Data supplement for Haller et al., Normalization of Fronto-Parietal Activation by Cognitive-Behavioral Therapy in Unmedicated Pediatric Patients With Anxiety Disorders. Am J Psychiatry (doi: 10.1176/appi.ajp.20220449)

**Recruitment, exclusion criteria, and additional information for each randomized controlled trial (RCT)**

Recruitment and exclusion criteria

Participants were recruited via direct mailings, advertisements in local news outlets and through community talks. Participants had to be aged 8-17, have an IQ > 70 as assessed by the WASI (1), be proficient in English and meet diagnostic criteria for any current primary anxiety disorder (i.e., social anxiety disorder, generalized anxiety disorder and/or separation anxiety disorder) and be medication free. Exclusion criteria included any serious medical condition, severe pervasive neurodevelopmental disorder, substance use, suicidal ideation, major depressive disorder, obsessive-compulsive disorder, and posttraumatic stress disorder; a lifetime history of psychosis, bipolar disorder, or extreme trauma was also exclusionary. Patients' other comorbid diagnoses included: Attention Deficit Hyperactivity Disorder (ADHD (n=2), Tic Disorder (n=1) and Enuresis (n=1). Stage of pubertal development was assessed via self-report ratings of schematic drawings of secondary sex characteristics. Drawings represent the five Tanner stages of pubertal development (2, 3). Parents provided consent for youth ratings. Participants' mean Tanner stage (4), averaged across the pubic hair and breast development in females and pubic hair and genital development in males, was *M*=3.19 *SD*=1.42 for females and *M*=3.47 *SD*=1.52 for males, respectively (N=17 youth did not have data available/chose not to respond).

Healthy controls (HC) had to be free of current or past history of psychopathology as determined by a semi-structured interview (5).

Additional information for each trial

The first of two RCTs included in the current study was completed in 2015 (ClinicalTrials.gov identifier: NCT00018057). The main aim of the study was to examine the efficacy of CBT-adjunct attention bias modification training (ABMT; 6) alongside assessing associations between pre-treatment neural seed-based connectivity profiles and treatment response. The primary outcome measures reported are the PARS and CGI-I. Treating therapists were two licensed psychologists with >5 years of experience delivering CBT to pediatric patients with anxiety disorders. Participants were randomized to receive either active or placebo ABMT, consisting of a 5-minute modified dot-probe task, delivered before and after the 5$^{th}$ to the 12$^{th}$ CBT session. For additional details, see (6).

The second trial was registered on September 14th, 2017 (Clinicaltrials.gov identifier: NCT03283930). The trial registration took place retrospectively; data acquisition started February 1st, 2017 and is ongoing. The main aim of the study is to examine the efficacy of CBT-adjunct ABMT. The main goal for the concurrently collected imaging data is to replicate previous findings regarding predicting treatment response from pre-treatment neural connectivity profiles. The ABMT training in this trial combines a modified dot-probe task with a visual search task to increase engagement and adherence; training takes around 10-15-minutes. The blinds were not broken for the purpose of the current analyses. Participants were randomized to either sham or active ABMT, delivered either before or after the CBT session. The primary outcome measures reported are the PARS and CGI-I. Treating therapists were two licensed psychologists with >5

years of experience delivering CBT to pediatric patients with anxiety disorders. For additional detail, please see the published protocol (7).

Clinician ratings across both trials were completed on average within a month of the fMRI scan ($M_{\text{pre-treatment}}$=30 days *SD*=24 days; $M_{\text{post-treatment}}$=10 days, *SD*=12 days).

Previous publications

Pre-treatment fMRI data has been previously published as part of two cross-sectional studies examining associations between threat processing and dimensionally assessed anxiety and irritability (8, 9). FMRI and clinical data from the first RCT has been published detailing the effects of ABMT (6) and examining temporal stability of task contrasts in HCs (10).

**fMRI threat processing task**

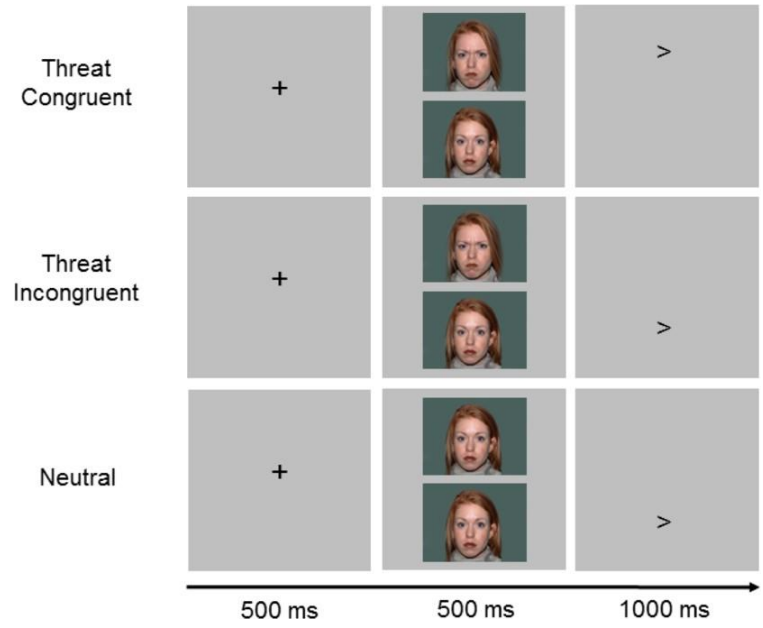


**FIGURE S1.** The task was provided by the Tel-Aviv University/NIMH Attention Bias Modification Treatment Initiative (http://people.socsci.tau.ac.il/mu/anxietytrauma/research/).

Each trial begins with a fixation cross (500ms) followed by the face pair (500ms) and concluded by the presentation of a probe (1000ms). In addition to the 80 congruent, incongruent and neutral trials, an additional 80 trials were fixation only trials to increase jittered intervals and provide an additional baseline.

**Secondary analyses on an at-risk (AR) sample to assess stability of anxiety associations**

Since we were unable to include a patient control arm of medication-free, anxious youth, we conducted a secondary analysis on published data (11). The aim was to test whether anxiety-associated differences in BOLD response remain stable over time, across development, in the absence of treatment.

**Participants**

Participants were drawn from a larger longitudinal community cohort of 291 developmentally healthy children selected at 4 months of age based on criteria for behavioral inhibition (BI), i.e., reactivity to novelty (12), a temperamental risk factor for anxiety. Behavioral inhibition was reassessed at age 2 and 3 and the cohort was median split into low and high BI samples. Exclusion criteria for the current study included psychotropic medication use or illicit substance use, contraindications to MRI, serious medical condition and an IQ<70 (1). Participants also needed to be free of acute psychopathology in need of immediate treatment, as determined by structured psychiatric interviews using the Schedule for Affective Disorders and Schizophrenia for School-Age Children - Present and Lifetime Version (5).

The final sample included 87 participants (IQ: *M*=117, *SD*=13 [IQ data were not available for 4 participants]; 61% female; 17% African American/Black, 2% Asian, 71% White, 8% Hispanic, 1% other; Maternal education: 18% high school graduate, 39% college graduate, 38% graduate school graduate, 4% other) who provided data at one of two time points: around age 10 (*M*=10.51 years, *SD*=0.43) and age 13 (*M*=13.04, *SD*=0.65) years. 61 participants provided data at age 10 (59% females; 30 youth high in BI) and 64 provided data at age 13 (67% females; 35 youth high in BI), for a total of 125 scans. Thirty- eight participants provided data at both time-points. Of the

87 participants who contributed data, four had a diagnosis of current anxiety disorder at age 10 (generalized anxiety disorder n=1, social anxiety disorder n=1, specific phobia n= 2). At age 13, nine participants had a diagnosis of current anxiety disorder (generalized anxiety disorder n= 3, social anxiety disorder n=2, generalized anxiety and social phobia n=1, specific phobia n=2, anxiety disorder not otherwise specified n=1).

Study procedures were approved by the National Institute of Mental Health and University of Maryland Institutional Review Boards. Informed consent and assent were obtained from parents and youth, respectively.

Associations between behavioral inhibition, anxiety symptoms, threat bias and connectivity were previously reported for this sample (11, 13).

## Measures

### Anxiety symptoms

Severity of anxiety was measured via parent- and child-reported Screen for Childhood Anxiety Related Emotional Disorders (SCARED) (14) within 6 weeks of each scan. The SCARED is a 41-item child- and parent-report measure of anxiety symptomology. Child- and parent-report were averaged for a single score reflecting current anxiety symptom severity. Anxiety at the first time point was $M_{SCARED}$=16.33, *SD*=8.68 and at the second $M_{SCARED}$=10.63, *SD*=6.95.

