Objective: Pediatric bipolar disorder involves poor social functioning, but the neural mechanisms underlying these deficits are not well understood. Previous neuroimaging studies have found deficits in emotional face processing localized to emotional brain regions. However, few studies have examined dysfunction in other regions of the face processing circuit. This study assessed hypoactivation in key face processing regions of the brain in pediatric bipolar disorder.

Method: Youth with a bipolar spectrum diagnosis (n = 20) were matched to a nonbipolar clinical group (n = 20), with similar demographics and comorbid diagnoses, and a healthy control group (n = 20). Youth participated in a functional magnetic resonance imaging (fMRI) scanning which employed a task-irrelevant emotion processing design in which processing of facial emotions was not germane to task performance.

Results: Hypoactivation, isolated to the fusiform gyrus, was found when viewing animated, emerging facial expressions of happiness, sadness, fearfulness, and especially anger in pediatric bipolar participants relative to matched clinical and healthy control groups.

Conclusions: The results of the study imply that differences exist in visual regions of the brain’s face processing system and are not solely isolated to emotional brain regions such as the amygdala. Findings are discussed in relation to facial emotion recognition and fusiform gyrus deficits previously reported in the autism literature. Behavioral interventions targeting attention to facial stimuli might be explored as possible treatments for bipolar disorder in youth. J. Am. Acad. Child Adolesc. Psychiatry, 2013;52(12):1314–1325.

Key Words: emotion, face processing, functional magnetic resonance imaging (fMRI), fusiform gyrus, pediatric bipolar disorder.
of psychopathology (e.g., attention-deficit/hyperactivity disorder [ADHD], oppositional defiant disorder [ODD], anxiety). They also have impairments in memory of faces and are more likely to misinterpret neutral faces as threatening. Furthermore, this effect has been found in those youth at familial risk for BPSD. Functional magnetic resonance imaging (fMRI) studies have also examined impairment in emotional face processing, mostly focusing on dysfunction in the amygdala, an important emotional brain region in the face-processing circuit. Rich et al. found amygdala hyperactivation, compared to that in control participants, when rating perceived fear of an emotional face and reduced functional connectivity to other regions in the face-processing circuit.

Although there is a wealth of literature indicating amygdala dysfunction during judgment of emotional faces, few, if any, studies have focused on other parts of the emotional face processing system in pediatric BPSD. One critical component of the brain’s face processing circuit is the “fusiform face area,” a region localized to the lateral fusiform gyri (FG) that is specialized for face perception. Limited evidence indicates that there may also be impairment in this area in pediatric BPSD. Although neural activity in the FG was not examined directly, Rich et al. found decreased connectivity between the hyperactivated amygdala and this region on an emotional face judgment task in children with BPSD compared to their healthy counterparts. This may indicate abnormal functioning in the FG itself or simply inefficient communication within the face processing circuit (e.g., hyperactivation of the amygdala as a compensatory attempt to communicate with the FG through impaired connections). Adleman et al. also found that faces that were remembered versus forgotten in youth with BPSD were encoded differently in the FG compared to those in healthy subjects.

Studies of adult patients with BPSD have examined structural changes in this region with inconclusive results. Adler et al. showed increased gray matter volume in the FG, whereas Moorhead et al. reported a decline in gray matter density, compared to that in control participants, over a 4-year period.

A second brain region of importance in the face processing circuit is the superior temporal sulcus (STS), which is noted for its role in the perception of biological motion and social action interpretation. The STS is particularly sensitive to the eye-gaze, head movement, and mouth motion. Very little research has focused on the STS in children with BPSD, likely due to the lack of evidence of impaired biological motion processing in this group. Pavuluri et al. did, however, find that pediatric BPSD patients showed elevated activation in the STS, in comparison to healthy children, while viewing angry faces.

The current study was designed to test the possibility of impairment in the entirety of the face processing system in youth with BPSD, which could potentially underlie deficits in emotional face judgment and social interaction. Given the previous literature indicating poor performance in emotional face judgment tasks, the instructions for the task used in the current study did not focus on emotion judgment, potentially enhancing our sensitivity to examine differences outside limbic regions. This allowed us to eliminate the possibility that findings might be due to higher-order differences in cognitive functioning rather than discrete visual processing of facial stimuli. Furthermore, given the evidence of impairments in FG and STS activity in other forms of psychopathology, we collected data from an additional group of children who were matched on demographic characteristics and comorbid psychopathology to the BPSD group (e.g., anxiety, disruptive behavior, ADHD), but who did not have BPSD, as well as a nonclinical healthy control group.

