Supplemental Methods

Face Emotion Labeling Task

Participants performed a face emotion labeling task during fMRI acquisition. Face stimuli were from 10 actors (4 male, 6 female) in the Pictures of Facial Affect set[1]. Using FantaMorph Deluxe software (www.fantamorph.com), angry, fearful, and happy faces were morphed with neutral faces to create 0% (i.e., neutral), 50%, 75%, and 100% intensity emotion faces (Figure S1). Across 4 runs of approximately 8.5 minutes each, there were 28 trials per emotion intensity condition (e.g., Angry 50%, Angry 75%, etc.), except for neutral faces (i.e., 0% intensity of each angry, fearful, and happy), which had 84 trials (28 trials x 3). Neutral faces were randomly assigned as 0% intensity faces for each of the 3 emotions (angry, fearful, and happy).

Before each trial, participants viewed a fixation cross for a variable amount of time (mean=1800 ms, ranging from 500 ms to 7000 ms). To obtain a robust estimate of baseline, timings unique to each participant for the fixation cross (i.e., inter-trial intervals) were generated using stim_analyze from Analysis of Functional Neuroimages (AFNI[2]). For each trial, a face was presented for 2000 ms. Then, four options to label the emotion ("angry", "fearful", "happy", or "neutral") appeared next to the face for an additional 2000 ms; participants responded via a button box attached to their right hand (MRI Devices, Milwaukee, WI). 50% intensity faces were counted as correct if they were labeled as the emotion (i.e., angry, fearful, or happy), not neutral. Presentation order of faces was randomized. Options of which buttons to press to label the emotion were presented in the same order each time and appeared next to

the face on each trial to reduce the working memory load. Of note, because participants were required to hold their responses until the options appeared next to the face, reaction times are not interpretable with this task.

fMRI Acquisition

MRI data were acquired on a 3T GE MR750 scanner with 32-channel head coil. Participants viewed stimuli projected onto a screen via mirrors. Blood oxygen level dependent (BOLD) images were acquired as 47 contiguous axial slices parallel to the AC-PC line, with whole brain coverage, using an echo planar single-shot gradient echo pulse sequence (matrix size=96x96, TR=2300 ms, TE=25 ms, flip angle=50°, FOV=240 mm, voxel size=2.5 x 2.5 x 2.6 mm). A high-resolution T1-weighed anatomical image was acquired for spatial normalization (124 1.2mm axial slices, flip angle=15°, matrix=256 x 256, FOV=240 mm).

fMRI Preprocessing

fMRI data were preprocessed using AFNI. The first four TRs of each run were discarded to allow the magnet to reach steady state, leaving 222 TRs for analysis in each of the four runs. Slice timing correction was performed. In addition, motion correction, affine alignment of the EPI to the T1 image and of the T1 image to the Talairach template were combined and applied as single per-volume transformations, resulting in a final voxel size of 2.5 x 2.5 x 2.5 mm. Images also underwent spatial smoothing, resulting in average full width at half maximum blur estimates xyz=6.14, 6.07, 5.65, and intensity scaling.

Supplemental Results

Additional analyses were performed to determine whether medication, mood state, anxiety comorbidity, global functioning, and age account for the main result, i.e., the amygdala cluster identified in the Diagnostic Group x Irritability x Emotion x Intensity whole brain analysis (Figure 1). (Mood state at time of scan was determined with the depression and mania ratings [see Table 1 in the main text]: Depressed=Mania≤12 and Depression [Child]≥40 or Depression[Adult]≥20, Hypomanic=Mania>12 and<26 and Depression[Child]<40 or Depression [Adult]<20, Manic=Mania≥26 and Depression [Child]<40 or Depression [Adult]<20, Mixed=Mania>12 and Depression [Child]≥40 or Depression [Adult]≥20, Euthymic=does not meet criteria for other mood states.) To constrain these additional analyses to this region of interest, values extracted and averaged from the cluster were exported to SPSS, and a Diagnostic Group x Irritability x Emotion x Intensity interaction was performed taking into account each of the potentially confounding factors. To summarize, these additional analyses suggest that the results are not primarily driven by these potentially confounding factors.