### Threat attention task

Participants completed a dot-probe task similar to the version used in the treatment-seeking sample. This version, however, included happy-neutral trials in addition to the angry-neutral and neutral-neutral trials used in the treatment-seeking sample. Participants were instructed to respond

via button press to indicate the direction of an arrow probe that followed a display of either angry-neutral, happy-neutral or neutral-neutral face pairs of the same actor. The task had five conditions: 1) threat/happy congruent trials, with probes presented in the angry/happy face location of angry-neutral or happy-neutral pairs; 2) threat/happy incongruent trials, with probes presented in the neutral face location of angry-neutral or happy-neutral pairs; and 3) neutral trials, with probes presented in either neutral face location of neutral-neutral pairs.

The task was composed of 48 congruent and 48 incongruent trials for each emotion (happy, angry) and 96 neutral-neutral trials, evenly distributed between four functional runs.

**Imaging data acquisition and preprocessing**

Functional image volumes with 47 contiguous interleaved axial slices were obtained with a T2*-weighted echo-planar sequence (TR = 2300ms; TE = 25ms; flip angle = 50°; field of view (FOV) = 240mm; matrix = 96 ξ 96; , slice thickness 3 mm). Functional data were anatomically localized and co-registered using a high-resolution T1-weighted whole-brain volumetric scan via a high-resolution magnetization prepared gradient echo sequence (MPRAGE; TE = min full; TI = 425ms; flip angle = 7°; FOV = 256 $mm^3$; matrix = 256 × 256 × 256).

Data were preprocessed with the same pipeline as the treatment-seeking/HC sample data. Data were analyzed using Analysis of Functional NeuroImages (AFNI; http://afni.nimh.nih.gov/afni/) (15) v18.3.03. Standard preprocessing included despiking, slice- timing correction, distortion correction, alignment of all volumes to a base volume, non-linear registration to the MNI template, spatial smoothing to 6.5 mm FWHM, masking, and intensity scaling. First-level models used a generalized least squares time series fit with restricted maximum likelihood estimation of the temporal autocorrelation structure with regressors for the five conditions (congruent, incongruent, neutral) and error trials per participant, modeled with a gamma hemodynamic response function.

Preprocessing and first-level general linear models (GLM) controlled for head motion. Six head motion parameters were included as nuisance regressors in the individual-level models. During preprocessing, any pair of successive TRs where the sum head displacement (Euclidean norm of the derivative of the translation and rotation parameters) between those TRs exceeded 1 mm were censored. TRs were also censored when more than 10% of voxels were timeseries signal outliers. Participants were excluded if the average motion per TR after censoring was more than 0.25 mm, or if more than 15% of TRs were censored for motion/outliers, or if behavioral performance accuracy was <70%

**Data analyses**

The analysis examined change in the association of anxiety and brain activation over the course of development. The analytic approach was consistent with the pre-post analyses conducted in the treatment-seeking sample, except that anxiety (SCARED total score) was included as a continuous variable. The behavioral and neuroimaging model included SCARED score as a continuous covariate, reflecting anxiety severity and two within-subjects factors, one for condition with three levels (congruent, incongruent, neutral) and one for time with two levels (first and second scan) and participant as a random factor. A whole-brain, voxel-wise linear mixed effects model (3dLMEr in AFNI; 16) was used with a focus on the anxiety-by-time interaction.

To correct for multiple comparisons, AFNI's 3dClustSim was used. Monte-Carlo simulations were performed within a whole-brain gray-matter mask (83,721), where at least 90% of individuals had data across the two time points. To estimate smoothness of the residual time series a Gaussian plus mono-exponential spatial autocorrelation function was used. Smoothness was estimated for each participant and then averaged, for an effective smoothness of 9.32 (acf parameters: a=0.589 b=3.429 c=10.759). For group analyses, two-sided thresholding was used with first nearest

neighbor clustering. All results were thresholded at a voxel-wise *p*-value of .001 and a cluster extent of $k$=19 for a whole brain family-wise error correction of $p$<.05.

A conjunction map was created to illustrate the overlap in brain regions that showed increased stable activation with anxiety across the two developmental time points (main effect of anxiety, cluster-corrected) in the absence of treatment and those regions that changed with treatment in the treatment-seeking group (group-by-time interaction, cluster-corrected). See main manuscript, Figure 1.

## Results

### Behavioral effects

A main effect for anxiety was observed for mean task response time ($F$(1,519)=19.22, $p$<.001). Reaction times were longer with increased anxiety. The anxiety-by-time-by-condition and the anxiety-by-time interactions were not significant (all *p*s>.05).

### Whole-brain analyses

For a summary of results see Table S1 and S2. Twenty-seven cluster showed a significant main effect of anxiety including clusters in the parietal and prefrontal cortex. Thirty-eight clusters emerged that showed a significant anxiety-by-time interaction. Clusters were largely located in the primary visual cortex and temporal lobe. No clusters emerged for the anxiety-by-condition interaction or the anxiety-by-condition-by-time interaction.

The conjunction analysis between the group-by-time interaction (changing activation patterns with treatment) and the main effect of anxiety in the AR sample (anxiety-associated differences across two time points in development) illustrates overlapping clusters in frontal and parietal cortex. This provides preliminary evidence that the treatment-related changes observed in the main analysis may be treatment specific, and not simply an effect of time.

**TABLE S1.** Stable associations of SCARED score with neural activation across all task conditions

| Region | k | mm3 | CM LR | CM PA | CM IS | Mean | SEM |
|---|---|---|---|---|---|---|---|
| R Superior Occipital Gyrus, R Middle Occipital Gyrus | 328 | 5125 | 26.1 | -78 | 28.6 | 23.09 | 0.51 |
| R Lingual Gyrus | 115 | 1797 | 20.2 | -82.5 | -6 | 23.80 | 0.78 |
| R Inferior Parietal Lobule | 105 | 1641 | 35.7 | -45.4 | 46.1 | 20.58 | 0.57 |
| R Supramarginal Gyrus | 101 | 1578 | 56 | -30.6 | 32.7 | 19.41 | 0.49 |
| L Calcarine Gyrus | 70 | 1094 | -11 | -64.3 | 7.3 | 20.80 | 0.70 |
| R Superior Frontal Gyrus | 70 | 1094 | 19.1 | 31.2 | 49.9 | 17.09 | 0.33 |
| R Middle Temporal Gyrus | 58 | 906 | 58.2 | -33.7 | -2.4 | 18.72 | 0.47 |
| L Middle Temporal Gyrus | 55 | 859 | -60.8 | -45.8 | 4.4 | 19.40 | 0.57 |
| L Superior Medial Gyrus | 52 | 813 | 1.4 | 54.9 | 32.3 | 17.32 | 0.44 |
| R Inferior Temporal Gyrus | 47 | 734 | 45.9 | -59 | -9.3 | 21.30 | 0.89 |
| R Middle Occipital Gyrus | 45 | 703 | 34.5 | -83.6 | 8.1 | 21.01 | 0.98 |
| L Inferior Temporal Gyrus | 42 | 656 | -45.4 | -54.3 | -8.9 | 17.74 | 0.47 |
| L Superior Frontal Gyrus | 36 | 563 | -18.7 | 32.9 | 46.6 | 17.12 | 0.48 |
| L Superior Occipital Gyrus | 34 | 531 | -22.5 | -87 | 22 | 23.89 | 1.25 |
| R Anterior Cingulate Cortex | 31 | 484 | 2 | 24.4 | 19.4 | 17.28 | 0.54 |
| R Cerebellum | 29 | 453 | 32.6 | -71.9 | -51.5 | 18.82 | 0.70 |
| R Superior Medial Gyrus | 28 | 438 | 9.4 | 60.7 | 9.5 | 17.42 | 0.61 |
| R Fusiform Gyrus | 27 | 422 | 40.6 | -61.6 | -19.6 | 19.30 | 0.88 |
| L Paracentral Lobule | 27 | 422 | -4.5 | -20.2 | 68.2 | 19.17 | 0.84 |
| R Mid Orbital Gyrus | 25 | 391 | 8.2 | 36.9 | -10.1 | 18.95 | 0.72 |
| L Superior Medial Gyrus | 23 | 359 | 1.7 | 45.1 | 43.5 | 20.36 | 1.06 |
| R Cuneus | 22 | 344 | 8.8 | -69.3 | 22.3 | 18.37 | 0.74 |
| L Postcentral Gyrus | 22 | 344 | -50.7 | -23.8 | 57.7 | 20.71 | 0.99 |
| R Calcerine Gyrus | 21 | 328 | 15.6 | -62 | 7.1 | 19.56 | 0.99 |
| R Middle Temporal Gyrus | 21 | 328 | 45.3 | -58.2 | 20.3 | 16.16 | 0.54 |
| L Superior Parietal Lobule | 21 | 328 | -24.4 | -54 | 64.5 | 18.65 | 0.89 |
| L Middle Frontal Gyrus | 19 | 297 | -41.7 | 22.4 | 36.1 | 20.22 | 1.39 |

*Note*. Cluster-corrected voxel-wise linear mixed-effects model results are presented here summarizing regions showing a main effect of group and those showing a group-by-time interaction. k=number of voxels in cluster, mm3=cluster volume, CM=center of mass of cluster, SEM=standard error of the mean, LR=left-right (x), PA=posterior-anterior (y), IS=inferior-superior (z), anatomical locations: Eickhoff-Zilles macro labels from N27 (MNI_ANAT space).