**METHOD**

All methods for recruitment and participant testing were approved by the institutional review boards of all participating universities.

**Participants**

Forty youth (9.89–16.88 years; mean = 13.61 years, SD = 1.95 years) who had been recruited into the original Longitudinal Assessment of Manic Symptoms Study (LAMS1) participated in the second phase of this project (LAMS2). LAMS1 was a multi-site study in which youth aged 6 to 12 years were recruited as they sought care for a variety of diagnoses in 9 mental health outpatient clinics associated with the 4 collaborating universities. The original goal of LAMS1 was to use longitudinal clinical assessment to determine early risk for BPSD. For the current study (LAMS2), beginning 5 years after the start of LAMS1, these youth returned to participate in neuroimaging procedures and clinical assessments (see below). From this sample, 20 youth who had obtained a bipolar spectrum diagnosis (BPSD) at or between entry into LAMS1 and start...
of LAMS2, and who had artifact-free neuroimaging data, were matched, to the best of our ability, to 20 clinic youth without a bipolar spectrum diagnosis (non-BPSD) on demographic characteristics (age, sex, IQ, socioeconomic status [SES]) and comorbid diagnoses. The non-BPSD clinical group was quite heterogeneous in diagnosis but did not statistically differ from the BPSD group in categories of comorbid diagnosis (Table 1). The BPSD group comprised 11 subjects with bipolar I disorder, 6 subjects diagnosed with bipolar NOS, and 3 subjects with cyclothymic disorder. No subjects in this sample had been diagnosed with bipolar II disorder. The amount of time between intake into the LAMS1 study and participation in the LAMS2 scan procedure ranged from 2.47 to 5.97 years (mean \( \bar{x} = 4.42 \) years, SD \( \sigma = 0.83 \) year). An additional, newly recruited sample of 20 mentally and physically healthy participants (controls) without personal or family history of mood disorders in first degree relatives and without family history of BPSD in second degree relatives were also included in the neuroimaging procedures. These participants were matched on demographic variables to the BPSD and non-BPSD participants (Table 1).

Participants were recruited from 3 of the LAMS sites: Case Western Reserve University (n = 24, 9 BPSD, 6 non-BPSD, 9 control); Cincinnati Children’s Hospital (n = 15, 7 BPSD, 7 non-BPSD, 1 control); and University of Pittsburgh Medical Center/Western Psychiatric Institute and Clinic (n = 21, 4 BPSD, 7 non-BPSD, 10 control). Parents/guardians provided written informed consent, and youth provided written informed assent before study participation. Participants received monetary compensation and a framed picture of their structural neuroimaging scan.

Exclusion criteria included severe systemic medical illnesses, neurological disorders, history of head trauma with loss of consciousness, use of medications that may produce central nervous system (CNS) effects (e.g., steroids), IQ < 70 (assessed by the Wechsler Abbreviated Scale of Intelligence), positive urine drug and/or salivary alcohol screen on the day of the scan, alcohol or substance abuse in the past 3 months (determined by the Schedule for Affective Disorders and Schizophrenia for School Age Children–Present and Life Version [K-SADS-PL-W]), visual disturbance (<20/40 Snellen visual acuity), being unable to complete questionnaires in English, parent-reported history of physical/sexual abuse, autism spectrum disorders or developmental delays, posttraumatic stress disorder, and taking more than 3 different psychotropic medications. Additional exclusion criteria for scanning included pregnancy, claustrophobia, or metal objects in the body. LAMS youth were permitted to use prescribed medication(s) before scanning, given the ethical problems with stopping medication for research purposes.

