Supplementary Online Content

Methods 1. Participant Assessment

All children were assessed using the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version (K-SADS-PL) (1). The K-SADS-PL, including an additional module for assessing disruptive mood dysregulation disorder (DMDD; available on request), was administered separately to children and parents by masters- or doctoral-level clinicians with good inter-rater reliability (kappa ≥ 0.7 for all diagnoses). Clinicians were trained to levels of acceptable reliability by rating video recordings and conducting supervised assessments. Children in each clinical group met criteria for primary diagnoses of DMDD, attention-deficit/hyperactivity disorder (ADHD), and anxiety disorders (ANX). Children with ADHD and ANX did not meet criteria for DMDD, but youth with DMDD may have co-occurring ADHD and ANX. Diagnoses were confirmed in consensus conferences chaired by a senior psychiatrist or clinical psychologist (coauthors EL, KT, MB, or DP).

Patients with ANX met criteria for generalized anxiety disorder, separation anxiety disorder, and/or social phobia. The ANX sample was drawn from a treatment study and were scanned at various points during treatment (some after treatment). Other inclusion criteria for ANX patients were: clinically significant anxiety on the Pediatric Anxiety Rating Scale (2) (score ≥ 10), impairment on the Children's Global Assessment Scale (3) (score < 60), and desire for weekly treatment. Additional exclusion criteria for ANX patients were current Tourette's syndrome, obsessive-compulsive disorder, major depressive disorder, post-traumatic stress disorder, current psychotropic medication use, or suicidal ideation.

Measures of irritability, anxiety, and ADHD symptoms were collected within 90 days of scan acquisition in >90% of participants (>70% were collected within one week of scanning). These measures have been shown to be highly stable across time periods longer than three months (4–6). Irritability was measured using the Affective Reactivity Index (ARI), a parent- and self-rated measure (4). This is a six-item scale assessing the frequency, duration, and threshold of irritability in the past 6 months (4). Each item was rated on a 3-point Likert scale from 0 (not true) to 2 (certainly true). Sample items include "get angry frequently," "stay angry for a long time," and "lose temper easily." Past research demonstrates good internal consistency (α's ≥ .80), test-retest reliability, and validity (i.e., construct validity and discriminant validity), as well as a single-factor structure in clinical and community samples (4, 7–9). Total scores for child- and parent-reported were averaged, resulting in a possible score of 0-12, treated as a continuous measure. Four participants had missing data on ARI (1 healthy volunteer [HV], 1 ADHD, 2 ANX). For these subjects, mean ARI from their respective diagnostic group was used.

Anxiety was measured by the Screen for Child Anxiety Related Emotional Disorders (SCARED), a validated, 41-item parent- and self-rated measure (10). Total scores of parent and child report averaged together ranged from 0.5-52. Five participants had missing data on SCARED (1 DMDD, 4 ADHD). For these subjects, mean SCARED from their respective diagnostic group was used.

ADHD symptoms were measured by the ADHD-Index subscale of the Conners' Parent Rating Scale-Revised: Long Form (CPRS-R: L) (11). It is a widely used rating scale for ADHD symptoms and has good reliability and validity (11). Age- and gender-referenced T-score was used in the analysis. The range of the ADHD-Index T-score in this sample was 40-90. Four participants had

missing data (3 HV, 1 DMDD). For these subjects, mean T-score from their respective diagnostic group was used.

Methods 2. Affective Posner 2 Paradigm

This task was adapted from the Affective Posner task used in previous studies (12–16). Prior to completing the task, participants underwent a 50-trial training during which they received accurate feedback on their performance but did not win or lose money. The task itself consisted of two non-frustration runs, a pilot frustration run, and two frustration runs (eFigure 1) and lasted for about 38 minutes. During the two non-frustration runs (eFigure 1), participants received accurate feedback on their performance, earning $0.50 for every correct response and losing $0.50 for every incorrect response. During the pilot frustration run and the two frustration runs, participants were instructed that they needed to respond both correctly and quickly in order to win money. The pilot frustration run of the task (eFigure 1) consisted of 32 trials and was used to acclimate participants to the "real" frustration runs; participants received rigged feedback on 10% of correct trials. Structural scans were acquired during this pilot frustration run. Behavioral data from the pilot frustration run were not included in the behavioral analyses. On trials with rigged feedback, participants were informed that they were "too slow," and lost money regardless of their actual reaction time. In the two "real" frustration runs (Figure 1 and eFigure 1), participants received rigged feedback on 60% of correct trials. After each run of the task, participants rated their feelings of unhappiness and frustration using 9-point Likert scales (i.e., 1 = "happy" or "not at all frustrated"; 9 = "sad" or "extremely frustrated"). Self-report questionnaires and follow-up interviews were used to assess participant deception following the scan. Participants who indicated that they experimented with button press to figure out if the task was rigged were excluded. No participant reported marked distress associated with the frustration manipulation or deception.

Methods 3. Image Acquisition and Preprocessing

For 61 participants, the echo-planar images (EPI) were acquired using a flip angle of 90°; for the rest of the sample (n=134), a flip angle of 75° was used. This switch from 90° to 75° flip angle was prompted by quality-control monitoring of the imaging data which determined that a 75° flip angle provided better image quality and contrast. Post-hoc analyses with Bonferroni correction showed no differences between the two flip angles in the extracted % blood-oxygen-level dependent (BOLD) signal change or functional connectivity measures.

T1-weighted magnetization-prepared 180 degrees radio-frequency pulses and rapid gradient-echo (MPRAGE) images were uniformity corrected and skull stripped with FreeSurfer (17). These processed anatomic images were then used to create alignment matrices to normalize the EPI. Normalization processes included EPI to EPI base, to each subject's individually processed anatomic image, affine registration to AFNI's TT_N27 Talairach-space template, and then non-linear registration to Talairach-space template (in a combined transformation), sampled on a 2.5-mm isotropic grid. EPI images were processed by excising the first four volumes, limiting each voxel's BOLD signal to four standard deviations from the

mean trend of its time series, correcting for slice timing, normalizing by the simultaneous application of four alignment matrices (as described above for normalization processes), smoothing using a 6-mm full width at half maximum Gaussian kernel, and intensity-scaling to a mean of 100.

Methods 4. Estimating Neural Activation and Functional Connectivity

Neural Activation

Processed, scaled EPI were entered into a general linear model (GLM) with the following parameters: a quadratic detrending polynomial, a regressor for each of six translational and rotational motion parameters, and 16 two-second block convolved gamma variate regressors for modeling task conditions. These 16 task regressors included 9 for the N+1 portion and 7 for the feedback portion. For the N+1 portion, the 9 regressors were validity [valid vs. invalid trials] x accuracy [correct vs. error] plus missed trials; valid correct trials were further partitioned into 5 types of conditions based on preceding feedback: trials following positive feedback, trials following rigged feedback, trials following all other feedback, the 1st trial of each run with no preceding feedback, and trials left un-partitioned [if any]). For the feedback portion, the 7 regressors were validity [valid vs. invalid] x feedback [positive, rigged, or error] plus missed trials). EPI volumes were censored from this GLM regression by three criteria: 1) motion shift, defined as movement exceeding a Euclidean distance of 1 mm from the preceding volume, 2) volumes immediately preceding such a motion shift, and 3) volumes with a large fraction (>10%) of outlying time points. Main analyses on the N+1 portion of the task focused on two regressors of interest: *trials following positive feedback* and *trials following rigged feedback*, whereas main analyses on the feedback portion of the task focused on *positive feedback* and *rigged feedback*. Parameter estimates for each task condition represent % BOLD signal change and mean neural activation to the task condition.

Functional Connectivity

We also conducted analyses to assess context-dependent functional connectivity between bilateral inferior frontal gyrus [IFG] and amygdala as the seed regions and the rest of the brain. The IFG was selected because of its central role in mediating responses on our fMRI attentional task (18) and in attention control (19–21); the amygdala was chosen given prior findings on a similar frustration task (12) and its role in processing emotionally-salient stimuli (22). IFG masks were created using a 10-mm sphere, centered at +50 +27 +6 (right IFG) and -50 +27 +6 (left IFG) in Talairach space (19,23). Amygdala masks were defined by the Talairach-Tournoux Daemon. We conducted generalized psychophysiological interactions (gPPI) (24) analyses with the following steps.

1. Processed, scaled EPI were prepared to estimate “task condition x seed” (IFG and amygdala, separately) timeseries regressors orthogonal to the main effects of task and nuisance variables. To do this, processed, scaled EPI were entered into the same GLM used to estimate neural activation (described above), except that censoring was not applied to the data. Censoring here would introduce temporal discontinuities that would interfere with partitioning IFG and amygdala BOLD response into task conditions.

Residuals from the GLM were then used to generate "task condition x seed" timeseries interaction regressors.

2. gPPI regressors for each task condition by each seed timeseries were generated. First, the mean timeseries from the residuals of the GLM in step 1 were calculated across all voxels in each seed mask. To finely partition the timeseries' variance to each task condition, the timeseries were up-sampled and then deconvolved with the hemodynamic response function (HRF; a 2-second block gamma variate) used in the GLM from steps 1 and 3. This generates a timeseries representing the strength of the seed's neural response to all stimuli. These were partitioned to each task condition by multiplying each deconvolved timeseries with a binary timeseries representing the presence (1) or absence (0) of each task condition stimulus. Next, these partitioned timeseries were reconvolved with the HRF to generate seed BOLD response for each condition, one for each seed timeseries by each task condition. There were 4 total, two for the N+1 trial (after rigged and after positive feedback) and two for the feedback portion (rigged and positive feedback). The regressors were then down-sampled to the scan acquisition TR timing to prepare for PPI regression.
3. For each seed, a generalized PPI regression was done. On the residual data from step 1, the same GLM used to estimate neural activation, including censoring with the same limits, was run with the 4 PPI regressors as well as the seed's mean timeseries as a main effect. The voxelwise parameter estimates for the seed timeseries and the 4 PPI regressors are interpreted as functional connectivity. Estimates of the 4 PPI regressors were used in the group-level analysis of functional connectivity (see eResults 6 & 7).

Results 1. Behavioral Results

Frustration and unhappiness ratings

We conducted two separate ARI x Age x Run (non-frustration vs. frustration) repeated-measures analysis of covariance (ANCOVA) to examine self-reported frustration and unhappiness ratings during the task. A main effect of Run emerged i.e., participants reported feeling more frustrated ($F_{1,189}$=295.70, p<.001, η_p^2=.61) and unhappy ($F_{1,189}$=135.44, p<.001, η_p^2=.42) during frustration vs. non-frustration runs. Also, main effects of ARI and Age were significant for both frustration and unhappiness ratings across non-frustration and frustration runs ($F_{1,189}$=4.52-8.83, p's<.05, η_p^2=.02-.05). Specifically, higher irritability was related to more frustration and unhappiness across the task, as was older age (r's=.15-.21). ARI × Age interaction was not significant. Neither frustration nor unhappiness were associated with the brain findings reported below.

