Supplemental materials for Gold et al.:

Age differences in the neural correlates of anxiety disorders:

An fMRI study of response to learned threat

Supplemental text: Supplemental Methods & Supplemental Results

Supplemental Tables: 2

Supplemental Figures: 4

SUPPLEMENT TEXT:

Supplemental Methods

Participants

Clinicians assessed pediatric patients' psychiatric symptoms and suitability for treatment on three separate occasions prior to study participation. First, a psychiatric nurse conducted a telephone screen. Second, pediatric patients and parents completed an in-person, standardized diagnostic assessment (Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (KSADS-PL;(1)) with a clinician trained to acceptable levels of reliability for all diagnoses assessed (kappa>0.75). This involved performance of ratings on an initial series of recorded interviews and subsequent monitoring of interview tapes and reassessments of patients. For adults, this second visit utilized the Structured Clinical Interview for DSM-IV-TR Axis I Disorders (SCID) (2). Finally, a senior psychiatrist confirmed anxiety diagnosis via independent assessment at a separate visit, and all pediatric patients agreed to enter treatment for their anxiety disorder. Adult patients also were assessed by the senior psychiatrist and offered treatment, but this was not a requirement for study participation.

Comorbid major depressive disorder (MDD) and additional anxiety disorders were permitted in the anxious group (See Table 1), but obsessive-compulsive disorder, posttraumatic stress disorder, bipolar disorder, disruptive mood dysregulation disorder, and other comorbid conditions were exclusionary. Patients were required to be medication-free for their current episode of anxiety, which translated in practice to a 6-month or greater medication-free period. All participants were free of MRI contraindications, physical health problems, inclusion of family relatives in the study, and were required to have IQ>70, based on the Vocabulary and Matrix Reasoning subscales of the Wechsler Abbreviated Scale of Intelligence (3).

Youth participants were administered the Screen for Child Anxiety Related Emotional Disorders (SCARED; (4)) within four months of the scan (maximum = 102 days between). As expected, both parent- and child-reported anxiety was greater in the anxious compared to the healthy youth, *p*s < .001. Pediatric patients with anxiety were administered the Pediatric Anxiety Rating Scales (PARS; (5)) as part of an ongoing treatment study reported elsewhere (6).

Procedures

Visit 1: Psychophysiology visit

Methods are identical to those reported previously (7). Psychophysiogical data from a subset of participants were reported previously (8), although neuroimaging data from no participant have been included in a prior publication. Visit 1 psychophysiology and self-report data from the current sample are reported in this Supplement to assess successful conditioning and extinction, based on measures observed prior to the fMRI visit.

Threat conditioning and extinction task. During the first visit in the psychophysiology lab, all participants completed the “Screaming Lady” threat conditioning and extinction task (panel A, Figure 1). Procedures were identical to those reported elsewhere (7, 8), and summarized briefly here.

Participants first completed a startle habituation phase, which presented six startle probes in the absence of any stimuli. The task then consisted of the following three phases: pre-acquisition, acquisition and extinction. Visual stimuli showing two women displaying neutral facial expressions were the conditioned stimuli (CS: CS+ and CS-). They were presented for 7-seconds, followed by a gray screen presented for 8-21 seconds (averaging 15 seconds).

During the pre-acquisition phase, each CS was presented four times to allow physiological responses to the novel stimuli to habituate and to provide a baseline for discriminative conditioning. The acquisition phase consisted of 10 trials per CS type, presented in pseudorandom order, in which the CS+ was paired with an unconditioned stimulus (US), i.e., the CS+ was followed by a 1-second image of the same stimulus identity as the CS+ displaying a fearful facial expression. This stimulus co-terminated with an aversive scream soundbit delivered at 95dB through headphones. The US was presented with an 80% reinforcement schedule, and prior to the task, participants were instructed that they could learn to predict when the US would occur. However, they were not explicitly informed of the CS+/CS- contingencies. The US was never paired with the CS- and never presented during pre-acquisition or extinction. The identities of the facial stimuli serving as the CS were counterbalanced across participants and groups. Finally, the extinction phase consisted of eight trials per CS type, in the absence of the US.

The PsyLab psychophysiological recording system (PsyLab SAM System, Contact Precision Instruments, London) was used to program and administer the visual and auditory task stimuli. In addition to self-report, skin conductance, startle electromyography (EMG) data, electrocardiography (ECG) data were also recorded, but not analyzed in the current report. To measure fear-potentiated startle (FPS) using EMG, startle probes (i.e., 40ms, 4-10 psi of compressed air delivered to the forehead) were presented during the CS trials (5-6 seconds post-stimulus onset) and during the inter-stimulus interval (ISI). All CS+, CS-, and ISI were presented in a blocked counterbalanced order.

Self-report data collection and processing. Participants completed self-reported fear ratings of the CS+ and CS- stimuli using a ten-point Likert scale (1=none, 10=extreme) before startle habituation, after acquisition, and after extinction. The rationale for collecting self-report data between, rather than during task phases was to avoid interference with threat and safety learning.

Psychophysiological recording and data processing. Skin conductance and startle electromyography (EMG) were recorded continuously at 1000Hz using PsyLab psychophysiological recording system (PsyLab SAM System Contact Precision Instruments, London, www.psylab.com).

Skin conductance. Skin conductance was recorded from two Ag/AgCl electrodes filled with non-saline gel and placed on the medial phalanx of the middle and ring fingers of the participant's non-dominant hand. Skin conductance response to each CS+ and CS- was determined by the square-root-transformed difference between base-to-peak amplitude within 5s after stimulus onset (7, 8). Four participants were removed from analyses due to technical errors during data collection, resulting in N=196 for skin conductance response analyses.