### TABLE 1  Demographic and Clinical Variables and Behavior Performance

<table>
<thead>
<tr>
<th></th>
<th>BPSD (n = 20)</th>
<th>Non-BPSD (n = 20)</th>
<th>Control (n = 20)</th>
<th>F/( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at scan, y, m (SD)</td>
<td>13.50 (2.04)</td>
<td>13.73 (1.90)</td>
<td>13.52 (2.13)</td>
<td>0.08</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>14 (70)</td>
<td>12 (60)</td>
<td>10 (50)</td>
<td>1.67</td>
</tr>
<tr>
<td>Female</td>
<td>6 (30)</td>
<td>8 (40)</td>
<td>10 (50)</td>
<td></td>
</tr>
<tr>
<td>IQ, m (SD)</td>
<td>104.04 (16.86)</td>
<td>104.60 (19.18)</td>
<td>105.59 (11.71)</td>
<td>0.07</td>
</tr>
<tr>
<td>SES, m/SD</td>
<td>3.75 0.79</td>
<td>4.05 0.69</td>
<td>4.05 1.0</td>
<td>0.86</td>
</tr>
<tr>
<td>Signal-to-noise, m/SD</td>
<td>168.61 11.82</td>
<td>162.27 14.91</td>
<td>165.39 10.59</td>
<td>1.28</td>
</tr>
<tr>
<td>Bipolar spectrum disorders, n (%)</td>
<td>2 (100)</td>
<td>0 (0)</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Anxiety, n (%)</td>
<td>2 (10)</td>
<td>2 (10)</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Depression, n (%)</td>
<td>—</td>
<td>3 (15)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>ADHD, n (%)</td>
<td>6 (30)</td>
<td>13 (65)</td>
<td>—</td>
<td>3.14\textsuperscript{b}</td>
</tr>
<tr>
<td>Disruptive behavior, n (%)</td>
<td>4 (20)</td>
<td>5 (25)</td>
<td>—</td>
<td>0.14</td>
</tr>
<tr>
<td>Antidepressants, n (%)</td>
<td>4 (20)</td>
<td>4 (20)</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Antipsychotics, n (%)</td>
<td>10 (50)</td>
<td>3 (15)</td>
<td>—</td>
<td>5.58\textsuperscript{a}</td>
</tr>
<tr>
<td>Mood stabilizers, n (%)</td>
<td>4 (20)</td>
<td>1 (5)</td>
<td>—</td>
<td>2.06</td>
</tr>
<tr>
<td>Stimulants, n (%)</td>
<td>7 (35)</td>
<td>11 (55)</td>
<td>—</td>
<td>1.62</td>
</tr>
<tr>
<td>Color labeling accuracy for all conditions, %, m (SD)</td>
<td>76 (20)</td>
<td>80 (19)</td>
<td>92 (8)</td>
<td>4.84\textsuperscript{a}</td>
</tr>
<tr>
<td>Reaction time during color labeling for conditions, ms, m (SD)</td>
<td>813 (154)</td>
<td>754 (205)</td>
<td>874 (258)</td>
<td>1.62</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Note: In post hoc comparisons, color labeling accuracy significantly differed between controls and patients BPSD patients (p = .004) and between controls and non-BPSD patients (p = .03), but not between BPSD and non-BPSD patients (p = .45). ADHD = attention-deficit/hyperactivity disorder; BPSD = bipolar spectrum disorder; SES = socioeconomic status.

\textsuperscript{b}Statistically significant differences between groups (p < .05).

\textsuperscript{a}Nonstatistically significant differences between groups (p < .10).
Symptom Assessment
Diagnostic assessments of the LAMS participants were performed annually by interviewing the caregiver and child using the K-SADS-PL-W. In addition, the mood modules of the K-SADS-PL-W interview were performed semianually. The number of months between the nearest K-SADS-PL-W assessment and scan date ranged between 3.9 months before the scan to 4.5 months after the scan. Based on this interview, participants were given a yes/no score for a lifetime history of each of 5 diagnostic categories (BPSD, depression, anxiety, ADHD, and disruptive behavior disorders [i.e., conduct disorder or oppositional defiant disorder]). No participant was given a lifetime history diagnosis of a substance use disorder (Table 1). BPSD and non-BPSD participants were matched on all diagnoses other than BPSD and on the demographic variables mentioned above.