Accuracy & Reaction Time: Non-frustration + frustration runs

An ARI × Age × Validity × Run ANCOVA tested whether accuracy and response time (RT) varied with irritability, age, trial type (valid vs. invalid trials), and run type (non-frustration vs. frustration). Results indicated that irritability was not related to accuracy or RT. Instead, Age × Validity predicted accuracy ($F_{1,189}$=4.26, p=.04, η_p^2=.02) and response time (RT) ($F_{1,187}$=27.46, p<.001, η_p^2=.13). That is, older age was related to higher accuracy for invalid vs. valid trials (i.e., response cost; r=.15, p=.04), and older age was related to faster RT for invalid vs. valid trials (r=-

.36, p<.001). Validity × Run also predicted accuracy ($F_{1,189}$=457.33, p<.001, η_p^2=.71) and RT ($F_{1,187}$=87.26, p<.001, η_p^2=.32). That is, accuracy was worse and RT was greater for invalid trials than for valid trials, particularly during frustration runs. Thus, the "validity effect" (i.e., response cost) was greater in the frustration relative to non-frustration runs.

Accuracy & Reaction Time: N+1 trial (frustration runs only)

An ARI × Age × Validity × Feedback ANCOVA evaluated whether accuracy and RT varied with irritability, age, trial type (valid vs. invalid trials), and preceding feedback (rigged vs. positive). This analysis used data from the frustration runs because only they contained both rigged and positive feedback. Results indicated that Validity × Feedback ($F_{1,189}$=4.77, p=.03, η_p^2=.03) predicted accuracy. Specifically, accuracy for valid trials was higher following rigged than following positive feedback (p=.03) but did not differ for invalid trials (p=.11). Main effects of validity ($F_{1,189}$=50.24, p<.001, η_p^2=.21) and age ($F_{1,189}$=12.35, p<.001, η_p^2=.06) were significant for RT. That is, RT was greater for invalid trials than for valid trials, and older age was related to faster RT (r=-.25). Of note, accuracy and RT during frustration runs were not associated with neural activations reported in the main text (p's>.05 Bonferroni corrected).

Accuracy and reaction time during either non-frustration or frustration runs were not associated with ADHD symptoms (measured by CPRS), after controlling for irritability and age. However, more anxiety symptoms (measured by SCARED) was related to poorer accuracy during frustration runs, after controlling for irritability and age (r=-.16, p=.02).

Results 2. Frustration and Unhappiness Ratings on Whole-Brain Activation

To examine the effect of "state irritability" (i.e., self-rated frustration and unhappiness during frustration runs) on whole-brain activation and the moderating effect of age, we conducted two repeated-measures ANCOVAs of Frustration x Age x Condition (with SCARED, CPRS, and motion as covariates), one for the N+1 trial and the other for the feedback portion. Similarly, we conducted two parallel ANCOVAs of Unhappiness x Age x Condition. Results are described below.

N+1 trial. No clusters for the 3-way Frustration (or Unhappiness) x Age x Condition interaction, or the 2-way Frustration (or Unhappiness) x Condition interaction, survived the whole-brain correction threshold at α=.05 and k≥703 mm^3 (voxelwise p=.001).

Feedback. Similarly, no clusters for the 3-way Frustration (or Unhappiness) x Age x Condition interaction, or the 2-way Frustration x Condition interaction, survived the whole-brain correction threshold at α=.05 and k≥703 mm^3 (voxelwise p=.001). However, there was a significant 2-way interaction of Unhappiness x Condition in the right superior temporal gyrus (xyz = 54, -16, 6, k = 984 mm^3). That is, during processing of rigged vs. positive feedback, higher self-rated unhappiness was related to more activation in the right superior temporal gyrus (r=.29, p<.001).

Taken together, the relative lack of significant findings with these "state" measures of irritability suggests that the neural activation patterns mediating transient, subjective feelings of frustration and unhappiness were not similar to those mediating trait irritability (reported in the main text). However, the self-rated frustration and unhappiness analyses have some

significant limitations. First, the ratings were only collected, retrospectively, at the end of each run. This yields 2 ratings for frustration and 2 ratings for unhappiness from the frustration runs; we used the average of these 2 ratings in the analyses described above. Had the ratings been collected more frequently e.g., after each trial, we might have captured more associations between the state measure and neural responses (importantly, the latter were estimated by the average of trial-by-trial neural activation). Unfortunately, it was not feasible to obtain a trial-by-trial rating of frustration and unhappiness because doing so would significantly lengthen the task, disrupt the flow of the attentional task, and potentially alter participants' momentary affective state. Second, the measures of 'state' irritability (ratings of frustration and unhappiness) and 'trait' irritability (ARI) were not as highly correlated as one might expect (r=.11, p=.13 between frustration and irritability and r=.17, p=.02 between unhappiness and irritability). This could be attributed to the unreliability of children's self-report, under-reporting due to embarrassment, or a lack of insight into one's own feelings and emotions. Low correlations between the state and trait measures may explain the lack of consistent neural findings across measures.

Results 3. Analysis by Primary Categorical Diagnoses and Dimensional Measure of Irritability

Whole-brain activation

We conducted two repeated-measures ANCOVAs of Diagnosis x ARI x Condition (with motion as a covariate), one for the N+1 trial and the other for the feedback portion. No clusters for the 3-way Diagnosis x ARI x Condition interaction, or the 2-way Diagnosis x Condition interaction, survived the whole-brain correction threshold at α=.05 and $k \geq 703$ mm^3 (voxelwise p=.001). This suggests that the neural activation pattern mediating dimensionally-measured trait irritability (as reported in the main text) did not differ across the diagnostic groups sampled in this study.

Functional Connectivity

We conducted eight repeated-measures ANCOVAs of Diagnosis x ARI x Condition (with motion as a covariate), four for the N+1 trial and the other four for the feedback portion (one for each seed region: left and right IFG, and left and right amygdala). Results are described below.

N+1 trial. No clusters for the 3-way Diagnosis x ARI x Condition interaction, or the 2-way Diagnosis x Condition interaction, survived the whole-brain correction threshold at α=.0125 (Bonferroni corrected) and $k \geq 1094$ mm^3 (voxelwise p=.001).

Feedback. Using the right IFG as the seed, functional connectivity analysis revealed a significant Diagnosis x ARI x Condition interaction in the right dorsolateral prefrontal cortex (xyz = 41, 26, 19, k = 1891 mm^3). Specifically, during processing of rigged vs. positive feedback, higher irritability was related to *decreased* connectivity between the right IFG and right dorsolateral prefrontal cortex in HV (r=-.42, p<.001) and youth with DMDD (r=-.37, p<.01); higher irritability was related to *increased* connectivity in youth with ADHD (r=.42, p<.01; the association was not significant in youth with ANX, r=.16, p=.31).

In addition, using the right amygdala as the seed, functional connectivity analysis during processing of rigged vs. positive feedback revealed a significant Diagnosis x ARI x Condition interaction in the left superior/middle temporal gyrus (xyz = -61, -46, 9, k = 4172 mm^3), left culmen (xyz = -4, -59, 1, k = 2047 mm^3), and right precentral gyrus (xyz = 46, -4, 46, k = 1906 mm^3). However, the only significant post-hoc analysis on the extracted functional connectivity estimates showed that ARI was related to functional connectivity between the right amygdala and left culmen in the DMDD group (r=-.34, p=.02).

Results 4. ARI × Age × Gender x Condition Analyses in Whole-Brain Activation

To examine the effect of gender on whole-brain activation, we conducted two repeated-measures ANCOVAs of ARI x Age x Gender x Condition (SCARED, CPRS, and motion as covariates), one for the N+1 trial and the other for the feedback portion. Results are described below.

N+1 trial. At the whole-brain correction threshold of α=.05 and k≥703 mm^3 (voxelwise p=.001), there was a significant 4-way interaction of ARI x Age x Gender x Condition in the left and right inferior parietal lobule (xyz = -19, -61, 54, k = 5875 mm^3 and xyz = 29, -54, 56, k = 5203 mm^3), right precentral gyrus (xyz = 19, -26, 66, k = 3313 mm^3), left postcentral gyrus (xyz = -51, -29, 26, k = 1609 mm^3), and left insula (xyz = -36, -9, 16, k = 922 mm^3). This interaction was driven by younger boys showing a positive association between irritability and activation across these regions (rs=.41-.53, ps≤.02) and older boys showing a negative association between irritability and activation in the right precentral gyrus and left insula (rs=-.44-.47, ps≤.03).

Feedback. No clusters for the 4-way interaction of ARI x Age x Gender x Condition survived the whole-brain correction threshold at α=.05 and k≥703 mm^3 (voxelwise p=.001). However, there was a significant 3-way interaction of ARI x Gender x Condition in the right declive (xyz = 34, -56, -14, k = 1141 mm^3). For boys, higher irritability was related to *increased* activation in this region (r=.37, p<.001). However, for girls, higher irritability was related to *decreased* activation in this region (r=-.35, p=.001).

Results 5. Region of Interest (ROI) Analysis in Amygdala and Striatum

Analysis

Amygdala and striatum (caudate, putamen, and nucleus accumbens), defined by the Talairach-Tournoux Daemon, were chosen for the ROI analyses, given findings from a previous version of this frustration paradigm (10). Mean signal intensity was extracted from each ROI for each event of interest. We applied Bonferroni adjustment for these eight ROI tests (bilateral amygdala, caudate, putamen, and nucleus accumbens), resulting in α=.05/8=.00625.

N+1 trial. The extracted mean signal intensity was subjected to an ARI × Age × Condition (after rigged vs. after positive feedback) ANCOVA in SPSS, with SCARED, CPRS, and motion as covariates.

Feedback. The extracted mean signal intensity was subjected to an ARI × Age × Condition (rigged vs. positive feedback) ANCOVA in SPSS, with SCARED, CPRS, and motion as covariates.

Results

N+1 trial. An ARI × Age × Condition interaction was found in right putamen ($F_{1,188}$=8.19, p=.005, η_p^2=.04) and right nucleus accumbens ($F_{1,188}$=9.90, p=.002, η_p^2=.05). Specifically, following rigged vs. positive feedback, higher irritability was related to more activation in right putamen in young children (aged 8-11.5 years; r=.37, p=.003). However, irritability was not related to activation in nucleus accumbens in young children (aged 8-11.5 years; r=.17, p=.191). Instead, higher irritability was related to more nucleus accumbens activation in younger adolescents (aged 11.5-14 years; r=.26, p=.046) and less activation in older adolescents (aged 14-18 years; r=-.32, p=.012).