Startle EMG. EMG was measured using two 6mm tin cup EMG electrodes filled with standard electrolyte solution placed under the subject's left eye. A ground electrode was attached to the subject's left forearm. Startle probes (4-10 psi of compressed air delivered to the forehead, 40ms) were presented during the CS trials (5-6 seconds post-stimulus onset) and during the inter-stimulus interval. FPS was measured by the eye blink reflex to the startle probes. EMG data were rectified and smoothed using moving averages with 20ms windows. The EMG response to the startle probe during each CS+, CS-, and ISI was calculated as the difference between the peak EMG response (within 150ms following the startle probe) and the baseline activity (50 ms prior to the startle probe). EMG was filtered using an amplifier bandwidth of 30-500Hz. FPS response data for each individual were standardized using a *T*-

score transformation. Four participants were removed from analyses due to technical errors during data collection, resulting in N=196 for the startle EMG analyses.

Data analysis. Skin conductance response and EMG data were averaged across trials and analyzed using separate repeated-measures analyses of covariance (ANCOVAs). Separate analyses were conducted for each of the three dependent variables (i.e., self-report, skin conductance response, and EMG). In each analysis, we tested the omnibus Anxiety × Age × Phase × Stimulus interaction effect, whereby Anxiety (healthy, anxious) and Age (continuous, in years) served as between-subjects factors, and Phase (pre-acquisition, acquisition, extinction) and Stimulus (CS-, CS+) served as within-subject factors. For EMG, ISI was included as an additional stimulus type. Higher-order ANCOVAs were decomposed using lower-order ANCOVAs and *t*-tests. Significance was determined using alpha=0.05 and two-tailed tests.

Visit 2: fMRI Visit

All participants (N=200) completed the initial psychophysiology visit. Participants completed the extinction recall task, adapted from (7). Task stimuli included the neutral photographs of the two identities presented in the threat conditioning and extinction task (i.e., CS+ and CS-). Nine additional neutral facial stimuli were presented. These stimuli included morphed images of the two identities in 10% increments ranging from 0% to 100% increasing resemblance to the CS+; i.e., 0% (CS-), 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100% (CS+). Participants completed three task runs, consisting of four blocks of each of the two attention conditions, threat appraisal and explicit memory, presented in random order. All morphs (11 images) and 3 blank images, providing an implicit baseline condition, were presented in random order within each block. Each trial was presented for 4000-milliseconds with an inter-stimulus interval averaging 1s. Twelve replicates were presented for each of the 22 regressors of interest (i.e., the 11 morphs presented in two attention conditions). E-prime

computer software and front-projection presented the task. Participants used a game controller device to record their responses to the questions presented in the two attention conditions: threat appraisal, “How afraid are you?”; explicit memory, “How likely was she to scream?”. Participants used both of their thumbs placed on the left and right buttons to maneuver their response (displayed in red highlight, See Figure 1B) along the 6-point Likert scale, ranging from 0 (Not at all) to 6 (Extremely). Participants were instructed to lock in their answers using additional buttons on top of the game controller device, at which point the responses were displayed with green highlight; these data were used for the reaction time responses analyzed in the current report.

Behavioral data analysis. Subjective rating and reaction time (RT) data were collected during extinction recall from the second visit in the fMRI environment. Rating and reaction time data were analyzed using linear mixed-effects modeling via the *nlme* package in R (Pinheiro J, Bates D, DebRoy S, Sarkar D and R Core Team (2018). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-137, https://CRAN.R-project.org/package=nlme). Rating and reaction time were each modeled by the interaction between Anxiety (healthy, anxious), Age (continuous, in years), Attention Condition (threat appraisal, explicit memory), and a linear or quadratic trend term for the morph factor (i.e., 11 facial morphs). The linear and quadratic trend terms were treated as random effects. Number of days between Visit 1 and Visit 2 was used as a fixed effect of nuisance. Additionally, reaction time measures were used in amplitude modulation fMRI analyses (described below and in main document). Given that trials with reaction time < 200ms do not reflect a physiologically valid response, such trials were replaced with values imputed from the mean reaction time for that condition.

***Imaging acquisition parameters*.** MRI scanning was conducted at the National Institute of Mental Health (NIMH) Functional Magnetic Resonance Imaging Core Facility (FMRIF) using standard sequences collected with a 3-Tesla MR750 General Electric scanner with a 32-channel head coil. Expandable pads were placed on each side of the participant’s head and

Coban Self Adherent Wrap (3M) was wrapped around the participant's head and the scanner's head cradle, in order to help minimize motion. Participants completed a high-resolution, T1-weighted magnetization-prepared rapid-conditioning gradient-echo (MPRAGE) anatomical scan of the whole brain, which was used for co-registration and normalization procedures for the fMRI data [176 axial slices, 256x256 matrix,1mm^3 isotropic slices; flip angle=7**°,** FOV=220 mm; repetition time (TR)=7.7ms, echo time (TE)=3.42s]. Three functional scan runs each included 272 functional image volumes, with 47 contiguous 3 mm interleaved axial slices, 96×96 matrix with a T2*-weighted echo-planar sequence (repetition time/echo time [TR/TE] = TR=2300 ms/TE=25 ms; field of view (FOV)=240 mm; flip angle=50^o, in-plane resolution = 2.5×2.5).