Paradigm
Participants completed a 13-minute emotional dynamic faces task during fMRI scanning. Participants were asked to use 1 of 3 fingers to press a button indicating the color of a semitransparent foreground color flash (orange, blue, or yellow) that appeared during the mid 200 to 650 milliseconds of a 1-second presentation of a dynamically changing background face (neutral to emotional) (Figure 1). These emotional faces occupied the majority of the screen (approximately 23 visual degrees) and were interleaved with a central crosshair that was jittered to remain on the screen for 2 to 3 seconds. Faces from the NimStim stimulus set were morphed in 5% increments, from neutral (0% emotion) to 100% emotion for 4 emotions: happy, sad, angry, and fearful. Participants were told to ignore the face, as it was not relevant to their task. Morphed faces were made into 1-second movies progressing from 0% to 100% emotion. In control trials, movies comprised a simple shape (dark oval) superimposed on a light-gray oval, which was subsequently morphed into a larger shape, approximating the movement shown by the morphed faces. There were 3 blocks for each of the above 4 types of emotion trial, with 12 stimuli per block (42 second blocks), and 12 control blocks, with 6 stimuli per block (21 second blocks). Control blocks were half the duration of emotional blocks to minimize the time difference between baseline and conditions for individual emotion contrasts and to lessen the overall time burden on young participants. Emotional blocks were presented in a pseudorandomized order so that no 2 blocks of any condition were presented sequentially. Control blocks were interleaved with emotional blocks. The task was slightly modified from previously published adult versions for use with children. A graphical reminder of which color was assigned to each button was visible on the right side of the screen for the duration of the experiment.

Data Acquisition
At the Pittsburgh site, neuroimaging data were collected on a 3.0 Tesla Siemens Trio (Munich, Germany), at the Cleveland site neuroimaging data were collected with a 3.0 Tesla Siemens Verio, and at the Cincinnati site neuroimaging data were collected with a Philips Achieva 3.0 Tesla X-series scanner (Eindhoven, The Netherlands). Structural 3D axial MPRAGE/PAR-REC images were acquired in the same session (repetition time [TR]/time to echo [TE] = 2200/3.29 ms, flip angle 9/14, field of view [FOV]: 256×192 mm², slice thickness 1 mm, matrix: 256×256, 192 continuous slices). Blood-oxygen-level-dependent (BOLD) images were then acquired with a gradient echo-planar imaging (EPI) sequence during approximately 13 minutes (386 successive brain volumes) covering 39 axial slices (3.2 mm thick, TR/TE = 2000/28 ms, FOV = 205×205 mm², matrix = 64×64; flip angle 90°). Imaging parameters were adjusted identically on all 3 scanners.

FIGURE 1 Dynamic faces task. Note: Graphic representation of a single happy trial of our emotional dynamic faces task. Over a 1-second duration, the face changed from neutral (0% emotion) to a happy, sad, angry, or fearful face (100% emotion). Participants were asked to identify the color flash presented in mid dynamic change.
Previous studies\textsuperscript{35,36} indicate the feasibility of combining structural and functional neuroimaging data from different sites, but emphasize the necessity of measuring, and controlling for intersite differences in scanner signal-to-noise ratio (SNR). To ensure intrasite reliability among scanners at our 3 study sites and to allow combination of neuroimaging data across sites, we used the recommended standards published by the Biomedical Informatics Research Network (BIRN; http://www.nbirn.net) for data acquisition and information sharing by measuring scanner SNR monthly with BIRN-recommended phantoms at each site. In addition, we explored the effects of site and monthly SNR in our final analyses (see below).

Data Analysis

Data were preprocessed and analyzed using BrainVoyager QX 2.4 (Brain Innovation, Maastricht, The Netherlands). Preprocessing included slice time correction (cubic spline interpolation), alignment of slice (cubic spline interpolation to the first nondiscarded scan time), 3-dimensional motion correction (trilinear interpolation), spatial smoothing (6-mm Gaussian kernel), linear trend removal, and temporal high-pass filtering (fast-Fourier transform based with a cutoff of 3 cycles per time course). The functional data sets were coregistered to the Talairach-transformed\textsuperscript{37} T1-weighted anatomical image series to create a 4-dimensional data representation.

Z-transformed participant movement was entered as a covariate of no interest at the individual participant level. No participants included in this analysis moved more than 3 mm from their starting head position, either from movement spikes or from slow drift, during the course of data collection. Groups did not significantly differ in mean frame displacement (F\textsubscript{2,57} = 0.49, p = .61).

