



Behavioral and neural responses during fear conditioning and extinction in a large transdiagnostic sample

Namik Kirlic^{a,b,*}, Rayus Kuplicki^a, James Touthang^a, Zsofia P. Cohen^a, Jennifer L. Stewart^{a,b}, Tulsa 1000 Investigators^a, Martin P. Paulus^{a,b}, Robin L. Aupperle^{a,b}

^a Laureate Institute for Brain Research, Tulsa, OK, USA

^b Department of Community Medicine, Oxley College of Health Sciences, The University of Tulsa, Tulsa, OK, USA

ABSTRACT

Background: Dysregulation of fear learning has been associated with psychiatric disorders that have altered positive and negative valence domain function. While amygdala-insula-prefrontal circuitry is considered important for fear learning, there have been inconsistencies in neural findings in healthy and clinical human samples. This study aimed to delineate the neural substrates and behavioral responses during fear learning in a large, transdiagnostic sample with predominantly depressive and/or anxious dysfunction.

Methods: Two-hundred and eighty-two individuals (52 healthy participants; 230 participants with depression and/or anxiety-related problems) from the Tulsa 1000 study, an ongoing, naturalistic longitudinal study based on a dimensional psychopathological framework, completed a Pavlovian fear learning task during functional magnetic resonance imaging. Linear mixed-effects analyses examined condition-by-time effects on brain activation (CS+, CS- across familiarization, conditioning, and extinction trials). A data-driven latent profile analysis (LPA) examined distinct patterns of behavioral and neural responses to threat across fear conditioning and extinction, while logistic regression analyses evaluated cognitive-affective predictors of latent profiles.

Results: Whole-brain analyses revealed a condition-by-time interaction in the anterior insula, postcentral gyrus, superior temporal gyrus, middle frontal gyrus, and cerebellum but not amygdala. The LPA identified distinct latent profiles across subjective and neural levels of measurement. Anterior insula profiles were characterized by marginal differences in age and state anxiety.

Conclusions: Our findings demonstrate that human fear learning recruits a distributed network of regions involved in interoceptive, cognitive, motivational, and psychomotor processes. Data-driven analyses identified distinct profiles of subjective and neural responses during fear learning that transcended clinical diagnoses, but no robust relationships to demographic or cognitive-affective variable were identified.

1. Introduction

Threat detection and defense response (fear learning) (LeDoux, 2014) is a cross-species, associative learning process considered crucial for adapting to one's environment. Fear learning models have been a mainstay of animal and human research attempting to better understand threat and defense behaviors and their underlying neurobiological mechanisms, as well pathophysiology of clinical fear and anxiety (Kindt, 2014). These paradigms typically involve both the acquisition (conditioning) and extinction of fear using Pavlovian conditioning (Pavlov, 2010), processes that extend beyond the bench and hold relevance for the development, maintenance, and treatment of fear and anxiety disorders. Pavlovian fear conditioning involves the pairing of a previously neutral stimulus with a salient, naturally noxious stimulus (unconditioned stimulus, US), such that the individual begins to exhibit a fear response to the previously neutral stimulus (conditioned stimulus; CS)

even when presented alone. On the other hand, fear extinction involves the presentation of the CS without subsequent negative outcomes (US), such that the fear response to the CS diminishes. Conditioning and extinction processes are considered crucial for survival, helping organisms adapt in order to respond appropriately to real and potential dangers in their environment (Ohman et al., 2008). Leveraging advanced neuroimaging techniques to understand this process in healthy and clinical populations may lead to development of novel targets for interventions and improve outcomes.

Functional neuroimaging has facilitated the examination of neural substrates of human fear learning in an attempt to translate animal findings and identify neural networks of importance for fear learning dysfunction in psychiatric disorders (Hermans et al., 2006). While animal studies have identified the importance of the amygdala, including its central and basolateral subregions (underlying the expression and inhibition of fear), hippocampus (contextualization or modulation of the

* Corresponding author at: Laureate Institute for Brain Research, 6655 South Yale Ave, Tulsa, OK 74136.

E-mail address: nkirlic@laureateinstitute.org (N. Kirlic).

<https://doi.org/10.1016/j.nicl.2022.103060>

Received 7 October 2021; Received in revised form 28 April 2022; Accepted 21 May 2022

Available online 25 May 2022

2213-1582/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

fear response), and medial prefrontal cortex (PFC; inhibition of the fear learning response, particularly during extinction) (LeDoux, 2000; Davis and Whalen, 2001; Milad et al., 2006; Milad and Quirk, 2012; Quirk and Mueller, 2008), human neuroimaging research has shown that fear learning studies activate a wider network of regions, including the amygdala, anterior insula (AI), dorsal anterior cingulate cortex (dACC), posterior cingulate cortex (PCC), orbitofrontal cortex (OFC), cerebellum, anterior thalamus, ventral putamen and pallidum, midbrain substantia nigra/ventral tegmentum, hippocampus, ventromedial prefrontal cortex (vmPFC), and dorsolateral PFC (dlPFC) (Etkin and Wager, 2007; Mechias et al., 2010; Sehlmeier et al., 2009; Marin et al., 2020; Savage et al., 2020). However, neural findings across these human studies have been inconsistent and limited by a number of factors including 1) relatively smaller sample sizes ($N = 11-114$) and 2) methodological differences across studies including number and timing of trials, reinforcement rates, instruction set (instructed vs. uninstructed threat), conditioning and extinction training approaches (e.g., cued vs. contextual; immediate vs. delayed extinction), and analyses of interest (e.g., $CS+ > CS-$; $CS+ < CS-$; $CS+ > CS-$, $CS+ < CS-$). In an effort to provide more uniform findings on neural underpinnings of fear learning, Fullana and colleagues implemented a *meta*-analytic functional magnetic resonance imaging (fMRI) approach involving anisotropic effect-size signed differential mapping capable of combining tabulated brain activation/deactivation results with actual empirical voxel-wise ‘brain maps’ of activations and deactivations across studies separately for fear conditioning (Fullana et al., 2016) and extinction (Fullana et al., 2018). For fear conditioning ($CS+ > CS-$, $CS+ < CS-$ during conditioning trials) (Fullana et al., 2016), their analysis identified consistent and robust evidence for the activation of “cingulofrontal cortex” regions including dACC, dorsomedial PFC, and AI, likely due to their importance in autonomic-interoceptive processing (Berntson and Khalsa, 2021; Smith et al., 2017) and integration of cognitive, affective, and physical states (McTeague et al., 2020). In corroboration of translational theories of fear learning, these studies highlighted evidence for deactivations (i.e., higher for $CS-$ than $CS+$) within default mode network regions (vmPFC, PCC), lateral PFC, and hippocampus, as potentially representing a “safety signal” network. With respect to fear extinction ($CS+ > CS-$ during extinction trials) (Fullana et al., 2018), the *meta*-analysis primarily implicated consistent activation of brain regions linked to threat appraisal and experience, including dACC and AI. Additionally, extinction recall (i.e., retrieval and expression of learned extinction memory following a delay) evidenced more dlPFC and vmPFC cortices and the hippocampus. Notably, these *meta*-analyses did not identify robust evidence for involvement of the amygdala in fear conditioning or extinction, which Fullana and colleagues argued could be due to the technical constraints of fMRI (perhaps failing to capture initial, transient responses that quickly habituate) or due to fear learning paradigms as conducted in humans not evoking the classic amygdala threat-detection response (Fullana et al., 2016; Fullana et al., 2018). Therefore, based on the extant evidence from neuroimaging studies, the role of amygdala in human fear learning remains uncertain.

Dysfunction in fear learning is considered to contribute to the development and maintenance of numerous psychiatric disorders, including not only anxiety disorders and posttraumatic stress disorder (Otto et al., 2014; Milad et al., 2014; Duits et al., 2015; Wicking et al., 2016), but also depression (Jovanovic et al., 2010; Sandi and Richter-Levin, 2009), borderline personality disorder (Krause-Utz et al., 2016), schizophrenia (Holt et al., 2012), and bipolar disorder (Acheson et al., 2015). Therefore, assessment of fear learning has become a pertinent approach to assessment and treatment of transdiagnostic psychopathology. However, small and relatively homogenous samples (e.g., single diagnostic group) have impeded delineation of variability in fear responses and the effect of individual differences across relevant psychological transdiagnostic domains in patients with most commonly-occurring complex diagnostic presentations. It would be therefore beneficial to identify cognitive-affective processes that determine fear

responses in large, transdiagnostic samples, in order to leverage fear learning to point toward strategies to improve patient outcomes. Previous research provides some evidence for state and trait anxiety (Alvarez et al., 2015; Sehlmeier et al., 2011; Dunsmoor et al., 2011; Fredrikson et al., 1995; Nitschke et al., 2006), trait fearfulness (Sylvester et al., 2011; Schmitz and Grillon, 2012; Panitz et al., 2018), perceived control (Alvarez et al., 2015), intolerance of uncertainty (Morris et al., 2015), anxiety sensitivity (Lueken et al., 2014), social anxiety (Pejic et al., 2013), emotional numbing (Wicking et al., 2016), symptom severity (Milad et al., 2013), negative affect (Kirlic et al., 2019), pain sensitivity (Kirlic et al., 2019; Yáguez et al., 2005), and catastrophizing and worrying (Kalisch and Gerlicher, 2014) as transdiagnostic contributors to differences in fear learning. However, these factors were examined in isolation from other relevant cognitive-affective processes, or other important moderators of fear learning, such as age (Schreurs et al., 2001; Ganella et al., 2018) and sex (Lebron-Milad et al., 2012; Merz et al., 2013). Examination of such factors in combination with identification of the potential moderating effects of age and sex on fear learning are necessary to a clearer understanding of the role of dysfunctional fear learning in psychiatric disorders (Verdi et al., 2021).

Previous research on fear learning has generally examined differences in $CS+$, or $CS+$ relative to $CS-$, during conditioning or extinction trials, as well as how these responses may relate to specific clinical variables of interest. Data-driven approaches to derive clinically meaningful subtypes of human neurobehavioral responses during various stages of fear learning have not been commonly employed despite the fact that such approaches contain several clinical benefits including identification of subtypes of fear learning in the context of complex sample and experimental designs, as well as evaluation of treatment outcomes (Verdi et al., 2021).

This study aimed to delineate the neural substrates and behavioral responses during fear learning in a large, transdiagnostic sample presenting with predominant positive and negative valence dysfunction, that is depression and/or anxiety disorders. The goals of the study were to 1) use a data-driven approach to characterize responses to threatening and non-threatening stimuli across multiple phases of fear learning (familiarization, conditioning, and extinction), 2) identify whether distinct patterns of fear learning existed across behavioral and neural domains, and 3) determine whether individual differences characterized these profiles. We hypothesized that fear learning would engage the fronto-cingulate and insular cortices, as well as portions of the limbic system, including the amygdala and hippocampus. We further hypothesized that distinct behavioral and neural profiles of responses during fear learning would be identified, distinguishing between participants demonstrating successful fear conditioning and extinction, exaggerated conditioning but successful extinction, exaggerated conditioning and impaired extinction, and finally impaired conditioning and impaired extinction. Finally, we hypothesized that these would relate to individual differences across cognitive and affective psychological measures.

We used a multi-modal data set (self-report, behavioral, and neural) collected on 282 participants from the first half of released participant data of the Tulsa 1000 (Victor et al., 2018) (T-1000) study, an observational study of a large, community-based sample, including healthy and treatment-seeking participants diagnosed with depression and/or anxiety disorders. The present study sample included predominantly participants with depression and/or anxiety disorder ($N = 230$) and their healthy comparisons ($N = 52$). Participants with substance use and eating disorders were excluded from present analyses given the particular relevance fear learning has in emergence, maintenance, and treatment of anxiety disorders and depression. Participants completed a range of cognitive-affective measures and underwent fMRI recording while completing a Pavlovian fear learning task. Subjective ratings and whole-brain analyses examined task responses in the entire sample, supplemented by region-of-interest (ROI) analyses based on clusters identified as critical for fear learning by *meta*-analytic work (Fullana

et al., 2016; Fullana et al., 2018). Next, we used a latent profile analysis approach to characterize unique trajectories of behavioral and neural threat responses during fear learning. Logistic regression analyses were utilized to identify transdiagnostic, individual-difference factors relating to the latent profiles. Although we included groups meeting distinct diagnostic criteria (healthy comparisons, depressed, anxious, and patients with comorbid depression and anxiety diagnoses), we referenced the Research Domain Criteria framework (Insel et al., 2010) and directed our focus on a dimensional evaluation and interpretation of relevant findings across all participants. We reasoned that this would allow the identification of transdiagnostic psychological constructs and their relationship with task-related responses, as well as have greater future treatment utility.

