Archival Report

Biological Psychiatry: GOS

Hippocampal Involvement in Safety Signal Learning Varies With Anxiety Among Healthy Adults

Paola Odriozola, Sahana Kribakaran, Emily M. Cohodes, Sadie J. Zacharek, Sarah McCauley, Jason T. Haberman, Lauren A. Quintela, Cristian Hernandez, Hannah Spencer, Luise Pruessner, Camila Caballero, and Dylan G. Gee

ABSTRACT

BACKGROUND: Safety signal learning (SSL), based on conditioned inhibition of fear in the presence of learned safety, can effectively attenuate threat responses in animal models and humans. Difficulty regulating threat responses is a core feature of anxiety disorders, suggesting that SSL may provide a novel mechanism for fear reduction. Cross-species evidence suggests that SSL involves functional connectivity between the anterior hippocampus and the dorsal anterior cingulate cortex. However, the neural mechanisms supporting SSL have not been examined in relation to trait anxiety or while controlling for the effect of novelty.

METHODS: Here, we investigated the neural mechanisms involved in SSL and associations with trait anxiety in a sample of 64 healthy (non-clinically anxious) adults (ages 18–30 years; 43 female, 21 male) using physiological, behavioral, and neuroimaging (functional magnetic resonance imaging) data collected during an SSL task.

RESULTS: During SSL, compared with individuals with lower trait anxiety, individuals with higher trait anxiety showed less fear reduction as well as altered hippocampal activation and hippocampal–dorsal anterior cingulate cortex functional connectivity, and lower inferior frontal gyrus and ventrolateral prefrontal cortex activation. Importantly, the findings show that SSL reduces threat responding, across learning and over and above the effect of novelty, and involves hippocampal activation.

CONCLUSIONS: These findings provide new insights into the nature of SSL and suggest that there may be meaningful variation in SSL and related neural correlates as a function of trait anxiety, with implications for better understanding fear reduction and optimizing interventions for individuals with anxiety disorders.

https://doi.org/10.1016/j.bpsgos.2023.05.007

Safety signal learning (SSL) refers to the ability to associate specific environmental stimuli with the non-occurrence of aversive events (1,2) and has been shown to effectively reduce threat responding in rodents (3), nonhuman primates (4), and healthy adult humans (1,5,6). Difficulty regulating threat responses and discriminating between threat and safety are core features of anxiety disorders (7–10), which are the most common psychiatric illnesses, affecting up to one-third of the population (1). The primary evidence-based behavioral treatment for anxiety, cognitive behavioral therapy, is based on principles of fear extinction learning and can be highly effective at reducing fear. However, up to 50% of patients do not benefit sufficiently (11,12), highlighting the need to augment fear reduction. Research that examines basic mechanisms underlying fear reduction, including investigations of normative anxiety in healthy humans, has the potential to inform optimization of current interventions for individuals with anxiety disorders. Past research in healthy humans (3,4,5,13,14), humans with posttraumatic stress disorder [(15); for a review, see (16)], and nonhuman animals (2,6) has suggested that SSL may serve as a novel mechanism for augmenting fear reduction. However, there is a crucial gap in the literature because the mechanisms that support SSL have not been examined in relation to normative anxiety.

In SSL, a stimulus is trained through Pavlovian conditioning to signal safety (or the absence of threat); as a result, this safety cue can inhibit the conditioned fear response. As a special class of conditioned inhibition, SSL requires that a safety cue inhibit the conditioned response as a result of learning (as opposed to the process by which a stimulus can inhibit the conditioned response without training, called external inhibition) (1). Research with rodents and nonhuman primates has relied on two procedures to test whether a stimulus acts as a conditioned inhibitor. During the summation test, the threat and safety cues are presented simultaneously as a compound stimulus (safety compound), yielding a reduction in threat-related behavior. During the retardation test, the safety cue is paired with the unconditioned stimulus (7). If a safety cue has been learned effectively, threat responding should be slower to emerge (relative to initial

ISSN: 2667-1743 Biological Psychiatry: Global Open Science January 2024; 4:155–164 www.sobp.org/GOS

conditioning). The current study focuses on the summation test of conditioned inhibition in humans. It is worth noting that while nonhuman tasks are typically conducted on the order of days, human studies are generally conducted on the order of hours, which may mean that the tasks rely on different memory systems (8).

The hippocampus is a highly interconnected region that contributes to the complex regulation of threat responding by segmenting information about an environmental stimulus and distributing this information to different regions of the brain (6,9). Given the central role of the hippocampus in contextual fear learning (10–12) and the fact that this region's projections modulate fronto-amygdala function by supplying information about the degree of threat or safety in the environment (9,17,18), the hippocampus has been hypothesized to be important for conditioned inhibition.

Indeed, recent cross-species evidence has demonstrated the involvement of the ventral hippocampus in rodents and the anterior hippocampus in humans during conditioned inhibition via learned safety (6). More specifically, ventral hippocampal neurons that project to the prelimbic cortex—but not to the infralimbic cortex or the basolateral amygdala (BLA)—showed higher activation during conditioned inhibition, and this activation was associated with lower freezing behavior (thought to be an index of threat responsivity) in mice. A corresponding distinction was observed in humans such that functional connectivity between the anterior hippocampus and dorsal anterior cingulate cortex (dACC)—but not hippocampal–anterior ventromedial prefrontal cortex (PFC) or hippocampal-BLA connectivity—was associated with conditioned inhibition. Importantly, this evidence suggests that the neural circuitry involved in conditioned inhibition (i.e., hippocampus-dACC connectivity in humans) differs from the regions that are typically involved in extinction (i.e., ventromedial PFC–amygdala connectivity in humans) (19–21). Previous research has shown that hippocampal inputs to the prelimbic cortex are capable of suppressing, or gating, fear expression (9), further highlighting this as a potential alternative pathway supporting fear reduction.

Several major gaps in knowledge about SSL exist that, if filled, could have clinical implications. First, while studies of nonhuman animals and behavioral studies of humans have included meaningful control conditions to evaluate the effects of SSL (2,3), the only neuroimaging investigation of SSL in humans (6) focused solely on comparing the safety compound with a threat cue (comprising a single stimulus) to more closely parallel the paradigm in rodents included in the same study. This design precluded the ability to rule out the possibility that the effects of safety signals were driven by external inhibition (e.g., due to the compound nature of the stimulus). Second, examining relationships between SSL and trait anxiety in a healthy sample is critical to advance the potential clinical utility of this work. While evidence across species has shown disrupted SSL in adult rodents (22) and human adults with posttraumatic stress disorder [(13); for a review, see (14)] and trauma exposure (13), the mechanisms that support SSL have not been examined in relation to normative anxiety. Lastly, studies on the neural bases of SSL have primarily relied on examinations of specific regions in isolation, such as lesions or fiber photometry in a specific brain region in nonhuman animals or region of interest (ROI) analyses in humans [for a review, see (2)]. While important findings have emerged from this body of work, such as the potential involvement of the ventral hippocampal–prelimbic cortex pathway (6), a more exploratory approach investigating the neural correlates of SSL across the whole brain has the potential to yield a richer understanding of other regions or circuits in the brain that may contribute to SSL (1,14).