**TABLE S2.** Changing associations between SCARED score and Time with neural activation across all task conditions

| Region | k | mm3 | CM LR | CM PA | CM IS | Mean | SEM |
|---|---|---|---|---|---|---|---|
| L Calcarine Gyrus | 151 | 2359 | 0.8 | -75.7 | 12.7 | 22.57 | 0.66 |
| L Rolandic Operculum | 121 | 1891 | -44.6 | -27.9 | 16.3 | 20.64 | 0.60 |
| R Superior Temporal Gyrus | 111 | 1734 | 53.6 | -20.2 | 11.1 | 19.06 | 0.44 |
| R Middle Temporal Gyrus | 86 | 1344 | 52.6 | -50.5 | 16.1 | 18.63 | 0.49 |
| L Middle Temporal Gyrus | 71 | 1109 | -49.4 | -69.9 | 10 | 19.92 | 0.54 |
| R Cuneus | 65 | 1016 | 15.7 | -86.8 | 24.7 | 24.57 | 0.96 |
| R Inferior Parietal Lobule | 54 | 844 | 54.7 | -41 | 48.2 | 20.63 | 0.71 |
| R Postcentral Gyrus | 53 | 828 | 48.1 | -20.9 | 39.1 | 21.56 | 0.87 |
| L SMA | 52 | 813 | 1.3 | -9.2 | 57.7 | 20.55 | 0.72 |
| L Cerebellum | 50 | 781 | -9.2 | -76.8 | -38.3 | 20.38 | 0.79 |
| R Cerebellum, R Lingual Gyrus | 47 | 734 | 11.2 | -69.7 | -13 | 20.73 | 0.84 |
| R Cerebellum (VI) | 46 | 719 | 34 | -50.5 | -23 | 20.77 | 0.87 |
| R Inferior Occipital Gyrus | 44 | 688 | 32.9 | -88.3 | -4.6 | 20.81 | 0.83 |
| L Inferior Occipital Gyrus | 39 | 609 | -42.1 | -73.8 | -14.1 | 23.62 | 1.41 |
| R Lingual Gyrus | 38 | 594 | 8.3 | -79.3 | -5 | 22.59 | 1.06 |
| L Fusiform Gyrus | 35 | 547 | -31.6 | -56.3 | -18.1 | 17.74 | 0.55 |
| R Lingual Gyrus | 35 | 547 | 18.9 | -84.1 | -4.8 | 22.24 | 0.94 |
| R Precuneus | 32 | 500 | 16.4 | -62.3 | 35.1 | 20.31 | 1.11 |
| R Hippocampus | 31 | 484 | 20.8 | -18.7 | -8.5 | 18.00 | 0.74 |
| L Supramarginal Gyrus | 31 | 484 | -56.3 | -46.5 | 25.3 | 23.72 | 1.28 |
| L Rolandic Operculum/L Insular Lobe | 28 | 438 | -38.8 | -7 | 13 | 17.58 | 0.61 |
| L Middle Temporal Gyrus | 28 | 438 | -32.2 | -85.6 | 21.1 | 19.08 | 0.85 |
| L Inferior Parietal Lobule | 27 | 422 | -58.2 | -37.4 | 46.8 | 20.87 | 1.14 |
| L SMA | 27 | 422 | -3.2 | 10.1 | 55.9 | 19.95 | 0.81 |
| R Fusiform Gyrus | 26 | 406 | 30.5 | -75.4 | -8.3 | 23.81 | 1.49 |
| L Precentral Gyrus | 26 | 406 | -28.9 | -28.3 | 62.9 | 18.88 | 0.82 |
| L Middle Temporal Gyrus | 25 | 391 | -55.8 | -12.8 | -14.9 | 18.74 | 0.91 |
| L Superior Temporal Gyrus | 24 | 375 | -55.1 | -45.6 | 13.3 | 21.26 | 1.28 |
| L Cerebellum (VII) | 22 | 344 | -33.3 | -70.9 | -54 | 17.92 | 0.57 |
| L Cerebellum (VIII) | 22 | 344 | -8.8 | -63.4 | -51.4 | 17.10 | 0.49 |
| R Lingual Gyrus | 21 | 328 | 18.8 | -85.5 | -15.6 | 19.72 | 1.06 |
| L Superior Occipital Gyrus | 21 | 328 | -23.4 | -83.9 | 33.6 | 20.92 | 1.14 |
| L Inferior Parietal Lobule | 20 | 313 | -30.6 | -47.1 | 49.8 | 18.78 | 1.08 |
| R Paracentral Lobule | 20 | 313 | 0.6 | -39.8 | 63.4 | 18.90 | 0.64 |
| R Fusiform Gyrus | 19 | 297 | 29.4 | -39.2 | -14.2 | 19.68 | 1.16 |
| R Inferior Temporal Gyrus | 19 | 297 | 54.6 | -41.8 | -13.5 | 19.59 | 1.26 |
| L Hippocampus | 19 | 297 | -22.4 | -26.5 | -4.3 | 18.12 | 0.69 |
| R SMA | 19 | 297 | 10.6 | 2.7 | 60.1 | 17.99 | 0.72 |

*Note*. Cluster-corrected voxel-wise linear mixed-effects model results are presented here summarizing regions showing a main effect of group and those showing a group-by-time interaction. k=number of voxels in cluster, mm3=cluster volume, CM=center of mass of cluster, SEM=standard error of the mean, LR=left-right (x), PA=posterior-anterior (y), IS=inferior-superior (z), anatomical locations: Eickhoff-Zilles macro labels from N27 (MNI_ANAT space).

## Supplemental Data Analyses

The main analysis in this report focused on activation patterns in youth with anxiety disorders and HC youth pre- to post-treatment. To supplement the main analyses, we examined 1) temporal stability of BOLD activation patterns during task contrasts in HC youth and in the AR sample, 2) effects of ABMT on neural activation patterns in the first, unblinded RCT, 3) changes from pre-to post-treatment per trial, 4) changes from pre-to post-treatment using the threat contrast, 5) changes from pre-to post-treatment, controlling for sex assigned at birth, 6) associations between reaction time and BOLD response and 7) associations between activation patterns pre-treatment and treatment response using the CGI-I.

### 1. Temporal stability of BOLD signal

### 1a. Temporal stability of BOLD signal in HC youth

Here, we completed analyses of test-retest reliability of the 62 HC youth's behavioral and imaging contrasts. These analyses supplement analyses previously published by White and colleagues (10); data from 38 participants was included in the sample by White and colleagues; data from 24 participants has not been published before.

#### Methods

To assess test-retest reliability, we used intraclass correlation coefficients (ICC[3,1]; 17), applying a linear mixed-effects model with a random effect for participant and a fixed effect for time. The ICC[3,1] formulation examines the consistency in the rank of value, which accounts for systematic changes over time, such as practice and repeated exposure effects. For both the reaction time and imaging contrasts, we applied a Bayesian framework (18) with a random effect for participant modeled with Gamma priors (shape = 2, rate = 0.5). This approach has several

advantages over the conventional technique, including the ability to deal with missing data, negative ICC values and confounding effects (18).

We tested the following task contrasts for the behavioral (RT) data: grand mean RT and incongruent vs. congruent trials. For the imaging data, we tested all events vs. fixation, as well as the two main task contrasts: incongruent vs. congruent trials and threat (including congruent and incongruent trials) vs. neutral trials.

Imaging ICC maps of participant-specific variance are represented in color bins of "poor" (ICC < 0.4), "fair" (ICC = .4–.6), "good" (ICC = .6–.75), and "excellent" (ICC > .75) test–retest reliability and cluster-corrected at *k*=21.

## Results

Behavioral data

The test–retest reliability for reaction time was good to excellent for overall grand mean RT (ICC=.73). The incongruent vs. congruent RT difference score exhibited poor reliability (ICC = .18).

Voxel-wise ICC maps

For the task vs. baseline contrast (Figure S2 and Table S3), reliable participant-specific variance was observed in visual, parietal, and prefrontal cortices.