2. Methods

2.1. Participants

Participants were 282 individuals from the first half of the released participant data from the Tulsa 1000 (T-1000) (Victor et al., 2018) study, an ongoing, naturalistic longitudinal study based on a dimensional psychopathological framework. Subsetting from the sample, the present study included healthy participants and those diagnosed with mood and/or anxiety-related disorders (N = 322). Of these, N = 19 participants did not complete the fear learning task, and N = 21 had excessive motion and were thus removed from the analysis, leaving a total of N = 282 for the present analysis (230 participants with mood and/or anxiety-related disorders, and 52 healthy comparison participants). Table 1 provides participant demographic and clinical characteristics.

Participants underwent a structured clinical diagnostic interview for DSM-IV (Mini International Neuropsychiatric Interview [MINI] (Heragueta et al., 1998) conducted by trained master's level clinicians who underwent reliability evaluations and consensus with a psychiatrist. In addition, participants completed a range of self-report clinical measures assessing positive and negative valence domains, as well as interoception, behavioral tasks assessing positive and negative valence, cognition and interoception, physiological measurements consisting of skin conductance, facial emotion expression monitoring, heart rate, respiration and eye-blink startle response, fMRI focusing on reward-related processing, fear conditioning and extinction, cognitive control and inhibition and interoceptive processing, biomarker assessment, microbiome assessment, and genetic as well as epigenetic assessments (Victor et al., 2018). For further T-1000 study description and inclusion/exclusion criteria, see the supplemental materials and Victor and colleagues (Victor et al., 2018).

2.2. Procedures

All study procedures were approved by the Western Institutional Review Board (WIRB; Study #1150412). Research was conducted in accordance with the World Medical Association Declaration of Helsinki. Participants underwent an extensive screening process and assessment procedures along cognitive and affective domains (Victor et al., 2018), as well as completed a fear conditioning and extinction (FC/FE) task while undergoing fMRI. All data acquisition and analyses were performed at the Laureate Institute for Brain Research.

2.3. The fear learning task

The fear conditioning/extinction (FC/FE) paradigm (Fig. 1) was based on Pavlovian conditioning and the task previously used in neuroimaging studies of individual differences in fear learning (Sehlmeyer et al., 2011; Sehlmeyer et al., 2009; Ball et al., 2017). The stimuli consisted of two neutral, non-social, abstract images as conditioned stimuli (CS), presented for 1.5 s at a time. The images designated as CS+ (paired

Table 1

Participant demographic and psychological characteristics.

Characteristic	N = 282 (HC, n = 52; MDD and/or ANX n = 230)
Age, M (SD)	34.97 (11.2)
Income, M (SD)	52329.15 (58395.29)
Education, M (SD)	6.49 (1.62)
Percent Body Fat, M (SD)	28.22 (5.19)
Psychotropic Medication Status, N (%)	164 (58.2)
Race, N (%)	
Asian	6 (2%)
African American	24 (9%)
Native American	43 (15.2%)
White	244 (86.5)
Pacific Islander	2 (<1%)
Other	9 (3.2%)
Ethnicity, N (%)	
Hispanic	11 (3.9%)
Non-Hispanic	271 (96.1%)
Diagnosis, N (%)	
Major Depressive Disorder	65 (23%)
Anxiety Disorder	146 (51.8%)
Comorbid Anxiety and Depression	19 (6.7%)
Childhood Trauma Exposure, M (SD)	
CTQ Total Score	45.1 (18.26)
Cognitive Function, M (SD)	
WRAT Reading	62.65 (5.05)
Personality, M (SD)	
BFI Neuroticism	27.05 (7.71)
Symptom and Trait-like Measures	
PHQ-9	10.68 (6.83)
OASIS	8.24 (4.63)
RRS Rumination	50.16 (14.73)
ASI-3 Total Score	21.77 (15.42)
BIS/BAS Inhibition	22.48 (4.15)
BIS/BAS Reward	16.99 (2.24)
PANAS-X Negative Affect	21.32 (7.88)
PANAS-X Positive Affect	27.26 (8.55)
STAI State Anxiety	42.21 (13.19)
WHODAS_Score	21.58 (8.52)

Note. ASI-3, Anxiety Sensitivity Scale - Third Edition; BFI, Big Five Inventory; BIS/BAS, Behavioral Inhibition System/Behavioral Activation System; RRS, Rumination Response Scale; CTQ, Childhood Trauma Questionnaire; HC, healthy comparison participants; M, mean; MDD and/or ANX, participants with depression and/or anxiety disorder; PANAS-X, The Positive and Negative Affective Schedule; OASIS, Overall Anxiety Severity and Impairment Scale; PHQ-9, Patient Health Questionnaire; SD, standard deviation; STAI, State-Trait Anxiety Inventory; WHODAS, World Health Organization Disability Assessment Schedule; WRAT, Wide-range Achievement Test.

with the unconditioned stimulus (US) during conditioning) and CS- (never paired with the US) were counter-balanced across participants. The US consisted of a loud scream beginning 500 ms after CS+ onset, lasting approximately 1 s, and presented at 108-120DBs with participants wearing silicone ear plugs providing 22DBs of attenuation.

To increase engagement and attention during the inter-trial interval, participants engaged in a continuous performance task requiring a right or left button press in response to right or left facing arrows in the 9–15 s between CS image presentations. Participants were told that for this task they would see arrows on the screen pointing either to the left of the right, and that their job was to push the button that corresponds to the arrow on the screen as quickly and as accurately as possible. They were further informed that during the task they may see an image on the screen, but that they did not need to respond to the image; instead, they were asked to look at the image and be prepared for the next set of arrows. Next, they were told that at times during the task, they would also hear a loud scream through the headphones and that no response was required to the scream. Finally, participants were told that this was a task of attention, and that their goal was to do their best to pay

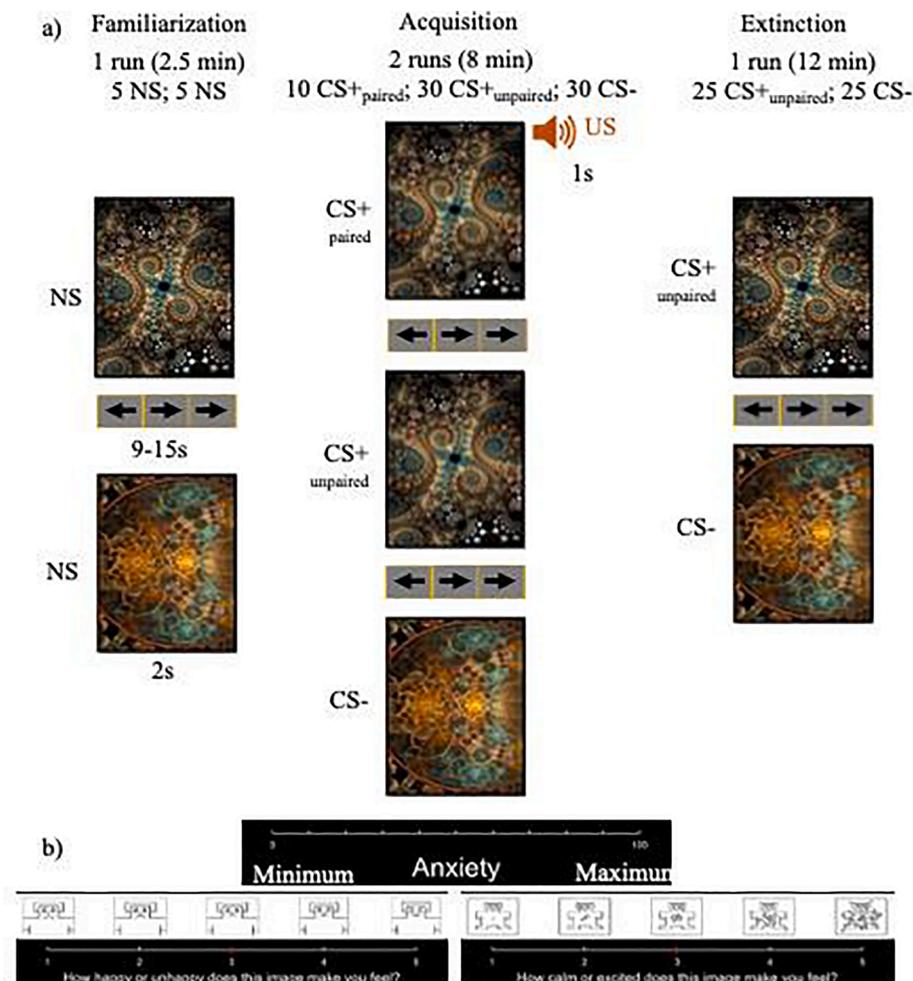


Fig. 1. Fear conditioning / extinction paradigm. (a) During the task, participants viewed two neutral, non-social, abstract images, one of which was paired with a loud scream. The task consisted of three consecutive phases: familiarization, conditioning, and extinction. (b) Following each run, participants rated their levels of anxiety, arousal, and image valence. Images were counterbalanced across participants. Trials were presented in a fixed, pseudo-randomized order. NS = neutral stimulus; CS⁺_{paired} = conditioned stimulus that was paired with the scream, paired trial; CS⁺_{unpaired} = conditioned stimulus that was paired with the scream, but not in this trial; US = unconditioned stimulus.

attention and push the button for the direction the arrows were pointing as quickly and as accurately as possible.

The task consisted of three phases: familiarization, fear conditioning, and fear extinction. The *familiarization phase* (lasting 2.5 min) involved five presentations of each CS with no instances of the US. Next, the *conditioning phase* involved 30 presentations of the CS⁻ and 40 presentations of the CS⁺ (10 with (CS⁺_{paired}) and 30 without (CS⁺_{unpaired}) the US, across two functional runs of eight minutes each. The 25% reinforcement rate is consistent with other fMRI studies of individual differences in fear conditioning, and also generates sufficient variability in extinction outcomes. CS⁺_{paired} trials were modeled in the individual-level deconvolution analysis but omitted from the group-level CS⁺ analysis to not confound processing of the CS⁺ with the reactivity to the US. This also allowed for an equal number of trials in the analysis. Finally, the *extinction phase* (lasting 12 min) followed immediately after conditioning and involved 25 presentations of each CS with no US reinforcement. Participants rated their anxiety level using a Likert scale (“On a scale from 0 = minimum anxiety to 100 = maximum anxiety, how anxious do you feel when you see this image?”), as well as valence (“How happy or unhappy does this image make you feel?”) and arousal (“How calm or excited does this image make you feel?”) levels using the Self-Assessment Manikin to each CS after each functional run. Trials were presented in a fixed, pseudo-randomized order, constrained so that no more than two identical trials occurred in a row.

2.4. Behavioral data analysis

Linear mixed effects (LME) analyses were employed to examine main

and interaction effects of FC/FE task Condition (CS⁺, CS⁻) and Time (familiarization, early and late conditioning, and extinction) on anxiety, valence, and arousal ratings. Post hoc two-tailed t-tests were used to further elucidate differences between threat conditions at each time point (please see the supplemental materials for arousal and valence ratings results).