fMRI processing. Neuroimaging processing and analysis utilized the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). Data were processed and analyzed using standard procedures within the Analysis of Functional NeuroImages (AFNI) software (9), implemented with afni_proc.py. The four initial acquisition images, prior to task onset, were discarded to reach longitudinal magnetization equilibrium, resulting in 268 TRs per each of the three runs. Pre-processing procedures included uniformity correction, AFNI's 3dDespike program, slice-time correction, co-registration, normalization, non-linear registration using the Talairach template (AFNI TTN27 template), and spatial smoothing with a 5mm full-width-half-maximum Gaussian kernel. Prior to individual-level general linear model (GLM) analysis, data were resampled to 2.5mm isotropic voxels and were scaled at the voxel level such that regression coefficients reflect percent signal change. Finally, motion-correction procedures censored TR pairs with a Euclidean norm motion derivative greater than 1mm or an outlier fraction greater than 0.1 Inclusion criteria required all participants to have no more than 20% of TRs censored. This resulted in the exclusion of six participants (2 healthy, 4 anxious) from the final dataset included in group-level analyses.

fMRI data analysis. For each subject, two types of GLMs were created using amplitude modulation and gPPI methods, separately.

Amplitude modulation by reaction time. First, we conducted an amplitude modulation analysis to test associations between trial-by-trial variation in BOLD response and variation in reaction time. Each individual-level GLM included baseline drift and motion (displacement: x, y, and z directions, and rotation: roll, pitch, and yaw), as well as regressors of interest. Regressors of interest included the 22 task conditions, corresponding to each morphed image (11) for each attention condition (2), as well as regressors for modulation by reaction time on each of these conditions. AFNI 3dDeconvolve was used for individual-level GLM analysis with the "stim_times_AM2" option. Two types of regressors were generated. One type of regressor examined task-related activation at average reaction time. The second type of regressor, i.e., reaction time-modulated regressors, assessed the strength of the RT-BOLD association, for which the coefficients for the amplitude modulation effects indicate the effect of reaction time in the unit of BOLD response (e.g., % signal change) when reaction time increases by one unit (e.g., one second).

Amygdala-based gPPI. Using standard gPPI methods in AFNI, individual-level GLMs were also created to investigate task-modulated functional connectivity. We used the left and right amygdala as seed regions given our prior finding linking amygdala functional connectivity with anxiety, age, and threat (10), as well as the ability to easily and objectively define the amygdala based on clear anatomical boundaries. Seed regions were anatomically-defined using the DKD_Desai_MPM atlas (11) available in AFNI, resampled to 2.5mm isotrophic, and masked by a gray matter volume comprised of voxels in which data were present in at least 90% of the sample. For each of the 22 task conditions, a PPI term was created as the product of the detrended and demeaned seed regressor and the event type; gPPI methods were based on (12) and implemented in AFNI as described elsewhere (13, 14). Individual-level GLMs using gPPI methods included the same regressors as in the amplitude modulation analyses, with the

addition of the seed time series (i.e., left or right amygdala) and the PPI terms for each of the 22 task conditions. Separate GLMs were created for the left and right amygdala.

Supplemental Results

Visit 1: Psychophysiology visit.

Self-report and psychophysiological results. Full statistics for the ANCOVAs testing the Phase × Stimulus × Anxiety × Age for the self-report, skin conductance response, and EMG data acquired during conditioning and extinction are reported in Table S1. See Figure S1 for a graphical depiction. Across all measures, a significant Phase × Stimulus was noted, providing evidence of discriminative conditioning across participants, independent of anxiety diagnosis or age. Follow-up analyses on the self-report data indicated no difference between CS- and CS+ prior to the pre-acquisition phase, $t(199)=0.64$, $p=0.524$. Following the acquisition phase, greater anxiety was reported in response to CS+ relative to CS-, $t(199)=10.53$, $p<0.001$; this effect was maintained following extinction, $t(199)=8.18$, $p<0.001$.

Follow-up analyses on skin conductance response data indicated no difference between CS- and CS+ during pre-acquisition, $t(193)=1.43$, $p=0.153$. During the acquisition phase, greater skin conductance response was noted in response to CS+ relative to CS-, $t(193)=5.48$, $p<0.001$. During extinction, there was no significant different between the CS types, reflective evidence of successful extinction across participants, $t(193)=0.15$, $p=0.882$.

Follow-up analyses on EMG data indicated a significant main effect of Stimulus during pre-acquisition, $F(2,390)=5.77$, $p=0.003$, driven by a greater response to CS- and CS+ relative to ISI, $ps<0.006$, but no significant difference between CS- and CS+, $p=0.978$. During acquisition, a main effect of Stimulus was also noted, $F(2,390)=37.63$, $p<0.001$, with a greater response to CS+ relative to CS- and ISI, $ps<0.001$, as well as CS- > ISI, $p<0.001$. The main effect of Stimulus was maintained during extinction, $F(2,390)=35.96$, $p<0.001$, with greater response to CS+ relative to CS- and ISI, $ps<0.001$, as well as CS- > ISI, $p<0.001$. Together,

these results indicate successful discriminative conditioning across all measures. Extinction of conditioned fear was noted in the primary physiological measure, skin conductance response, while self-report and EMG data showed some maintenance of the acquired contingency as evidenced by heightened responses to the CS+ relative to CS-. Several additional effects were noted (see Table S1) but are not interpreted in the current context due to the focus on fMRI data collected during the second, extinction recall visit. Of note, anxiety diagnosis was not associated with any interaction effects, demonstrating that anxiety was unrelated to discriminative conditioning or extinction across the psychophysiological or self-report measures.

Visit 2: fMRI visit.