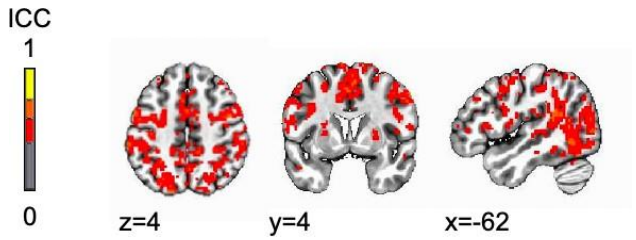


**FIGURE S2.** Regions showing reliable participant-specific variance at ICC>.4

**TABLE S3.** Regions showing reliable participant-specific variance at ICC>.4

| Region | k | mm3 | CM LR | CM PA | CM IS | MI LR | MI PA | MI IS | Mean | SEM |
|---|---|---|---|---|---|---|---|---|---|---|
| L Middle Occipital Gyrus | 12047 | 188234 | -3.5 | -58.6 | 17.8 | -45 | -69.2 | -7 | 0.5013 | 6.60E-04 |
| R Precentral Gyrus | 1547 | 24172 | 22.8 | -0.1 | 49.7 | 57.5 | 0.8 | 45.5 | 0.4919 | 0.0017 |
| L Middle Frontal Gyrus | 105 | 1641 | -32.3 | 37.9 | 31.1 | -27.5 | 38.2 | 30.5 | 0.5161 | 8.10E-03 |
| R Inferior Frontal | 58 | 906 | 23.6 | 6.1 | 5.2 | 22.5 | 5.8 | 3 | 0.4485 | 0.0047 |
| R Putamen | 58 | 906 | 51.2 | 19.6 | 10.6 | 47.5 | 18.2 | 5.5 | 0.4639 | 0.0058 |
| L Heschls Gyrus | 45 | 703 | -32.7 | -25.9 | 13.9 | -32.5 | -24.2 | 13 | 0.4488 | 0.0057 |
| R Superior Temporal Gyrus | 39 | 609 | 48.4 | -29.3 | 1.1 | 47.5 | -26.8 | 3 | 0.4585 | 0.008 |
| L Middle Cingulate Cortex | 35 | 547 | -9.9 | -23 | 41.8 | -12.5 | -24.2 | 38 | 0.4543 | 0.0081 |
| R Middle Frontal Gyrus | 34 | 531 | 39.5 | 31.5 | 39.9 | 35 | 30.8 | 50.5 | 0.4542 | 0.0075 |
| R Precentral Gyrus | 33 | 516 | 38.8 | -26 | 61.7 | 40 | -26.8 | 65.5 | 0.4472 | 0.0089 |
| L Anterior Cingulate Cortex | 21 | 328 | -0.9 | 40.9 | 14.5 | -5 | 43.2 | 13 | 0.4515 | 0.0079 |

*Note*. Cluster-corrected voxel-wise linear mixed-effects model results are presented here summarizing regions showing a main effect of group and those showing a group-by-time interaction. k=number of voxels in cluster, mm3=cluster volume, CM=center of mass of cluster, SEM=standard error of the mean, LR=left-right (x), PA=posterior-anterior (y), IS=inferior-superior (z), anatomical locations: Eickhoff-Zilles macro labels from N27 (MNI_ANAT space).

Next, two task contrasts of interest were examined. No region was reliable at ICC > .4 for the incongruent vs. congruent contrast. For the threat vs. neutral contrast, no reliable clusters emerged at the voxel-wise level at ICC > .4.

### 1b. Temporal stability of BOLD signal in HC youth compared to youth in the AR sample

To examine whether HC youth and youth from the AR sample showed differential stability, a conjunction map was created for the task vs. baseline contrast ICC map from each sample. Regions exhibiting temporal stability of ICC > .4 were largely consistent across samples.

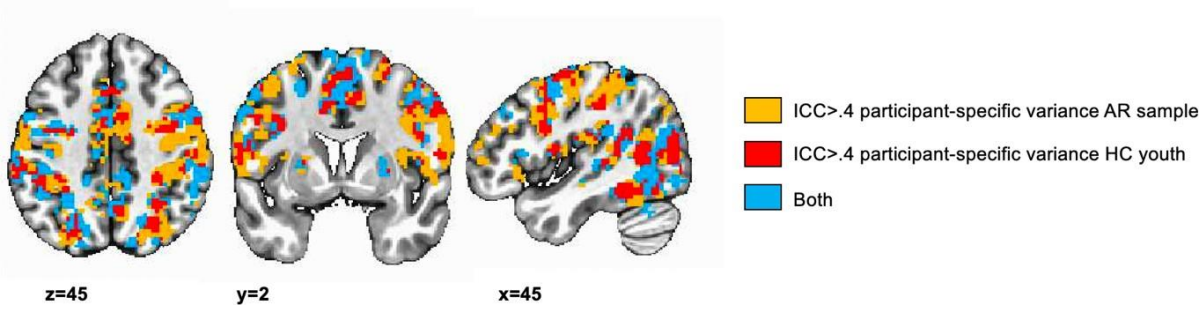


**FIGURE S3.** Conjunction map of regions showing reliable participant-specific variance at ICC>.4 for the HC participants and AR sample

## 2. Effects of ABMT on BOLD activation in the first RCT

We compared change over time between those who received active ABMT (*n*=16) vs. those who received sham ABMT (*n*=15) in the first RCT.

### Methods

The model included group (active, sham) as a two-level between-subjects factor and two within-subjects factors, one for task condition with three levels (congruent, incongruent, neutral) and one for time with two levels (first and second scan, pre- and post-treatment of CBT + active or CBT + sham ABMT) and participant as a random factor. The analyses were whole-brain, voxel-wise linear mixed effects models (3dLMEr in AFNI; 16). For the purpose of this analysis, we focus on group differences, particularly the group-by-condition-by-time interaction. We expect change in activation patterns to be specific to the threat bias (incongruent - congruent) or the threat orienting contrast (threat – neutral) in the active group, who received training to orient away from threat faces. A conjunction map was created to further explore any the overlap in brain regions that showed group (active vs. sham)-specific change (cluster-corrected) and those regions that showed a change in activation with CBT (group-by-time interaction, cluster-corrected).

We used AFNI's 3dClustSim to correct for multiple comparisons within a whole-brain gray-matter mask (80,193), where at least 90% of individuals had data across the two time points. All results were thresholded at a voxel-wise *p*-value of .001 and a cluster extent of *k*=21 for a whole-brain family-wise error correction of $p<.05$.

**Results**

For a summary of results see Table S4. No regions showed the expected significant group- by-condition-by-time interaction. Fifteen clusters emerged that showed a significant group-by- time interaction. Critically, a conjunction map illustrates that there is minimal overlap between these clusters and the regions emerging in the group-by-time interaction in youth with ANX undergoing treatment. No clusters emerged for the group-by-condition interaction or the main effect of group.

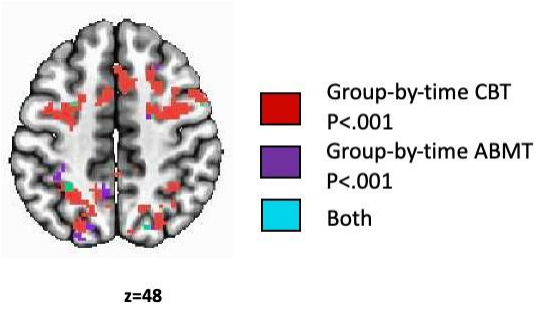


**FIGURE S4.** Conjunction map illustrating the overlap between brain regions showing a change in group activation patterns with Active vs. Sham treatment (group-by-time interaction, cluster-corrected) and those regions that changed with treatment in the treatment-seeking group (group-by-time interaction, cluster-corrected)

**TABLE S4.** Changing ABMT group differences (Active vs. Sham) across time

| Region | k | mm3 | CM LR | CM PA | CM IS | Mean | SEM |
|---|---|---|---|---|---|---|---|
| R Middle Temporal Gyrus | 84 | 1313 | 57.3 | -49.3 | 3.1 | 19.72 | 0.62 |
| R Postcentral Gyrus | 57 | 891 | 54.7 | -21.7 | 35.8 | 17.32 | 0.42 |
| R Cerebellum (IX) | 51 | 797 | 10.3 | -43 | -29.1 | 17.07 | 0.34 |
| L Inferior Parietal Lobule | 47 | 734 | -33.4 | -42 | 45 | 18.36 | 0.52 |
| L Posterior Cingulate Cortex | 32 | 500 | -0.3 | -40.4 | 21.8 | 20.07 | 0.69 |
| L Superior Parietal | 30 | 469 | -14.2 | -75.9 | 52.4 | 18.32 | 0.62 |
| L Superior Occipita Gyrus | 27 | 422 | -16.7 | -66.4 | 27.8 | 18.19 | 0.59 |
| R Mid Orbital Gyrus | 25 | 391 | 13.1 | 64.3 | -6.1 | 18.89 | 0.91 |
| L Thalamus | 25 | 391 | -0.4 | -21.9 | 19.3 | 17.78 | 0.81 |
| L Cuneus | 24 | 375 | -2.7 | -85.8 | 30.7 | 19.54 | 0.92 |
| R Superior Temporal Gyrus | 23 | 359 | 57.9 | 0.2 | -14 | 18.82 | 0.90 |
| L Insular Lobe | 23 | 359 | -44.2 | -2.7 | -0.9 | 17.25 | 0.48 |
| L Middle Temporal Gyrus | 23 | 359 | -47.5 | -71.9 | 17.1 | 17.89 | 0.69 |
| R Precentral Gyrus | 22 | 344 | 58.6 | 5.7 | 24.6 | 20.58 | 1.03 |
| L SMA | 21 | 328 | -8.6 | -9.9 | 75.4 | 17.51 | 0.65 |

*Note*. Cluster-corrected voxel-wise linear mixed-effects model results are presented here summarizing regions showing a main effect of group and those showing a group-by-time interaction. k=number of voxels in cluster, mm3=cluster volume, CM=center of mass of cluster, SEM=standard error of the mean, LR=left-right (x), PA=posterior-anterior (y), IS=inferior-superior (z), anatomical locations: Eickhoff-Zilles macro labels from N27 (MNI_ANAT space).