First, medication usage is quite high in pediatric psychiatric illnesses and could potentially influence brain activation findings. To address potential effects of psychotropic medication, the model was rerun covarying number of medications. When removing variance associated with number of medications, the result was confirmed: Diagnostic Group x Irritability x Emotion x Intensity is still significant for the amygdala cluster ($F_{12, 360}$=3.25, p<.001). To address the potential effects of each class of medication, a series of analyses was performed in which individuals on each class of medication were removed from the analysis. Even with decreased numbers of participants, the results still stood when excluding individuals on antidepressants ($F_{12, 270}$=3.26, p<.001), anti-epileptics ($F_{12, 282}$=3.36, p<.001), stimulants ($F_{12, 276}$=6.61, p<.001), and nonstimulant anti-ADHD medications ($F_{12, 300}$=5.41, p<.001). However, excluding individuals on atypical antipsychotics ($F_{12, 240}$=1.57, p=.101) reduced the significance

level to a trend, which may be due to the reduced sample size (only 3 youth with bipolar disorder, 21 youth with DMDD, and 22 healthy youth included in the analysis). Additionally, we reran the post-hoc analyses for the DMDD group excluding youths on *any* medication. With n=10 medication-free youths with DMDD, the results still stood (happy, p=.03; angry, p=.02; fearful, p=.06). Post-hocs were not re-run for medication-free youths with bipolar disorder due to low numbers (n=3).

Second, because irritability may differ depending on bipolar mood state, we performed the analysis excluding all youths who were not euthymic (1 youth with BP in a depressed episode, 5 youths with BP in a hypomanic episode). Diagnostic Group x Irritability x Emotion x Intensity is still significant when including only euthymic youth ($F_{12,354}$=3.19, p<.001). Additionally, irritability levels were similar between DMDD and bipolar groups whether non-euthymic youth were included (t_{47}=.94, p=.35) or not (t_{41}=1.28, p=.21).

Third, the patient groups in this study did not differ on comorbid anxiety diagnoses nor anxiety symptoms (Table 1). Nevertheless, because anxiety has been associated with alterations in face emotion processing brain circuitry (e.g.,[3]), we performed additional analyses to address potential effects of anxiety on our results. When excluding all participants with a comorbid anxiety diagnosis, the Diagnostic Group x Irritability x Emotion x Intensity interaction is significant ($F_{12, 246}$=5.62, p<.001). Similarly, when covarying for anxiety symptoms, as measured by SCARED mean child- and parent-report, the main results still stand (Diagnostic Group x Irritability x Emotion x Intensity, $F_{12, 110}$=2.84, p=.002).

Fourth, youth with bipolar disorder had lower global functioning (CGAS/GAF) scores than youth with DMDD (Table 1). Thus, to address the possibility that overall impairment is driving our results, we reran the significant post-hoc analyses from the main result, covarying global functioning scores. (Note that because healthy youth did not receive a CGAS/GAF score, the full four-way interaction was unable to be recreated while covarying global functioning scores.) When covarying global functioning scores,

there was still a significant cubic relationship between irritability and happy ($F_{1, 19}$=4.71, *p*=.04), fearful ($F_{1, 19}$=10.95, *p*=.004), and angry ($F_{1, 19}$=17.10, *p*=.001) faces for youth with DMDD and fearful ($F_{1, 19}$=7.77, *p*=.01) faces for youth with bipolar disorder.

Fourth, significant neural development occurs during adolescence, which may influence our findings. Thus, we performed a follow-up analysis to examine whether age-related differences in brain function were primarily driving our results. When covarying for age, the Diagnostic Group x Irritability x Emotion x Intensity is still significant for the amygdala ($F_{12,384}$ = 5.97, p < .001). This suggests that age is not primarily driving our findings. Further, this is consistent with the fact that our DMDD, bipolar, and healthy groups do not differ in age (Table 1) and that irritability scores do not correlate with age (r = -.11, p = .38).

Lastly, follow-up analyses on the main result were performed with irritability and diagnostic group separately to contrast with the results from the analysis that included both. Neither irritability nor diagnostic group separately identified the full extent of amygdala dysfunction related to irritability in the diagnostic groups. Using values from the amygdala cluster identified in the Diagnostic Group x Irritability x Emotion x Intensity interaction (Figure 1), the analysis was first re-run without Diagnostic Group as a factor. In this analysis, the Irritability x Emotion x Intensity interaction is significant ($F_{2,68}$=4.60, *p*=.01). False discovery rate corrected post-hoc analyses indicate that this is driven by happy faces (*p*=.006), not fearful (*p*=.73) or angry (*p*=.73) faces. Next, the analysis was re-run without Irritability in the model. In this analysis, the Diagnostic Group x Emotion x Intensity interaction is not significant ($F_{12, 128}$<1).