In addition, an ARI × Condition interaction was found in left and right caudate (Left: $F_{1,188}$=9.37, p=.003, η_p^2=.05; Right: $F_{1,188}$=12.67, p<.001, η_p^2=.06) and left putamen ($F_{1,188}$=12.70, p<.001, η_p^2=.06). Specifically, higher irritability was related to more activation in bilateral caudate and left putamen after receiving rigged vs. positive feedback (r's=.22-.25).

Feedback. Irritability, age, or their interaction was not associated with activation during processing of rigged vs. positive feedback.

We also conducted repeated-measures ANCOVAs of Diagnosis x Condition (motion as a covariate) to evaluate the effect of diagnosis on amygdala and striatal activation. There were no significant Diagnosis x Condition interactions. Together, these ROI findings suggest that irritability is not related to amygdala activation during processing of a frustrating event or during attention immediate after frustration.

Results 6. ARI × Age × Condition Analyses in Functional Connectivity

Analysis

Voxelwise p value threshold was set at .001. We applied Bonferroni adjustment for four gPPI tests (bilateral IFG and amygdala), resulting in α=.05/4=.0125 and cluster size ≥ 1094 mm^3.

N+1 trial. Four separate ARI × Age × Condition (after rigged vs. after positive feedback) ANCOVAs were conducted for the gPPI estimate of each voxel for each seed region (left and right IFG and amygdala).

Feedback. Similarly, four separate ARI × Age × Condition (rigged vs. positive feedback) ANCOVAs were conducted (left and right IFG and amygdala).

Results

N+1 trial. Using left IFG as the seed, the functional connectivity analysis revealed a significant ARI × Age × Condition interaction in the left putamen (eTable 5 & eFigure 5). Specifically, following rigged vs. positive feedback, higher irritability was related to decreased connectivity between left IFG and left putamen in younger adolescents (aged 11.5-14 years) at a trend level (eFigure 5). In addition, there was a significant ARI × Condition interaction in the functional connectivity between left IFG and periaqueductal gray (extending to culmen; eTable 5 & eFigure 6). That is, following rigged vs. positive feedback, higher irritability was related to decreased connectivity between left IFG and periaqueductal gray (eFigure 6). There were no

findings with the amygdala seeds.

Feedback. Irritability, age, or their interaction was not associated with functional connectivity during processing of rigged vs. positive feedback.

Results 7. ARI x SCARED x Condition Analyses in Functional Connectivity

Analysis

We also conducted analyses to examine the joint, interacting effects of irritability and anxiety on functional connectivity, given that a recent study demonstrates that irritability and anxiety interact to influence neural connectivity while processing social threat (25). Voxelwise *p* value threshold was set at .001. We applied Bonferroni adjustment for four gPPI tests (bilateral IFG and amygdala), resulting in α=.05/4=.0125 and cluster size ≥ 1094 mm^3.

N+1 trial. Four separate ARI × SCARED × Condition (after rigged vs. after positive feedback) ANCOVAs were conducted for the gPPI estimate of each voxel for each seed region (left and right IFG and amygdala).

Feedback. Similarly, four separate ARI × SCARED × Condition (rigged vs. positive feedback) ANCOVAs were conducted (left and right IFG and amygdala).

Results

N+1 trial. Using left and right IFG (separately) as the seed, the functional connectivity analysis revealed an ARI × SCARED × Condition interaction in insula, posterior cingulate cortex, superior and middle temporal gyrus, and precentral gyrus (eTable 6, eFigure 7, and eFigure 8). For those with high irritability and high anxiety, there was increased connectivity after receiving rigged vs. positive feedback. In contrast, for those with high irritability and low anxiety, there was decreased connectivity (eFigure 7). There were no findings with the amygdala seeds.

Feedback. Irritability, anxiety, or their interaction was not associated with functional connectivity during processing of rigged vs. positive feedback.

Results 8. Threshold-Free Cluster Enhancement Analyses

We conducted a supplementary whole-brain voxelwise analysis using Threshold-Free Cluster Enhancement (TFCE) (26). The general linear model at the 2^{nd} level comprised of Condition, interactions of Condition x ARI, Condition x Age, and Condition x ARI x Age, as well as subject-specific intercepts as nuisance variables (accounting for the main effects of Age and ARI as well as any other subject-specific factors such as motion, SCARED, CPRS that were constant for both conditions). Inference was performed using the tool Permutation Analysis of Linear Models (PALM) (27; available at https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/PALM). We used 5000 permutations, further enriched by the approximation of the tail of the empirical distribution using a generalized Pareto distribution (GPD), from which familywise error rate corrected p-values were obtained (28). Correction considered all 8 contrasts that were investigated, namely positive and negative effects of ARI x Age x Condition, ARI x Condition, Age x Condition, and Condition.

Significant results for the 3-way interaction of ARI x Age x Condition and 2-way interaction of ARI x Condition largely replicated the findings reported in the main text using AFNI's 3dMVM with Monte Carlo cluster-size simulation to correct for multiple testing (as opposed to a permutation test). Two small clusters for the ARI × Condition effect (left caudate and pre-central gyrus) from the original analysis in AFNI became non-significant (right caudate remains significant; see eFigure 9).

Table S1. Excluded Participants

Excluded reasons	HV (n=7)	DMDD (n=12)	ADHD (n=12)	ANX (n=4)	Total (n=35)
Not deceived[a]	0	1	1	2	4
Excessive motion[b]	1	6	7	1	15
Poor behavioral data[c]	2	1	1	0	4
Incomplete task due to frustration	1	1	1	1	4
Incomplete NOT due to frustration	1	1	0	0	2
Technical difficulties/errors[d]	2	2	0	0	4
Abnormal clinical scan	0	0	2	0	2

ADHD = attention-deficit/hyperactivity disorder; ANX = anxiety disorders; DMDD = disruptive mood dysregulation disorder; HV = healthy volunteer.

[a] Defined as participants who experimented with button press to try to figure out if the task was rigged.

[b] More than 25% of echo-planar imaging volumes were above the censor threshold and the average Euclidean distance of volume-to-volume shift was > 0.25 mm for all volumes that remain after censoring.

[c] Accuracy lower than 60% during either non-frustration or frustration runs and/or low numbers of positive feedback trials during frustration runs (< 20).

[d] Scanner issues or task administration errors.

Table S2. Correlations among Symptom Dimensions and Sample Characteristics

	ARI	SCARED	CPRS	Age	Gender[a]	IQ[b]	SES[c]	Motion[d]
ARI	–							
SCARED	.24**	–						
CPRS	.54***	.21**	–					
Age	-.08	-.04	-.12	–				
Gender	-.02	.10	-.04	.09	–			
IQ	-.07	.03	-.05	-.12	.03	–		
SES	.09	.10	.13	.02	.00	-.11	–	
Motion	.16*	-.04	.14	-.30***	-.21**	.04	-.05	–

ARI = Affective Reactivity Index; CPRS = Conners' Parent Rating Scale; SCARED = Screen for Child Anxiety Related Emotional Disorders; SES = Socioeconomic status.

*p<.05, **p<.01, ***p<.001

[a] Coded as 0 (male) and 1 (female). n (%) is for the male.

[b] Measured by the Wechsler Abbreviated Scale of Intelligence. Missing data for 1 participant.

[c] Measured by the Hollingshead 2-factor index. Missing data for 25 participants.

[d] Calculated as the mean Euclidean distance of framewise volume shift after censoring.

Table S3. Post-hoc Analyses of Medication Effects on ARI x Age x Condition[a] Activation Results

Regions	Analysis		
	F value	*p* value	η_p^2
R cuneus			
Full Sample (n=195)	25.25	<.001	.12
Exclude Stimulants (n=150)	26.66	<.001	.16
Exclude Anti-depressants (n=164)	19.80	<.001	.11
Exclude Anti-psychotics (n=185)	26.10	<.001	.13
R superior parietal lobule			
Full Sample (n=195)	28.51	<.001	.13
Exclude Stimulants (n=150)	31.72	<.001	.18
Exclude Anti-depressants (n=164)	24.37	<.001	.13
Exclude Anti-psychotics (n=185)	31.21	<.001	.15
L precuneus/cuneus			
Full Sample (n=195)	22.91	<.001	.11
Exclude Stimulants (n=150)	21.97	<.001	.13
Exclude Anti-depressants (n=164)	16.87	<.001	.10
Exclude Anti-psychotics (n=185)	22.38	<.001	.11
L medial frontal gyrus/ACC			
Full Sample (n=195)	23.94	<.001	.11
Exclude Stimulants (n=150)	27.50	<.001	.16
Exclude Anti-depressants (n=164)	24.44	<.001	.14
Exclude Anti-psychotics (n=185)	25.27	<.001	.12
R pre- and post-central gyrus			
Full Sample (n=195)	25.62	<.001	.12
Exclude Stimulants (n=150)	26.59	<.001	.16
Exclude Anti-depressants (n=164)	20.54	<.001	.12
Exclude Anti-psychotics (n=185)	26.36	<.001	.13
L precuneus			
Full Sample (n=195)	25.15	<.001	.12
Exclude Stimulants (n=150)	28.49	<.001	.17
Exclude Anti-depressants (n=164)	20.20	<.001	.11
Exclude Anti-psychotics (n=185)	27.18	<.001	.13
L middle frontal gyrus			
Full Sample (n=195)	20.80	<.001	.10
Exclude Stimulants (n=150)	26.91	<.001	.16
Exclude Anti-depressants (n=164)	17.54	<.001	.10
Exclude Anti-psychotics (n=185)	19.19	<.001	.10
R middle occipital gyrus			
Full Sample (n=195)	25.18	<.001	.12
Exclude Stimulants (n=150)	21.45	<.001	.13
Exclude Anti-depressants (n=164)	17.39	<.001	.10
Exclude Anti-psychotics (n=185)	25.21	<.001	.12