## 3. Pre- to post-treatment change analysis conducted for each RCT separately

The primary analysis examining change over the course of CBT is repeated here for each RCT separately.

### Methods

We replicated this model for each RCT separately. This analysis examined change over the course of CBT, comparing change over time between ANX and HC youth within whole-brain voxel-wise linear mixed effects model (3dLMEr in AFNI; 16). Each model included group (ANX, HC) as a two-level between-subjects factor and two within-subjects factors, one for task condition with three levels (congruent, incongruent, neutral) and one for time with two levels (first and second scan, pre- and post-CBT for ANX youth) and participant as a random factor. We focus on the same group differences as the main analysis, particularly the group-by-time interaction.

Using AFNI's 3dClustSim, Monte-Carlo simulations were performed to correct for multiple comparisons. Results for both trials were thresholded at a voxel-wise $p$-value of .001 and a cluster extent of $k$=21 for a whole brain family-wise error correction of $p$<.05.

### Results

For a summary of results see Table S5 and S6. As illustrated in Figure S4, the group-by- time interaction yielded several overlapping clusters across both RCTs.

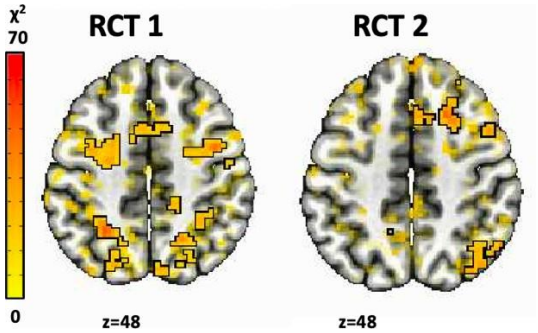


**FIGURE S4.** Group-by-time interaction illustrated separately for each RCT

**TABLE S5.** Changing group differences (ANX vs. HC) across time in the first RCT

| Region | k | mm3 | CM LR | CM PA | CM IS | Mean | SEM |
|---|---|---|---|---|---|---|---|
| R Lingual Gyrus, L Calcarine Gyrus, L Cuneus | 606 | 9469 | 1.5 | -74.9 | 9.5 | 20.45 | 0.24 |
| R Middle Temporal Gyrus, R Inferior Occipital Gyrus | 587 | 9172 | 43.3 | -65 | 0.2 | 23.04 | 0.37 |
| R Superior Parietal Lobule | 401 | 6266 | 20.9 | -59.7 | 52.1 | 19.60 | 0.26 |
| R Rolandic Operculum, R Postcentral Gyrus | 366 | 5719 | 48.4 | -13.6 | 19.5 | 20.92 | 0.31 |
| L Precentral Gyrus | 283 | 4422 | -27.4 | -4.6 | 54.2 | 20.86 | 0.33 |
| R Superior Frontal Gyrus | 277 | 4328 | 32.4 | -1.4 | 56.6 | 22.88 | 0.62 |
| L Middle Occipital Gyrus | 222 | 3469 | -45.8 | -70.8 | 0.3 | 22.21 | 0.42 |
| L Middle Frontal Gyrus, L Inferior Frontal Gyrus | 211 | 3297 | -38.8 | 47.1 | 1.9 | 21.71 | 0.48 |
| L Postcentral Gyrus | 183 | 2859 | -55.5 | -8.3 | 24.8 | 19.75 | 0.39 |
| R Inferior Frontal Gyrus | 151 | 2359 | 47 | 29.4 | 18.5 | 20.02 | 0.41 |
| L SMA | 136 | 2125 | 1.3 | 7 | 47.7 | 18.74 | 0.37 |
| L Inferior Parietal Lobule | 127 | 1984 | -26.6 | -53 | 48 | 23.29 | 0.82 |
| L Superior Parietal Lobule | 119 | 1859 | -16.1 | -71.2 | 53.7 | 20.24 | 0.52 |
| R Putamen | 108 | 1688 | 27.6 | 6.7 | 1.7 | 18.92 | 0.42 |
| R Inferior Frontal Gyrus | 93 | 1453 | 38.9 | 39.3 | -10.8 | 19.14 | 0.45 |
| L Middle Occipital Gyrus | 84 | 1313 | -27 | -74.1 | 28.2 | 21.68 | 0.85 |
| L Middle Orbital Gyrus | 64 | 1000 | 1.1 | 60.5 | -1.5 | 17.97 | 0.39 |
| L Middle Temporal Gyrus | 59 | 922 | -63.8 | -42.4 | -2.7 | 19.01 | 0.65 |
| Cerebellar Vermis, L Cerebellum | 55 | 859 | -3.1 | -53.1 | -21.2 | 18.65 | 0.53 |
| R Superior Temporal Gyrus | 50 | 781 | 65.1 | -36.6 | 22.2 | 20.83 | 0.92 |
| R Calcarine Gyrus | 47 | 734 | 7.7 | -71.1 | 14.2 | 22.52 | 1.18 |
| R Postcentral Gyrus | 47 | 734 | 44.3 | -37.2 | 61 | 21.32 | 0.98 |
| L Putamen | 46 | 719 | -27.6 | -0.2 | 7 | 17.82 | 0.48 |
| R Paracentral Lobule, R Middle Cingulate Cortex | 43 | 672 | 16.2 | -37.8 | 48.6 | 18.63 | 0.51 |
| R Precentral Gyrus | 40 | 625 | 29.3 | -21.2 | 68.1 | 21.32 | 0.88 |
| R Superior Orbital Gyrus | 35 | 547 | 21.9 | 61.8 | -7.2 | 19.70 | 0.80 |
| L Cerebellum (VI) , L Fusiform Gyrus | 33 | 516 | -30.7 | -65.3 | -18.1 | 20.42 | 0.96 |
| L Superior Temporal Gyrus | 31 | 484 | -42.4 | -37.5 | 17.2 | 21.61 | 0.98 |
| L Precuneus | 28 | 438 | -2.6 | -56.8 | 26.7 | 16.67 | 0.41 |
| L Superior Temporal Gyrus | 26 | 406 | -59.4 | -48.9 | 22.4 | 17.27 | 0.48 |
| L Cerebellum (IX) | 25 | 391 | -2.3 | -48.5 | -36.9 | 22.29 | 1.43 |
| L Mid Orbital Gyrus | 24 | 375 | -0.7 | 51.1 | -14.6 | 17.26 | 0.53 |
| L Lingual Gyrus | 23 | 359 | -11.3 | -67.2 | -2 | 18.62 | 1.03 |
| L Middle Occipital Gyrus | 22 | 344 | -27.5 | -87.8 | 6.9 | 22.37 | 1.44 |
| Cerebellar Vermis (8) | 21 | 328 | 1.8 | -58.9 | -30.1 | 19.07 | 0.66 |
| L Middle Occipital Gyrus | 21 | 328 | -39.2 | -86.2 | 1.6 | 19.80 | 0.99 |
| R Insular Lobe | 21 | 328 | 41.9 | -6.2 | 3.2 | 19.16 | 0.83 |

*Note*. Cluster-corrected voxel-wise linear mixed-effects model results are presented here summarizing regions showing a main effect of group and those showing a group-by-time interaction. k=number of voxels in cluster, mm3=cluster volume, CM=center of mass of cluster, SEM=standard error of the mean, LR=left-right (x), PA=posterior-anterior (y), IS=inferior-superior (z), anatomical locations: Eickhoff-Zilles macro labels from N27 (MNI_ANAT space).

