 [0]	0 [0]	$X^2 = 0$	1.0
Animal CS+ n [%]	12 [52.2]	12 [52.2]	$X^2 = 0$	1.0
Right Handed n [%]	22 [95.6]	21 [91.3]	$X^2 = 3.02$	0.22

\*Trauma chronicity is based on information provided by participants on the Trauma Life Events Questionnaire (#24) to indicate the number of years since the last occurrence of the trauma deemed to be the most distressing [23].

Abbreviations: SD= Standard Deviation, AA=African American, AUDIT = Alcohol Use Disorders Test, BDI-II=Beck Depression Inventory II, CES = Combat Exposure Scale; CAPS = Clinician Administered PTSD Scale for DSM-5, CTQ = Childhood Trauma Questionnaire, CS+ = Conditioned Stimulus, DAST = Drug Abuse Screening Test, TLEQ=Trauma Life Events Questionnaire.

## Conceptual fear generalization in PTSD

**Table 2. Brain regions involved in conceptual fear learning (CS+ > CS-) with greater activation in the PTSD group than the trauma-exposed control group.**

Sub/Cortical Region	Number Voxels 3.8-mm <sup>3</sup>	Max Z- statistic	Peak 3-D Coordinates		
			<b>x</b>	<b>y</b>	<b>z</b>
<i>rostral cingulate/vmPFC right</i>	389	4.05	42	60	57
<i>putamen left</i>	194	3.43	56	73	38
<i>planum temporale right</i>	191	3.59	17	56	39
<i>frontal pole right</i>	161	3.45	40	89	32
<i>posterior cingulate cortex left</i>	148	3.33	48	40	50
<i>lingual gyrus/ precuneus right</i>	144	4.6	36	34	36
<i>postcentral right</i>	144	3.33	31	47	67
<i>putamen right</i>	140	3.68	29	59	34
<i>central operculum right</i>	79	3.7	20	60	46
<i>precuneus left</i>	51	3.47	58	35	41
<i>lateral occipital cortex left</i>	41	3.01	68	24	47
<i>amygdala right</i>	43	3.42	35	62	22
<i>intracalcarine cortex right</i>	39	3.29	37	21	39
<i>right postcentral</i>	39	3.26	28	52	57
<i>supramarginal right</i>	35	3.04	21	51	64
<i>dorsal ACC left</i>	30	3.51	49	89	47
<i>subgenual ACC right</i>	24	3.61	41	71	31
<i>precentral right</i>	162	3.43	44	49	68
<i>temporo-occipital fusiform left</i>	23	3.17	65	33	37
<i>cuneus right</i>	22	3.08	39	21	51
<i>amygdala right</i>	21	3.81	30	62	29

All clusters are significant at  $p < 0.05$  with whole brain correction using FLAME-2.