2.5. fMRI data processing and analysis

Imaging data were acquired using two identical GE MR750 3 T scanners and an 8-channel phased-array coil. The following parameters were used for all EPI data: TR/TE = 2000/27 ms, FOV/slice = 240/2.9 mm, 128x128 matrix, 39 axial slices, and varied numbers of TRs depending on the functional run (familiarization = 79, conditioning 1/2 = 260 each, and extinction = 368). For normalization to standard space, a high-resolution Magnetization-Prepared Rapid Acquisition with Gradient Echo sequence was also acquired with the following parameters: TR/TE = 5/2.012 ms, FOV/slice = 240x192/0.9 mm, and 186 axial slices.

The Analysis of Functional NeuroImages (AFNI; <https://afni.nimh.nih.gov>) software package was used for all first-level neuroimaging data analyses (Cox, 1996). Processing steps included: removal of the first three volumes, despiking, slice timing correction, co-registration with anatomical volumes, motion correction, 4 mm full width at half maximum (FWHM) Gaussian smoothing, regression with the task design matrix, scaling to percent signal change, normalization to Montreal Neurological Institute (MNI) space using an affine transformation and resampling to 2 mm isometric voxels. Censoring was applied at the

regression step using a Euclidean norm motion threshold of 0.3. Nuisance regressors were included for a 5th order polynomial at each functional run and the estimated motion parameters. Task-relevant regressors were included for response (any time the subject responded to an arrow), and US, as well as the primary threat conditions of interest: CS+ and CS-. CS+ and CS- regressors were fit for 5 different timepoints: familiarization (run1), early and late conditioning (runs 2 and 3, respectively), and early and late extinction (run 4 divided into two epochs).

The R statistical package (version 3.5.1) (Core Team R. R., 2017) was used to fit voxel-wise linear mixed effects models including subject as the random factor and fixed effects for condition (CS+ or CS-), time (coded categorically), and the condition-by-time interaction. The smoothness of the group level error terms was estimated with AFNI's 3dFWHMx using the spatial autocorrelation function (acf) (Cox et al., 2017) and used with 3dClustSim to produce cluster size thresholds controlling the family-wise error rate (-acf a, b, c parameters: 0.56, 3.22, 9.14). Significance criterion for the whole-brain analysis was set at a corrected rate of $p < .05$ (cluster size ≥ 83 voxels) and thresholded per-voxel at $p < .005$. For a meaningful separation of clusters, the main effect of time was thresholded per voxel at $p < .001$ (cluster size ≥ 33 voxels). Average percent signal change (PSC) was extracted per individual from surviving clusters of activation of interest for condition-by-time interactions and submitted to follow-up analyses. In addition, we conducted whole brain analyses using more traditional approaches in two different ways. Specifically, we first used the CS+ vs. CS- contrast as the dependent variable and fit the effect of timepoint. Second, we ran voxelwise one-sample t-tests for the CS+ vs. CS- contrast at each timepoint. Smoothness and significance criterion thresholding were conducted as above.

Region of Interest Analyses. Research has implicated several other regions in conditioning and extinction of aversive stimuli that in this study did not survive the current recommendation for thresholding in whole brain analyses (Milad and Quirk, 2012; Milad et al., 2007; Etkin and Wager, 2007; Mechias et al., 2010; Fullana et al., 2016; Fullana et al., 2018). Driven by these past findings, we explored task effects from both hemispheres of the following ROIs as defined by the Brainnetome atlas (Fan et al., 2016): medial and lateral amygdala, dACC (corresponding to anterior dACC [adACC] or perigenual [pACC] as defined previously (Etkin et al., 2011; Vogt et al., 2003) anterior midcingulate cortex (aMCC; corresponding to posterior dACC [pdACC] or anterior MCC [aMCC] as defined previously (Etkin et al., 2011; Vogt et al., 2003); rostral and caudal hippocampus, and vmPFC (Figure S1). Average PSC for both threat conditions and across all time points were extracted from these ROIs and subjected to LME and follow-up analyses. We also examined the imaging data using trial-wise regressors (via AFNI's -stim_times_IM) to model the trial-by-trial BOLD response to CS+ and CS- from the four amygdala ROIs in the Brainnetome atlas (left/right lateral/medial). We modeled the extracted timeseries as $\beta \sim \text{condition} * \text{time} + \text{condition} * \text{time}^2$ to test for evidence of conditioning/extinction.

2.6. Latent profile analysis

We employed a data-driven latent profile analysis (LPA) to identify distinct patterns of behavioral and neural responses to threats (CS+) across each experimental condition (i.e., familiarization, early and late conditioning, extinction [early and late for brain data]) for regions identified in the whole brain analysis and regions from the ROI analysis that showed significant condition by time interactions. LPA was implemented with Gaussian-mixture modeling. The Bayesian information criterion (BIC) and bootstrap likelihood ratio test (LRT) were used to compare models with the number of components varying from 1 to 5. The model with the lowest BIC, or where additional components did not significantly improve the model, was selected as the optimal description of latent components in the data. The analyses were implemented with R's Mclust package (Scrucca et al., Aug 2016).

2.7. Logistic regression analysis

LPA results were subjected to multiple logistic regression models to ascertain the effects of participant characteristic, clinical symptom, and psychological trait variables on the likelihood that participants fell into a latent group of behavioral and neural responses of interest. Participant characteristic variables included age, sex, percent body fat, stable dose (i.e., 6 weeks prior to beginning the study) of psychotropic medication status, and academic skill achievement. Clinical symptom variables included the Patient Health Questionnaire-9 (PHQ-9) (Kroenke et al., 2001) and Overall Anxiety Severity and Impairment Scale (OASIS) (Norman et al., 2006), while psychological variables included the Neuroticism scale from the Big Five Inventory (BFI) (John and Srivastava, 1999), Inhibition and Reward sensitivity subscales from the Behavioral Inhibition/Behavioral Activation System (BIS/BAS) (Carver and White, 1994), Anxiety Sensitivity Inventory-3 (ASI-3) total score (Taylor et al., 2007); Positive and Negative Affect (PANAS-X) subscales (Watson et al., 1988), Rumination Response Scale (RRS) (Treynor et al., 2003), and State Anxiety subscale from State-Trait Anxiety Inventory (STAI) (Spielberger, 1983). The logistic regressions were estimated using the generalized linear model (glm) function. The overall effect of the model was tested using the chi-square on the difference between null and residual deviance and their respective degrees of freedom. Exponentiated coefficients were used to calculate odds-ratios. The analyses were implemented with R's Mlogit package (Croissant, 2020).

3. Results

3.1. Behavioral data

LME analyses (Fig. 2, Table S1) revealed main effects of condition and time on state ratings of anxiety [condition: $F_{(1,1967)} = 133.30$, $p = 2.2^{-16}$, time: $F_{(3,1967)} = 66.82$, $p = 2.2^{-16}$], valence [condition: $F_{(1,1967)} = 116.68$, $p = 2.2^{-16}$, time: $F_{(3,1967)} = 69.84$, $p = 2.2^{-16}$], and arousal [condition: $F_{(1,1967)} = 96.20$, $p = 2.2^{-16}$, time: $F_{(3,1967)} = 36.27$, $p = 2.2^{-16}$]. These were qualified by a condition-by-time interaction [anxiety: $F_{(1,1967)} = 23.26$, $p < 8.5^{-15}$, $BIC = 19484.85$; valence: $F_{(3,1967)} = 14.73$, $p = 1.7^{-09}$, $BIC = 5441.38$; arousal: $F_{(3,1967)} = 9.12$, $p = 5.4^{-06}$, $BIC = 5948.27$]. Post hoc analyses indicated that anxiety ratings for CS+ were significantly increased relative to CS- during both conditioning [early: $t_{(281)} = 7.46$, $p = 1.1^{-12}$; late: $t_{(281)} = 9.03$, $p = 2.2^{-16}$] and extinction [$t_{(281)} = 6.59$, $p = 2.2^{-10}$]. There was a significant decrease in anxiety ratings for CS+ from conditioning to extinction [$t_{(281)} = 5.57$, $p = 5.9^{-08}$]. For post hoc analyses on arousal and valence ratings, see the supplemental materials.

4. Imaging results

Whole-brain analyses. Results (Fig. 2; Table 2) revealed a main effect of condition in the bilateral AI, posterior MCC, right middle frontal gyrus (MFG), left precuneus, and left inferior parietal lobe (IPL). A main effect of time was evidenced in the bilateral dACC/superior medial gyrus, insula, precuneus, posterior MCC, MFG, right inferior frontal gyrus, right superior frontal gyrus, right thalamus, right putamen, and cerebellum. These main effects were qualified by a condition-by-time interaction most notably in the right postcentral gyrus (PG), bilateral superior temporal gyrus (STG), right AI, right MFG, and cerebellum. The results of the traditional analyses are shown in Figure S2 and S3 and Tables S2 and S3. The data are qualitatively similar across analysis approaches.

Region of interest analyses. ROI analyses were performed on the subregions of amygdala, hippocampus, dACC, mACC, and vmPFC (Table S1). Both the medial and lateral amygdala showed a main effect of time [medial: $F_{(4,2529)} = 3.76$, $p = .0047$; lateral: $F_{(4,2529)} = 3.86$, $p = .0039$], but not condition [medial: $F_{(1,2529)} = 0.16$, $p = .68$; lateral: $F_{(1,2529)} = 0.03$, $p = .84$] or condition-by-time interaction [medial: $F_{(4,2529)} = 1.10$, $p = .35$, $BIC = 26374.35$; lateral: $F_{(4,2529)} = 0.40$, $p = .81$, $BIC = -28375.32$; Figure S4]. The analysis modeling trial-by-trial

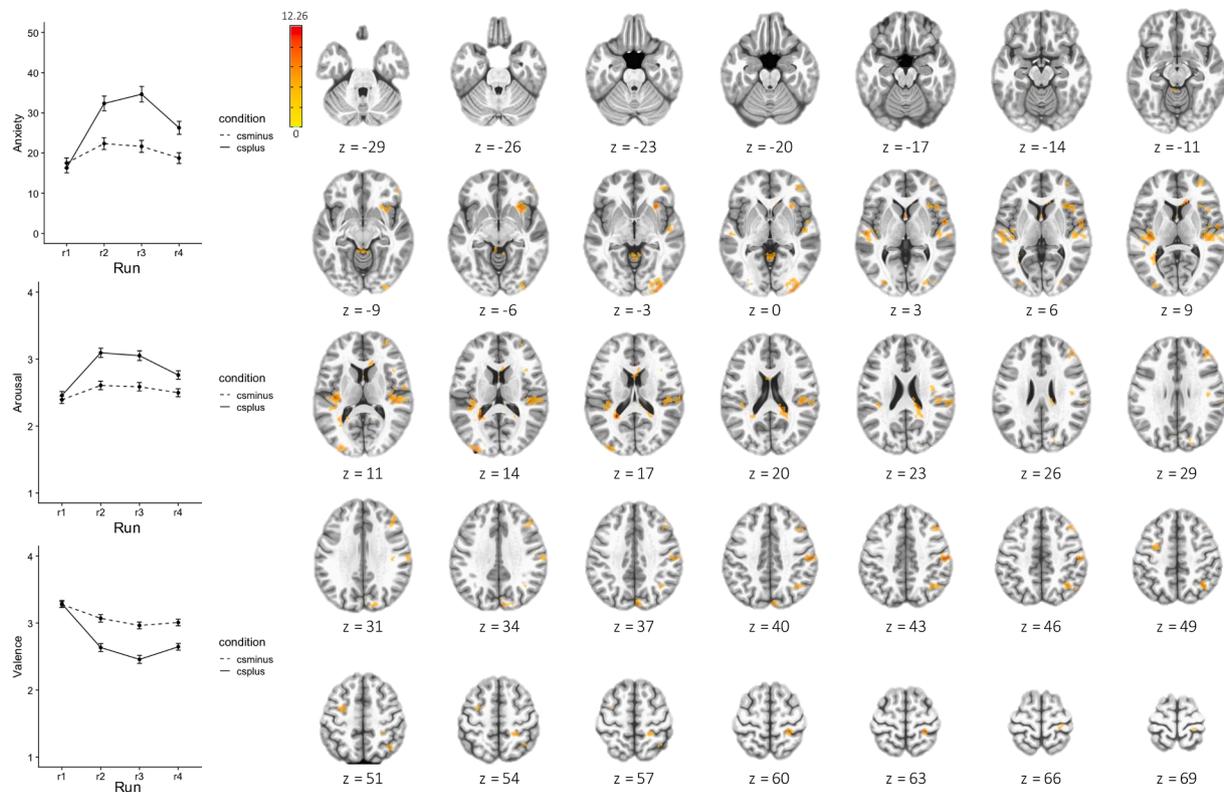


Fig. 2. Task ratings and activation network for condition (CS+, CS-) by time (Familiarization [R1], Early Conditioning [R2], Late Conditioning [R3], and Extinction [R4; split in two equal timepoints for fMRI data]) linear mixed effect model. The activation maps are projected on the MNI152 standard-space T1-weighted average structural template. Left is left.