TABLE S1. Accuracy Including Low Performers. Predicting accuracy to label face emotion, when including scanned participants with behavioral performance too poor for MRI analysis (8 youths with DMDD, 3 youths with bipolar disorder, 3 healthy youths with <62 TRs/condition).

	F	df	p
Emotion	19.68	2, 78	<.001
Emotion x Irritability	4.19	2, 78	0.019
Emotion x Group	4.35	4, 158	0.002
Emotion x Group x Irritability	2.56	4, 158	0.04
Intensity	63.54	3, 77	<.001
Intensity x Irritability	1.11	3, 77	0.35
Intensity x Group	0.54	6, 156	0.78
Intensity x Group x Irritability	0.37	6, 156	0.90
Emotion x Intensity	6.58	6, 74	<.001
Emotion x Intensity x Irritability	1.29	6, 74	0.27
Emotion x Intensity x Group	2.39	12, 150	0.008
Emotion x Intensity x Group x Irritability	1.32	12, 150	0.21

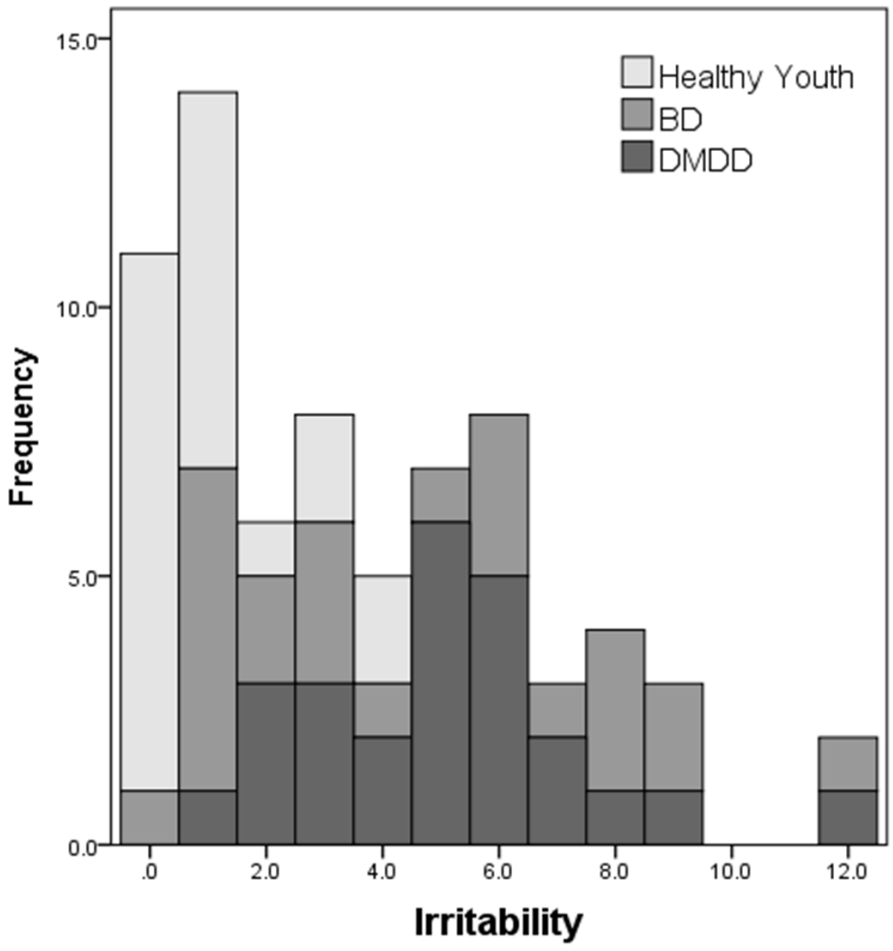

FIGURE S1. Distribution of irritability for each diagnostic group. Skewness values for each group (bipolar disorder=0.51; DMDD=0.81; healthy youth=1.51) were within the range (+/- 2) considered acceptable for normal distribution[4]