Behavioral results: Rating and reaction time measures. Mean ratings and reaction time for the threat appraisal and explicit memory condition by morph, separately for healthy vs. anxious and younger vs. older participants (based on median split) are presented in Figure S2.

Linear mixed-effects modeling of task ratings indicated a main effect of Attention Condition, $F(1,4180)=62.75$, $p<0.001$, with higher mean ratings of explicit memory relative to threat appraisal. A main effect of the linear trend was also noted, $F(1,4180)=6.16$, $p=0.013$, with ratings increasing as the proportion of CS+ in the morphs linearly increased. These main effects were qualified by an Age $\times$ Anxiety interaction; however, follow-up analyses in each anxiety group revealed no effect of age in either group, $ps>0.062$. Ratings also showed an Age x Attention Condition interaction, $F(1,4180)=23.65$, $p<.0001$. Follow-up tests indicated that the age effect in each attention condition was not statistically significant, $ps>0.08$. In addition, we noted a Anxiety $\times$ Attention Condition interaction, $F(1,4180)=5.77$, $p=0.016$; however, when tested separately within each attention condition, the effect of anxiety was not significant in either condition, $ps>0.488$. In addition, a linear trend $\times$ Attention Condition interaction was noted, $F(1,4180)=91.30$, $p<0.001$. Follow-up analyses within each attention condition indicated a effect of the linear trend in the threat appraisal condition, $F(1,1992)=7.32$, $p=0.007$; this trend

was likewise observed in the explicit memory condition, only with a larger magnitude, $F(1,1992)=71.87$, $p<0.001$, indicating greater rating in the latter with increasing CS+ proportion. These effects were further qualified by a Anxiety × Age × Attention Condition interaction, $F(1,4180)=18.40$, $p<0.001$. Age differentially moderated the anxiety group differences based on attention condition, where the extent to which age moderated the anxiety differences was greater for the threat appraisal relative to explicit memory conditions (See Figure S3). Follow-up analyses within each attention condition indicated a trend towards significance for the Anxiety × Age interaction in the threat appraisal condition, $F(1,195)=3.50$, $p=0.063$, and a non-significant Anxiety x Age interaction in the explicit memory condition, $F(1,195)=0.17$, $p=0.682$. No other significant effects on task ratings were noted, all ps>0.05.

Linear mixed-effects modeling of reaction time indicated a main effect of Attention Condition, $F(1,4180)=85.13$, $p<0.001$, with greater mean reaction time for explicit memory relative to threat appraisal. This effect was qualified by a quadratic trend × Attention Condition interaction, $F(1,4180)=4.23$, $p=0.04$. Although follow-up analyses indicated the quadratic trend to be non-significant in both conditions, ps>0.189, different quadratic trends associated with the reaction time measure across morphs emerged as a function of attention condition (See Figure S2). These patterns suggest that threat appraisal and explicit memory tap into different psychological processes. No other significant effects on reaction time were noted, all ps>0.07; of note, neither anxiety diagnosis nor age were associated with reaction time with sufficient statistical evidence in terms of main effects or interactions.

Reaction time-modulated fMRI response. A four-way interaction emerged for slope, reflecting the linear trend of reaction time-modulated activation across stimuli. Thus, this manifested across morphed stimuli (one factor) as a function of anxiety (second factor), age (third factor), and attention condition (fourth factor) in the right dlPFC (Figure S4; Talairach coordinates, X,Y,Z: 36, 16, 51; 69 voxels; peak: $F(1,4180)=17.62$). To decompose this

interaction, post-hoc analyses compared coefficients for the linear trends of reaction time effects across morphed stimuli based on anxiety diagnosis, age, and attention condition. While analyses modeled age as a continuous regressor, post-hoc resulted are plotted based on median split (median=17.29 years) to illustrate the patterns of age-related differences. For simplicity, participants whose age is above the median are referred to as adults, and those below the median as youths.

One slope arose in the threat-appraisal condition in healthy adults versus adults with anxiety. For these events in healthy adults, reaction time during threat appraisal, compared to explicit memory, scaled more linearly with levels of threat in facial morphs for activation in the right dlPFC ($t(59)=2.74$, $p=.008$). In adults with anxiety disorders, in contrast, no such reaction time-related modulation occurred ($ps>.33$), thereby explaining the healthy-versus-anxious interaction in adults. That is, healthy and anxious adults showed significantly different activation during the threat appraisal ($t(98)=2.16$, $p=.03$, Cohen's $d = .43$) but not explicit memory conditions ($t(98)=1,35$, $p=.18$, Cohen's $d = .27$).

A distinct slope arose in youth. In healthy youth, as in healthy adults, reaction time during threat appraisal, compared to explicit memory, also did scale differentially with levels of threat in facial morphs for activation in right dlPFC ($t(46)=-3.71$, $p=.001$). However, relative to healthy adults, healthy youth showed an opposite pattern of relationship between dlPFC activation and reaction time during both the threat appraisal task ($t(105)=-3.73$, $p<.001$, Cohen's $d =.71$ and the explicit memory task ($t(105)=2.55$, $p=.012$, Cohen's $d =.48$). Finally, youths with anxiety failed with enough statistical evidence to show differential scaling of reaction time with facial morphs across the two attention states.

Additionally, the 3-way reaction time-modulated interaction of anxiety, age, and attention task resulted in significant activation the right cerebellum, Talairach coordinates (X,Y,Z): 9, -49, -41; cluster size=88 voxels; and peak: $F(1,4180)=20.50$. No significant clusters emerged for any other reaction time-modulated interaction tests involving anxiety diagnosis.