responses to condition for amygdala ROIs yielded no effects of time, time², or their interactions with condition [(all p's > 0.1) Figure S5]. The rostral hippocampus evidenced a main effect of time [$F_{(4, 2529)} = 3.14, p = .014$], but not condition [$F_{(1, 2529)} = 0.09, p = .76$], qualified by condition-by-time interaction [$F_{(4, 2529)} = 3.74, p = .0049$; $BIC = -27725.74$]. The caudal hippocampus did not show a main effect of time [$F_{(4, 2529)} = 1.75, p = .14$], condition [$F_{(1, 2529)} = 0.17, p = .69$], nor condition-by-time interaction [$F_{(4, 2529)} = 1.10, p = .35$; $BIC = -29601.44$]. With respect to the vmPFC, there was a main effect of time [$F_{(4, 2529)} = 12.33, p = 6.2 \cdot 10^{-10}$], but not condition [$F_{(1, 2529)} = 0.13, p = .72$], qualified by condition-by-time interactions [$F_{(4, 2529)} = 3.20, p = .012$; $BIC = -25715.69$]. Finally, dorsal subregions of the dACC showed a main effect of time [$F_{(4, 2529)} = 9.51, p = 1.2 \cdot 10^{-7}$], but not condition [$F_{(1, 2529)} = 2.81, p = .093$], nor condition-by-time interaction [$F_{(4, 2529)} = 2.02, p = .089$; $BIC = -27040.49$], while the aMCC evidenced a main effect of time [$F_{(4, 2529)} = 3.29, p = .011$], but not condition [$F_{(1, 2529)} = 0.37, p = .54$], qualified by a condition-by-time interaction [$F_{(4, 2529)} = 2.43, p = .046, BIC = -26792.08$].

4.1. Latent profile analysis

The fit indices are reported in Table 3 and shown in Fig. 3. For subjective anxiety, the BIC suggested that a model with 5 latent subgroups fit the data best [$LRT_{(33)} = 33.42, p = .06, BIC = -7156.13$]; however, the non-significant bootstrapping results for 5 components indicated no further improvement of an additional component. Models with 3 and 4 components resulted in identical BIC: [3-component model: $LRT_{(21)} = 197.11, p = .001, BIC = -7182.43$]. A 2-component model was chosen for logistic regression analyses [$LRT_{(15)} = 341.71, p = .001, BIC = -7244.50$]. Regarding imaging data, the BIC suggested that a 2-component model was optimal and included right AI [$LRT_{(20)} = 113.18, p = .001, BIC = 11200.61$], right PG [$LRT_{(20)} = 33.14, p = .001, BIC = 11130.97$], right dIPFC [$LRT_{(20)} = 113.18, p = .001,$

$BIC = -10596.03$], bilateral STG [$LRT_{(20)} = 76.88, p = .001, BIC = 10378.37$], rostral hippocampus [$LRT_{(20)} = 77.89, p = .001, BIC = 11144.06$], vmPFC [$LRT_{(20)} = 91.92, p = .001, BIC = 10467.06$], and aMCC [$LRT_{(20)} = 84.79, p = .001, BIC = 10770.40$]. For additional information, please see the supplemental materials.

4.2. Logistic regression analyses

Table 4 shows regression coefficients, odds ratios, and the 95% confidence intervals for odds ratios for each predictor. Only the model distinguishing latent profile responses to CS+ in the right AI and STG were significant [$\chi^2(16, N = 228) = 42.86, p = .0003, BIC = 386.71$; $R^2 = 0.13$ and $\chi^2(16, N = 228) = 28.20, p = .0299, R^2 = 0.08, BIC = 415.12$; respectively]. Participant age ($B = -0.07, z = -4.63, p = 3.7 \cdot 10^{-6}$) and state anxiety ($B = -0.04, z = -2.03, p = .0427$) significantly predicted response profiles for right AI, while the response profiles for right STG were predicted by age only ($B = 0.04, z = 2.99, p = .0028$). See the supplemental materials for additional results from the logistic regression analyses.

5. Discussion

We examined behavioral and neural responses during Pavlovian fear learning in a large sample of healthy and individuals with predominant positive and negative valence dysfunction, that is anxiety disorders and/or depression. Taking a transdiagnostic, dimensional approach, we aimed to better understand psychological processes and neural substrates contributing to how fear is acquired and extinguished in these populations. Thus, in addition to examining behavioral and neural responses to fear learning, we derived latent profile models to threat (CS+) responses and examined whether various cognitive-affective processes distinguish between profiles.

The results are in line with past findings and show differentiated

Table 2

Regions of the brain showing differences in the hemodynamic response for main and interaction effects of condition and time.

Hemisphere / Location	Peak coordinates in MNI			F	Volume (mm ³)
Condition Main Effects					
R Inferior Temporal Gyrus	55	-67	-11	27.71	692
L Fusiform Gyrus / Cerebellum	-29	-67	-13	24.61	403
R Superior Occipital Gyrus	35	-73	47	15.15	295
R Anterior Insula	39	25	-3	32.56	293
L Inferior Parietal Lobule	-53	-29	37	17.99	247
L Anterior Insula	-29	17	-11	18.75	235
L and R Middle Cingulate Cortex (posterior)	-5	-21	31	16.38	166
L Calcarine Gyrus	1	-95	-1	16.43	149
L Precuneus	-9	-73	39	26.78	146
R Middle Frontal Gyrus	25	45	5	14.00	138
R Supramarginal Gyrus	63	-41	37	25.01	131
L Supramarginal Gyrus	-57	-49	31	15.63	89
Time Main Effects					
L and R Fusiform / Occipital Gyrus	29	-55	-7	52.27	13,253
L and R Superior Medial Gyrus / Anterior Cingulate Cortex	-1	61	11	16.10	2336
R Putamen / Caudate Nucleus	19	9	-9	12.19	1777
L Angular Gyrus	-47	-67	31	14.62	1584
R Angular Gyrus	53	-61	43	14.70	1390
L Inferior Frontal Gyrus / Insula	-37	21	7	16.56	1242
L and R Precuneus	7	-55	25	11.91	1095
L and R Middle Cingulate Cortex (posterior)	3	-21	41	12.94	564
L Inferior Frontal Gyrus / Postcentral Gyrus	65	5	17	10.15	219
L Middle Temporal Gyrus	-41	-57	11	8.46	181
R Inferior Frontal Gyrus / Precentral Gyrus	45	7	29	9.34	139
L Middle Temporal Gyrus	-41	-41	3	10.82	127
L Postcentral Gyrus	-49	-41	59	9.27	95
R Superior Frontal Gyrus	21	53	35	8.26	81
L Superior Parietal Lobule	-33	-43	25	8.61	78
L Precentral Gyrus	-47	9	31	7.45	76
R Thalamus / Hippocampus	23	-29	-1	11.01	73
R Middle Frontal Gyrus	41	47	3	8.11	69
L Postcentral Gyrus	-47	-21	51	7.98	69
L Middle Frontal Gyrus	-43	23	41	9.33	62
L Postcentral Gyrus	-59	-17	47	7.51	46
R Parahippocampal Gyrus	39	-43	-3	8.51	44
L Cerebellum	-47	-75	-23	7.08	40
L Cerebellum	-15	-53	-25	6.66	37
R Posterior Insula	35	-13	15	7.58	37
R Inferior Frontal Gyrus	45	31	17	6.30	37
R Postcentral Gyrus	47	-29	45	6.20	36
L Inferior Temporal Gyrus	-59	-25	-23	8.69	34
R Lingual Gyrus	11	-67	3	6.52	34
R Supplementary Motor Area	7	9	59	7.64	33
Condition by Time Interaction Effects					
R Postcentral Gyrus	61	-19	43	10.97	722
L Superior/Middle Temporal Gyrus	-55	-17	1	8.77	425
R Anterior Insula	31	21	-3	9.90	267
R Middle Occipital Gyrus	33	-99	-1	11.01	227
R Middle Frontal Gyrus	43	43	29	7.74	179
R Angular Gyrus	37	-57	49	7.63	179
L Middle Occipital Gyrus	-31	-93	13	10.04	155
R Thalamus-Temporal	17	-33	23	7.89	152
R Postcentral Gyrus	23	-41	57	7.47	126
R Cuneus	7	-85	39	7.32	117
R Superior Temporal Gyrus	55	-1	3	8.47	113
Cerebellum	-3	-41	-11	6.68	108
L Precuneus	-25	-49	17	12.26	86
L Middle Frontal Gyrus	-29	-3	51	6.85	83

Note: A voxel-wide threshold of $p < .005$ was set for the main effect of condition and condition by time interaction, and a voxel-wide threshold of $p < .001$ for the main effect of time. All significant activations passed a cluster size correction for multiple comparisons at $\alpha < 0.05$. L, left; R, right.

Table 3

Comparison of Gaussian mixture models on task behavioral and neural responses.

Measure	Solution	Parameters	BIC	LRT	p
Subjective Anxiety	1 component	9	-7384.26		
	2 component	15	-7244.50	341.71	0.001
	3 component	21	-7182.43	197.11	0.001
	4 component	21	-7182.43	79.46	0.001
	5 component	33	-7156.13	33.42	0.064
Right Anterior Insula	1 component	14	11144.93		
	2 component	20	11200.61	113.18	0.001
	3 component	20	11200.61	17.90	0.060
	4 component	20	11200.61		
	5 component	20	11200.61		
Right dlPFC	1 component	14	10444.93		
	2 component	20	10596.03	113.18	0.001
	3 component	20	10596.03	17.90	0.074
	4 component	20	10596.03		
	5 component	20	10596.03		
Right Postcentral Gyrus	1 component	14	11112.62	79.24	
	2 component	20	11130.97	33.14	0.001
	3 component	20	11130.97		0.054
	4 component	20	11130.97		
	5 component	20	11130.97		
Right Superior Temporal Gyrus	1 component	14	10363.49		
	2 component	20	10378.37	76.88	0.001
	3 component	20	10378.37	27.14	0.006
	4 component	20	10378.37	20.17	0.069
	5 component	20	10378.37		
Left Superior Temporal Gyrus	1 component	14	11297.58		
	2 component	20	11329.6	94.38	0.001
	3 component	20	11329.6	33.55	0.059
	4 component	20	11329.6		
	5 component	20	11329.6		
Rostral Hippocampus ROI	1 component	14	11135.41		
	2 component	20	11144.06	77.89	0.001
	3 component	20	11144.06	13.22	0.261
	4 component	20	11144.06		
Ventromedial PFC ROI	1 component	14	10430.2		
	2 component	20	10467.06	91.92	0.001
	3 component	20	10467.06		
	4 component	20	10467.06		
	5 component	20	10467.06		

(continued on next page)

Table 3 (continued)

Measure	Solution	Parameters	BIC	LRT	p
Anterior MCC ROI	2 component				
	3 component	20	10467.06	8.35	0.622
	4 component	20	10467.06		
	5 component	20	10467.06		
	1 component	14	10745.11		
	2 component	20	10770.4	84.79	0.001
	3 component	20	10770.4	16.12	0.133
	4 component	20	10770.4		
	5 component	20	10770.4		
	component				

Note: BIC, Bayesian information criterion; aMCC, anterior midcingulate cortex; dlPFC, dorsolateral prefrontal cortex; LRT, likelihood ration test; MCC, midcingulate cortex; PFC, prefrontal cortex.

subjective anxious responses to CS+ and CS- during fear conditioning and extinction periods, including a reduction in anxious responses to CS+ from conditioning to extinction trials. The LPA showed three distinct data-driven profiles of subjective anxious responses to CS+. While all profiles evidenced successful fear conditioning, the three profiles were distinguished from each other by the severity of subjective anxiety. Participants who reported the least amount of anxiety during conditioning trials, although showing evidence of fear conditioning relative to familiarization, did not evidence fear extinction. This pattern may be characteristic of impaired or reduced fear learning. The remaining two profiles likely evidenced what is considered typical fear learning, including extinction, with differences between the two profiles

observed in the intensity of anxious responses across all trials. Previous studies have identified differentiated profiles of subjective anxious responses characterized by elevations in state and trait anxiety, anxiety disorder diagnoses, and poor treatment response (Leen et al., 2021; Duits et al., 2021). The effort to identify demographic and affective descriptors accounting for differentiation in subjective anxious responses to CS+ in the present sample did not yield significant results and was therefore not consistent with previous studies.