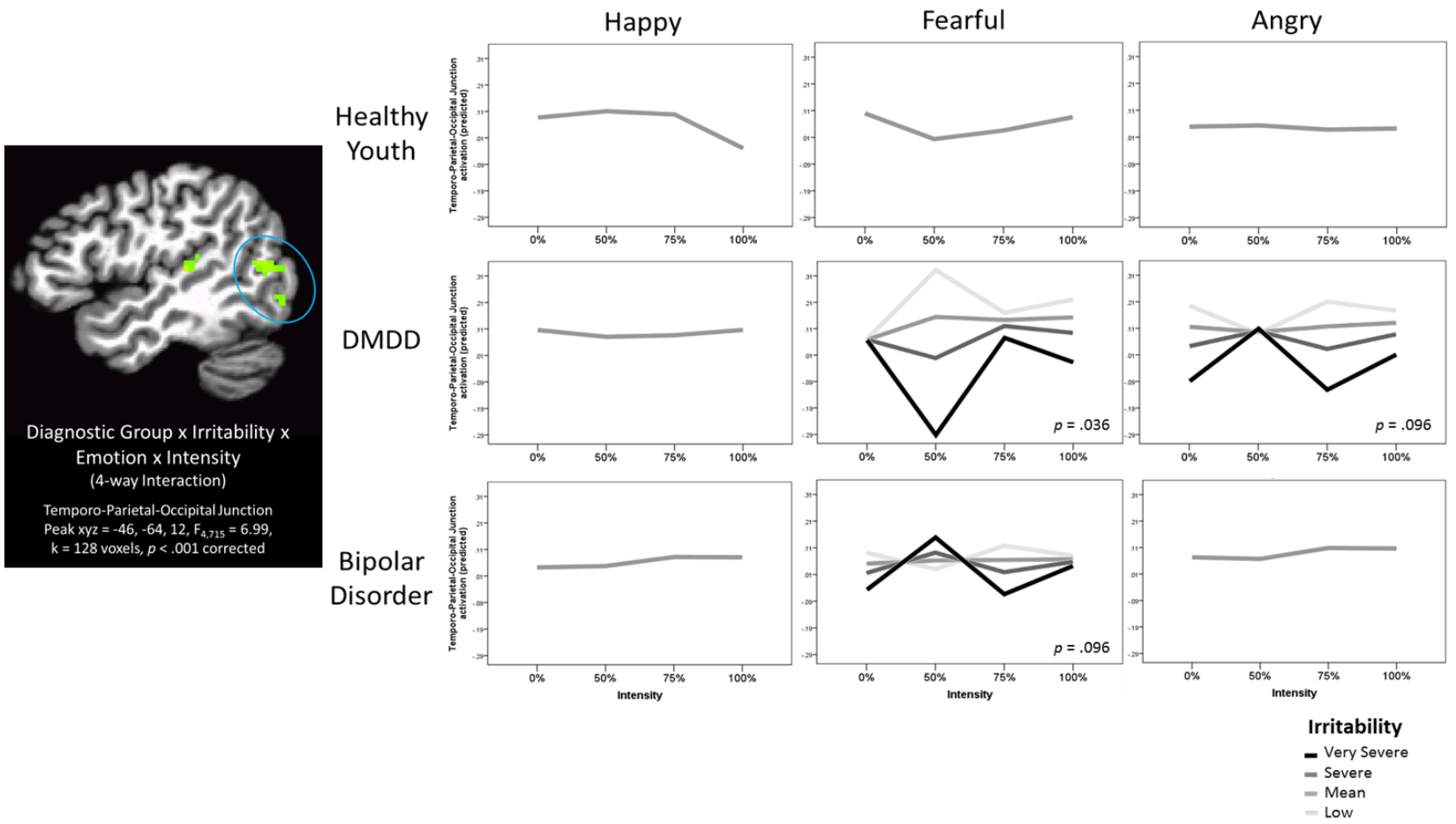

FIGURE S2. Temporo-Parietal-Occipital Junction. For Figures S2-5, brain image depicts cluster significant at whole-brain corrected false probability rate of $p<.05$ in Diagnostic Group x Irritability x Emotion x Intensity contrast. Plots depict predicted cluster activation based on selected levels of irritability to illustrate significant Group x Irritability x Emotion x Intensity interaction, with intensity modeled cubically. Irritability was used as a continuous variable in analyses, but for illustrative purposes, selected irritability levels (low=0 (i.e., ~1 SD below mean), mean=3.4, severe=6.4 (i.e., ~1 SD above mean), very severe=12 (i.e., maximum of scale)) are shown in the plot. False discovery rate corrected post-hocs examined association between irritability and intensity levels modeled cubically. Plots for non-significant post-hocs display temporo-parietal-occipital junction response across intensity levels at mean irritability.