Supplemental References

1. Kaufman J, Birmaher B, Brent D, et al.: Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 1997; 36:980–988
2. First MB, Spitzer RL, Gibbon M, et al.: Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P). New York, New York State Psychiatric Institute, Biometrics Research Department, 1995
3. Wechsler D: Wechsler Abbreviated Scale of Intelligence (WASI). San Antonio, TX, Harcourt Assessment, 1999
4. Birmaher B, Khetarpal S, Brent D, et al.: The Screen for Child Anxiety Related Emotional Disorders (SCARED): scale construction and psychometric characteristics. J Am Acad Child Adolesc Psychiatry 1997; 36:545–553
5. The Research Units On Pediatric Psychopharmacology Anxiety Study Group: The Pediatric Anxiety Rating Scale (PARS): Development and Psychometric Properties. J Am Acad Child Adolesc Psychiatry 2002; 41:1061–1069
6. White LK, Sequeira S, Britton JC, et al.: Complementary features of attention bias modification therapy and cognitive-behavioral therapy in pediatric anxiety disorders. Am J Psychiatry 2017; 174
7. Britton JC, Grillon C, Lissek S, et al.: Response to learned threat: An fMRI study in adolescent and adult anxiety. Am J Psychiatry 2013; 170:1195–1204
8. Shechner T, Britton JC, Ronkin EG, et al.: Fear condition and extinction in anxious and nonanxious youth and adults: examining a novel developmentally-appropriate fear-conditioning task. Depress Anxiety 2015; 32:277–288
9. Cox RW: AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 1996; 29:162–73
10. Gold AL, Shechner T, Farber MJ, et al.: Amygdala–Cortical Connectivity: Associations

with Anxiety, Development, and Threat. Depress Anxiety 2016; 33

11. Desikan RS, Ségonne F, Fischl B, et al.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 2006; 31:968–80
12. McLaren DG, Ries ML, Xu G, et al.: A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. Neuroimage 2012; 61:1277–1286
13. Tseng W-L, Deveney CM, Stoddard J, et al.: Brain Mechanisms of Attention Orienting Following Frustration: Associations With Irritability and Age in Youths. Am J Psychiatry 2019; 176:67–76
14. Stoddard J, Tseng W-L, Kim P, et al.: Association of Irritability and Anxiety With the Neural Mechanisms of Implicit Face Emotion Processing in Youths With Psychopathology. JAMA Psychiatry 2017; 74:95

Supplemental Figure Legends

Figure S1. **Fear conditioning and extinction: Self-report and psychophysiological measures**. Mean (A) self-reported fear (scale 0-10), (B) skin conductance response (in microSiemens), and (C) electromyography (*T*-scores) measures in response to the conditioned stimuli (CS-, CS+) by task phase (pre-acquisition, acquisition, extinction), by anxiety group (healthy, anxious).

Note: Error bars represent one standard error of the mean. Abbreviations: Pre-acq, pre-acquisition; ISI, inter-stimulus interval; ***, $p<0.001$;

Figure S2. Extinction recall: Rating and reaction time measures. (A) mean ratings of threat appraisal and explicit memory of the CS morphs (from 0% CS+ to 100% CS+), and (B) reaction time for making these ratings (in milliseconds), by anxiety group (healthy, anxious) and age group (adult, youth).

Abbreviations: CS+, reinforced conditioned stimulus; ms, millisecond.

Figure S3. Extinction recall: Interaction of age, anxiety diagnosis, and attention condition on behavioral ratings. A three-way interaction of age, anxiety diagnosis, and attention condition was observed for the behavioral task ratings collected during the fMRI, $F(1,4180)=18.40$, $p<0.001$. Mean behavioral task ratings (y-axis: Likert scale 0 [lowest] to 6 [highest]) during fMRI scanning are plotted separately for the two attention conditions, i.e., (a) threat appraisal (“How afraid are you?”) and (b) explicit memory (“How likely was she to scream?”) conditions, and anxiety diagnosis groups (healthy: red, anxious: blue), based on age (x-axis: continuous, in years).

Figure S4. Reaction time-modulated activation in dorsolateral prefrontal cortex (dlPFC). Whole-brain analyses of reaction time-modulation revealed a significant interaction of anxiety diagnosis, age, attention condition, and linear trend in the right dlPFC. (a) Images are shown in neurological convention (i.e., left is left) and thresholded at $F>10.76$, $p<.001$, cluster size > 57 voxels (890.625mm^3), alpha = .05, Bonferroni-corrected for 28 *F*-tests. To decompose these complex interaction effects, mean extracted values for the right dlPFC cluster are plotted separately by attention condition and group (four groups: healthy and anxious groups by median split of age, i.e., adults and youths). (b) In the graphs, the y-axis shows the linear trend of reaction time effect, in which the extracted reaction time effect reflect the effect magnitude in the unit of BOLD response (e.g., % signal change) when reaction time increases by one unit (e.g., one second). Positive values on the y-axis show an increasing reaction time effect across morphs when the morph stimulus becomes similar to the CS+ (i.e., 100% morph). In contrast, negative values reflect a decreasing relationship of reaction time effect across morphs with the morph stimulus becoming similar to the CS+."

*$p<.05$. Error bars represent ± standard deviation.

Abbreviations: dlPFC, dorsolateral prefrontal cortex

Table S1. Psychophysiology visit findings: ANCOVA results for Phase × Stimuli × Anxiety × Age for Self-report, Skin conductance, & EMG.