Neuroimaging studies have consistently reported that fear learning paradigms activate a distributed network of brain regions, commonly known as the fear network (Sehlmeyer et al., 2009; Büchel and Dolan, 2000; Kim and Jung, 2006). This network includes the amygdala, AI, and ACC, with less robust findings for the vmPFC and hippocampus. Consistent with previous findings, our results show evidence for the role of AI, dlPFC, STG, dmPFC and vmPFC, aMCC, rostral hippocampus, somatosensory cortex, thalamus, precuneus, and cerebellum in fear learning.

Contrary to previous studies (Öhman, 2009; Phelps et al., 2004), we did not find evidence for the role of the amygdala in fear learning. Specifically, both whole-brain and ROI analyses did not show evidence of amygdala change in BOLD signal across conditioning and extinction trials as a function of threat conditions (CS+, CS-). Although it is widely believed that amygdala plays a central role in fear learning, data are inconsistent across studies. Experimental factors (e.g., type of stimuli, type of conditioning, patient populations, and analysis approach) may, in part, account for discordant findings (Sehlmeyer et al., 2009; Greco and Liberzon, 2016). It has also been argued that the amygdala is important for initial conditioning of fear, following which its activity reduces with time (Schiller et al., 2008; LaBar et al., 1998). However, others have observed the opposite, that is, recruitment of the amygdala during late conditioning phases (Sehlmeyer et al., 2011). While our data support the notion that the amygdala activity habituates as the experiment continues, we did not observe CS+/CS- differences, even in the early conditioning trials. In fact, amygdala activation appeared to be

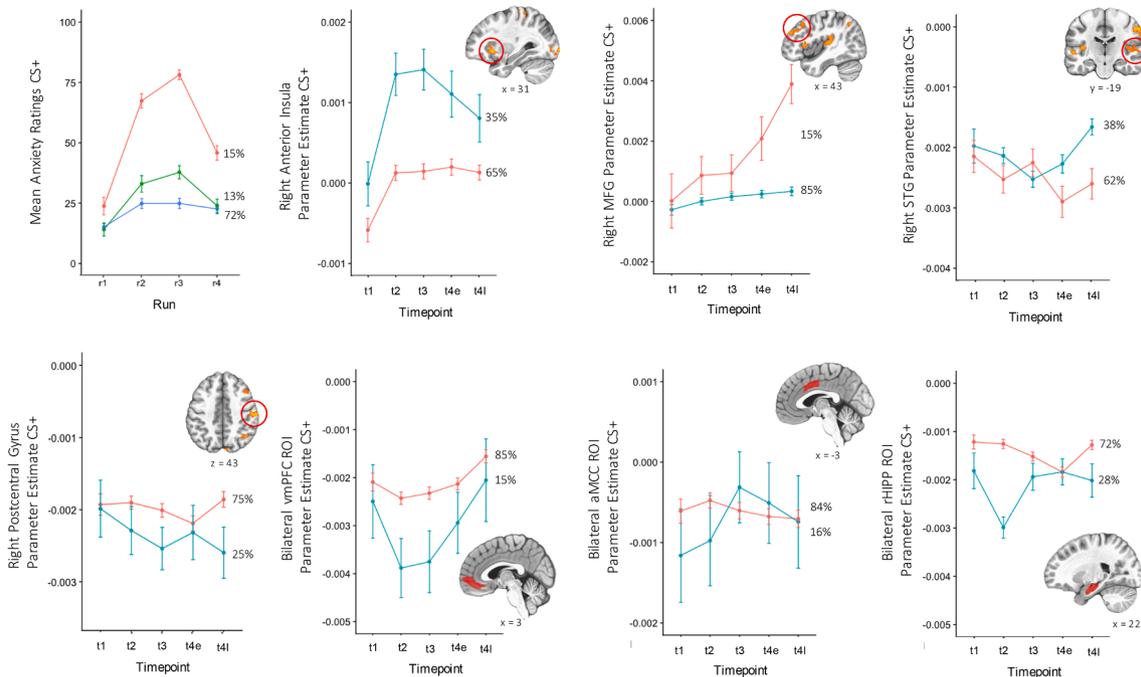


Fig. 3. Results of the latent profile analysis on behavioral and neural responses of interest to the conditioned stimulus (CS+). Lines represent distinct response profiles for models with optimal significant number of components. Behavioral runs include Familiarization (r1), Early Conditioning (r2), Late Conditioning (r3), and Extinction (r4). fMRI timepoints include Familiarization (t1), Early Conditioning (t2), Late Conditioning (t3), and Extinction (early [t4e] and late [t4I]). The activation maps for the whole brain analysis and Brainnetome atlas defined regions of interests are projected on the MNI152 standard-space T1-weighted average structural template. Left is left. aMCC, anterior midcingulate cortex; MFG, medial frontal gyrus; rHIPP, rostral hippocampus; superior temporal gyrus, (STG); vmPFC, ventromedial prefrontal cortex.

Table 4
Logistic regression analysis.

Model / Variable	B (SE)	Odds Ratio	95 % Confidence Interval for Odds Ratio	
			Lower	Upper
Subjective Anxiety				
Age	-0.01 (0.02)	0.99	0.95	1.03
Sex	-0.33 (0.52)	0.72	0.25	1.95
Psychotropic Medication Status	0.93 (0.46)	2.54	1.06	6.54
Body Percent Fat	-0.02 (0.02)	0.98	0.94	1.03
WRAT Reading	-0.02 (0.04)	0.98	0.91	1.07
PHQ-9	-0.06 (0.05)	0.95	0.86	1.03
OASIS	-0.03 (0.06)	0.97	0.86	1.10
Neuroticism	0.09 (0.05)	1.10	1.00	1.21
Rumination	-0.03 (0.02)	0.97	0.93	1.01
State Anxiety	0.02 (0.02)	1.02	0.98	1.07
Anxiety Sensitivity	0.01 (0.01)	1.01	0.98	1.04
Inhibition	-0.04 (0.07)	0.96	0.83	1.11
Reward Responsivity	0.19 (0.11)	1.21	0.99	1.51
Negative Affect	0.01 (0.04)	1.01	0.94	1.09
Positive Affect	0.01 (0.03)	1.01	0.95	1.08
Childhood Trauma	-0.01 (0.01)	0.99	0.97	1.02
Right Insula				
Age	-0.07 (0.02)	0.93	0.90	0.96
Sex	0.43 (0.38)	1.53	0.72	3.28
Psychotropic Medication Status	-0.20 (0.33)	0.82	0.43	1.57
Body Percent Fat	0.02 (0.01)	1.02	0.99	1.06
WRAT Reading	-0.01 (0.03)	0.99	0.94	1.06
PHQ-9	-0.03 (0.04)	0.97	0.90	1.04
OASIS	-0.02 (0.05)	0.98	0.89	1.08
Neuroticism	0.02 (0.04)	1.02	0.95	1.10
Rumination	-0.01 (0.02)	0.99	0.96	1.03
State Anxiety	-0.04 (0.02)	0.96	0.93	1.00
Anxiety Sensitivity	-0.01 (0.01)	0.99	0.97	1.02
Inhibition	0.08 (0.06)	1.08	0.96	1.22
Reward Responsivity	0.13 (0.08)	1.13	0.97	1.33
Negative Affect	0.05 (0.03)	1.05	0.99	1.11
Positive Affect	-0.05 (0.03)	0.95	0.91	1.00
Childhood Trauma	-0.00 (0.01)	1.00	0.98	1.02
Right Medial Frontal Gyrus				
Age	-0.04 (0.02)	0.96	0.92	0.99
Sex	-0.50 (0.49)	0.61	0.22	1.56
Psychotropic Medication Status	0.15 (0.43)	1.16	0.50	2.74
Body Percent Fat	-0.01 (0.02)	0.99	0.94	1.03
WRAT Reading	0.07 (0.05)	1.07	0.98	1.18
PHQ-9	-0.03 (0.05)	0.97	0.87	1.06
OASIS	-0.03 (0.06)	0.97	0.86	1.11
Neuroticism	-0.04 (0.05)	0.96	0.88	1.05
Rumination	-0.02 (0.02)	0.98	0.94	1.02
State Anxiety	0.01 (0.02)	1.01	0.96	1.06
Anxiety Sensitivity	0.00 (0.02)	1.00	0.97	1.03
Inhibition	-0.05 (0.07)	0.95	0.83	1.09
Reward Responsivity	0.23 (0.11)	1.26	1.02	1.58
Negative Affect	0.06 (0.04)	1.07	0.99	1.15
Positive Affect	-0.03 (0.03)	0.97	0.92	1.03
Childhood Trauma	0.01 (0.01)	1.01	0.99	1.04
Right Postcentral Gyrus				
Age	0.01 (0.02)	1.01	0.98	1.04
Sex	-0.64 (0.40)	0.53	0.24	1.17
Psychotropic Medication Status	0.10 (0.35)	1.10	0.56	2.16
Body Percent Fat	-0.02 (0.02)	0.98	0.95	1.02
WRAT Reading	-0.02 (0.03)	0.99	0.92	1.05
PHQ-9	-0.02 (0.04)	0.98	0.91	1.06
OASIS	0.05 (0.05)	1.05	0.95	1.17
Neuroticism	0.05 (0.04)	1.06	0.98	1.14
Rumination	-0.01 (0.02)	0.99	0.96	1.02
State Anxiety	0.01 (0.02)	1.01	0.97	1.04
Anxiety Sensitivity	0.00 (0.01)	1.00	0.98	1.03
Inhibition	-0.17 (0.06)	0.84	0.74	0.95
Reward Responsivity	0.06 (0.08)	1.06	0.90	1.25
Negative Affect	-0.01 (0.03)	0.99	0.93	1.05
Positive Affect	-0.04 (0.03)	0.96	0.91	1.01
Childhood Trauma	-0.01 (0.01)	0.99	0.98	1.01
Right Superior Temporal Gyrus				
Age	0.04 (0.01)	1.04	1.01	1.07

Table 4 (continued)