R post-central gyrus			
Full Sample (n=195)	22.25	<.001	.11
Exclude Stimulants (n=150)	22.22	<.001	.14
Exclude Anti-depressants (n=164)	14.98	<.001	.09
Exclude Anti-psychotics (n=185)	24.66	<.001	.12
R superior temporal gyrus			
Full Sample (n=195)	20.33	<.001	.10
Exclude Stimulants (n=150)	19.14	<.001	.12
Exclude Anti-depressants (n=164)	15.87	<.001	.09
Exclude Anti-psychotics (n=185)	22.60	<.001	.11
R superior frontal gyrus			
Full Sample (n=195)	21.70	<.001	.10
Exclude Stimulants (n=150)	28.11	<.001	.16
Exclude Anti-depressants (n=164)	22.87	<.001	.13
Exclude Anti-psychotics (n=185)	22.51	<.001	.11
R lingual/fusiform gyrus			
Full Sample (n=195)	21.18	<.001	.10
Exclude Stimulants (n=150)	26.22	<.001	.16
Exclude Anti-depressants (n=164)	27.25	<.001	.15
Exclude Anti-psychotics (n=185)	21.49	<.001	.11
L pre-central gyrus			
Full Sample (n=195)	21.34	<.001	.10
Exclude Stimulants (n=150)	21.58	<.001	.13
Exclude Anti-depressants (n=164)	16.97	<.001	.10
Exclude Anti-psychotics (n=185)	20.21	<.001	.10
R middle/superior frontal gyrus			
Full Sample (n=195)	19.23	<.001	.09
Exclude Stimulants (n=150)	24.26	<.001	.15
Exclude Anti-depressants (n=164)	13.59	<.001	.08
Exclude Anti-psychotics (n=185)	18.72	<.001	.10
R middle frontal gyrus			
Full Sample (n=195)	19.75	<.001	.10
Exclude Stimulants (n=150)	24.96	<.001	.15
Exclude Anti-depressants (n=164)	16.64	<.001	.10
Exclude Anti-psychotics (n=185)	20.33	<.001	.10
R fusiform gyrus			
Full Sample (n=195)	19.41	<.001	.09
Exclude Stimulants (n=150)	21.24	<.001	.13
Exclude Anti-depressants (n=164)	21.10	<.001	.12
Exclude Anti-psychotics (n=185)	20.84	<.001	.11
L superior parietal lobule			
Full Sample (n=195)	20.58	<.001	.10
Exclude Stimulants (n=150)	26.42	<.001	.16
Exclude Anti-depressants (n=164)	20.14	<.001	.11

Exclude Anti-psychotics (n=185)	23.01	<.001	.11
L medial frontal gyrus			
Full Sample (n=195)	19.04	<.001	.09
Exclude Stimulants (n=150)	21.63	<.001	.13
Exclude Anti-depressants (n=164)	14.62	<.001	.09
Exclude Anti-psychotics (n=185)	18.95	<.001	.10
L superior frontal gyrus			
Full Sample (n=195)	20.03	<.001	.10
Exclude Stimulants (n=150)	26.24	<.001	.16
Exclude Anti-depressants (n=164)	15.49	<.001	.09
Exclude Anti-psychotics (n=185)	19.69	<.001	.10

ACC = Anterior Cingulate Cortex; ARI = Affective Reactivity Index; L = Left; R = Right.

Iterative analyses excluding subjects by medication class are shown. The first row in a region summarizes results of the full sample as reported in the main text. Subsequent rows summarize results excluding subjects on stimulants, anti-depressants, and anti-psychotics, respectively. Medication data were collected within 30 days of the scan, except for 4 subjects. All the significant results remained after excluding subjects on each class of medication.

[a] The Condition effect refers to the attention portion of the trial immediately after receiving rigged vs. positive feedback.

Table S4. Post-hoc Analyses of Medication Effects on ARI x Condition[a] Activation Results

Regions	Analysis			Correlation
	F value	*p* value	η_p^2	(*r*)[b]
L & R cingulate gyrus, R superior frontal gyrus				
Full Sample (n=195)	34.81	<.001	.16	.40
Exclude Stimulants (n=150)	36.74	<.001	.20	.45
Exclude Anti-depressants (n=164)	25.86	<.001	.14	.38
Exclude Anti-psychotics (n=185)	37.06	<.001	.17	.42
R middle frontal gyrus				
Full Sample (n=195)	31.63	<.001	.14	.38
Exclude Stimulants (n=150)	31.92	<.001	.18	.43
Exclude Anti-depressants (n=164)	29.76	<.001	.16	.40
Exclude Anti-psychotics (n=185)	31.12	<.001	.15	.39
L middle frontal gyrus				
Full Sample (n=195)	25.85	<.001	.12	.35
Exclude Stimulants (n=150)	25.22	<.001	.15	.39
Exclude Anti-depressants (n=164)	17.28	<.001	.10	.32
Exclude Anti-psychotics (n=185)	25.89	<.001	.13	.36
R caudate, thalamus				
Full Sample (n=195)	27.72	<.001	.13	.36
Exclude Stimulants (n=150)	27.25	<.001	.16	.40
Exclude Anti-depressants (n=164)	20.45	<.001	.12	.34
Exclude Anti-psychotics (n=185)	30.98	<.001	.15	.39
R dorsolateral prefrontal cortex				
Full Sample (n=195)	29.23	<.001	.14	.37
Exclude Stimulants (n=150)	29.44	<.001	.17	.41
Exclude Anti-depressants (n=164)	22.66	<.001	.13	.36
Exclude Anti-psychotics (n=185)	31.88	<.001	.15	.39
R cuneus				
Full Sample (n=195)	23.65	<.001	.11	.33
Exclude Stimulants (n=150)	20.89	<.001	.13	.36
Exclude Anti-depressants (n=164)	14.33	<.001	.08	.29
Exclude Anti-psychotics (n=185)	22.83	<.001	.11	.34
R precuneus				
Full Sample (n=195)	25.68	<.001	.12	.35
Exclude Stimulants (n=150)	21.69	<.001	.13	.36
Exclude Anti-depressants (n=164)	19.66	<.001	.11	.33
Exclude Anti-psychotics (n=185)	27.87	<.001	.14	.37
L middle frontal gyrus				
Full Sample (n=195)	22.33	<.001	.11	.33
Exclude Stimulants (n=150)	24.29	<.001	.15	.38
Exclude Anti-depressants (n=164)	19.60	<.001	.11	.33
Exclude Anti-psychotics (n=185)	23.71	<.001	.12	.34

R inferior frontal gyrus				
Full Sample (n=195)	28.29	<.001	.13	.36
Exclude Stimulants (n=150)	24.69	<.001	.15	.38
Exclude Anti-depressants (n=164)	19.60	<.001	.11	.33
Exclude Anti-psychotics (n=185)	31.10	<.001	.15	.39
L pre- and post-central gyrus				
Full Sample (n=195)	20.81	<.001	.10	.32
Exclude Stimulants (n=150)	19.48	<.001	.12	.35
Exclude Anti-depressants (n=164)	17.96	<.001	.10	.32
Exclude Anti-psychotics (n=185)	19.84	<.001	.10	.32
L parahippocampal gyrus				
Full Sample (n=195)	23.86	<.001	.11	.34
Exclude Stimulants (n=150)	21.09	<.001	.13	.36
Exclude Anti-depressants (n=164)	20.05	<.001	.11	.34
Exclude Anti-psychotics (n=185)	24.44	<.001	.12	.35
L caudate				
Full Sample (n=195)	21.53	<.001	.10	.32
Exclude Stimulants (n=150)	20.01	<.001	.12	.35
Exclude Anti-depressants (n=164)	16.37	<.001	.09	.31
Exclude Anti-psychotics (n=185)	23.19	<.001	.12	.34
R superior temporal gyrus				
Full Sample (n=195)	21.37	<.001	.10	.32
Exclude Stimulants (n=150)	24.18	<.001	.15	.38
Exclude Anti-depressants (n=164)	19.74	<.001	.11	.33
Exclude Anti-psychotics (n=185)	22.66	<.001	.11	.34
R pre-central gyrus				
Full Sample (n=195)	21.29	<.001	.10	.32
Exclude Stimulants (n=150)	20.81	<.001	.13	.36
Exclude Anti-depressants (n=164)	17.75	<.001	.10	.32
Exclude Anti-psychotics (n=185)	23.20	<.001	.12	.34
L pre-central gyrus				
Full Sample (n=195)	20.85	<.001	.10	.32
Exclude Stimulants (n=150)	20.43	<.001	.13	.35
Exclude Anti-depressants (n=164)	17.79	<.001	.10	.32
Exclude Anti-psychotics (n=185)	21.11	<.001	.11	.33
L cingulate gyrus				
Full Sample (n=195)	20.77	<.001	.10	.32
Exclude Stimulants (n=150)	21.77	<.001	.13	.36
Exclude Anti-depressants (n=164)	18.90	<.001	.11	.33
Exclude Anti-psychotics (n=185)	19.26	<.001	.10	.31
L superior frontal gyrus				
Full Sample (n=195)	19.68	<.001	.10	.31
Exclude Stimulants (n=150)	23.48	<.001	.14	.38
Exclude Anti-depressants (n=164)	15.76	<.001	.09	.30

Exclude Anti-psychotics (n=185)	24.72	<.001	.12	.35

ARI = Affective Reactivity Index; L = Left; R = Right.

Iterative analyses excluding subjects by medication class are shown. The first row in a region summarizes results of the full sample as reported in the main text. Subsequent rows summarize results excluding subjects on stimulants, anti-depressants, and anti-psychotics, respectively. Medication data were collected within 30 days of the scan, except for 4 subjects. All the significant results remained after excluding subjects on each class of medication.

[a] The Condition effect refers to the attention portion of the trial immediately after receiving rigged vs. positive feedback.

[b] Correlations between ARI and the difference in brain activation after receiving rigged vs. positive feedback (rigged minus positive), after adjusting for motion and symptoms of anxiety and attention-deficit/hyperactivity disorder.

Table S5. ARI × Age × Condition[a] Analyses in IFG-based Functional Connectivity

Regions of Co-activation[b]	Size (mm^3)	Peak (x, y, z)[c]	Analysis[d]		
			$F_{1,187}$	*p* value	η_p^2
ARI × Age × Condition					
L putamen	1344	(-24, -4, 9)	6.89	.009	.04
ARI × Condition					
Periaqueductal gray, culmen	8484	(-9, -24, -14)	26.19	<.001	.12

ARI = Affective Reactivity Index; IFG = inferior frontal gyrus; L = Left.

[a] The Condition effect refers to the attention portion of the trial immediately after receiving rigged vs. positive feedback (i.e., the N+1 trial).

[b] Region comprising the greatest portion of the cluster extent. Clusters were determined using voxelwise p=.001. At this threshold, clusters ≥ 1094 mm^3 survive whole-brain correction at α=.05/4=.0125 (correcting for 4 tests for each of the seed regions: left and right inferior frontal gyrus and amygdala).

[c] Coordinates are in Talairach space.

[d] One outlier was excluded from the post-hoc analysis.