	Self-Report			Skin conductance			EMG		
Effect	*df*	*F*-value	*p*-value	*df*	*F*-value	*p*-value	*df*	*F*-value	*p*-value
Task									
Phase	2,394	**8.54**	<.001	2,386	**7.47**	<.001	2,386	2.28	.104
Stimulus	1,197	**20.65**	<.001	1,193	**7.81**	.006	2,386	**13.83**	<.001
Phase × Stimulus	2,394	**13.56**	<.001	2,386	**12.75**	.001	4,772	**2.81**	.025
Age									
Age	1,197	3.24	.074	1,193	**53.40**	<.001	1,193	2.16	.143
Phase × Age	2,394	1.15	.319	2,386	1.86	.157	2,386	1.98	.140
Stimulus × Age	1,197	0.66	.417	1,193	3.69	.056	2,386	1.33	.265
Phase × Stimulus × Age	2,394	0.22	.799	2,386	**5.21**	.006	4,772	0.87	.480
Anxiety									
Anxiety	1,197	**7.61**	.006	1,193	**4.46**	.036	1,193	0.34	.562
Age × Anxiety	1,196	0.01	.999	1,192	0.22	.641	1,192	0.02	.876
Phase × Anxiety	2,394	1.25	.287	2,386	0.05	.951	2,386	0.53	.590
Stimulus × Anxiety	1,197	0.20	.657	1,193	0.01	.925	2,386	1.28	.280
Phase × Stimulus × Anxiety	2,394	2.40	.092	2,386	0.84	.434	4,772	1.66	.157
Phase × Stimulus × Age × Anxiety	2,392	0.43	.654	2,384	0.10	.901	8,768	0.87	.482

Note: For self-report and skin conductance response, the Stimulus factor included responses to the CS- and CS+; for EMG, the Stimulus factor included responses during CS-, CS+, and ISI. Abbreviations: EMG, electromyography; df, degrees of freedom; CS, conditioned stimulus; ISI, inter-stimulus interval. Significant results ($p<.05$) are indicated in **bold** font.

Table S2. Whole-brain voxelwise results: task-related activation from linear mixed-effects model testing interactions of anxiety diagnosis, age, attention condition, and linear and quadratic trends across morph stimuli[a]

	Talairach Coordinates			Cluster size	*F* (1, 4180)
	x	**y**	**z**	**# voxels**	
Four-way interaction of diagnosis, age, attention condition, and quadratic trend					
None					
Four-way interaction of diagnosis, age, attention condition, and linear trend					
None					
Three-way interaction of diagnosis, attention condition, and quadratic trend					
None					
Three-way interaction of diagnosis, age and quadratic trend					
None					
Three-way interaction of diagnosis, attention condition, and linear trend					
None					
Three-way interaction of diagnosis, age, and linear trend					
None					
Three-way interaction of diagnosis, age, and attention condition					
Ventromedial PFC	-9	49	-1	667	47.63
Right inferior temporal gyrus	51	-59	-11	263	42.05
	31	-54	-19	65	28.20
Two-way interaction of diagnosis and quadratic trend					
None					
Two-way interaction of diagnosis and linear trend					
None					
Two-way interaction of diagnosis and attention condition					
Ventromedial PFC	4	29	1	161	26.34
Left angular gyrus	-39	-71	34	90	35.51
Left lingual gyrus	-14	-81	-6	77	34.05
Right cerebellum	4	-54	-41	64	24.44

Left anterior temporal cortex	-49	6	-21	61	24.63

Abbreviations: PFC, prefrontal cortex; Whole-brain corrected threshold: $p<.001$, cluster size > 57 voxels (890.625mm^3), alpha = .05 Bonferroni-corrected for 28 *F*-tests).

[a]Only findings from the omnibus interaction tests for anxiety diagnosis are reported, given this study's focus on anxiety pathophysiology.