The current study aimed to address these gaps in existing findings on the physiological and neural correlates of SSL in healthy adult humans by 1) directly comparing physiological reactivity and neural activation between the safety compound and a compound control condition to isolate the effect of the safety signal itself on conditioned inhibition and investigating learning over the course of the task; 2) examining physiological and neural associations between SSL and individual differences in trait anxiety; and 3) more broadly investigating the neural substrates of SSL in healthy human adults. First, we hypothesized that safety signals would reduce physiological reactivity (as measured via skin conductance). Based on a previous study in humans (6), we also hypothesized that SSL would engage the right anterior hippocampus and left hippocampal connectivity with the dACC over the course of learning, and that these effects would be significant over and above the effect of novelty. More specifically, we tested the effect of the safety signal over time, the safety signal versus threat condition, and the safety signal versus novelty on skin conductance response (SCR), hippocampal activation, and hippocampal connectivity. We hypothesized that there would be a reduction in SCR and an increase in hippocampal activation and functional connectivity over the course of learning in response to the safety signal and that these would be associated with anxiety. We also hypothesized that the effect of the safety signal on SCR, hippocampal activation, and connectivity would occur over and above the effect of novelty and the effect of the threat cue after learning had occurred. Second, when examining relations between SSL and individual differences in trait anxiety, we hypothesized that safety signals would be less effective for reducing physiological reactivity among individuals with higher anxiety and that individuals with higher anxiety would show lower hippocampal activation and lower hippocampal-dACC functional connectivity during SSL. Finally, we used exploratory whole-brain approaches with activation and functional connectivity to investigate the potential involvement of broader neural circuitry in fear reduction via SSL, particularly among individuals with higher levels of anxiety.

METHODS AND MATERIALS

Participants and Study Procedures

The study sample comprised 64 adults ages 18 to 30 years (Table 1) with no magnetic resonance imaging (MRI) contraindications and no current or past psychiatric diagnoses or use of psychotropic medications. All procedures were approved by the institutional review board at Yale University, and all participants provided written informed consent. The State-Trait Anxiety Inventory (STAI) (23,24) was used to assess trait anxiety. Additional details for all methods are provided in the Supplement.

Table 1. Demographic Characteristics in the Total Neuroimaging Sample, *n* = 64

Variable	Mean (SD), *n* (%), or *n*	Range
Age, Years	23.15 (3.38)	18–30
Sex at Birth, Female:Male	43:21	–
Trait Anxiety, STAI-T	36.13 (8.76)	22–54
Race/Ethnicity[a]		
Asian	13 (20.3%)	–
Black or African American	11 (17.2%)	–
Hawaiian/Pacific Islander	0 (0%)	–
Hispanic/Latino	7 (10.9%)	–
White, non-Hispanic	35 (54.7%)	–
Other/unknown	1 (1.6%)	–
Motion-Related Measures, mm[b]		
Absolute motion	0.52 (0.35)	0.09–1.70
Relative motion	0.08 (0.03)	0.03–0.14
Framewise displacement	0.13 (0.04)	0.06–0.24
Outlier time points, No.	16.39 (11.91)	0–57
Outlier time points, %	3.15% (2.29%)	0%–10.96%

STAI-T, State-Trait Anxiety Inventory–Trait scale (24).
[a]Includes mixed race within each category.
[b]Following motion exclusion (*n* = 64).

SSL Task Design

The SSL task (6) was adapted from the AX+/BX− task of conditioned inhibition (3,5) designed to be used specifically with children and adolescents in related studies. Conditioned stimuli were neutral geometric shapes of different colors; the unconditioned stimulus was an aversive metallic noise (25) delivered at 95 to 100 dB through MRI-safe noise-canceling headphones (Figure 1). Participants completed this task in the scanner while functional MRI (fMRI) and SCR data were acquired.

MRI Acquisition Parameters

Participants were scanned on a 3T Siemens Magnetom Prisma scanner (Siemens Medical Solutions). Scan parameters were based on the Adolescent Brain Cognitive Development (ABCD) Study (26). A whole-brain high-resolution T1-weighted anatomical scan magnetization-prepared rapid acquisition gradient-echo was acquired for each subject. During the SSL task, high spatial and temporal resolution multiband echo planar imaging fMRI scans were collected across a total of 5 runs.

Acquisition and Analysis of Physiological Data

SCR data were collected during the SSL task in the MRI scanner using an MRI-compatible Biopac system and AcqKnowledge software (biopac.com/product/acqknowledge-software/). Following collection, 2 independent coders visually inspected SCR data and excluded some data for poor quality (see the Supplement). This resulted in a subsample of *n* = 27 participants for the SCR analyses (see Table S1 for subsample demographics). SCR data were analyzed using PsychoPhysiological Modeling software (pspm.sourceforge.net), and the resulting reconstructed SCR values for each condition using early and late measurements of SCR within each task phase were entered into SPSS (version 28; IBM Corp.) for statistical analyses. These values were subjected to a repeated-measures analysis of variance for each task phase. A separate repeated-measures analysis of covariance was conducted including trait anxiety scores as a covariate of interest (see the Supplement for more details).

Analysis of fMRI Data

fMRI Preprocessing. Raw neuroimaging data for the *n* = 64 sample were converted to Brain Imaging Data Structure (27) using heudiconv (github.com/nipy/heudiconv) and preprocessed with the HCP (Human Connectome Project) minimal