Model / Variable	B (SE)	Odds Ratio	95 % Confidence Interval for Odds Ratio	
			Lower	Upper
Sex	-0.44 (0.37)	0.64	0.31	1.32
Psychotropic Medication Status	0.16 (0.32)	1.17	0.63	2.19
Body Percent Fat	-0.02 (0.02)	0.98	0.95	1.02
WRAT Reading	0.03 (0.03)	1.03	0.97	1.09
PHQ-9	0.04 (0.03)	1.05	0.98	1.12
OASIS	0.05 (0.05)	1.05	0.96	1.15
Neuroticism	-0.02 (0.03)	0.98	0.92	1.05
Rumination	0.00 (0.02)	1.00	0.97	1.03
State Anxiety	0.03 (0.02)	1.03	0.99	1.06
Anxiety Sensitivity	-0.00 (0.01)	1.00	0.97	1.02
Inhibition	0.00 (0.05)	1.00	0.91	1.12
Reward Responsivity	-0.02 (0.07)	0.98	0.85	1.13
Negative Affect	-0.04 (0.03)	0.96	0.91	1.01
Positive Affect	-0.02 (0.02)	0.98	0.94	1.03
Childhood Trauma	-0.01 (0.01)	0.99	0.97	1.00
Left Superior Temporal Gyrus				
Age	-0.01 (0.01)	0.99	0.96	1.01
Sex	0.76 (0.36)	2.13	1.06	4.38
Psychotropic Medication Status	-0.27 (0.31)	0.76	0.41	1.42
Body Percent Fat	0.01 (0.02)	1.01	0.98	1.04
WRAT Reading	0.04 (0.03)	1.04	0.98	1.10
PHQ-9	-0.01 (0.03)	0.99	0.93	1.06
OASIS	0.04 (0.05)	1.04	0.95	1.14
Neuroticism	-0.07 (0.03)	0.93	0.87	1.00
Rumination	0.02 (0.02)	1.02	0.99	1.05
State Anxiety	-0.01 (0.02)	0.99	0.96	1.02
Anxiety Sensitivity	0.01 (0.01)	1.01	0.98	1.03
Inhibition	0.11 (0.05)	1.12	1.01	1.25
Reward Responsivity	0.03 (0.07)	1.03	0.89	1.19
Negative Affect	0.00 (0.03)	1.00	0.95	1.06
Positive Affect	0.03 (0.02)	1.03	0.99	1.08
Childhood Trauma	0.01 (0.01)	1.01	0.99	1.03
Bilateral vmPFC				
Age	-0.00 (0.02)	1.00	0.96	1.03
Sex	-0.58 (0.47)	0.56	0.22	1.41
Psychotropic Medication Status	0.24 (0.43)	1.27	0.54	2.99
Body Percent Fat	0.03 (0.02)	1.03	0.98	1.08
WRAT Reading	-0.06 (0.05)	0.94	0.85	1.02
PHQ-9	0.02 (0.05)	1.02	0.93	1.13
OASIS	-0.06 (0.07)	0.94	0.83	1.08
Neuroticism	-0.00 (0.05)	1.00	0.91	1.10
Rumination	-0.01 (0.02)	0.99	0.95	1.04
State Anxiety	-0.01 (0.02)	0.99	0.94	1.03
Anxiety Sensitivity	0.01 (0.02)	1.01	0.98	1.04
Inhibition	0.04 (0.07)	1.04	0.90	1.20
Reward Responsivity	-0.13 (0.10)	0.88	0.72	1.06
Negative Affect	-0.04 (0.04)	0.96	0.89	1.04
Positive Affect	0.03 (0.03)	1.03	0.97	1.10
Childhood Trauma	0.02 (0.01)	1.02	0.99	1.05
Bilateral aMCC				
Age	0.01 (0.02)	1.01	0.98	1.05
Sex	-0.10 (0.47)	0.91	0.35	2.26
Psychotropic Medication Status	-0.70 (0.42)	0.50	0.22	1.14
Body Percent Fat	0.02 (0.02)	1.02	0.98	1.07
WRAT Reading	-0.00 (0.04)	1.00	0.92	1.08
PHQ-9	-0.04 (0.04)	0.96	0.88	1.05
OASIS	0.02 (0.06)	1.02	0.91	1.15
Neuroticism	0.01 (0.04)	1.01	0.93	1.10
Rumination	0.02 (0.02)	1.02	0.98	1.06
State Anxiety	0.01 (0.02)	1.01	0.96	1.06
Anxiety Sensitivity	-0.02 (0.02)	0.98	0.95	1.01
Inhibition	-0.05 (0.06)	0.95	0.84	1.08
Reward Responsivity	0.19 (0.11)	1.21	0.99	1.50
Negative Affect	-0.01 (0.04)	0.99	0.92	1.06
Positive Affect	0.01 (0.03)	1.01	0.95	1.07
Childhood Trauma	0.00 (0.01)	1.00	0.98	1.02
Bilateral Rostral Hippocampus				
Age	0.01 (0.01)	1.01	0.98	1.04
Sex	-1.13 (0.38)	0.32	0.15	0.68
Psychotropic Medication Status	0.28 (0.34)	1.32	0.67	2.58
Body Percent Fat	-0.02 (0.02)	0.98	0.95	1.02
WRAT Reading	-0.02 (0.03)	0.98	0.92	1.04

(continued on next page)

Table 4 (continued)

Model / Variable	B (SE)	Odds Ratio	95 % Confidence Interval for Odds Ratio	
			Lower	Upper
PHQ-9	0.00 (0.04)	1.00	0.93	1.08
OASIS	0.06 (0.05)	1.07	0.97	1.18
Neuroticism	-0.03 (0.04)	0.97	0.90	0.04
Rumination	-0.02 (0.02)	0.98	0.95	1.02
State Anxiety	0.03 (0.02)	1.03	0.99	1.07
Anxiety Sensitivity	0.01 (0.01)	1.01	0.98	1.04
Inhibition	0.01 (0.05)	1.01	0.91	1.13
Reward Responsivity	-0.01 (0.008)	0.99	0.84	1.15
Negative Affect	-0.02 (0.03)	0.99	0.93	1.05
Positive Affect	-0.02 (0.02)	0.98	0.94	1.03
Childhood Trauma	-0.01 (0.01)	0.99	0.97	1.00

Abbreviations: OASIS, Overall Anxiety Severity and Impairment Scale; PHQ-9, Patient Health Questionnaire – 9 items; WRAT, Wide Range Achievement Test.

most elevated during the familiarization phase. This finding is consistent with previous research characterizing the role of amygdala in novelty detection. Studies have found increased amygdala responses to novel presentation of a range of visual stimuli, and particularly those that are less familiar and more ambiguous in their predicted outcomes (Blackford et al., 2010). This is the case for both emotionally and neutrally valenced stimuli (Balderston et al., 2011; Balderston et al., 2013; Pedersen et al., 2017). Similarly, previous research has also shown rapid amygdala habituation over time (Yin et al., 2018; Büchel et al., 1999; Büchel et al., 1998). Finally, it is possible that the continuous performance task within our paradigm exerted a significant cognitive load on the participant, thus suppressing amygdala response, a phenomenon which has been previously reported (Van Dillen et al., 2009; Kellermann et al., 2012). It is also noteworthy that the present sample uniquely consisted of predominantly patients with depressive and/or anxious dysfunction. While we did not find differences in fear learning response for individuals with depression and/or anxiety disorders versus healthy controls, further research is warranted to delineate how amygdala response may or may not relate to specific aspects of mental health symptoms or specific diagnoses. Nevertheless, we propose that fear learning in humans is largely an explicit learning process that extends beyond amygdala-governed novelty detection and associations between salient stimuli, in turn involving complex cognitive and emotional processes that rely on regions such as the AI and dlPFC.

Our results are in line with recent meta-analyses on fear conditioning (Fullana et al., 2016) and extinction (Fullana et al., 2018), which point to a network of regions that represent autonomic, interoceptive, cognitive, motivation, and psychomotor processes. Specifically, we found robust activations during fear learning in the somatosensory cortex, STG, and AI, likely reflecting representation of body states and generation of emotional states during presentation of aversive stimuli. The human insula is believed to play a role in bottom-up detection of salient events, integration with other large-brain networks to gain attention and working memory resources for further processing of salient events, modulation of consequent autonomic reactivity, and facilitation of appropriate behavioral responses to salient events via its strong functional connections with the aMCC (Menon and Uddin, 2010; Deen et al., 2011). Specific insular subdivisions map onto discrete functions related to these processes, such that the AI is thought to be involved in the integration of affective states with past knowledge of salient stimuli, context information, and expected impact and outcomes of salient events (Uddin et al., 2017; Craig and Craig, 2009; Kurth et al., 2010). Overall, AI activation evidenced fear conditioning responses, while the LPA analysis showed two distinct profiles for AI responses to threat during fear learning. Specifically, while both profiles evidenced greater signal during conditioning and extinction trials relative to familiarization, a profile consisting of 35% of participants exhibited greater AI

activation across all time points, with evident relative differentiation in activation between conditioning and extinction trials. Participants exhibiting this pattern were relatively younger (i.e., 31 vs. 37 years of age on average) and reported lower, albeit clinically negligible, levels of state anxiety. Although greater AI activation relating to lower state anxiety is a somewhat an unexpected result, this specific response profile resembled the typical fear conditioning and extinction response pattern.

Previous studies have implicated dlPFC in associative learning (Fletcher et al., 2001; Corlett et al., 2004). Compared to healthy control participants, patients with dlPFC lesions have been shown to successfully acquire conditioned threat but exhibit impairments in regulating subjective fear (Kroes et al., 2019). Taken together, this implicates dlPFC as a critical region in regulating subjective anxiety responses to CS+ and CS- during fear conditioning and extinction periods. Meta-analytic results implicate dlPFC particularly during extinction learning and recall (Fullana et al., 2018), likely reflecting effortful, top-down regulation of subcortical structures and thereby affective autonomic and behavioral responses (Ochsner et al., 2012; Delgado et al., 2008). Our results evidenced greater recruitment of the dlPFC during late extinction relative to conditioning trials in response to CS+, particularly pronounced in a subgroup of 15% of participants. Although our findings corroborate the role of dlPFC in fear extinction, we did not identify cognitive-affective predictors of latent profiles of threat responses during fear learning for the dlPFC.

The results of the present study evidenced two distinct profiles of responses during threat (CS+) processing for vmPFC and rostral hippocampus, characterized by differences particularly during conditioning trials. Furthermore, the pattern of CS+ responses in the vmPFC shows relatively greater recruitment during extinction trials. These two regions have been implicated in inhibition of fear responses during extinction learning and recall, and notably during contextual retrieval of conditioned stimuli (Milad et al., 2007; Schiller et al., 2008; Hartley and Phelps, 2010; Kalisch et al., 2006; Pennington et al., 2017; Harrison et al., 2017; Schiller and Delgado, 2010). Specifically, vmPFC and hippocampus may be crucial for distinguishing between threatening and non-threatening stimuli, thereby allowing for successful safety learning. However, similar to dlPFC, we did not identify cognitive-affective predictors of latent profiles of threat responses during fear learning for the vmPFC and rostral hippocampus.

Data-driven analyses have been proposed as promising in discerning mechanisms of fear learning. While we clearly identified distinct profiles of fear learning, current analyses did not establish robust relationships between these profiles and demographic, clinical, and other self-report variables. Therefore, our results are inconclusive in explaining which individual characteristics may contribute to the onset and persistence of fear-learning related psychopathology. Nevertheless, this does not preclude the potential for such individual differences in fear learning to be meaningful in understanding clinical outcomes, such as serving as predictors of mechanisms of treatment outcome or response to acute stress and trauma. Indeed, the identified brain networks suggest that fear learning in humans involves complex processes related to sensorimotor appraisal and responses, as well as integration across cognitive, affective, and motivational domains. It must also be recognized that diagnosis and treatment of mood, stress, and anxiety disorders involve not only the count of unconditioned and conditioned events, their habituation, and successful identification of safety signals, but also evaluation and modulation of beliefs, contexts, and values among others. Therefore, experimental paradigms that account for some or all of these processes may be better at relating neuroimaging findings to self-report data.

6. Limitations

First, while cross-sectional studies with large transdiagnostic samples are valuable in increasing understanding of mechanisms underlying fear learning, longitudinal studies are better suited to delineate how

variations in fear learning bring about psychopathology and potential effects of treatment. Second, the present analysis was based on 282 individuals from the first half of the released participant data from the T-1000 study. Future analyses will allow for replication of these findings on the second 500 participants. Third, although inclusive respective to depression and anxiety disorders, the sample did not consist of diagnostic categories for which fear learning is also relevant, including obsessive-compulsive, bipolar, and personality disorders. Broader inclusion of psychiatric presentations, as well as differentiating fear and anxious responses, would increase the clinical utility of fear learning processes. Fourth, we did not directly assess the aversiveness of the US, and only inferred it through assessment of image valence. Finally, the paradigm did not include extinction recall trials, thereby leaving out an important aspect of fear learning from analysis.