Table S6. ARI × SCARED × Condition[a] Analyses in IFG-based Functional Connectivity

Regions of Co-activation[b]	Size (mm^3)	Peak (x, y, z)[c]	Analysis		
			$F_{1,189}$	*p* value	η_p^2
		R Inferior Frontal Gyrus (seed)			
L superior temporal gyrus, precentral gyrus	2984	(-49, -1, 6)	25.64	<.001	.12
R insula, precentral gyrus	2594	(46, -4, 9)	23.52	<.001	.11
R insula	2172	(39, -19, 9)	23.38	<.001	.11
L superior temporal gyrus	1891	(-44, -34, 4)	24.57	<.001	.12
R superior temporal gyrus	1484	(56, -29, 4)	20.50	<.001	.10
L middle temporal gyrus	1453	(-41, -64, 16)	20.22	<.001	.10
L posterior cingulate cortex	1375	(-14, -54, 11)	19.85	<.001	.10
		L Inferior Frontal Gyrus (seed)			
L superior temporal gyrus	14719	(-54, -24, 6)	36.23	<.001	.16
R superior temporal gyrus	9813	(44, -29, 9)	28.03	<.001	.13
R posterior cingulate cortex	4828	(19, -44, 4)	26.79	<.001	.12
L posterior cingulate cortex	4578	(-14, -71, 6)	27.27	<.001	.13

ARI = Affective Reactivity Index; IFG = inferior frontal gyrus; L = Left; R = Right; SCARED = Screen for Child Anxiety Related Emotional Disorders.

[a] The Condition effect refers to the attention portion of the trial immediately after receiving rigged vs. positive feedback (i.e., the N+1 trial).

[b] Region comprising the greatest portion of the cluster extent. Clusters were determined using voxelwise *p*=.001. At this threshold, clusters ≥ 1094 mm^3 survive whole-brain correction at α=.05/4=.0125 (correcting for 4 tests for each of the seed regions: left and right inferior frontal gyrus and amygdala).

[c] Coordinates are in Talairach space.

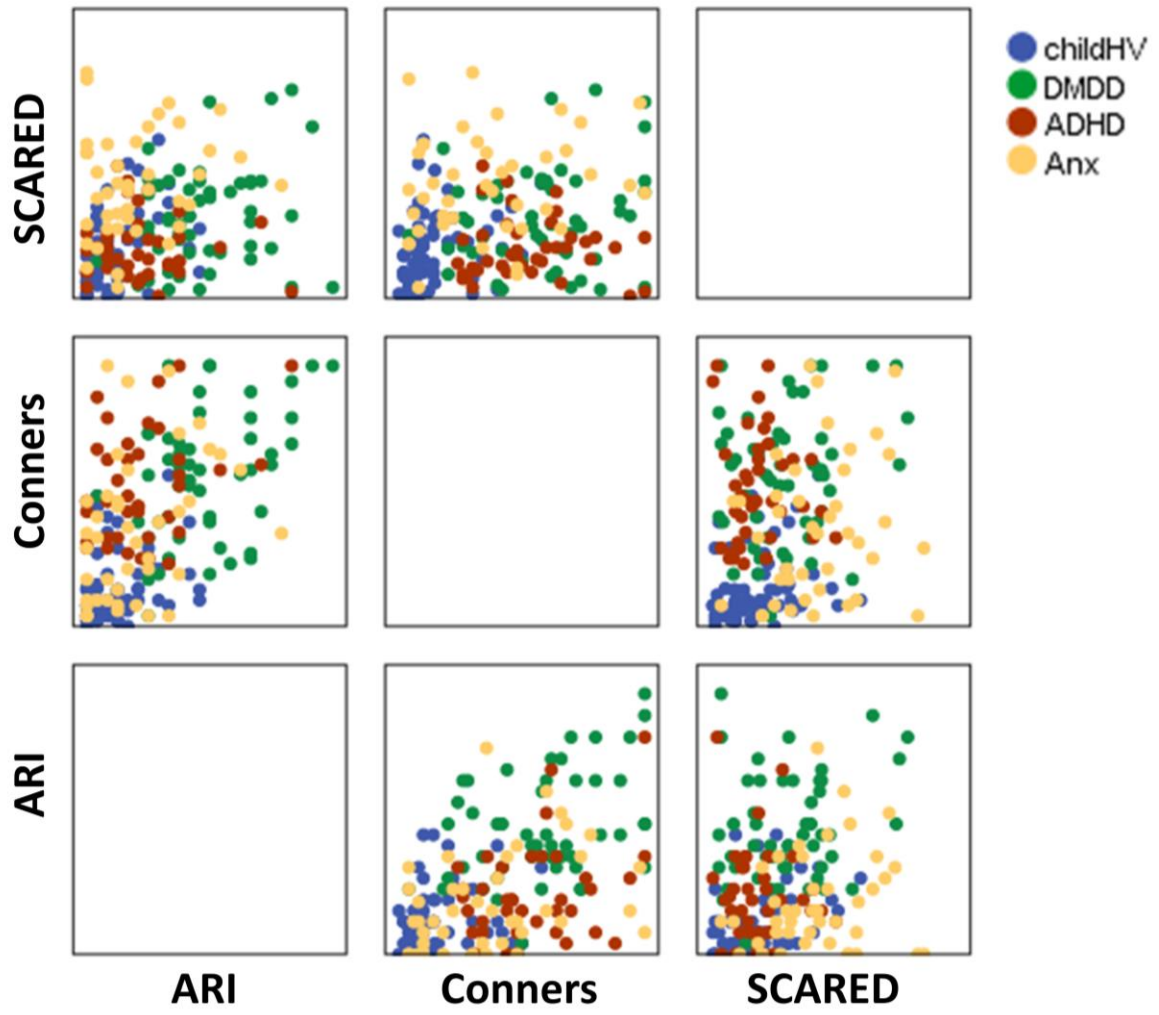

Figure S1. Scatterplot Matrices Depicting Participant Distribution along Irritability, Anxiety, and ADHD Dimensions by Diagnosis

ADHD = Attention-deficit/hyperactivity disorder; Anx = Anxiety Disorder; ARI = Affective Reactivity Index; HV = Healthy Volunteers; Conners = Conners' Parent Rating Scale; DMDD = Disruptive Mood Dysregulation Disorder; SCARED = Screen for Child Anxiety Related Emotional Disorders.