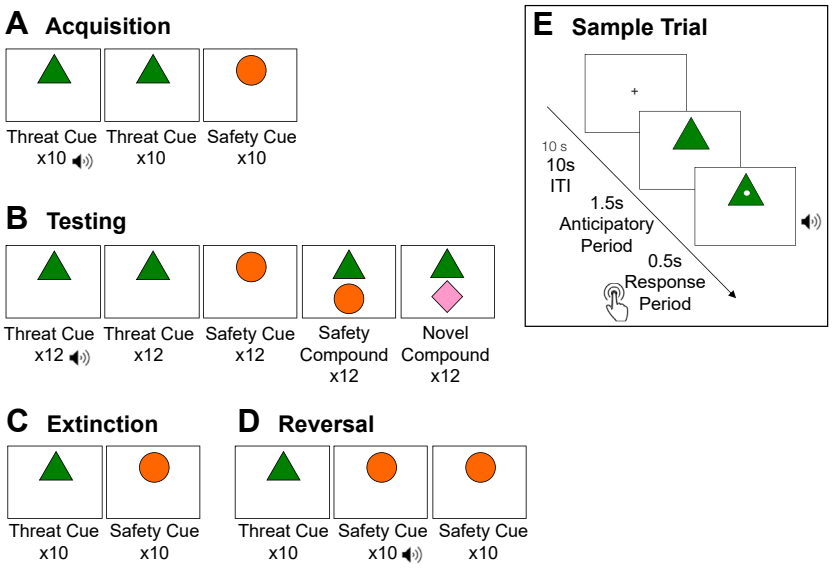

Figure 1. Safety signal learning task design. **(A)** The acquisition phase of the task included a threat cue, which was reinforced by the unconditioned stimulus (US) (an aversive sound) on 50% of trials, and a safety cue that was never reinforced. **(B)** The testing phase included the following cues: reinforced threat, nonreinforced threat, safety, safety compound (i.e., paired threat and safety cues), and novel compound (i.e., paired threat and novel cues). **(C)** In the extinction phase, participants were presented with repeated exposures of the previously learned threat cue without the US and continued to see the previously learned safety cue without the US. **(D)** In the reversal phase, the roles of the shapes were reversed such that the previously learned safety cue was reinforced by the US on 50% of trials and the previously learned threat cue was presented without a US. The current study focused on the testing phase (see the Supplement for more details on the extinction and reversal phases). **(E)** Timing of each trial. Participants were instructed to make a button press when the dot appeared on each trial. The US onset co-occurred with the response period on reinforced trials. Geometric shapes were counterbalanced across conditions between participants. ITI, intertrial interval.

preprocessing pipeline (28) using the HCP Pipelines Brain Imaging Data Structure app (github.com/BIDS-Apps/HCPPipelines) version 3.17.14. See the Supplement for details on data quality assessment.

Individual-Level fMRI Analyses. fMRI analyses were completed using FEAT version 6.00, which is part of FSL (fsl.fmrib.ox.ac.uk/fsl/fslwiki) version 5.11.

Activation: ROI Analyses. Mean percent signal change values were extracted for each subject and each task condition using anatomical masks for the right and left anterior hippocampus (29). These values were then subjected to separate repeated-measures analysis of variances in SPSS. Follow-up analyses for functional activation and functional connectivity focused on testing our hypotheses regarding differences in the safety compound condition over the course of learning, the safety compound versus threat contrast, and the safety compound versus novel compound contrast. See the Supplement for more details.

Activation: Whole-Brain Analysis. To analyze the testing phase at the whole-brain level, higher-level analyses were conducted in FEAT to compute the group mean for each contrast of interest. A separate higher-level analysis was also conducted with a regressor for trait anxiety (mean-centered STAI-Trait scale scores). Final statistical images were thresholded in FEAT ($z > 3.1$, cluster $p < .05$, the most up-to-date FSL defaults).

Functional Connectivity Analyses. To analyze task-evoked functional connectivity between the hippocampus and dACC, a generalized psychophysiological interaction model (30) was conducted with FILM autocorrelation correction (fsl.fmrib.ox.ac.uk/fsl/fslwiki/FEAT).

RESULTS

Skin Conductance Response

Time-Related Reduction in Reactivity Specific to SSL. There was a significant effect of time (i.e., early vs. late trials during the testing phase) on SCR ($F_{1,26} = 6.39$, $p = .018$) but no main effect of condition ($F_{2.15,55.92} = 2.13$, $p = .125$). As predicted, planned pairwise comparisons revealed a significant difference between late and early trials of the safety compound condition (i.e., SCR to the safety compound decreased over time) ($t_{26} = -3.14$, false discovery rate–corrected p [p_{FDR}] = .006). Furthermore, there was a significant difference between late trials of the safety compound and novel compound conditions such that SCR to the safety compound was lower than SCR to the novel compound ($t_{26} = -2.28$, $p_{FDR} = .031$). Lastly, SCR was lower during the late trials of the safety compound versus the threat condition ($t_{26} = -2.56$, $p_{FDR} = .006$) (Figure 2).

Individual Differences in Trait Anxiety During SSL. There was a significant quadratic contrast for the interaction between condition and trait anxiety on SCR ($F_{1,25} = 7.23$, $p_{FDR} = .030$). There was no main effect of trait anxiety ($ps > .05$) and no significant interaction between time and trait anxiety or between condition, time, and trait anxiety ($p > .05$). Follow-up analyses for the interaction of condition and trait anxiety revealed a significant positive correlation between trait anxiety and SCR to the safety compound [$r_{25} = 0.402$, $p = .038$] such that individuals with higher anxiety showed higher reactivity to the safety compound than individuals with lower anxiety. There was also a significant positive correlation between trait anxiety and SCR to the difference in SCR during the safety compound versus threat [$r_{25} = 0.392$, $p = .043$] such that relative to individuals with lower anxiety, individuals with higher anxiety showed higher reactivity to the safety compound compared with the threat condition. There was no correlation of anxiety with the safety compound versus the novel compound ($p > .05$) (Figure 3).

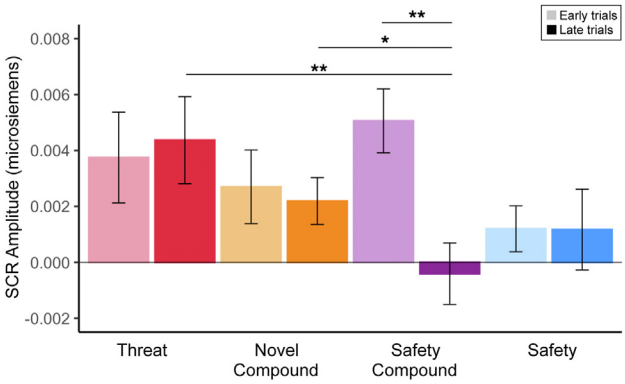

Figure 2. Skin conductance response (SCR) during early and late trials of the safety signal learning task. The SCR reconstructed response amplitude values (in microsiemens) were plotted for each condition during early (first 3) and late (last 3) trials within the first run of the testing phase of the safety signal learning task. There was a significant main effect of time ($p = .018$). SCR was significantly lower during the late vs. early trials of the safety compound, during late trials of the safety compound vs. the novel compound, and during late trials of the safety compound vs. the threat condition. *False discovery rate–corrected $p < .05$. **False discovery rate–corrected $p < .01$.

A Priori ROI Analyses of Activation

Hippocampal Activation. For the right hippocampus, there was a significant main effect of condition ($F_{2.59,163.19} = 2.948$, $p = .042$) and a main effect of time ($F_{2.68,168.93} = 7.80$, $p < .001$). To follow up on the main effect of condition in the overall model, we conducted pairwise comparisons for the overall mean right hippocampal activation for the safety compound versus threat and safety compound versus the novel compound. These comparisons revealed higher mean right hippocampal engagement during the safety compound compared with the threat condition ($t_{63} = 2.31$, $p = .024$) but no significant difference between the safety compound and novel compound conditions ($t_{63} = 0.42$, $p = .678$) (Figure 4).