7. Conclusions

The present study demonstrates that fear learning activates a distributed network of regions involved in interoceptive, cognitive, motivational, and psychomotor processes in this predominantly depressed and/or anxious sample. We did not find support for the role of amygdala in fear learning in this transdiagnostic sample, further evidencing that development and maintenance of depressive and anxious psychopathology involves complex cognitive and emotional processes. Data-driven analyses identified distinct profiles of subjective and neural responses during fear learning. Robust relationships between response profiles and cognitive-affective variables of interest were not identified, though longitudinal designs are needed to explore the relevance of these profiles to clinical prediction or treatment mechanisms. Future studies aiming to understand mechanisms underlying disrupted fear-learning in patient populations may also be improved by including clinically relevant features into experimental designs.

CRedit authorship contribution statement

Namık Kirlic: Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Rayus Kuplicki:** Methodology, Investigation, Data curation, Formal analysis, Validation. **James Touthang:** Data curation, Formal analysis. **Zsofia P. Cohen:** Writing – original draft, Writing – review & editing. **Jennifer L. Stewart:** Writing – original draft, Writing – review & editing. **Martin P. Paulus:** Supervision, Funding acquisition, Writing – original draft, Writing – review & editing. **Robin L. Aupperle:** Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has been supported in part by The William K. Warren Foundation and the National Institute of General Medical Sciences Center Grant Award Number P20GM121312. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The ClinicalTrials.gov identifier for the clinical protocol associated with data published in the current paper is NCT02450240, “Latent Structure of Multi-level Assessments and Predictors of Outcomes in Psychiatric Disorders”. The Tulsa 1000 Investigators include the following contributors: *Robin Aupperle, Ph.D., Jerzy Bodurka, Ph.D., Salvador Guinjoan, M.D., Sahib S. Khalsa, M.D., Ph.D., Rayus Kuplicki, Ph.D., Martin P. Paulus, M.D., Jonathan Savitz, Ph.D., Jennifer Stewart, Ph.D., and Teresa A. Victor, Ph.D.*

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nicl.2022.103060>.

References

- Acheson, D.T., Eyler, L.T., Resovsky, J., Tsan, E., Risbrough, V.B., 2015. Fear extinction memory performance in a sample of stable, euthymic patients with bipolar disorder. *J. Affect. Disord.* 185, 230–238.
- Alvarez, R.P., Kirlic, N., Misaki, M., Bodurka, J., Rhudy, J.L., Paulus, M.P., Drevets, W.C., 2015. Increased anterior insula activity in anxious individuals is linked to diminished perceived control. *Transl Psychiatry.* 5 (6), e591.
- Balderston, N.L., Schultz, D.H., Helmstetter, F.J., 2011. The human amygdala plays a stimulus specific role in the detection of novelty. *Neuroimage.* 55 (4), 1889–1898.
- Balderston, N.L., Schultz, D.H., Helmstetter, F.J., Krueger, F., 2013. The effect of threat on novelty evoked amygdala responses. *PLoS ONE* 8 (5), e63220.
- Ball, T.M., Knapp, S.E., Paulus, M.P., Stein, M.B., 2017. Brain activation during fear extinction predicts exposure success. *Depression Anxiety.* 34 (3), 257–266.
- Berntson, G.G., Khalsa, S.S., 2021. Neural circuits of interoception. *Trends Neurosci.* 44 (1), 17–28.
- Blackford, J.U., Buckholz, J.W., Avery, S.N., Zald, D.H., 2010. A unique role for the human amygdala in novelty detection. *Neuroimage.* 50 (3), 1188–1193.
- Büchel, C., Dolan, R.J., 2000. Classical fear conditioning in functional neuroimaging. *Curr. Opin. Neurobiol.* 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.
- Büchel, C., Dolan, R.J., Armony, J.L., Friston, K.J., 1999. Amygdala–hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *J. Neurosci.* 19 (24), 10869–10876.
- Carver, C.S., White, T.L., 1994. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *J. Pers. Soc. Psychol.* 67 (2), 319.
- Core Team R. R., 2017. A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/> [Google Scholar], Vienna, Austria.
- Corlett, P.R., Aitken, M.R., Dickinson, A., et al., 2004. Prediction error during retrospective reevaluation of causal associations in humans: fMRI evidence in favor of an associative model of learning. *Neuron* 44 (5), 877–888.
- Cox, R.W., Chen, G., Glen, D.R., Reynolds, R.C., Taylor, P.A., 2017. FMRI clustering in AFNI: false-positive rates redux. *Brain Connect.* 7 (3), 152–171.
- Craig, A.D., Craig, A., 2009. How do you feel–now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10 (1).
- Croissant Y. Estimation of Random Utility Models in R: The mlogit Package. *Journal of Statistical Software.* 10/07 2020;95(11):1 - 41. doi:10.18637/jss.v095.i11.
- Davis, M., Whalen, P.J., 2001. The amygdala: Vigilance and emotion. *Mol. Psychiatry* 6 (1), 13–34. <https://doi.org/10.1038/sj.mp.4000812>.
- Deen, B., Pitskel, N.B., Pelphrey, K.A., 2011. Three systems of insular functional connectivity identified with cluster analysis. *Cereb. Cortex* 21 (7), 1498–1506.
- Delgado, M.R., Nearing, K.I., Ledoux, J.E., Phelps, E.A., 2008. Neural circuitry underlying the regulation of conditioned fear and its relation to extinction. *Neuron* 59 (5), 829–838. <https://doi.org/10.1016/j.neuron.2008.06.029>.
- Duits, P., Cath, D.C., Lissek, S., Hox, J.J., Hamm, A.O., Engelhard, I.M., van den Hout, M.A., Baas, J.M.P., 2015. Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression Anxiety.* 32 (4), 239–253.
- Duits, P., Baas, J.M.P., Engelhard, I.M., Richter, J., Huisman - van Dijk, H.M., Limberg-Thiesen, A., Heitland, I., Hamm, A.O., Cath, D.C., 2021. Latent class growth analyses reveal overrepresentation of dysfunctional fear conditioning trajectories in patients with anxiety-related disorders compared to controls. *J. Anxiety Disord.* 78, 102361.
- Dunsmoor, J.E., Prince, S.E., Murty, V.P., Kragel, P.A., LaBar, K.S., 2011. Neurobehavioral mechanisms of human fear generalization. *Neuroimage.* 55 (4), 1878–1888.
- Etkin, A., Egner, T., Kalisch, R., 2011. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cognitive Sci.* 15 (2), 85–93. <https://doi.org/10.1016/j.tics.2010.11.004>.
- Etkin, A., Wager, T.D., 2007. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am. J. Psychiatry.* 164 (10), 1476–1488.
- Fan, L., Li, H., Zhuo, J., Zhang, Y.u., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A.R., Fox, P.T., Eickhoff, S.B., Yu, C., Jiang, T., 2016. The human brainnetome atlas: a new brain atlas based on connective architecture. *Cereb. Cortex* 26 (8), 3508–3526.
- Fletcher, P.C., Anderson, J.M., Shanks, D.R., Honey, R., Carpenter, T.A., Donovan, T., Papadakis, N., Bullmore, E.T., 2001. Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nat. Neurosci.* 4 (10), 1043–1048.
- Fredrikson, M., Wik, G., Fischer, H., Andersson, J., 1995. Affective and attentive neural networks in humans: a PET study of Pavlovian conditioning. *Neuroreport: An Int. J. Rapid Commun. Res. Neurosci.* 7 (9), 977–1011.
- Fullana, M.A., Harrison, B.J., Soriano-Mas, C., Vervliet, B., Cardoner, N., Ávila-Parcet, A., Radua, J., 2016. Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Mol. Psychiatry* 21 (4), 500–508.
- Fullana, M.A., Albajes-Eizaguirre, A., Soriano-Mas, C., Vervliet, B., Cardoner, N., Benet, O., Radua, J., Harrison, B.J., 2018. Fear extinction in the human brain: a