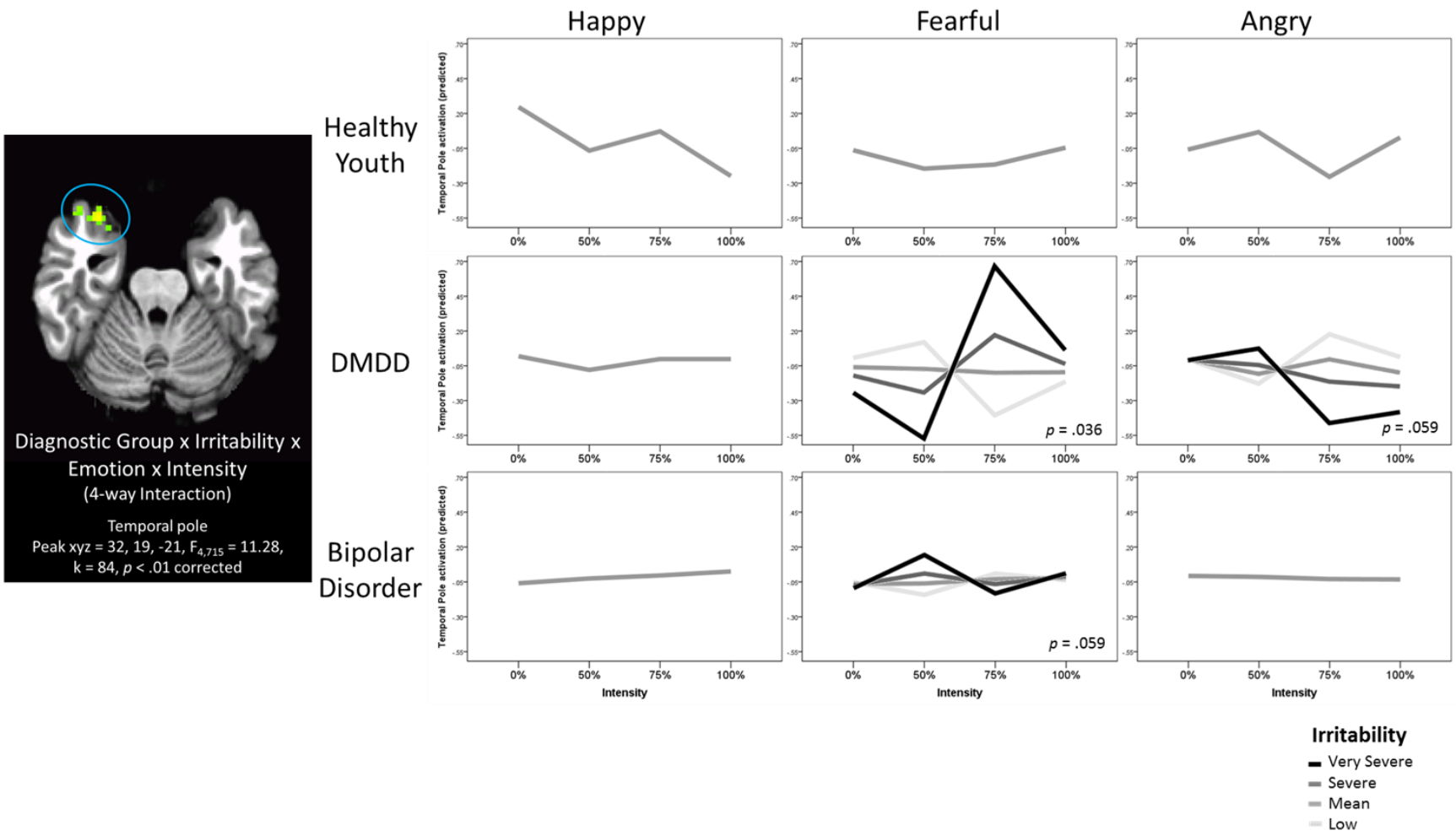

FIGURE S3. Temporal pole. See Figure S2 for brain image and plot information.