Based on a priori hypotheses, planned follow-up analysis of variances were conducted for each contrast of interest (i.e., safety compound over time, safety compound vs. threat, and safety compound vs. novel compound). For the model including only the safety compound, there was a main effect of time ($F_{3,189} = 5.25$, $p_{FDR} = .003$) such that hippocampal engagement increased over the course of learning. For the

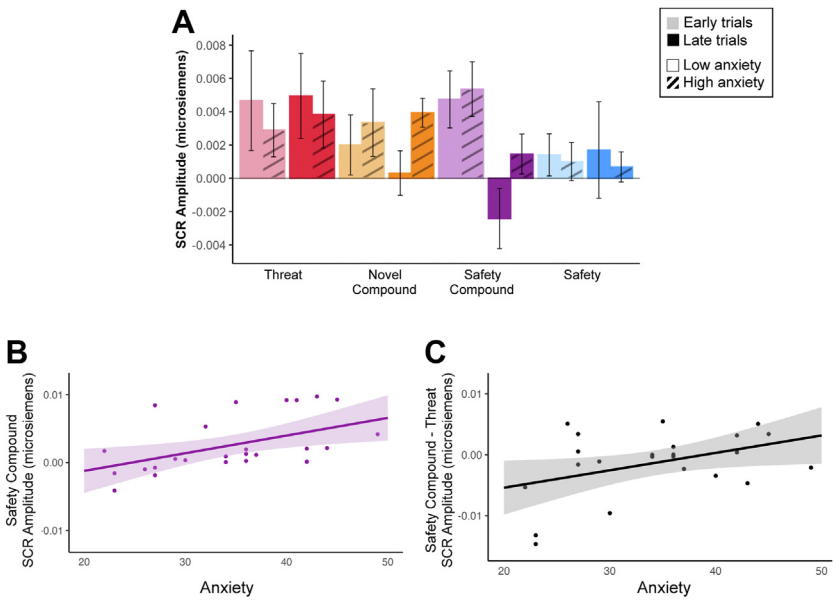

Figure 3. Skin conductance response (SCR) during early and late trials of the safety signal learning task by trait anxiety. **(A)** There was a significant quadratic contrast of a condition and trait anxiety interaction (false discovery rate–corrected $p = .030$). Participants were separated into high and low groups using a median split on trait anxiety for visualization purposes only. **(B)** There was a significant positive correlation between trait anxiety and SCR during the safety compound condition ($p = .038$). **(C)** There was a significant positive correlation between trait anxiety and SCR during the safety compound vs. threat conditions ($p = .043$).

model comparing the safety compound and threat, there was a significant main effect of time ($F_{3,189} = 6.58$, $p_{FDR} = .003$) but no effect of condition ($F_{1,63} = 4.69$, $p_{FDR} = .102$) and no interaction between condition and time ($p_{FDR} > .05$). Lastly, for the model comparing the safety compound and novel compound, there was a significant effect of time ($F_{2.61,164.63} = 5.14$, $p_{FDR} = .003$) but no main effect of condition or interaction between condition and time ($ps_{FDR} > .05$).

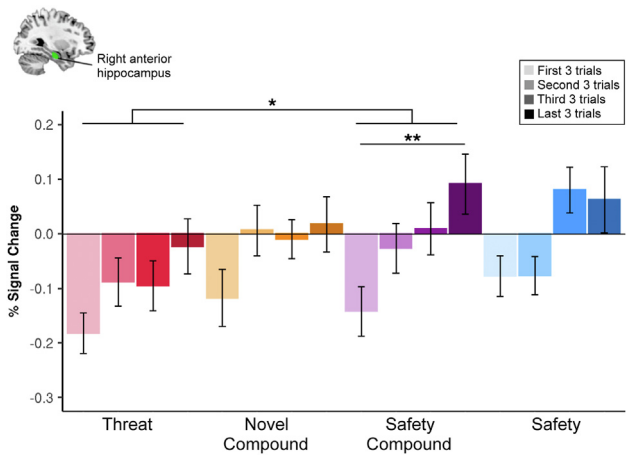

Figure 4. Right hippocampal activation during early and late trials of the safety signal learning task. Percent signal change (parameter estimates) is shown on the y-axis for each quarter of the safety signal learning task (i.e., one measurement every 3 trials of the task). There was a significant main effect of condition and a main effect of time. Planned follow-up comparisons revealed a significant positive linear effect of time on hippocampal engagement during the safety compound condition. Hippocampal activation was also higher during the safety compound condition than the threat condition when averaged across the entire task. *False discovery rate–corrected $p < .05$. **False discovery rate–corrected $p < .01$.

Variation in Hippocampal Activation as a Function of Trait Anxiety. There was a significant quadratic contrast of a condition × time × trait anxiety interaction ($F_{1,62} = 9.99$, $p_{FDR} = .018$) (Figure 5). Planned follow-up analyses used bivariate correlations between trait anxiety and each contrast of interest (i.e., safety compound over time, safety compound vs. threat, and safety compound vs. novel compound). There was a significant positive correlation between hippocampal activation during late versus early trials of the safety compound condition (i.e., last 3 trials minus first 3 trials) [$r_{63} = 0.23$, $n = 64$, $p = .042$] such that individuals with higher anxiety showed a greater increase in hippocampal activation than individuals with lower anxiety. However, there was no correlation when examining late trials of the safety compound versus late trials of the threat cue ($p = .075$) or the novel compound ($p = .312$).

A Priori ROI-ROI Functional Connectivity

Hippocampal-dACC Functional Connectivity. Contrary to our hypotheses, there were no significant effects of condition, time, or interactions between condition and time when examining left or right hippocampal–bilateral dACC functional connectivity ($p > .05$) (Figure S5).

Associations Between Hippocampal-dACC Functional Connectivity and Anxiety. There was a significant interaction between trait anxiety, condition, and time on left hippocampal–bilateral dACC functional connectivity ($F_{3,183} = 5.10$, $p = .002$) but no main effect of anxiety or interaction between anxiety and condition or between anxiety and time ($ps > .05$). Planned bivariate correlations revealed a significant relationship between trait anxiety and left hippocampus–dACC functional connectivity to the late trials of safety compound versus novel compound conditions [$r_{62} = -0.306$, $p = .015$] but no relationship between trait anxiety and early versus late trials

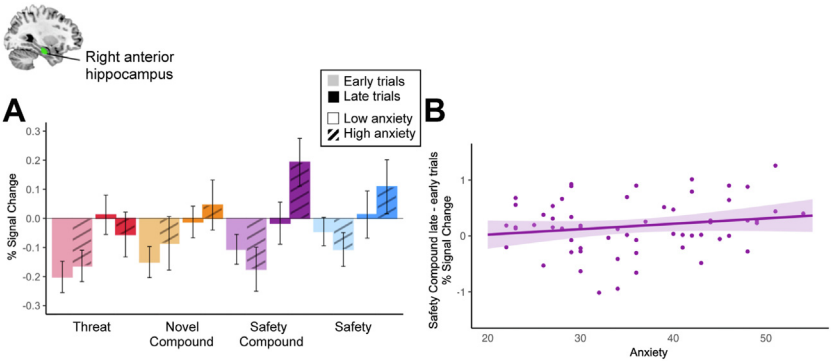

Figure 5. Right hippocampal activation by trait anxiety. **(A)** Percent signal change (parameter estimates) is shown on the y-axis for early and late trials of the safety signal learning task. There was a significant quadratic contrast of the interaction between condition, time, and anxiety (false discovery rate–corrected p = .018). Participants were separated into high and low groups using a median split on trait anxiety for visualization purposes only. **(B)** There was a significant positive correlation between trait anxiety and activation in the right hippocampus during late vs. early trials of the safety compound condition (p = .042).

of the safety compound or late trials of the safety compound versus threat (ps $>$ .05) (Figure 6).