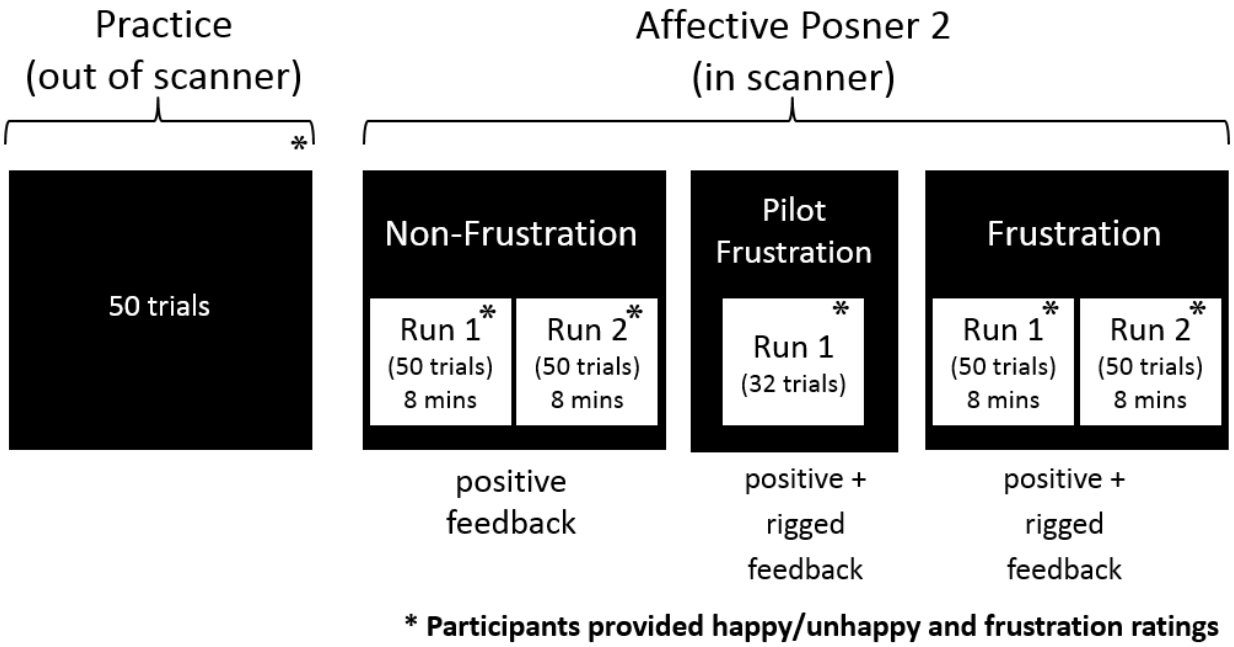

Figure S2. Task Structure of the Affective Posner 2

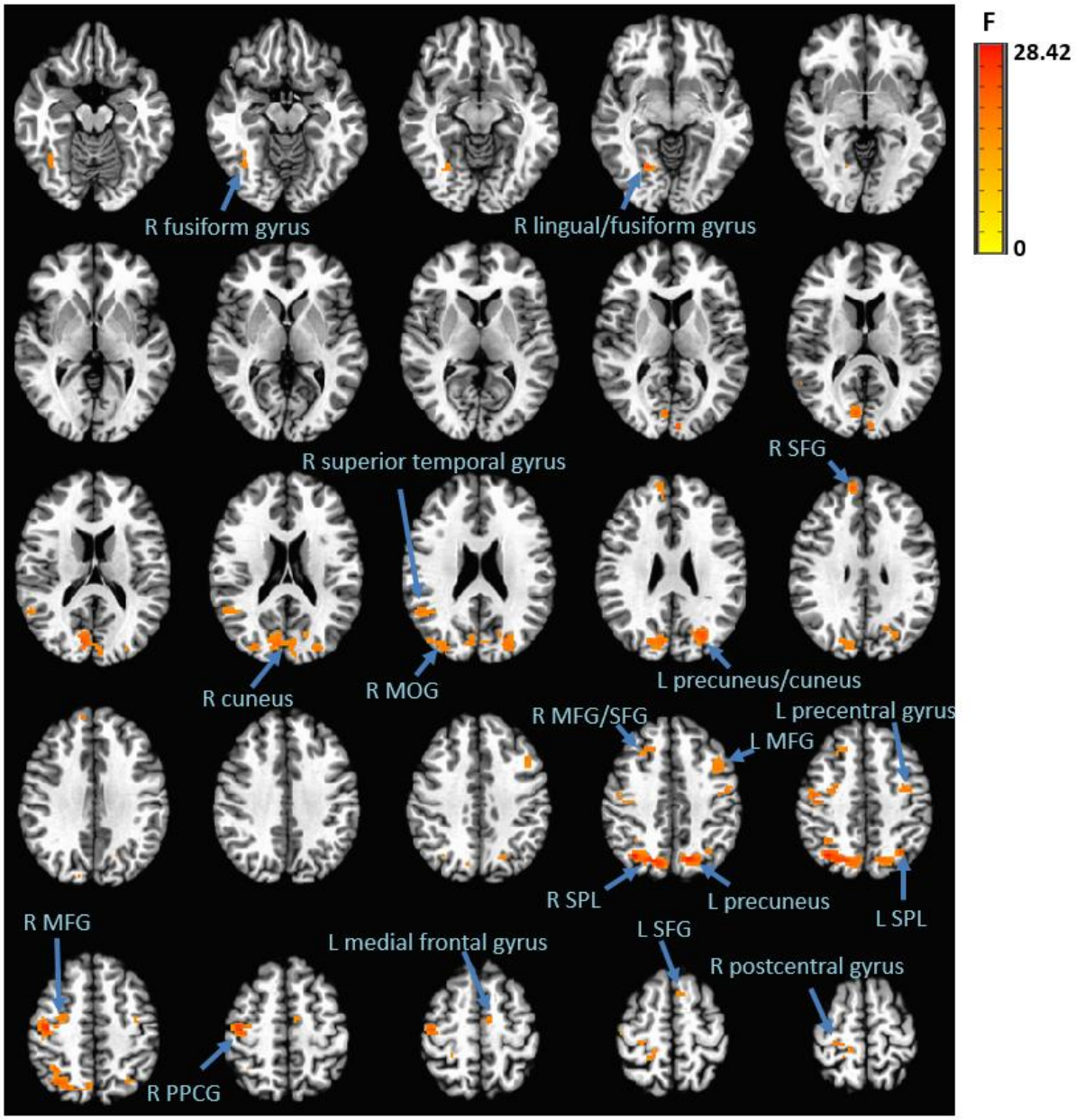

Figure S3. ARI × Age × Condition on Activation in Regions Not Depicted in Main Text

Activation in these regions after receiving rigged vs. positive feedback varied depending on ARI score and age. Overall, higher irritability was more strongly related to increased activation in young children (aged 8-11.5 years) than in younger adolescents (aged 11.5-14 years); irritability was not related to activation in older adolescents (aged 14-18 years).

Clusters were thresholded at voxelwise p=.0001 and $k \geq 203$ mm^3 (whole-brain corrected at α=.05).

ARI = Affective Reactivity Index; L = left; MFG = middle frontal gyrus; MOG = middle occipital gyrus; PPCG = pre- and post-central gyrus; R = right; SFG = superior frontal gyrus; SPL = superior parietal lobule.

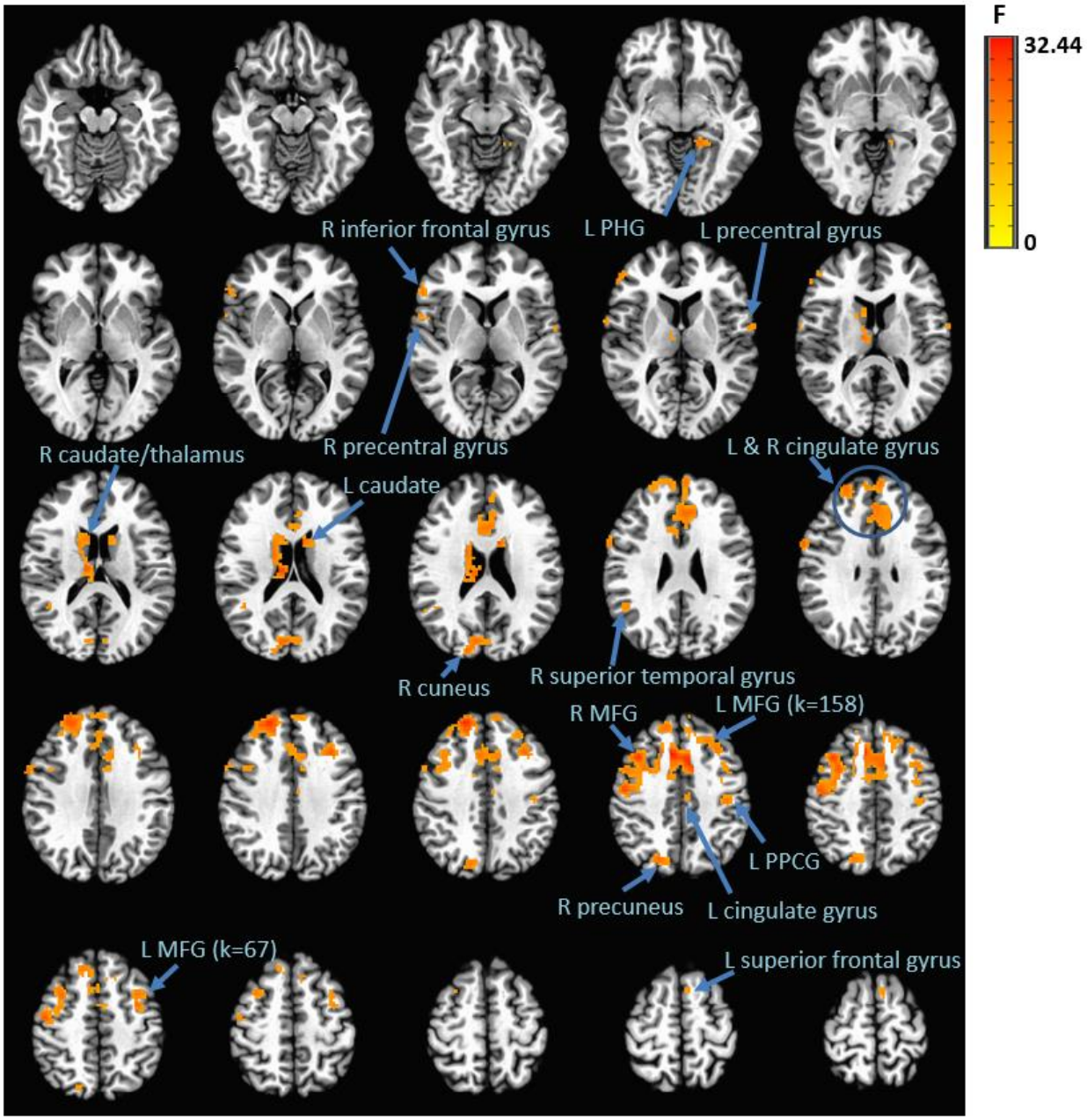

Figure S4. ARI × Condition on Activation in Regions Not Depicted in Main Text

Activation in these regions after receiving rigged vs. positive feedback varied depending on ARI score. That is, higher irritability was related to *more* activation after receiving rigged relative to positive feedback.

Clusters were thresholded at voxelwise p=.0001 and $k \geq 203$ mm^3 (whole-brain corrected at α=.05).

ARI = Affective Reactivity Index; L = left; MFG = middle frontal gyrus; PHG = parahippocampal gyrus; PPCG = pre- and post-central gyrus; R = right.

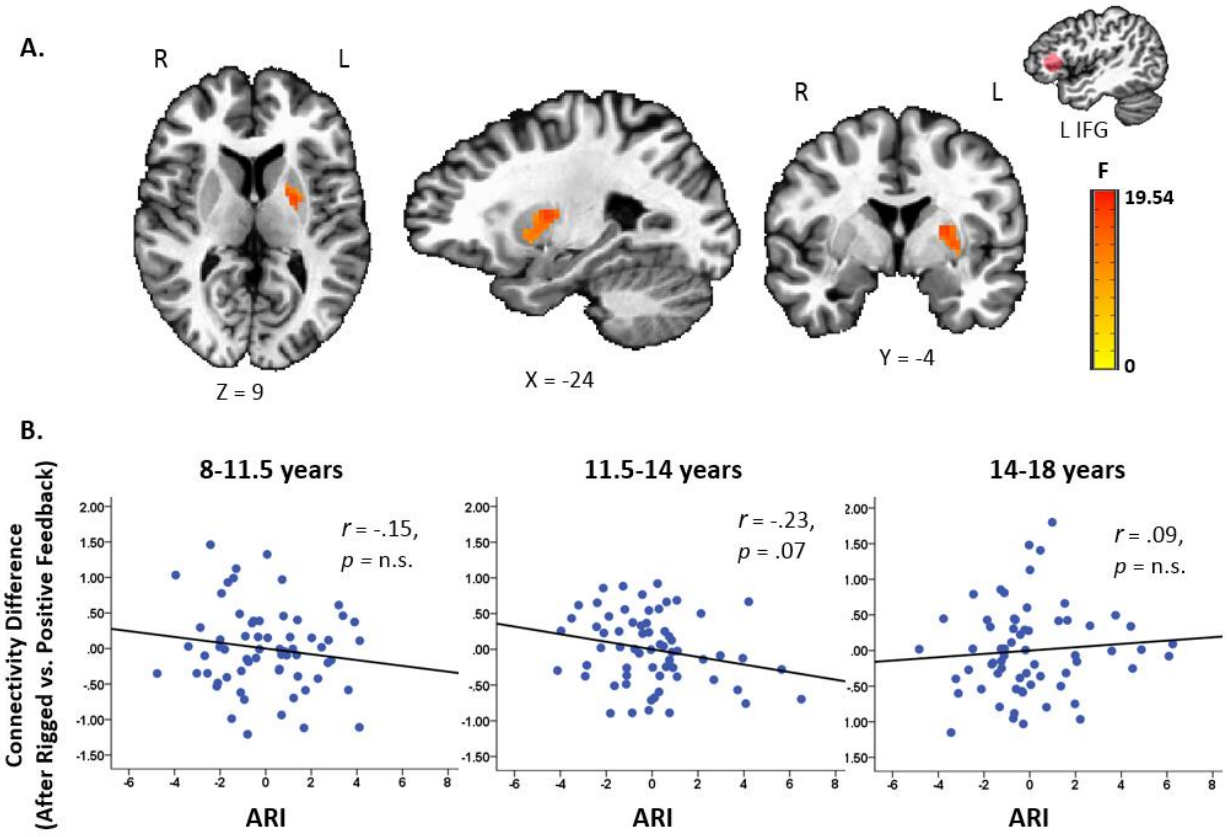

Figure S5. ARI × Age × Condition on IFG-based Functional Connectivity

A. Functional connectivity between the left inferior frontal gyrus (L IFG; seed region) and the left putamen (cluster size = 1344 mm^3) on the N+1 trial (i.e., after receiving rigged vs. positive feedback) varied with levels of irritability (Affective Reactivity Index, ARI, scores) and age. Clusters were thresholded at voxelwise p=.001 and $k \geq 1094$ mm^3 (whole-brain corrected at α=.0125).