**TABLE S6.** Changing group differences (ANX vs. HC) across time in the second RCT

| Region | k | mm3 | CM LR | CM PA | CM IS | Mean | SEM |
|---|---|---|---|---|---|---|---|
| R Superior Frontal Gyrus, R Middle Frontal Gyrus | 300 | 4688 | 30.9 | 15.7 | 52.2 | 19.41 | 0.28 |
| R Anterior Cingulate Cortex | 178 | 2781 | 5 | 36.2 | 25.5 | 18.99 | 0.34 |
| R Superior Frontal Gyrus | 170 | 2656 | 28.6 | 57.1 | 3.5 | 18.56 | 0.30 |
| R Angular Gyrus | 168 | 2625 | 39.6 | -63.3 | 45.2 | 20.22 | 0.47 |
| R Superior Frontal Gyrus | 164 | 2563 | 20.9 | 46.5 | 35.7 | 21.29 | 0.58 |
| L Cerebellum | 139 | 2172 | -23.3 | -81.6 | -32.8 | 18.78 | 0.39 |
| R Superior Medial Gyrus, L Superior Medial Gyrus, L Mid Orbital Gyrus | 132 | 2063 | 1.8 | 62.8 | 5.7 | 19.40 | 0.42 |
| R Caudate Nucleus | 100 | 1563 | 15 | -2.6 | 20.6 | 22.90 | 0.87 |
| L Anterior Cingulate Cortex | 88 | 1375 | 2.2 | 44.9 | 1.2 | 18.46 | 0.42 |
| R Fusiform Gyrus, R Lingual Gyrus | 86 | 1344 | 24.8 | -81.1 | -8.9 | 20.40 | 0.50 |
| R Middle Cingulate Cortex | 75 | 1172 | 1.8 | -32.3 | 28.9 | 17.56 | 0.33 |
| L Caudate Nucleus | 54 | 844 | -18.6 | 0.7 | 22.4 | 20.80 | 0.92 |
| L Precuneus | 51 | 797 | -12 | -44.8 | 61.9 | 17.50 | 0.44 |
| R Middle Frontal Gyrus | 47 | 734 | 39.9 | 19.7 | 33.8 | 19.89 | 0.78 |
| R Caudate Nucleus | 46 | 719 | 18.5 | 20.8 | 11.2 | 19.00 | 0.54 |
| L Precuneus | 46 | 719 | 1.7 | -62.9 | 33 | 17.54 | 0.38 |
| R Cuneus | 45 | 703 | 9.6 | -82.8 | 31.1 | 19.16 | 0.78 |
| R Middle Temporal Gyrus | 43 | 672 | 65.9 | -38 | -6 | 20.61 | 0.91 |
| R Cerebellum | 40 | 625 | 26.2 | -85.6 | -29.5 | 20.93 | 0.85 |
| L Precuneus | 40 | 625 | -12 | -50.3 | 37.2 | 18.37 | 0.60 |
| L Superior Medial Gyrus | 39 | 609 | -3.4 | 53.6 | 36.3 | 18.75 | 0.63 |
| R Paracentral Lobule | 39 | 609 | 4.1 | -31.3 | 61 | 18.66 | 0.66 |
| R Precentral Gyrus | 39 | 609 | 28.3 | -27 | 64.4 | 19.64 | 0.82 |
| L Caudate Nucleus | 38 | 594 | -17.6 | -16.8 | 18.9 | 20.09 | 0.82 |
| R SMA | 37 | 578 | 5 | 17.1 | 47.8 | 18.01 | 0.60 |
| L Precuneus | 35 | 547 | -28.6 | -56.5 | 3.6 | 20.26 | 0.94 |
| L Cerebellum (Crus 1) | 33 | 516 | -40.6 | -69 | -34.6 | 17.26 | 0.54 |
| R Supramaginal Gyrus | 33 | 516 | 53.2 | -43.1 | 39.4 | 16.82 | 0.35 |
| L Lingual Gyrus | 32 | 500 | -13.8 | -74.4 | -9.5 | 20.73 | 0.84 |
| R Inferior Frontal Gyrus | 28 | 438 | 45.5 | 33.9 | 17.8 | 15.65 | 0.36 |
| R SMA | 28 | 438 | 10 | -0.9 | 73.8 | 18.81 | 0.79 |
| R Inferior Temporal Gyrus | 27 | 422 | 47.5 | -51.5 | -9.8 | 18.37 | 0.80 |
| R Superior Temporal Gyrus | 27 | 422 | 48.9 | -43.3 | 17.4 | 17.48 | 0.50 |
| L Inferior Frontal Gyrus | 26 | 406 | -49.7 | 26.4 | 29 | 18.71 | 0.64 |
| R Mid Orbital Gyrus | 25 | 391 | 5 | 58.9 | -12.6 | 19.07 | 0.97 |
| L Lingual Gyrus | 25 | 391 | -25.2 | -54.6 | -5.4 | 17.45 | 0.68 |
| R Superior Temporal Gyrus | 25 | 391 | 63.1 | -33.3 | 21 | 18.02 | 0.48 |
| R Inferior Frontal Gyrus | 24 | 375 | 47.9 | 17.6 | 3.6 | 19.06 | 0.89 |
| L Calcarine Gyrus | 24 | 375 | -6.4 | -62 | 20.4 | 19.25 | 0.94 |
| R Thalamus | 23 | 359 | 24 | -29.5 | 4.5 | 18.42 | 0.81 |
| L Superior Medial Gyrus | 23 | 359 | -13.6 | 14 | 38.8 | 22.33 | 1.72 |
| R Middle Cingulate Cortex | 22 | 344 | 8.1 | -43.5 | 34 | 17.81 | 0.69 |
| L Mid Orbital Gyrus | 21 | 328 | -9.7 | 49.8 | -5.7 | 17.12 | 0.68 |

*Note*. Cluster-corrected voxel-wise linear mixed-effects model results are presented here summarizing regions showing a main effect of group and those showing a group-by-time interaction. k=number of voxels in cluster, mm3=cluster volume, CM=center of mass of cluster, SEM=standard error of the mean, LR=left-right (x), PA=posterior-anterior (y), IS=inferior-superior (z), anatomical locations: Eickhoff-Zilles macro labels from N27 (MNI_ANAT space).

**4. Pre- to post-treatment change analysis using the threat contrast**

We examined change over the course of CBT in terms of change in activation to the threat contrast (incongruent – congruent trials) only, instead of including condition as a multi-level factor. A whole-brain voxel-wise linear mixed effects model (3dLMEr in AFNI; 16) included group (ANX, HC) as a two-level between-subjects factor and one within-subjects factor: time with two levels (first and second scan, pre- and post-CBT for ANX youth). Instead of an additional within-subjects factor coding task condition, the contrast values for incongruent-congruent were entered in the analysis.

Associations between pre-treatment activation patterns and symptom improvement in ANX youth undergoing treatment were explored using whole-brain voxel-wise multivariate models (3dMVM in AFNI; 19). Post-treatment PARS total scores were entered as a continuous variable, controlling for baseline anxiety using pretreatment PARS total score as a covariate. Instead of an additional within-subjects factor coding task condition, the contrast values for incongruent-congruent were entered in the analysis.

Thresholding was consistent with the main analyses ($p<.001$, $k=21$).

**Results**

No clusters survived adequate thresholding for either analysis.

## 5. Pre- to post-treatment change analysis controlling for sex assigned at birth

The first analysis examined change over the course of CBT controlling for sex assigned at birth, as well as age and cohort/head coil. A whole-brain voxel-wise linear mixed effects model (3dLMEr in AFNI; 16) included group (ANX, HC) as a two-level between-subjects factor and one within-subjects factor: time with two levels (first and second scan, pre- and post-CBT for ANX youth) and participant as a random factor. We focus on the same group differences as the main analysis, particularly the group-by-time interaction.

Associations between pre-treatment activation patterns and symptom improvement in ANX youth undergoing treatment were explored using whole-brain voxel-wise multivariate models (3dMVM in AFNI; 19). Post-treatment PARS total scores were entered as a continuous variable, controlling for baseline anxiety using pretreatment PARS total score as a covariate. Age, cohort/head coil and sex were included as covariates. Thresholding was consistent with the main analyses ($p$<.001, $k$=21).

**Results**

Controlling for sex assigned at birth left results virtually unchanged. See table S7 and S8 for a summary of results.

No clusters survived adequate thresholding for the second analysis examining associations between pre-treatment activation patterns and symptom improvement.

**TABLE S7.** Stable group differences (ANX vs. HC) across time

| Region | k | mm3 | CM LR | CM PA | CM IS | Mean | SEM |
|---|---|---|---|---|---|---|---|
| R Precentral Gyrus | 87 | 1359 | 40.7 | -21.6 | 55.7 | 17.35 | 0.29 |
| L Precentral Gyrus, L Postcentral Gyrus | 80 | 1250 | -35.3 | -28.3 | 63.1 | 16.30 | 0.23 |
| R Hippocampus | 68 | 1063 | 22.8 | -10.7 | -17.8 | 16.95 | 0.30 |
| L Postcentral Gyrus | 35 | 547 | -44.9 | -19.7 | 49 | 16.69 | 0.41 |
| R Superior Temporal Gyrus | 34 | 531 | 56 | -9.1 | -10.9 | 17.71 | 0.44 |
| R Inferior Frontal Gyrus Triangularis | 34 | 531 | 46 | 19.3 | 26.4 | 16.27 | 0.35 |
| L Inferior Frontal Gyrus | 22 | 344 | -26.5 | 28.7 | -13.9 | 16.01 | 0.42 |

*Note*. Cluster-corrected voxel-wise linear mixed-effects model results are presented here summarizing regions showing a main effect of group and those showing a group-by-time interaction. k=number of voxels in cluster, mm3=cluster volume, CM=center of mass of cluster, SEM=standard error of the mean, LR=left-right (x), PA=posterior-anterior (y), IS=inferior-superior (z), anatomical locations: Eickhoff-Zilles macro labels from N27 (MNI_ANAT space).