Exploratory Whole-Brain Activation

Differences in Neural Activation During SSL. There were significant differences in activation to the safety compound compared with the threat condition in visual regions, including the primary visual cortex, the lateral occipital cortex, and fusiform gyrus, as well as in motor and somatosensory regions, including the pre- and postcentral gyri, and in the left superior temporal sulcus, supramarginal gyrus, and right amygdala ($p < .001$; cluster $p < .05$) (Figure 7; Table 2). There were no significant whole-brain differences in the contrast comparing activation to the safety compound versus the novel compound.

Activation in the Ventrolateral PFC Correlates With Anxiety During SSL. Trait anxiety was negatively correlated with activation in the ventrolateral PFC (vlPFC), partially overlapping with the left inferior frontal gyrus, for the safety compound compared with the novel compound ($p < .001$; cluster $p < .05$; peak Montreal Neurological Institute coordinates: −46, 48, 0) (Figure 8). Thus, individuals with higher anxiety showed lower vlPFC activation in response to the safety compound than to the novel compound. This region was labeled using meta-analyses from neurosynth (neurosynth.org), which showed overlap between the peak local maxima for the resulting thresholded cluster and the following terms: "IFG," "ventrolateral prefrontal," and "vlpfc." There were no significant clusters of activation for the contrast of the safety compound versus the threat condition.

DISCUSSION

Anxiety disorders are extremely common, yet many individuals do not benefit sufficiently from current evidence-based treatments that are based on the principles of fear extinction. Previous empirical and theoretical research suggests that the judicious incorporation of safety signals into cognitive behavioral therapy–based interventions could provide a means to optimize existing interventions for individuals with anxiety (2,6,31,32). The current findings extend previous cross-species evidence of SSL (3–6,33) in 3 notable ways. First, the current study indicates that safety signals effectively reduce physiological reactivity in adult humans across time, in a manner that is specific to the inclusion of the safety signal, and over and above an effect of novelty. This study also builds on a growing literature highlighting the role of the anterior hippocampus in assessing the degree of threat or safety in the environment (6,9,18,34). Second, we identified anxiety-related differences in fear reduction and hippocampal involvement during SSL, thereby providing novel insight into the potential clinical relevance of this approach to fear reduction. Third, whole-brain exploratory analyses revealed the involvement of regions, including the right BLA and the superior temporal gyrus during

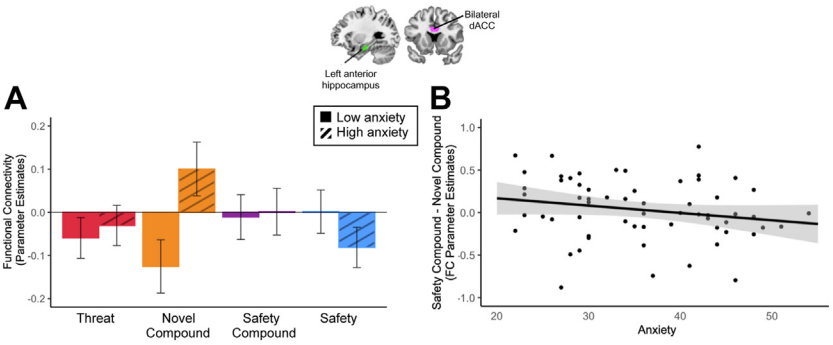

Figure 6. Left hippocampal functional connectivity (FC) with the bilateral dorsal anterior cingulate cortex (dACC) is correlated with trait anxiety during safety signal learning. **(A)** There was a significant condition × time × anxiety interaction (p = .002). Participants were separated into high and low groups using a median split on trait anxiety for visualization purposes only. **(B)** FC between the left hippocampus and the bilateral dACC during the late trials of the safety compound vs. novel compound conditions of the safety signal learning task was negatively correlated with trait anxiety such that individuals with higher anxiety showed lower FC to the safety compound compared with the novel compound condition (p = .032).

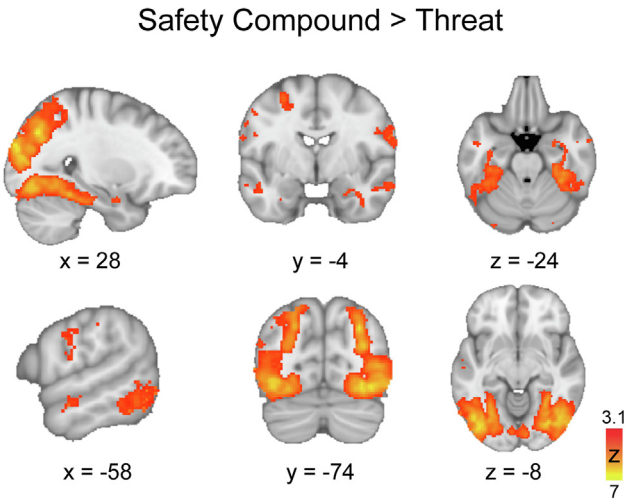

Figure 7. Differential activation during safety compound vs. threat (exploratory whole-brain results). Activation to the safety compound was higher than to the threat condition in the right amygdala, left superior temporal gyrus, bilateral fusiform gyrus, primary visual cortex, and lateral occipital cortex (Table 2) ($z > 3.1$, cluster $p < .05$; color bar indicates z score).

SSL, as well as lower activation of the vlPFC associated with higher trait anxiety. These findings have implications for optimizing interventions for individuals with anxiety, suggesting that SSL may be a means to augment fear reduction for individuals for whom the efficacy of existing exposure-based therapies is limited.

Existing evidence suggests that SSL reduces threat-related behavior and reactivity in rodents and humans (1,5,6). Notably, the current findings indicate that this reduction in threat-related reactivity may occur over and above the effect of novelty. That is, the reduction in threat-related reactivity was greater for the safety compound than for the novel compound. This result highlights the specificity of the inclusion of the conditioned safety cue in the safety compound condition (as opposed to a nonconditioned novel stimulus) in reducing threat-related reactivity, even in the presence of the learned threat cue.