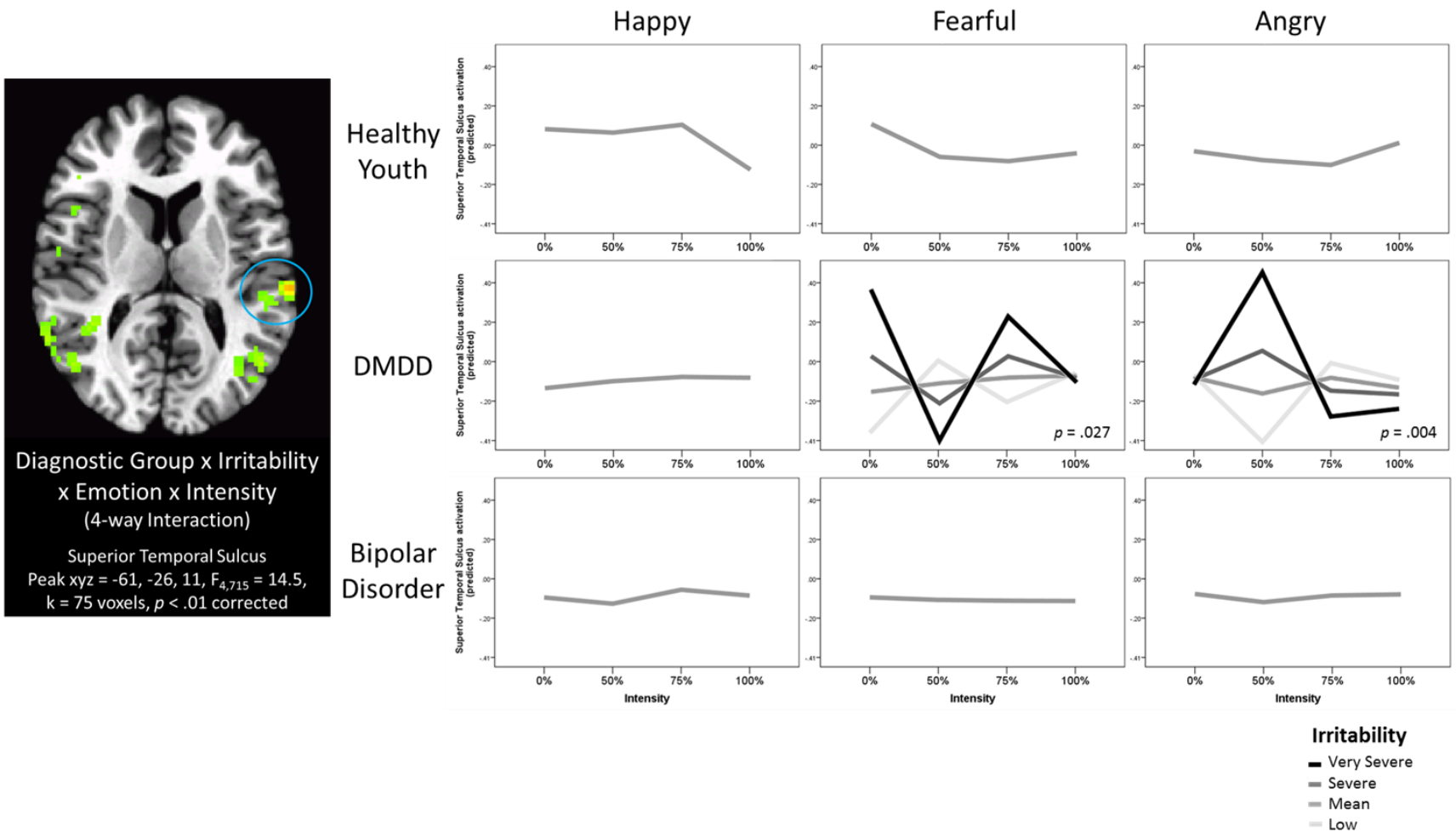

FIGURE S4. Superior Temporal Sulcus. See Figure S2 for brain image and plot information.