Figure S1

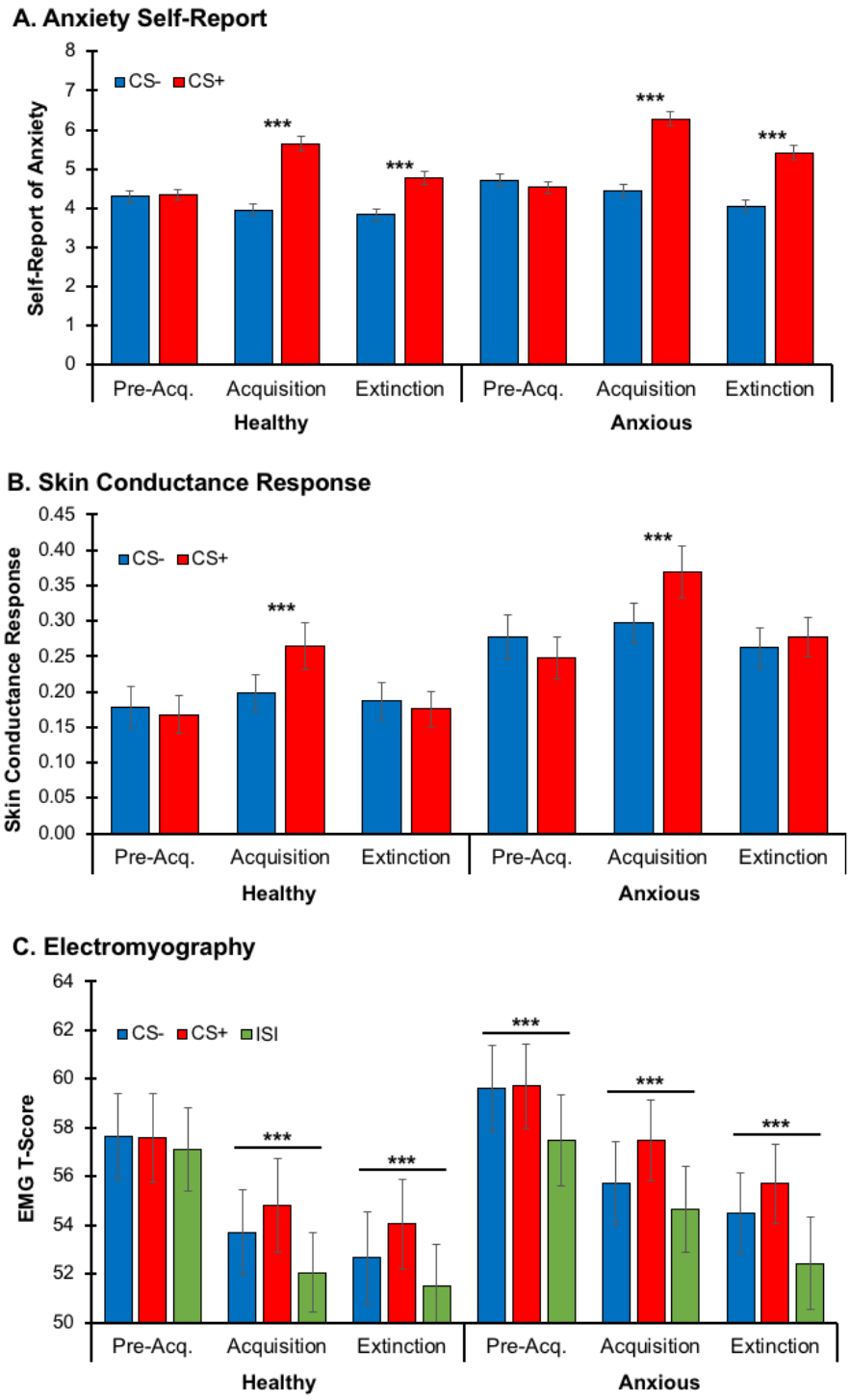

A. Anxiety Self-Report
Self-Report of Anxiety
CS- CS+
***
Pre-Acq.
Acquisition
Extinction
Healthy
Anxious
B. Skin Conductance Response
Skin Conductance Response
C. Electromyography
EMG T-Score
CS- CS+ ISI

Figure S2

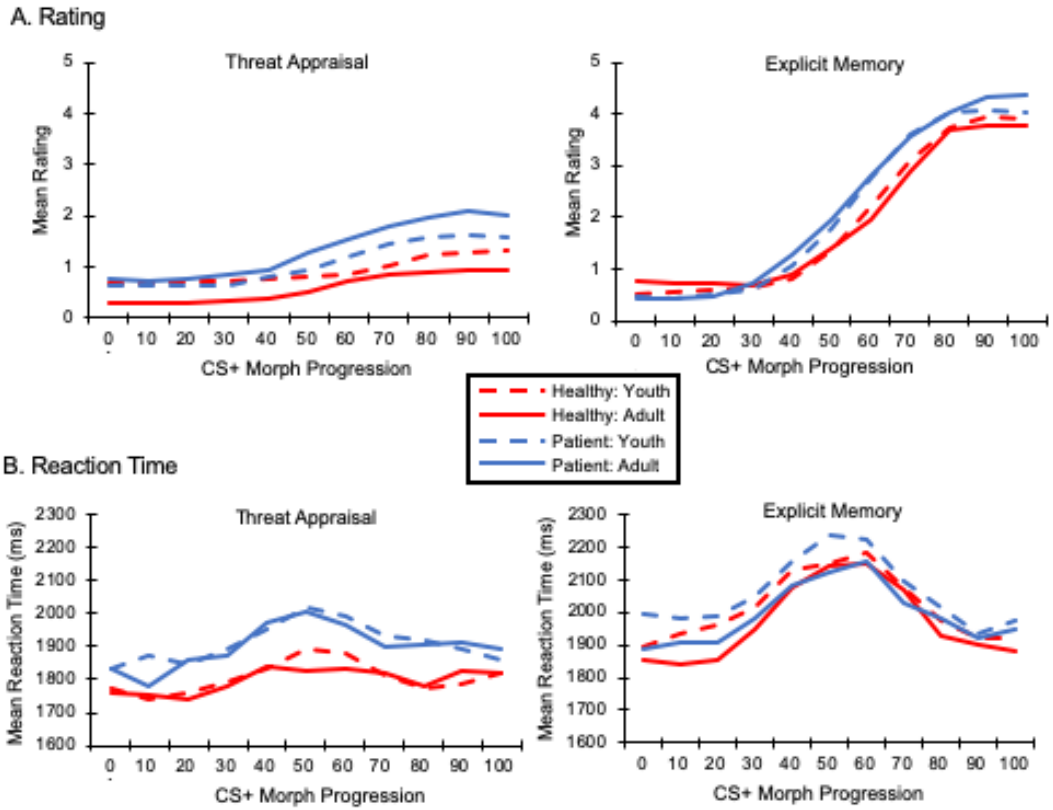

A. Rating
Threat Appraisal
Explicit Memory
Mean Rating
CS+ Morph Progression
Healthy: Youth
Healthy: Adult
Patient: Youth
Patient: Adult
B. Reaction Time
Mean Reaction Time (ms)
CS+ Morph Progression