Identifying anxiety-related differences in SSL is crucial to determining the potential relevance of SSL to interventions for fear reduction. The current findings showed that normative variation in trait anxiety was associated with differential patterns of physiological reactivity and hippocampal activation during SSL. More specifically, individuals with higher anxiety showed higher reactivity to the safety compound than individuals with lower anxiety. Furthermore, while individuals with lower anxiety showed the more expected pattern of lower reactivity to the safety compound than to threat, individuals with higher anxiety showed the opposite pattern overall. Although participants with higher anxiety showed less of a reduction in threat reactivity during SSL as compared with participants with lower anxiety, participants with higher anxiety did seem to show reduced reactivity via SSL over the course of learning, suggesting that this technique could still be effective for augmenting fear reduction. However, future research with individuals with clinically impairing anxiety is needed to better understand this relationship. In parallel, compared with individuals with lower anxiety, individuals with higher anxiety showed higher hippocampal activation during SSL. One possible interpretation of this pattern of findings is that individuals with higher anxiety may rely on recruiting the anterior hippocampus to a greater extent to attain even some reduction in physiological reactivity.

The anterior hippocampus is a highly interconnected region that contributes to the complex regulation of threat responding (9), with previous evidence suggesting specific involvement of hippocampal-dACC circuitry in SSL (6). Here, we found that individuals with higher (vs. lower) trait anxiety showed a greater increase in hippocampal activation to the safety compound over the course of learning. These findings raise the possibility that future research could begin to identify which individuals may be poised to benefit most from approaches that center on SSL, which previous work suggests may facilitate extinction learning by reducing postextinction recovery of fear (35). Furthermore, contrary to our hypotheses based on earlier work (6), we did not observe heightened hippocampal-dACC functional connectivity during SSL in the overall sample. This analysis differed from earlier work in that the current study compared the safety compound directly with a novel compound condition (rather than the threat cue alone) and

Table 2. Differential Activation During Safety Signal Learning (Whole-Brain Exploratory Results)

Brain Region	Max z Score	Cluster Size, Voxels	Brodmann Area	Peak MNI Coordinates		
				x	y	z
Bilateral Occipital Pole, LOC	8.09	19,133	18, 17	12	−100	12
Bilateral Fusiform Gyrus	6.84	–	37	32	−44	−19
Right Amygdala	3.91	67[a]	–	26	−4	−24
Left Middle Frontal Gyrus, Precentral Gyrus, Postcentral Gyrus	5.08	478	44, 1	−48	8	34
Left Precuneus, Posterior Cingulate	4.81	301	7	2	−56	24
Right Postcentral Gyrus, Precentral Gyrus	4.87	222	1, 3, 4	64	−4	24
Left Middle Temporal Gyrus, STG	4.67	173	21, 22	−58	−8	−14
Right STG, Middle Temporal Gyrus	3.91	87	21, 38	52	−2	−20

Results are shown for safety compound vs. threat contrast (Figure 7).

LOC, lateral occipital cortex; MNI, Montreal Neurological Institute; STG, superior temporal gyrus.

[a]Overlap with amygdala calculated using overlap with whole amygdala mask from Juelich atlas (59).

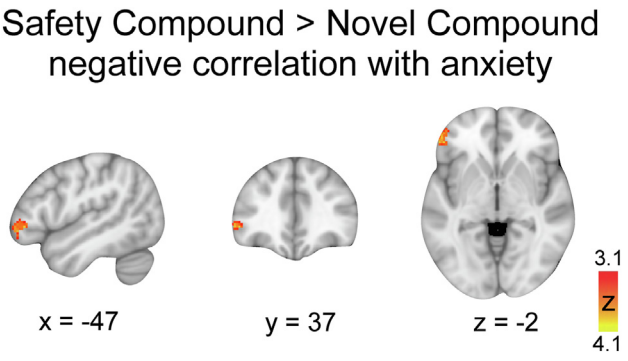

Figure 8. Activation in ventrolateral prefrontal cortex is negatively correlated with trait anxiety during safety compound vs. novel compound conditions (exploratory whole-brain results). Trait anxiety was negatively correlated with activation in a region in the left ventrolateral prefrontal cortex, partially overlapping with the left inferior frontal gyrus, during the safety compound vs. novel compound ($z > 3.1$, cluster $p < .05$; color bar indicates z score).

examined a different bilateral dACC ROI due to advances in parcellation and segmentation methods (36,37). However, we did observe that individuals with higher (relative to lower) trait anxiety showed lower functional connectivity to the safety compound than to the novel compound condition. Interestingly, this effect seems to be driven by individual differences in hippocampal-dACC connectivity during the novel compound condition. Future studies with clinically anxious individuals are needed to better understand the relationship of this neural pathway in conditioned inhibition, as opposed to external inhibition, across higher levels of anxiety.

Building upon past work that has solely relied on ROI analyses [(3); for a review, see (17,36–39)], and whole-brain exploratory analyses in the current study revealed the involvement of regions including the right BLA and the superior temporal gyrus in discriminating between the safety compound and threat during SSL. Although recent neuroimaging work has highlighted inconsistencies in the involvement of the amygdala in threat and safety learning, a new and highly powered study by Wen *et al.* (40) determined the critical role of the BLA in associative learning. Moreover, rodent studies have suggested that a subpopulation of neurons in the BLA may be selectively responsive to safety cues during SSL (41). However, amygdala lesion studies have not found an effect on the summation test (42), suggesting that pathways independent of the amygdala can also support SSL via conditioned inhibition. Our current findings indicate that such alternate pathways may include the anterior hippocampus.

In addition, at the whole-brain level, individuals with higher anxiety showed lower activation of the vlPFC to the safety compound than to the novel compound during early trials of the task (although this difference was not observed during later trials of the task). Given the involvement of the vlPFC in processes that include cognitive control, implicit emotion regulation, and affect labeling (43–47), this difference may be consistent with weaker regulation in the affective domain that has been shown in previous work on anxiety (48–52). We did not observe differences in recruitment of the vlPFC later in the task, which could suggest either that this region was no longer recruited across participants or that there was no longer a relationship to anxiety during the later trials. Additional research is needed to better understand the involvement of these prefrontal regions in SSL.

The current study had several strengths, including the characterization of anxiety-related differences during SSL, the direct comparison of learned safety with a condition that controlled for novelty, and the inclusion of exploratory whole-brain analyses. However, there are several limitations of the current study that should be addressed in future research. First, as has been well-documented to be the case for most studies using SCR (6,53), only a subsample of participants could be included in the SCR analyses due to challenges such as low overall SCR signal and lack of sustained learning. Future research would benefit from a larger sample size that provides greater statistical power in this domain, as well as from the incorporation of other physiological measurements that are not as prone to data loss. Additionally, the use of a relatively mild unconditioned stimulus (i.e., an aversive noise) may have promoted more rapid habituation. In the current task design, the safety compound is contrasted with the pairing of a nonconditioned stimulus with the conditioned threat cue (i.e., the novel compound condition). However, the novel cue is only completely new to the participant during the first trial of the testing phase. Future research may benefit from using a probe trial design as in Myers and Davis (3) to further investigate the differences between conditioned and external inhibition. Past research has highlighted the potential for amygdala activation to be confounded by stimulus-correlated signal in veins draining distant brain regions (54). Our whole-brain results do not preclude this possibility, and future research could use auditory or olfactory cues to better avoid potential visual region spillover into the amygdala and nearby regions. Lastly, the current study relied on trait anxiety as a measure of normative levels of anxiety in a nonclinical population. Given the potential clinical relevance of SSL and its implications for individuals with anxiety disorders, it is important that future studies investigate SSL in adults with clinical levels of anxiety.