**TABLE S8.** Changing group differences (ANX vs. HC) across time

| Region | k | mm3 | CM LR | CM PA | CM IS | Mean | SEM |
|---|---|---|---|---|---|---|---|
| R Lingual Gyrus, L Cuneus, R Calcarine Gyrus | 535 | 8359 | 3.6 | -77.8 | 10.2 | 21.42 | 0.30 |
| R Middle Frontal Gyrus, R Superior Frontal Gyrus | 434 | 6781 | 33.3 | 2.2 | 53.5 | 22.07 | 0.40 |
| R Superior Parietal Lobule | 264 | 4125 | 19.2 | -69.2 | 53.4 | 19.85 | 0.31 |
| R Superior Temporal Gyrus | 256 | 4000 | 55.7 | -41.8 | 20.8 | 22.61 | 0.59 |
| R Inferior Temporal Gyrus | 215 | 3359 | 43.7 | -64.4 | -7.9 | 19.08 | 0.31 |
| L Precentral Gyrus | 189 | 2953 | -31.5 | -1.4 | 53.9 | 18.96 | 0.35 |
| L Inferior Parietal Lobule | 180 | 2813 | -21.7 | -55.2 | 47.7 | 21.35 | 0.53 |
| L Middle Occipital Gyrus | 171 | 2672 | -46.4 | -70.3 | 2.2 | 19.87 | 0.33 |
| L SMA | 135 | 2109 | 1.3 | 10.6 | 46 | 18.93 | 0.40 |
| R Putamen | 123 | 1922 | 31.1 | 8 | -0.7 | 19.37 | 0.38 |
| L Middle Occipital Gyrus | 98 | 1531 | -28 | -72.2 | 29.1 | 20.71 | 0.62 |
| R Caudate Nucleus | 89 | 1391 | 15 | -5.1 | 17.5 | 20.84 | 0.55 |
| R Middle Frontal Gyrus | 61 | 953 | 29.6 | 44.3 | 34.3 | 19.40 | 0.57 |
| L Middle Frontal Gyrus | 57 | 891 | -35.8 | 48.8 | 2 | 19.29 | 0.65 |
| L Middle Cingulate Cortex | 54 | 844 | -10.1 | 12.7 | 36.1 | 18.80 | 0.63 |
| R Middle Occipital Gyrus | 53 | 828 | 34.1 | -68.7 | 36.5 | 18.75 | 0.49 |
| R Precuneus | 53 | 828 | 6.8 | -44.1 | 53.3 | 19.33 | 0.75 |
| R Inferior Parietal Lobule | 52 | 813 | 31.7 | -54.4 | 47.6 | 16.75 | 0.35 |
| R Rolandic Operculum | 49 | 766 | 46.4 | -12.6 | 12.3 | 18.41 | 0.62 |
| L Superior Parietal Lobule | 43 | 672 | -21 | -73 | 45.8 | 20.46 | 0.80 |
| L Precuneus/L Superior Parietal Lobule | 42 | 656 | -14.4 | -71.4 | 59.1 | 21.68 | 1.15 |
| R Thalamus | 40 | 625 | 6.6 | -26.6 | 16.2 | 16.97 | 0.47 |
| R Middle Occipital Gyrus | 40 | 625 | 36.7 | -77.9 | 18.7 | 19.24 | 0.70 |
| R Superior Frontal Gyrus | 38 | 594 | 28.4 | 56.7 | 15.1 | 19.96 | 0.77 |
| L Superior Temporal Grus | 37 | 578 | -59.2 | -45.7 | 19.4 | 19.55 | 0.59 |
| L Cerebellum (VII) | 36 | 563 | -1 | -74.9 | -39.8 | 19.81 | 0.79 |
| L Inferior Occipital Gyrus | 34 | 531 | -24 | -95.9 | -6.1 | 21.00 | 1.01 |
| Cerebellar Vermis (4/5) | 31 | 484 | 1.8 | -61.1 | -14.9 | 20.15 | 1.17 |
| R SMA | 30 | 469 | 11 | 0.4 | 72.4 | 21.00 | 1.11 |
| R Superior Temporal Gyrus | 29 | 453 | 45.1 | -31 | 13.7 | 23.76 | 1.44 |
| R Middle Cingulate Cortex | 29 | 453 | 6.6 | 24.1 | 30.8 | 16.46 | 0.42 |
| R Postcentral Gyrus | 29 | 453 | 46.4 | -34.9 | 59.8 | 19.21 | 0.77 |
| R Hippocampus | 28 | 438 | 24.4 | -32.9 | 6.2 | 20.11 | 0.99 |
| R Fusiform Gyrus | 27 | 422 | 43.3 | -39.3 | -18.3 | 20.79 | 0.92 |
| L Superior Temporal Gyrus | 26 | 406 | -41.6 | -37.4 | 19.3 | 19.09 | 0.83 |
| L Lingual Gyrus | 22 | 344 | -8.5 | -67.2 | -3 | 22.03 | 1.64 |
| Cerebellar Veris (8) | 21 | 328 | -0.9 | -57.8 | -29.2 | 19.98 | 1.29 |

*Note*. Cluster-corrected voxel-wise linear mixed-effects model results are presented here summarizing regions showing a main effect of group and those showing a group-by-time interaction. k=number of voxels in cluster, mm3=cluster volume, CM=center of mass of cluster, SEM=standard error of the mean, LR=left-right (x), PA=posterior-anterior (y), IS=inferior-superior (z), anatomical locations: Eickhoff-Zilles macro labels from N27 (MNI_ANAT space).

## 6. Variability in reaction time and associations between reaction time and brain activation

Attention bias variability (ABV) measures trial-wise attentional fluctuations between vigilance and avoidance within a session. Here, we examine attentional variability via ABV and precisely quantify trial-wise relationships between brain activity and RT using amplitude modulation.

### Behavioral data

ABV was calculated as reported in previous studies (20). We examined effects of group and time using a linear mixed-effects model with group (ANX, HC) as a two-level between- subjects factor and time as a within-subjects factor with two levels (first and second scan, pre- and post-CBT for ANX youth).

### Imaging data

RTs were mean centered by task condition. A lower mean standardized RT (MSRT) indicate a faster RT relative to one's condition-wise mean RT, and higher MSRT indicate slower relative RT. We applied amplitude modulation by trial-wise MSRT, which allows us to examine trial-by-trial associations between fluctuations in BOLD response and fluctuations in RT. Regression coefficients for the amplitude modulation indicate the strength of the RT-BOLD association, with differences indicating increases/decreases in percentage of signal change per second increase/decrease in participants' mean RT.

The first analysis examined change over the course of CBT controlling for age and cohort/head coil. A whole-brain voxel-wise linear mixed effects model (3dLMEr in AFNI; 16) included group (ANX, HC) as a two-level between-subjects factor and two within-subjects factors: task

condition with three levels (congruent, incongruent, neutral) and time with two levels (first and second scan, pre- and post-CBT for ANX youth) and participant as a random factor. We focus on the same group differences as the main analysis.

Associations between pre-treatment RT-BOLD patterns and symptom improvement in ANX youth undergoing treatment were explored using whole-brain voxel-wise multivariate models (3dMVM in AFNI; 19). Post-treatment PARS total scores were entered as a continuous variable, controlling for baseline anxiety using pretreatment PARS total score as a covariate. Age, cohort/head coil and sex were included as covariates. Thresholding was consistent with the main analyses ($p<.001$, $k=21$).

**Results Behavioral effects**

A main effect of group was observed for ABV ($F(1,644)=43.75$, $p<.001$). Treatment- seeking youth showed increased ABV compared to HC youth across both time points. Hence, hypervigilance-avoidance patterns as measured by ABV characterize ANX youth pre-treatment, which do not change with treatment.

**Whole brain analyses**

No clusters survived adequate thresholding for the group-by-time-by-condition or the group-by-time interaction. Several clusters showed a main effect of group, with significantly increased RT-BOLD coefficients for ANX youth stable across the two time points (see Table S9).

No clusters survived adequate thresholding for the second analysis examining associations between pre-treatment RT-BOLD patterns and symptom improvement.

**TABLE S9.** Stable group differences (ANX vs. HC) in RT-BOLD associations across time

| Region | k | mm3 | CM LR | CM PA | CM IS | Mean | SEM |
|---|---|---|---|---|---|---|---|
| L Inferior Parietal Lobule | 39 | 609 | -35.8 | -53.8 | 40.3 | 16.339 | 0.3209 |
| R Middle Frontal Gyrus | 25 | 391 | 34.1 | 33.3 | 31.8 | 17.817 | 0.6231 |
| L SMA | 24 | 375 | 0.2 | 7.4 | 52.7 | 16.32 | 0.3952 |

*Note*. Cluster-corrected voxel-wise linear mixed-effects model results are presented here summarizing regions showing a main effect of group and those showing a group-by-time interaction. k=number of voxels in cluster, mm3=cluster volume, CM=center of mass of cluster, SEM=standard error of the mean, LR=left-right (x), PA=posterior-anterior (y), IS=inferior-superior (z), anatomical locations: Eickhoff-Zilles macro labels from N27 (MNI_ANAT space).