- meta-analysis of fMRI studies in healthy participants. *Neurosci. Biobehav. Rev.* 88, 16–25.
- Ganella, D.E., Drummond, K.D., Ganella, E.P., Whittle, S., Kim, J.H., 2018. Extinction of conditioned fear in adolescents and adults: a human fMRI study. *Front. Hum. Neurosci.* 11, 647.
- Greco, J.A., Liberzon, I., 2016. Neuroimaging of fear-associated learning. *Neuropsychopharmacology*. 41 (1), 320–334.
- Harrison, B.J., Fullana, M.A., Via, E., Soriano-Mas, C., Vervliet, B., Martínez-Zalacain, I., Pujol, J., Davey, C.G., Kircher, T., Straube, B., Cardoner, N., 2017. Human ventromedial prefrontal cortex and the positive affective processing of safety signals. *Neuroimage*. 152, 12–18.
- Hartley, C.A., Phelps, E.A., 2010. Changing fear: The neurocircuitry of emotion regulation. *Neuropsychopharmacology*. 35 (1), 136–146. <https://doi.org/10.1038/npp.2009.121>.
- Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J. Clin. Psychiatry*. 59 (Suppl 20), 2233.
- Hermans, D., Craske, M.G., Mineka, S., Lovibond, P.F., 2006. Extinction in human fear conditioning. *Biol. Psychiatry* 60 (4), 361–368.
- Holt, D.J., Coombs, G., Zeidan, M.A., Goff, D.C., Milad, M.R., 2012. Failure of neural responses to safety cues in schizophrenia. *Arch. Gen. Psychiatry* 69 (9), 893–903.
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D.S., Quinn, K., Sanislow, C., Wang, P., 2010. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am. J. Psychiatry*. 167 (7), 748–751.
- John, O.P., Srivastava, S., 1999. The Big-Five trait taxonomy: History, measurement, and theoretical perspectives, vol 2. University of California Berkeley.
- Jovanovic, T., Norrholm, S.D., Blanding, N.Q., Davis, M., Duncan, E., Bradley, B., Ressler, K.J., 2010. Impaired fear inhibition is a biomarker of PTSD but not depression. *Depression Anxiety*. 27 (3), 244–251.
- Kalisch, R., Gerlicher, A.M., 2014. Making a mountain out of a molehill: On the role of the rostral dorsal anterior cingulate and dorsomedial prefrontal cortex in conscious threat appraisal, catastrophizing, and worrying. *Neurosci. Biobehav. Rev.* 42, 1–8.
- Kalisch, R., Korenfeld, E., Stephan, K.E., Weiskopf, N., Seymour, B., Dolan, R.J., 2006. Context-dependent human extinction memory is mediated by a ventromedial prefrontal and hippocampal network. *J. Neurosci.* 26 (37), 9503–9511.
- Kellermann, T.S., Sternkopf, M.A., Schneider, F., Habel, U., Turetsky, B.I., Zilles, K., Eickhoff, S.B., 2012. Modulating the processing of emotional stimuli by cognitive demand. *Social Cognitive Affective Neurosci.* 7 (3), 263–273.
- Kim, J.J., Jung, M.W., 2006. Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. *Neurosci. Biobehav. Rev.* 30 (2), 188–202.
- Kindt, M., 2014. A behavioural neuroscience perspective on the aetiology and treatment of anxiety disorders. *Behav. Res. Ther.* 62, 24–36.
- Kirlic, N., Aupperle, R.L., Rhudy, J.L., Misaki, M., Kuplicki, R., Sutton, A., Alvarez, R.P., 2019. Latent variable analysis of negative affect and its contributions to neural responses during shock anticipation. *Neuropsychopharmacology*. 44 (4), 695–702.
- Krause-Utz, A., Keibel-Mauchnik, J., Ebner-Priemer, U., Bohus, M., Schmahl, C., 2016. Classical conditioning in borderline personality disorder: an fMRI study. *Eur. Arch. Psychiatry Clin. Neurosci.* 266 (4), 291–305.
- Kroenke, K., Spitzer, R.L., Williams, J.B., 2001. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 16 (9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>.
- Kroes, M.C.W., Dunsmoor, J.E., Hakimi, M., Oosterwaal, S., Meager, M.R., Phelps, E.A., 2019. Patients with dorsolateral prefrontal cortex lesions are capable of discriminatory threat learning but appear impaired in cognitive regulation of subjective fear. *Social Cognitive Affective Neurosci.* 14 (6), 601–612.
- Kurth, F., Zilles, K., Fox, P.T., Laird, A.R., Eickhoff, S.B., 2010. A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Struct. Funct.* 214 (5–6), 519–534.
- 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.
- Lebrun-Milad, K., Abbs, B., Milad, M.R., Linnman, C., Rougemont-Bücking, A., Zeidan, M.A., Holt, D.J., Goldstein, J.M., 2012. Sex differences in the neurobiology of fear conditioning and extinction: a preliminary fMRI study of shared sex differences with stress-arousal circuitry. *Biology Mood Anxiety Disorders*. 2 (1).
- LeDoux, J.E., 2000. Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23 (1), 155–184.
- LeDoux, J.E., 2014. Coming to terms with fear. *Proc. Natl. Acad. Sci.* 111 (8), 2871–2878.
- Leen, N.A., Duits, P., Baas, J.M.P., 2021. Trajectories of fear learning in healthy participants are able to distinguish groups that differ in individual characteristics, chronicity of fear and intrusions. *J. Behav. Ther. Exp. Psychiatry* 72, 101653.
- Lueken, U., Straube, B., Reinhardt, I., Maslowski, N.I., Wittchen, H.-U., Ströhle, A., Wittmann, A., Pfeleiderer, B., Konrad, C., Ewert, A., Uhlmann, C., Arolt, V., Jansen, A., Kircher, T., 2014. Altered top-down and bottom-up processing of fear conditioning in panic disorder with agoraphobia. *Psychol. Med.* 44 (2), 381–394.
- Marin, M.-F., Hammoud, M.Z., Klumpp, H., Simon, N.M., Milad, M.R., 2020. Multimodal categorical and dimensional approaches to understanding threat conditioning and its extinction in individuals with anxiety disorders. *JAMA Psychiatry*. 77 (6), 618–627.
- McTeague, L.M., Rosenberg, B.M., Lopez, J.W., Carreon, D.M., Huemer, J., Jiang, Y., Chick, C.F., Eickhoff, S.B., Etkin, A., 2020. Identification of common neural circuit disruptions in emotional processing across psychiatric disorders. *Am. J. Psychiatry* 177 (5), 411–421.
- Mechias, M., Etkin, A., Kalisch, R., 2010. A meta-analysis of instructed fear studies: Implications for conscious appraisal of threat. *Neuroimage*. 49 (2), 1760–1768. <https://doi.org/10.1016/j.neuroimage.2009.09.040>.
- Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214 (5–6), 655–667.
- Merz, C.J., Wolf, O.T., Schweckendiek, J., Kluckner, T., Vaitl, D., Stark, R., 2013. Stress differentially affects fear conditioning in men and women. *Psychoneuroendocrinology*. 38 (11), 2529–2541.
- Milad, M.R., Quirk, G.J., 2012. Fear extinction as a model for translational neuroscience: ten years of progress. *Annu. Rev. Psychol.* 63, 129–151. <https://doi.org/10.1146/annurev.psych.121208.131631>.
- 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. *Biol. Psychol.* 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. *Biol. Psychiatry* 62 (5), 446–454.
- Milad, M.R., Furtak, S.C., Greenberg, J.L., Keshaviah, A., Im, J.J., Falkenstein, M.J., Jenike, M., Rauch, S.L., Wilhelm, S., 2013. Deficits in conditioned fear extinction in obsessive-compulsive disorder and neurobiological changes in the fear circuit. *JAMA Psychiatry*. 70 (6), 608.
- Milad, M.R., Rosenbaum, B.L., Simon, N.M., 2014. Neuroscience of fear extinction: implications for assessment and treatment of fear-based and anxiety related disorders. *Behav. Res. Ther.* 62, 17–23.
- Morriss, J., Christakou, A., Van Reekum, C.M., 2015. Intolerance of uncertainty predicts fear extinction in amygdala-ventromedial prefrontal cortical circuitry. *Biology of Mood & Anxiety Disorders*. 5 (1), 1–13.
- 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.
- Norman, S.B., Hami Cissell, S., Means-Christensen, A.J., Stein, M.B., 2006. Development and validation of an overall anxiety severity and impairment scale (OASIS). *Depression Anxiety*. 23 (4), 245–249.
- Ochsner, K.N., Silvers, J.A., Buhle, J.T., 2012. Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Ann. N. Y. Acad. Sci.* 1251, E1–E24. <https://doi.org/10.1111/j.1749-6632.2012.06751.x>.
- Ohman, A., 2008. Fear and anxiety: Overlaps and dissociations. In: Lewis, M., Haviland-Jones, J.M., Barrett, L.F. (Eds.), *Handbook of emotions*, (3rd ed). Guilford Press, pp. 709–728.
- Öhman, A., 2009. Human fear conditioning and the amygdala. The Guilford Press.
- Otto, M.W., Moshier, S.J., Kinner, D.G., Simon, N.M., Pollack, M.H., Orr, S.P., 2014. De novo fear conditioning across diagnostic groups in the affective disorders: evidence for learning impairments. *Behav. Ther.* 45 (5), 619–629.
- Panitz, C., Sperl, M.F., Hennig, J., Klucken, T., Hermann, C., Mueller, E.M., 2018. Fearfulness, neuroticism/anxiety, and COMT Val158Met in long-term fear conditioning and extinction. *Neurobiol. Learn. Mem.* 155, 7–20.
- Pavlov, P.I., 2010. Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. *Annals Neurosci.* 17 (3), 136.
- Pedersen, W.S., Balderston, N.L., Miskovich, T.A., Belleau, E.L., Helmstetter, F.J., Larson, C.L., 2017. The effects of stimulus novelty and negativity on BOLD activity in the amygdala, hippocampus, and bed nucleus of the stria terminalis. *Social Cognitive Affective Neurosci.* 12 (5), 748–757.
- Pejic, T., Hermann, A., Vaitl, D., Stark, R., 2013. Social anxiety modulates amygdala activation during social conditioning. *Social Cognitive Affective Neuroscience*. 8 (3), 267–276.
- Pennington, Z.T., Anderson, A.S., Fanselow, M.S., 2017. The ventromedial prefrontal cortex in a model of traumatic stress: fear inhibition or contextual processing? *Learning Memory*. 24 (9), 400–406.
- 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.
- Quirk, G.J., Mueller, D., 2008. Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology*. 33 (1), 56–72.
- Sandi, C., Richter-Levin, G., 2009. From high anxiety trait to depression: A neurocognitive hypothesis. *Trends Neurosci.* 32 (6), 312–320. <https://doi.org/10.1016/j.tins.2009.02.004>.
- Savage, H.S., Davey, C.G., Fullana, M.A., Harrison, B.J., 2020. Clarifying the neural substrates of threat and safety reversal learning in humans. *Neuroimage*. 207, 116427.
- Schiller, D., Delgado, M.R., 2010. Overlapping neural systems mediating extinction, reversal and regulation of fear. *Trends Cognitive Sci.* 14 (6), 268–276.
- 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. *J. Neurosci.* 28 (45), 11517–11525.
- Schmitz, A., Grillon, C., 2012. Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test). *Nat. Protoc.* 7 (3), 527–532.
- Schreurs, B.G., Bahro, M., Molchan, S.E., Sunderland, T., McIntosh, A.R., 2001. Interactions of prefrontal cortex during eyeblink conditioning as a function of age. *Neurobiol. Aging* 22 (2), 237–246.
- Scrucca, L., Fop, M., Murphy, T.B., Raftery, A.E., Aug 2016. mclust 5: clustering, classification and density estimation using gaussian finite mixture models. *R Journal*. 8 (1), 289–317.
- Sehlmeier, C., Schöning, S., Zwitserlood, P., Pfeleiderer, B., Kircher, T., Arolt, V., Konrad, C., Gendelman, H.E., 2009. Human fear conditioning and extinction in neuroimaging: a systematic review. *PLoS ONE* 4 (6), e5865.
- Sehlmeier, C., Dannlowski, U., Schöning, S., Kugel, H., Pyka, M., Pfeleiderer, B., Zwitserlood, P., Schifflauer, H., Heindel, W., Arolt, V., Konrad, C., 2011. Neural correlates of trait anxiety in fear extinction. *Psychol. Med.* 41 (4), 789–798.

- Smith, R., Thayer, J.F., Khalsa, S.S., Lane, R.D., 2017. The hierarchical basis of neurovisceral integration. *Neurosci. Biobehav. Rev.* 75, 274–296.
- Spielberger, C.D., 1983. Manual for the state-trait anxiety inventory STAI (form Y) (“self-evaluation questionnaire”). Consulting Psychology Press.
- Sylvers, P., Lilienfeld, S.O., LaPrairie, J.L., 2011. Differences between trait fear and trait anxiety: Implications for psychopathology. *Clinical Psychology Review.* 31 (1), 122–137.
- Taylor, S., Zvolensky, M.J., Cox, B.J., Deacon, B., Heimberg, R.G., Ledley, D.R., Abramowitz, J.S., Holaway, R.M., Sandin, B., Stewart, S.H., Coles, M., Eng, W., Daly, E.S., Arrindell, W.A., Bouvard, M., Cardenas, S.J., 2007. Robust dimensions of anxiety sensitivity: Development and initial validation of the Anxiety Sensitivity Index-3. *Psychol. Assess.* 19 (2), 176–188.
- Treynor, W., Gonzalez, R., Nolen-Hoeksema, S., 2003. Rumination reconsidered: A psychometric analysis. *Cognitive Therapy Research.* 27 (3), 247–259.
- Uddin, L.Q., Nomi, J.S., Hébert-Seropian, B., Ghaziri, J., Boucher, O., 2017. Structure and function of the human insula. *J. Clinical Neurophysiology: Official Publication American Electroencephalographic Society.* 34 (4), 300–306.
- Van Dillen, L.F., Heslenfeld, D.J., Koole, S.L., 2009. Tuning down the emotional brain: an fMRI study of the effects of cognitive load on the processing of affective images. *Neuroimage.* 45 (4), 1212–1219.
- Verdi, S., Marquand, A.F., Schott, J.M., Cole, J.H., 2021. Beyond the average patient: how neuroimaging models can address heterogeneity in dementia. *Brain* 144 (10), 2946–2953.
- Victor, T.A., Khalsa, S.S., Simmons, W.K., et al., 2018. Tulsa 1000: a naturalistic study protocol for multilevel assessment and outcome prediction in a large psychiatric sample. *BMJ open.* 8 (1), e016620.
- Vogt, B.A., Berger, G.R., Derbyshire, S.W., 2003. Structural and functional dichotomy of human midcingulate cortex. *Eur. J. Neurosci.* 18 (11), 3134–3144.
- Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. *J. Pers. Soc. Psychol.* 54 (6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>.
- Wicking, M., Steiger, F., Nees, F., Diener, S.J., Grimm, O., Ruttorf, M., Schad, L.R., Winkelmann, T., Wirtz, G., Flor, H., 2016. Deficient fear extinction memory in posttraumatic stress disorder. *Neurobiol. Learn. Mem.* 136, 116–126.
- Yágüez, L., Coen, S., Gregory, L.J., Amaro, E., Altman, C., Brammer, M.J., Bullmore, E.T., Williams, S.C.R., Aziz, Q., 2005. Brain response to visceral aversive conditioning: a functional magnetic resonance imaging study. *Gastroenterology* 128 (7), 1819–1829.
- Yin S, Liu Y, Petro NM, Keil A, Ding M. Amygdala adaptation and temporal dynamics of the salience network in conditioned fear: a single-trial fMRI study. *Eneuro.* 2018;5 (1).