A core feature of anxiety disorders is difficulty regulating fear (55–58), which may stem from difficulty in learning about or incorporating cues that signal safety. Behavioral and neuroscientific studies using conditioned inhibition paradigms have shown that safety cues can effectively reduce fear and prevent the onset of new fears in animals (3,4) and that safety cues are effective for actively inhibiting fear in humans (3,5). Here, we showed that SSL reduced threat responding across time and over and above the effect of novelty and that SSL involved activation in the right anterior hippocampus. Moreover, we showed that individuals with higher trait anxiety may exhibit less fear reduction, lower vlPFC activation, and altered hippocampal activation and hippocampal-dACC functional connectivity during SSL. Finally, exploratory findings suggest that the right amygdala and right superior temporal gyrus may also be recruited during SSL. Taken together, these results add to a growing body of literature suggesting that targeting a distinct neural pathway via safety signals may provide a means to augment fear reduction for individuals for whom the efficacy of existing exposure-based therapies is limited.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the National Institutes of Health Director's Early Independence Award (Award No. DP5OD021370 [to DGG]), Brain & Behavior Research Foundation (National Alliance for Research on Schizophrenia and Depression) Young Investigator Award (to DGG), Jacobs Foundation Early Career Research Fellowship (to DGG), National Science Foundation Graduate Research Fellowship Program Award (Grant No. DGE1122492 [to PO]), a National Institute of Mental Health National Research Service Award (Grant No. F30MH124271 [to SK]), and a Scholar Award granted by the International Chapter of the Philanthropic Educational Organization (to PO).

We are immensely grateful to each of our participants. We gratefully acknowledge Jeffrey Mandell for contributions to the analytic pipeline; Emma Goodman, Amy Kwarteng, Rob Colgate, Neida Moreno, Alissa Wong, Zoe Hopson, and Zhiliang Fang for assistance with data collection and data quality assessment; and Lucinda Sisk, Hopewell Hodges, Cristina Nardini, Janeen Thomas, Beatriz Rios, Sophie Rader, Sarah Bakirci, Uma Raul, and Monica Krah for assistance with data collection. We also gratefully acknowledge Dr. B.J. Casey and Dr. Nim Tottenham for their generous support and feedback during study conceptualization and design.

The data included in this paper partially overlap with those published in Meyer *et al.* (6) and Kribakaran *et al.* (13).

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Psychology, Yale University, New Haven, Connecticut.

Address correspondence to Dylan G. Gee, Ph.D., at dylan.gee@yale.edu.

Received Nov 16, 2022; revised Mar 15, 2023; accepted May 31, 2023.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsgos.2023.05.007.

REFERENCES

1. Christianson JP, Fernando ABP, Kazama AM, Jovanovic T, Ostroff LE, Sangha S (2012): Inhibition of fear by learned safety signals: A mini-symposium review. J Neurosci 32:14118–14124.
2. Odriozola P, Gee DG (2021): Learning about safety: Conditioned inhibition as a novel approach to fear reduction targeting the developing brain. Am J Psychiatry 178:136–155.
3. Myers KM, Davis M (2004): AX+, BX- discrimination learning in the fear-potentiated startle paradigm: Possible relevance to inhibitory fear learning in extinction. Learn Mem 11:464–475.
4. Kazama AM, Schauder KB, McKinnon M, Bachevalier J, Davis M (2013): A novel AX+/BX- paradigm to assess fear learning and safety-signal processing with repeated-measure designs. J Neurosci Methods 214:177–183.
5. Jovanovic T, Keyes M, Fiallos A, Myers KM, Davis M, Duncan EJ (2005): Fear potentiation and fear inhibition in a human fear-potentiated startle paradigm. Biol Psychiatry 57:1559–1564.
6. Meyer HC, Odriozola P, Cohodes EM, Mandell JD, Li A, Yang R, *et al.* (2019): Ventral hippocampus interacts with prelimbic cortex during inhibition of threat response via learned safety in both mice and humans. Proc Natl Acad Sci USA 116:26970–26979.
7. Rescorla RA (1969): Conditioned inhibition of fear resulting from negative CS-US contingencies. J Comp Physiol Psychol 67:504–509.
8. Haaker J, Maren S, Andreatta M, Merz CJ, Richter J, Richter SH, *et al.* (2019): Making translation work: Harmonizing cross-species methodology in the behavioural neuroscience of Pavlovian fear conditioning. Neurosci Biobehav Rev 107:329–345.
9. Sotres-Bayon F, Sierra-Mercado D, Pardilla-Delgado E, Quirk GJ (2012): Gating of fear in prelimbic cortex by hippocampal and amygdala inputs. Neuron 76:804–812.
10. Fanselow MS, Dong HW (2010): Are the dorsal and ventral hippocampus functionally distinct structures? Neuron 65:7–19.
11. Kalisch R, Korenfeld E, Stephan KE, Weiskopf N, Seymour B, Dolan RJ (2006): Context-dependent human extinction memory is mediated by a ventromedial prefrontal and hippocampal network. J Neurosci 26:9503–9511.
12. Maren S (2001): Neurobiology of Pavlovian fear conditioning. Annu Rev Neurosci 24:897–931.
13. Kribakaran S, Odriozola P, Cohodes EM, McCauley S, Zacharek SJ, Hodges HR, *et al.* (2022): Neural circuitry involved in conditioned inhibition via safety signal learning is sensitive to trauma exposure. Neurobiol Stress 21:100497.
14. Laing PAF, Steward T, Davey CG, Felmingham KL, Fullana MA, Vervliet B, *et al.* (2022): Cortico-striatal activity characterizes human safety learning via Pavlovian conditioned inhibition. J Neurosci 42:5047–5057.
15. Jovanovic T, Norrholm SD, Fennell JE, Keyes M, Fiallos AM, Myers KM, *et al.* (2009): Posttraumatic stress disorder may be associated with impaired fear inhibition: Relation to symptom severity. Psychiatry Res 167:151–160.
16. Jovanovic T, Kazama A, Bachevalier J, Davis M (2012): Impaired safety signal learning may be a biomarker of PTSD. Neuropharmacology 62:695–704.
17. Fanselow MS (2000): Contextual fear, gestalt memories, and the hippocampus. Behav Brain Res 110:73–81.
18. Ji J, Maren S (2007): Hippocampal involvement in contextual modulation of fear extinction. Hippocampus 17:749–758.
19. Kim MJ, Whalen PJ (2009): The structural integrity of an amygdala-prefrontal pathway predicts trait anxiety. J Neurosci 29:11614–11618.
20. LeDoux JE (2000): Emotion circuits in the brain. Annu Rev Neurosci 23:155–184.
21. Sotres-Bayon F, Quirk GJ (2010): Prefrontal control of fear: More than just extinction. Curr Opin Neurobiol 20:231–235.
22. Meyer HC, Gerhard DM, Amelio PA, Lee FS (2021): Pre-adolescent stress disrupts adult, but not adolescent, safety learning. Behav Brain Res 400:113005.
23. Barnes LLB, Harp D, Jung WS (2002): Reliability generalization of scores on the Spielberger State-Trait Anxiety Inventory. Educ Psychol Meas 62:603–618.
24. Spielberger CD (1970): Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press.
25. Neumann DL, Waters AM, Westbury HR (2008): The use of an unpleasant sound as the unconditional stimulus in aversive Pavlovian conditioning experiments that involve children and adolescent participants. Behav Res Methods 40:622–625.
26. Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM, *et al.* (2018): The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Dev Cogn Neurosci 32:43–54.
27. Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, *et al.* (2016): The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data 3:160044.
28. Glasser MF, Smith SM, Marcus DS, Andersson JLR, Auerbach EJ, Behrens TEJ, *et al.* (2016): The Human connectome Project's neuroimaging approach. Nat Neurosci 19:1175–1187.
29. Hindy NC, Turk-Browne NB (2016): Action-based learning of multistate objects in the medial temporal lobe. Cereb Cortex 26:1853–1865.
30. McLaren DG, Ries ML, Xu G, Johnson SC (2012): A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. Neuroimage 61:1277–1286.
31. Blakey SM, Abramowitz JS (2016): The effects of safety behaviors during exposure therapy for anxiety: Critical analysis from an inhibitory learning perspective. Clin Psychol Rev 49:1–15.
32. Milosevic I, Radomsky AS (2013): Incorporating the judicious use of safety behavior into exposure-based treatments for anxiety disorders: A study of treatment acceptability. J Cogn Psychother 27:155–174.
33. Winslow JT, Noble PL, Davis M (2008): AX+/BX− discrimination learning in the fear-potentiated startle paradigm in monkeys. Learn Mem 15:63–66.
34. Padilla-Coreano N, Bolkan SS, Pierce GM, Blackman DR, Hardin WD, Garcia-Garcia AL, *et al.* (2016): Direct ventral hippocampal-prefrontal input is required for anxiety-related neural activity and behavior. Neuron 89:857–866.