B. Partial regression plots by age tertiles depict individual data points and the correlations between mean-centered irritability and the differences between connectivity measures after rigged vs. after positive feedback. Following rigged vs. positive feedback, higher irritability was associated with decreased connectivity in younger adolescents (aged 11.5-14 years) at a trend level.

ARI = Affective Reactivity Index; IFG = inferior frontal gyrus; L = left; R = right.

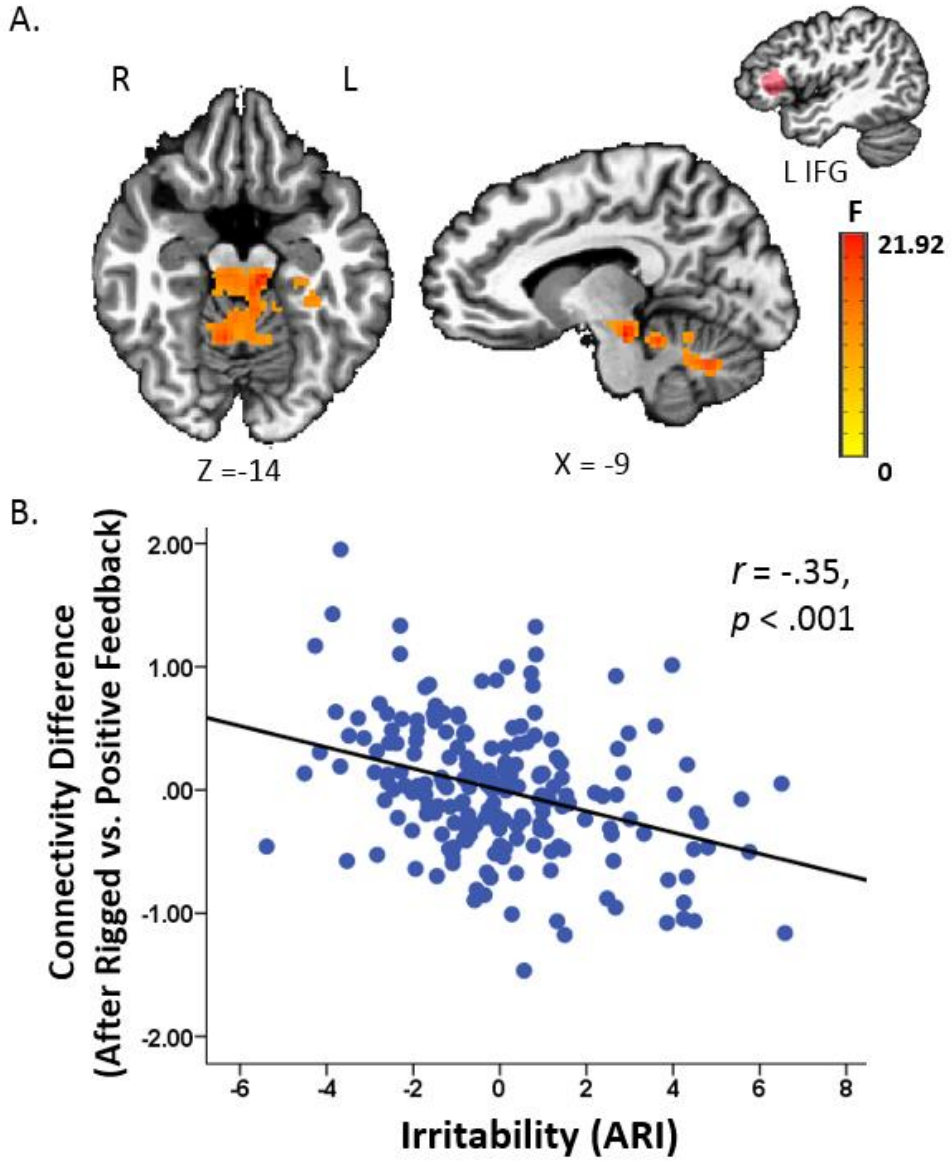

Figure S6. ARI × Condition on IFG-based Functional Connectivity

A. Functional connectivity between the left inferior frontal gyrus (L IFG; seed region) and periaqueductal gray/culmen (cluster size = 8484 mm^3) on the N+1 trial (i.e., after receiving rigged vs. positive feedback) varied with levels of irritability (Affective Reactivity Index, ARI, scores). Clusters were thresholded at voxelwise p=.001 and $k \geq 1094$ mm^3 (whole-brain corrected at α=.0125).
B. Partial regression plot depicts individual data points and the correlation between mean-centered irritability and the difference between connectivity measures after rigged vs. after positive feedback. Following rigged vs. positive feedback, higher irritability was associated with decreased connectivity.

ARI = Affective Reactivity Index; IFG = inferior frontal gyrus; L = left; R = right.

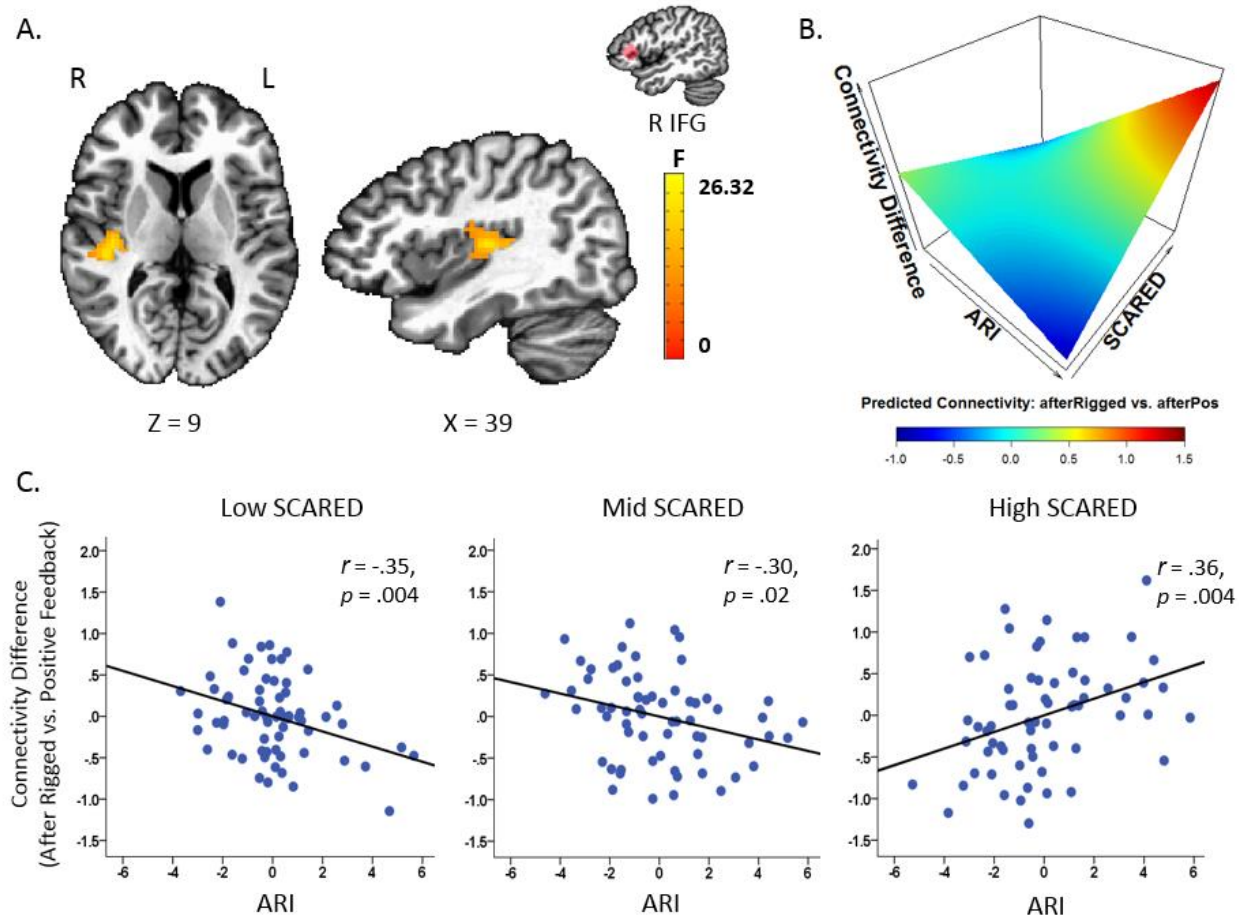

Figure S7. ARI × SCARED × Condition on Functional Connectivity between IFG and Insula

A. Functional connectivity between right inferior frontal gyrus (R IFG; seed region) and right insula (cluster size = 2172 mm^3) on the N+1 trial (i.e., after receiving rigged vs. positive feedback) varied with levels of irritability (ARI scores) and anxiety (SCARED scores). Clusters were thresholded at voxelwise p=.001 and k≥1094 mm^3 (whole-brain corrected at α=.0125).
B. Associations among ARI, SCARED, and connectivity. Functional connectivity measures were extracted for each condition (the N+1 trial after rigged and after positive feedback) for each subject. These values were entered in the same ANCOVA model (controlling for ADHD symptoms and motion) as in the main analysis, and predicted connectivity measures were generated. The differences between the predicted connectivity measures after rigged vs. after positive feedback were plotted. The 3-D graph shows that, after receiving rigged vs. positive feedback, youth with high irritability and high anxiety exhibited increased IFG-insula connectivity, whereas youth with high irritability and low anxiety exhibited decreased connectivity.
C. Partial regression plots by SCARED tertiles depict individual data points and the correlations between mean-centered irritability and the connectivity difference on the contrast.

ARI = Affective Reactivity Index; IFG = inferior frontal gyrus; L = left; R = right; SCARED = Screen for Child Anxiety Related Emotional Disorders.