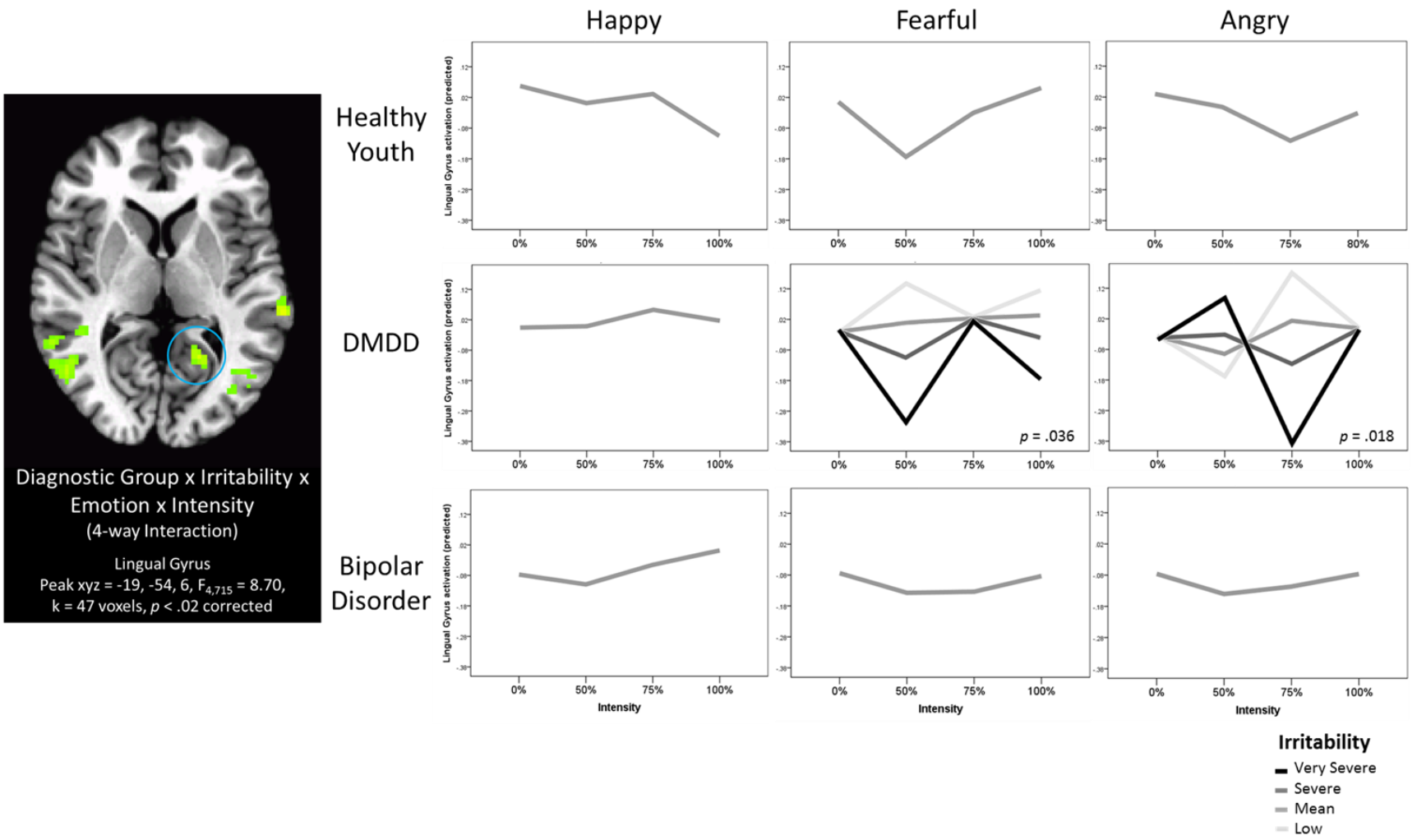

FIGURE S5. Lingual Gyrus. See Figure S2 for brain image and plot information.

Supplemental References

1. Ekman P, Friesen WV. *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press; 1976.

2. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and biomedical research, an international journal.* Jun 1996;29(3):162-173.

3. Monk CS, Telzer EH, Mogg K, et al. Amygdala and ventrolateral prefrontal cortex activation to masked angry faces in children and adolescents with generalized anxiety disorder. *Arch Gen Psychiat.* 2008;65:568-576.

4. George D, Mallery P. *SPSS for Windows Step by Step: A Simple Guide and Reference, 18.0 Update.* Pearson Education; 2010.