Figure S3

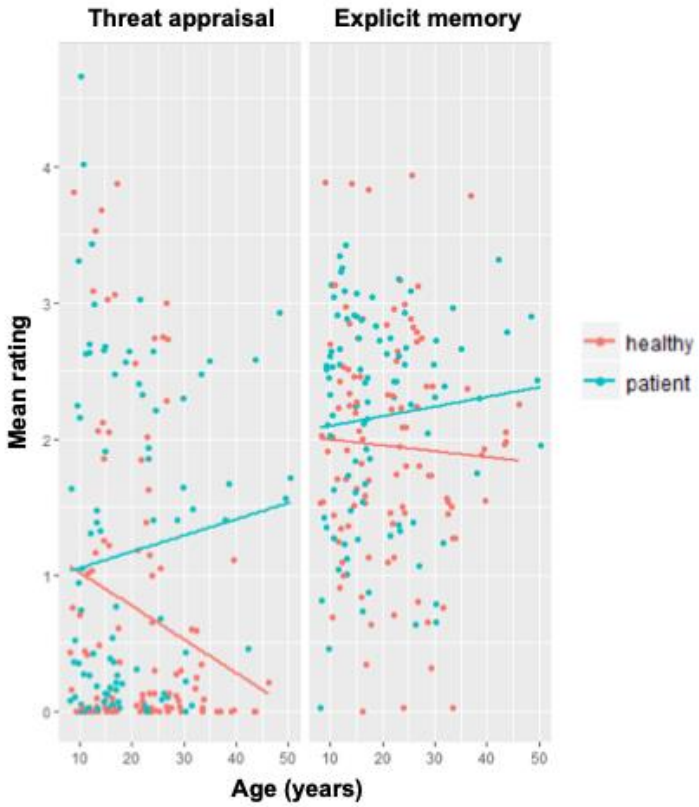
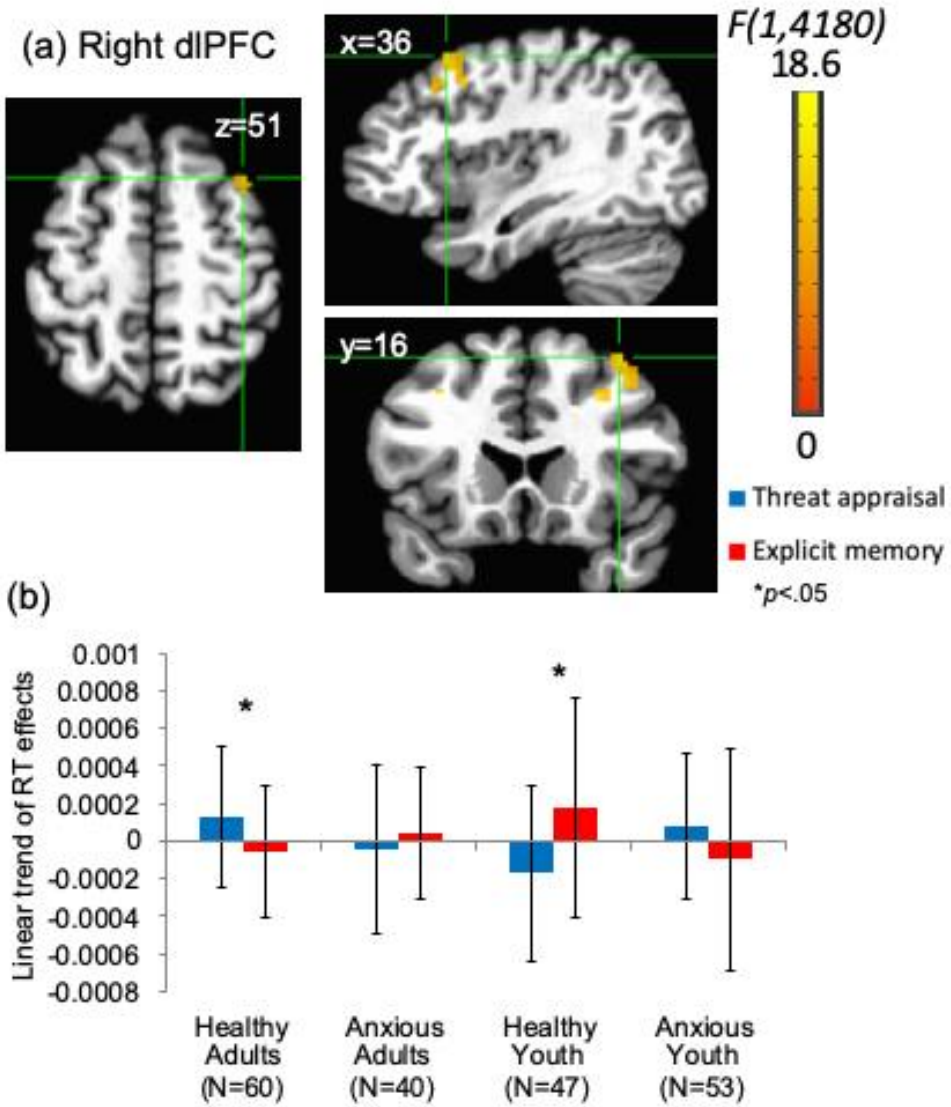

Threat appraisal
Explicit memory
Mean rating
Age (years)
healthy
patient

Figure S4


(a) Right dlPFC
z=51
x=36
y=16
F(1,4180)
18.6
0
Threat appraisal
Explicit memory
*p<.05
(b)
Linear trend of RT effects
0.001
0.0008
0.0006
0.0004
0.0002
0
-0.0002
-0.0004
-0.0006
-0.0008
*
*
Healthy Adults (N=60)
Anxious Adults (N=40)
Healthy Youth (N=47)
Anxious Youth (N=53)