## 7. Associations between activation patterns pre-treatment and treatment response using the CGI-I

Associations between pre-treatment activation patterns and symptom improvement in ANX youth undergoing treatment were explored using whole-brain voxel-wise multivariate models (3dMVM in AFNI; 19). Post-treatment CGI-I total scores were entered as a continuous variable, controlling for baseline anxiety using pretreatment CGI-I total score as a covariate. Age, cohort/head coil and sex were included as covariates. Thresholding was consistent with the main analyses ($p$<.001, $k$=21).

**Results**

No clusters survived adequate thresholding for analyses examining associations between pre-treatment activation patterns and symptom improvement using the CGI-I.

**8. Stability and change in amygdala responses over time in all groups (HC vs. ANX vs. AR)**

Regions of interest were each of the amygdalae extracted from the Harvard-Oxford subcortical structural atlas at 50% population probability intersected with a mask where at least 90% of individuals had data across the two time points. For each participant, we extracted the mean percent BOLD signal change of all voxels in this volume. We were interested in examining similarities in amygdala response across time between ANX and non-ANX youth (i.e., test whether anxiety-prone AR youth also show elevated stable amygdala activation across task conditions).We divided AR youth into AR-high and AR-low anxious groups based on a median split on the SCARED total score. We extracted average amygdala activity across all conditions. A linear mixed effects model with time (first scan, second scan), group (ANX, HV, AR-high and AR-low) and amygdala laterality (right, left) as fixed factors, and participant as a random factor.

**Results**

A significant main effect of group emerged ($F(3, 541)=6.18$, $p<.001$, see Table S10 and S11 for a summary of statistical results). Post-hoc Tukey's pairwise comparisons indicated significant differences in BOLD activation between ANX and HC youth ($t(216)=2.96$, $p=.02$) and ANX and AR-low youth ($t(216)=-4.05$, $p<.001$). AR-high youth were not significantly different from either HCs ($t(216)=0.71$, $p=.89$) or ANX youth ($t(216)=-2.06$, $p=.17$).

**TABLE S10.** Mixed effects model for amygdala BOLD response by group, time and laterality

| | numDF | denDF | *F*-value | *p*-value |
|---|---|---|---|---|
| (Intercept) | 1 | 541 | 11.212375 | 0.0009 |
| Time | 1 | 541 | 3.046507 | 0.0815 |
| Group | 3 | 541 | 6.180921 | 0.0004 |
| Amygdala Laterality | 1 | 541 | 2.426653 | 0.1199 |
| Time:Group | 3 | 541 | 1.421854 | 0.2354 |
| Time:Amygdala Laterality | 1 | 541 | 0.091764 | 0.7621 |
| Group:Amy | 3 | 541 | 0.354059 | 0.7862 |
| Time:Group:Amygdala Laterality | 3 | 541 | 0.982164 | 0.4008 |

**TABLE S11.** Post-hoc comparisons between groups

| Comparison | Estimate | SE | df | *t*-value | *p*-value |
|---|---|---|---|---|---|
| AR-high - AR-low | 0.01701 | 0.00774 | 542 | 2.197 | 0.1255 |
| AR-high - ANX | -0.01654 | 0.00804 | 216 | -2.057 | 0.1706 |
| AR-high - HC | 0.00585 | 0.00823 | 216 | 0.711 | 0.8926 |
| AR-low - ANX | -0.03355 | 0.00828 | 216 | -4.052 | 0.0004 |
| AR-low - HC | -0.01116 | 0.00846 | 216 | -1.318 | 0.5524 |
| Anx - HC | 0.02239 | 0.00758 | 216 | 2.955 | 0.0181 |

## 9. Comparing changes with treatment to developmental change in BOLD-anxiety relationships

A conjunction map was created to illustrate the overlap in brain regions that showed change with anxiety across the two developmental time points (SCARED-by-time interaction, cluster-corrected), and those regions that changed with treatment in the treatment-seeking group (Group-by-time, cluster-corrected; see Figure S5).

### Results

The conjunction analysis between the group-by-time interaction in the treatment seeking/HC sample (changing activation patterns with treatment) and the group-by-time interaction in the AR sample (anxiety-associated changes across early adolescence) illustrates minimal overlap.

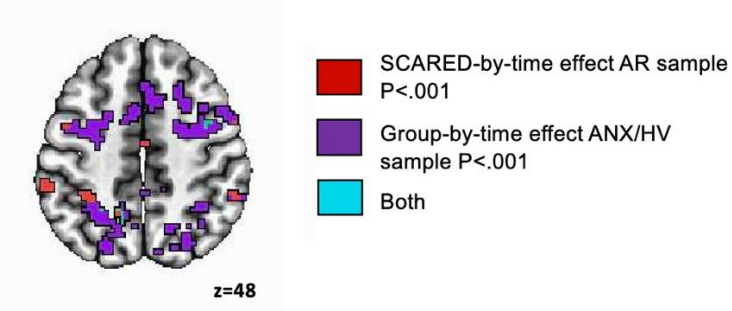


**FIGURE S5.** Conjunction map illustrating the overlap between brain regions showing a change in activation patterns with CBT vs. no treatment (group-by-time interaction, cluster-corrected) and those regions that changed with anxiety in the AR sample (SCARED-by-time interaction, cluster-corrected)

## References


1. Wechsler D. Manual for the Wechsler abbreviated intelligence scale (WASI). San Antonio, TX: The Psychological Corporation. 1999.

2. Marshall WA, Tanner JM. Variations in pattern of pubertal changes in girls. Archives of disease in childhood. 1969;44(235):291.

3. Marshall WA, Tanner JM. Variations in the pattern of pubertal changes in boys. Archives of disease in childhood. 1970;45(239):13-23.

4. Tanner JM. Sequence, tempo, and individual variation in the growth and development of boys and girls aged twelve to sixteen. Daedalus. 1971:907-30.

5. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, 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(7):980-8.

6. White LK, Sequeira S, Britton JC, Brotman MA, Gold AL, Berman E, et al. Complementary Features of Attention Bias Modification Therapy and Cognitive-Behavioral Therapy in Pediatric Anxiety Disorders. Am J Psychiatry. 2017;174(8):775-84.

7. Linke JO, Jones E, Pagliaccio D, Swetlitz C, Lewis KM, Silverman WK, et al. Efficacy and mechanisms underlying a gamified attention bias modification training in anxious youth: protocol for a randomized controlled trial. BMC Psychiatry. 2019;19(1):246.

8. Kircanski K, White LK, Tseng WL, Wiggins JL, Frank HR, Sequeira S, et al. A Latent Variable Approach to Differentiating Neural Mechanisms of Irritability and Anxiety in Youth. JAMA Psychiatry. 2018;75(6):631-9.

9. Naim R, Haller SP, Linke JO, Jaffe A, Stoddard J, Jones M, et al. Context-dependent amygdala–prefrontal connectivity during the dot-probe task varies by irritability and attention bias to angry faces. Neuropsychopharmacology. 2022;47(13):2283-91.

10. White LK, Britton JC, Sequeira S, Ronkin EG, Chen G, Bar-Haim Y, et al. Behavioral and neural stability of attention bias to threat in healthy adolescents. Neuroimage. 2016;136:84- 93.

11. Abend R, Swetlitz C, White LK, Shechner T, Bar-Haim Y, Filippi C, et al. Levels of early-childhood behavioral inhibition predict distinct neurodevelopmental pathways to pediatric anxiety. Psychol Med. 2020;50(1):96-106.

12. Fox NA, Henderson HA, Rubin KH, Calkins SD, Schmidt LA. Continuity and discontinuity of behavioral inhibition and exuberance: Psychophysiological and behavioral influences across the first four years of life. Child development. 2001;72(1):1-21.

13. White LK, Degnan KA, Henderson HA, Pérez-Edgar K, Walker OL, Shechner T, et al. Developmental relations among behavioral inhibition, anxiety, and attention biases to threat and positive information. Child development. 2017;88(1):141-55.

14. Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, Kaufman J, et al. The Screen for Child Anxiety Related Emotional Disorders (SCARED): scale construction and psychometric characteristics. J Am Acad Child Adolesc Psychiatry. 1997;36(4):545-53.

15. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance

neuroimages. Comput Biomed Res. 1996;29(3):162-73.

16. Chen G, Saad ZS, Britton JC, Pine DS, Cox RW. Linear mixed-effects modeling approach to FMRI group analysis. Neuroimage. 2013;73:176-90.

17. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86(2):420-8.

18. Chen G, Taylor PA, Haller SP, Kircanski K, Stoddard J, Pine DS, et al. Intraclass correlation: Improved modeling approaches and applications for neuroimaging. Hum Brain Mapp. 2018;39(3):1187-206.

19. Chen G, Adleman NE, Saad ZS, Leibenluft E, Cox RW. Applications of multivariate modeling to neuroimaging group analysis: a comprehensive alternative to univariate general linear model. Neuroimage. 2014;99:571-88.

20. Naim R, Abend R, Wald I, Eldar S, Levi O, Fruchter E, et al. Threat-related attention bias variability and posttraumatic stress. American Journal of Psychiatry. 2015;172(12):1242-50.