35. Dunsmoor JE, Campese VD, Ceceli AO, LeDoux JE, Phelps EA (2015): Novelty-facilitated extinction: Providing a novel outcome in place of an expected threat diminishes recovery of defensive responses. Biol Psychiatry 78:203–209.
36. Ashburner J, Friston KJ (2005): Unified segmentation. Neuroimage 26:839–851.
37. Zhang YJ (2001): A review of recent evaluation methods for image segmentation. In: Proceedings of the Sixth International Symposium on Signal Processing and Its Applications (Cat.No.01EX467). Kuala Lumpur, Malaysia: IEEE, 148–151.
38. Falls WA, Davis M (1995): Lesions of the central nucleus of the amygdala block conditioned excitation, but not conditioned inhibition of fear as measured with the fear-potentiated startle effect. Behav Neurosci 109:379–387.
39. Heldt SA, Coover GD, Falls WA (2002): Posttraining but not pretraining lesions of the hippocampus interfere with feature-negative discrimination of fear-potentiated startle. Hippocampus 12:774–786.
40. Wen Z, Raio CM, Pace-Schott EF, Lazar SW, LeDoux JE, Phelps EA, Milad MR (2022): Temporally and anatomically specific contributions of the human amygdala to threat and safety learning. Proc Natl Acad Sci USA 119:e2204066119.
41. Sangha S, Chadick JZ, Janak PH (2013): Safety encoding in the basal amygdala. J Neurosci 33:3744–3751.
42. Kazama AM, Heuer E, Davis M, Bachevalier J (2012): Effects of neonatal amygdala lesions on fear learning, conditioned inhibition, and extinction in adult macaques. Behav Neurosci 126:392–403.
43. Aron AR, Robbins TW, Poldrack RA (2004): Inhibition and the right inferior frontal cortex. Trends Cogn Sci 8:170–177.
44. Berboth S, Morawetz C (2021): Amygdala-prefrontal connectivity during emotion regulation: A meta-analysis of psychophysiological interactions. Neuropsychologia 153:107767.
45. Egner T (2011): Right ventrolateral prefrontal cortex mediates individual differences in conflict-driven cognitive control. J Cogn Neurosci 23:3903–3913.
46. Krawczyk DC (2002): Contributions of the prefrontal cortex to the neural basis of human decision making. Neurosci Biobehav Rev 26:631–664.
47. Viinikainen M, Jääskeläinen IP, Alexandrov Y, Balk MH, Autti T, Sams M (2010): Nonlinear relationship between emotional valence and brain activity: Evidence of separate negative and positive valence dimensions. Hum Brain Mapp 31:1030–1040.
48. Klemanski DH, Curtiss J, McLaughlin KA, Nolen-Hoeksema S (2017): Emotion regulation and the transdiagnostic role of repetitive negative thinking in adolescents with social anxiety and depression. Cognit Ther Res 41:206–219.
49. McLaughlin KA, Mennin DS, Farach FJ (2007): The contributory role of worry in emotion generation and dysregulation in generalized anxiety disorder. Behav Res Ther 45:1735–1752.
50. Mennin DS, Holaway RM, Fresco DM, Moore MT, Heimberg RG (2007): Delineating components of emotion and its dysregulation in anxiety and mood psychopathology. Behav Ther 38:284–302.
51. Mennin DS, McLaughlin KA, Flanagan TJ (2009): Emotion regulation deficits in generalized anxiety disorder, social anxiety disorder, and their co-occurrence. J Anxiety Disord 23:866–871.
52. Michalska KJ, Benson B, Ivie EJ, Sachs JF, Haller SP, Abend R, *et al.* (2023): Neural responding during uncertain threat anticipation in pediatric anxiety. Int J Psychophysiol 183:159–170.
53. Lonsdorf TB, Menz MM, Andreatta M, Fullana MA, Golkar A, Haaker J, *et al.* (2017): Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. Neurosci Biobehav Rev 77:247–285.
54. Boubela RN, Kalcher K, Huf W, Seidel EM, Derntl B, Pezawas L, *et al.* (2015): fMRI measurements of amygdala activation are confounded by stimulus correlated signal fluctuation in nearby veins draining distant brain regions. Sci Rep 5:10499.
55. American Psychiatric Association (2013): Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Washington, DC: American Psychiatric Press.
56. Amstadter A (2008): Emotion regulation and anxiety disorders. J Anxiety Disord 22:211–221.
57. Britton JC, Lissek S, Grillon C, Norcross MA, Pine DS (2011): Development of anxiety: The role of threat appraisal and fear learning. Depress Anxiety 28:5–17.
58. Graham BM, Milad MR (2011): The study of fear extinction: Implications for anxiety disorders. AJP 168:1255–1265.
59. Amunts K, Kedo O, Kindler M, Pieperhoff P, Mohlberg H, Shah NJ, *et al.* (2005): Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: Intersubject variability and probability maps. Anat Embryol (Berl) 210:343–352.