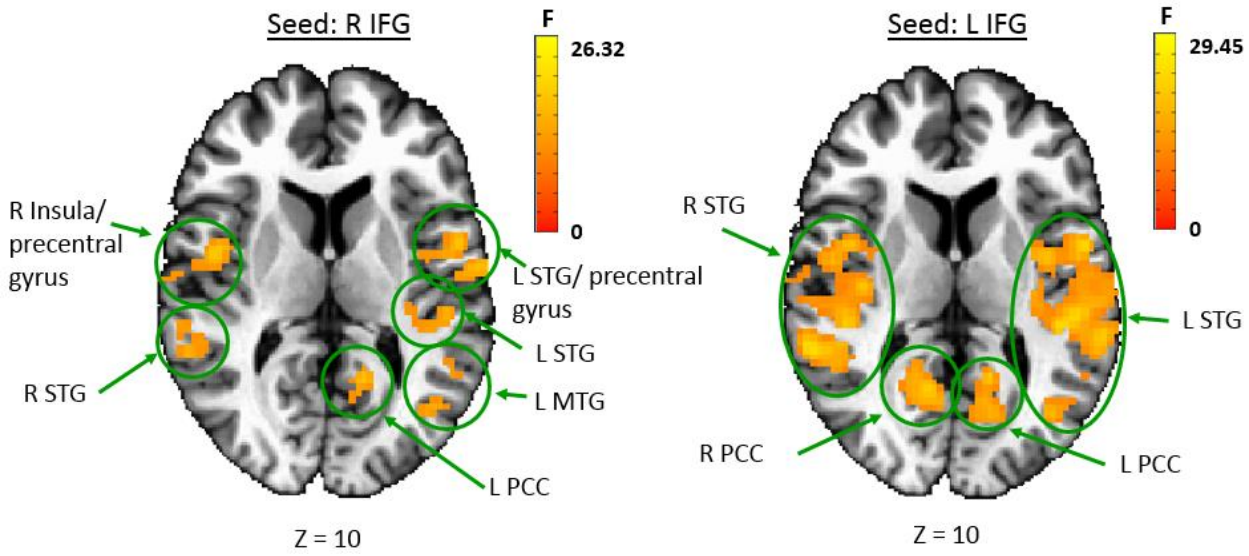

Figure S8. ARI × SCARED × Condition on IFG-based Functional Connectivity in Other Regions

Functional connectivity between the inferior frontal gyrus (IFG; seed region) and other regions after receiving rigged vs. positive feedback varied depending on ARI and SCARED scores. Across these regions, after receiving rigged vs. positive feedback, youth with high irritability and high anxiety exhibited increased connectivity, whereas youth with high irritability and low anxiety exhibited decreased connectivity.

Clusters were thresholded at voxelwise p=.001 and $k \geq 1094$ mm^3 (whole-brain corrected at α=.0125).

ARI = Affective Reactivity Index; IFG = inferior frontal gyrus; L = left; MTG = middle temporal gyrus; PCC = posterior cingulate cortex; R = right; SCARED = Screen for Child Anxiety Related Emotional Disorders; STG = superior temporal gyrus.

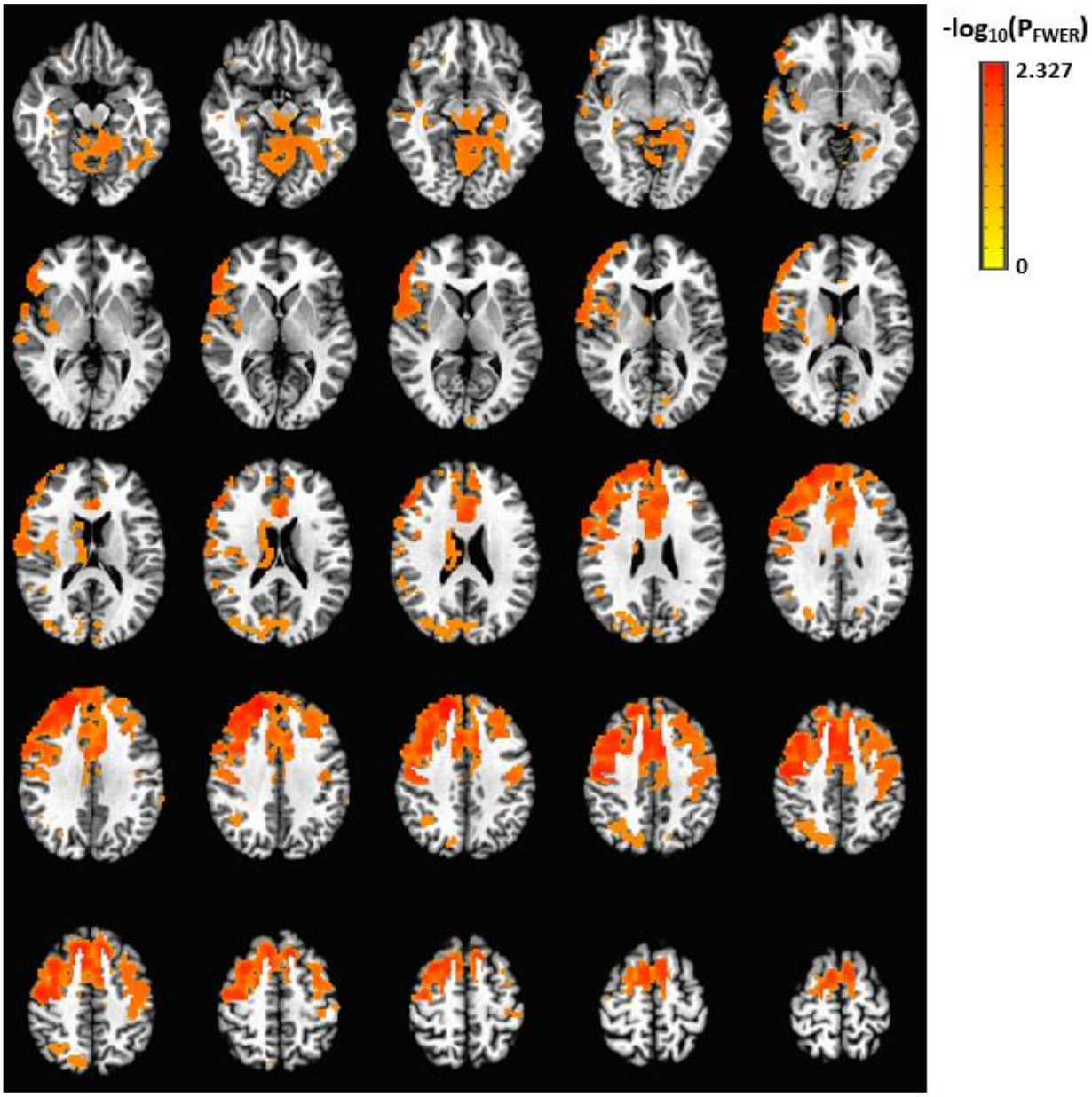

Figure S9. ARI × Condition Effect from Threshold-Free Cluster Enhancement Analysis

Activation in these regions after receiving rigged vs. positive feedback varied depending on ARI score. That is, higher irritability was related to *more* activation after receiving rigged relative to positive feedback.

Clusters were thresholded at $P_{FWER}<.05$ [$-\log_{10}(P_{FWER})=1.301$].

ARI = Affective Reactivity Index.

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. 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
3. Shaffer D, Gould MS, Brasic J, et al: A children's global assessment scale (CGAS). Arch Gen Psychiatry 1983; 40:1228–1231
4. Stringaris A, Goodman R, Ferdinando S, et al: The Affective Reactivity Index: A concise irritability scale for clinical and research settings. J Child Psychol Psychiatry 2012; 53:1109–1117
5. Hale WW, Raaijmakers Q, Muris P, et al: Developmental trajectories of adolescent anxiety disorder symptoms: A 5-year prospective community study. J Am Acad Child Adolesc Psychiatry 2008; 47:556–564
6. Glow RA, Glow PH, Rump EE: The stability of child behavior disorders: A one year test-retest study of Adelaide versions of the Conners teacher and parent rating scales. J Abnorm Child Psychol 1982; 10:33–59
7. Tseng W-L, Moroney E, Machlin L, et al: Test-retest reliability and validity of a frustration paradigm and irritability measures. J Affect Disord 2017; 212:38–45
8. DeSousa DA, Stringaris A, Leibenluft E, et al: Cross-cultural adaptation and preliminary psychometric properties of the Affective Reactivity Index in Brazilian Youth: implications for DSM-5 measured irritability. Trends Psychiatry Psychother 2013; 35:171–180
9. Mulraney MA, Melvin GA, Tonge BJ: Psychometric properties of the Affective Reactivity Index in Australian adults and adolescents. Psychol Assess 2014; 26:148–155
10. Birmaher B, Brent DA, Chiappetta L, et al: Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): a replication study. J Am Acad Child Adolesc Psychiatry 1999; 38:1230–1236
11. Conners CK, Sitarenios G, Parker JDA, et al: The revised Conners' Parent Rating Scale (CPRS-R): Factor structure, reliability, and criterion validity. J Abnorm Child Psychol 1998; 26:257–68
12. Deveney CM, Connolly ME, Haring CT, et al: Neural Mechanisms of Frustration in Chronically Irritable Children. Am J Psychiatry 2013; 170:1186–1194
13. Rich BA, Carver FW, Holroyd T, et al: Different neural pathways to negative affect in youth with pediatric bipolar disorder and severe mood dysregulation. J Psychiatr Res 2011; 45:1283–1294

14. Rich BA, Holroyd T, Carver FW, et al: A preliminary study of the neural mechanisms of frustration in pediatric bipolar disorder using magnetoencephalography. Depress Anxiety 2010; 27:276–286

15. Rich BA, Schmajuk M, Perez-Edgar KE, et al: The impact of reward, punishment, and frustration on attention in pediatric bipolar disorder. Biol Psychiatry 2005; 58:532–539

16. Rich BA, Schmajuk M, Perez-Edgar KE, et al: Different psychophysiological and behavioral responses elicited by frustration in pediatric bipolar disorder and severe mood dysregulation. Am J Psychiatry 2007; 164:309–317

17. Fischl B: FreeSurfer. NeuroImage 2012; 62:774–781

18. Corbetta M, Shulman GL: Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 2002; 3:201–215

19. Sylvester CM, Barch DM, Corbetta M, et al: Resting state functional connectivity of the ventral attention network in children with a history of depression or anxiety. J Am Acad Child Adolesc Psychiatry 2013; 52:1326–1336.e5.

20. Aron AR, Robbins TW, Poldrack RA: Inhibition and the right inferior frontal cortex. Trends Cogn Sci 2014; 8:170-177

21. Hampshire A, Chamberlain SR, Monti MM, et al: The role of the right inferior frontal gyrus: inhibition and attentional control. Neuroimage 2010; 50:1313–1319

22. Anderson AK, Phelps EA: Lesions of the human amygdala impair enhanced perception of emotionally salient events. Nature 2001; 411:305–309

23. Power JD, Cohen AL, Nelson SM, et al: Functional Network Organization of the Human Brain. Neuron 2011; 72:665–678

24. 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

25. Stoddard JS, Tseng WL, 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; 74:95-103

26. Smith SM, Nichols TE: Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localization in cluster inference. Neuroimage 2009; 44:83-98

27. Winkler AM, Ridgway GR, Webster MA, et al: Permutation inference for the general linear model. Neuroimage 2014; 92:381-397

28. Winkler AM, Ridgway GR, Douaud G, et al: Faster permutation inference in brain imaging. Neuroimage 2016; 141:502-516