Although we were primarily interested in the face processing system, our primary analysis strategy used a whole-brain, conservative approach to allow for unexpected findings in other brain regions. A secondary analysis restricted our regions of interest to those of the face processing system (amygdala, FG, STS), using a more liberal threshold. For both analyses, a multi-participant statistical analysis was performed by multiple linear regression of the time course of the BOLD response in each voxel across the whole brain. Regressors were generated to represent the design matrix of the experiment, and a general linear model was computed to fit these regressors to each participant’s z-normalized volume time courses. Model predictors were defined by convolving an ideal boxcar response with a γ-function model of the hemodynamic response.\textsuperscript{38} Boxcar values were equal to 1 during the emotion morph blocks and 0 during shape morph blocks.

First, we computed a 3 (group: BPSD, non-BPSD, control) × 4 (condition: happy, sad, anger, fear) analysis of variance (ANOVA), with shapes as the baseline condition, to examine the main effects of condition and group and the group × condition interaction across the whole brain. Activation maps were visualized on a Talairach-transformed template brain, and displayed at a resolution of 1 mm\textsuperscript{3}, and all p values for all analyses were subjected to a whole-brain threshold of p < .001 and a 20 3 mm\textsuperscript{3} voxel extent, which is slightly more conservative than the p < .005 and a 10 voxel extent threshold that has been proposed to balance type I and type II errors.\textsuperscript{39} β values were extracted from regions displaying the effects of interest and plotted to examine differences between conditions and groups. Finally, a replication of the analysis described above was conducted using a region of interest mask of the face-processing system (bilateral amygdala, FG, and STS). Maps for this analysis were subjected to a more liberal threshold of p < .001, with no voxel extent threshold.

To examine the effects of behavioral, demographic, and clinical variables on our neuroimaging effects, we computed the contribution of these variables to our findings using SPSS version 20 (IBM Software). Across all groups, relationships were examined between accuracy and reaction time during color labeling in emotion blocks, as well as demographic variables, and β values for extracted activation differences from baseline. In both patient groups, relationships were examined between comorbid diagnoses and medication usage and these β values. A problem for all neuroimaging studies of BPSD is the potential confounding effect of psychotropic medication, as it is difficult to recruit medication-free participants into such studies.\textsuperscript{40} Thus, variables representing the taking versus not taking of each psychotropic medication class (antipsychotic, antidepressant, mood stabilizer, and stimulant) were examined. All analyses were computed by univariate analysis of covariance (ANCOVA) with extracted BOLD signal as the dependent variable, group (BPSD, non-BPSD, control) as the independent variable, and the variable of interest as a covariate.

RESULTS

Supplement 1 and Figures S1 and S2, available online, provide a discussion of data combination across multiple sites.

Task Performance

Color labeling accuracy and reaction times for emotion blocks were calculated based on individual participant task performance for each of the 5 conditions (fearful, angry, sad, happy, and shapes). Overall, color labeling accuracy was 76%, 80%, and 92% for BPSD, non-BPSD, and control, respectively. We computed a repeated-measures ANOVA for accuracy and reaction time with 3 groups (BPSD, non-BPSD, control) × 5
conditions (happy, sad, anger, fear, shapes). For task accuracy, we found a main effect of group ($F_{2,56} = 4.68, p = .013$) and a main effect of condition ($F_{4,224} = 4.35, p = .002$), but no group × condition interaction. Post hoc tests of group differences, with a least significant difference (LSD) correction, revealed decreased accuracy of both the BPSD group (mean difference = 0.16, $p = .005$) and non-BPSD group (mean difference = 0.12, $p = .03$) compared to the control group, but the accuracy difference between the 2 patient groups was not significant. Pairwise comparisons for each condition, with an LSD correction, revealed decreased accuracy for the happy condition compared to the sad (mean difference = -0.03, $p = .01$) and angry (mean difference = -0.04, $p = .001$) conditions and an increased accuracy for the anger condition compared to the fear condition (mean difference = -0.03, $p = .01$), but not between any other conditions. Although there was no significant group × condition interaction for accuracy, we note that scores differed significantly between the BPSD and control groups for all conditions (smallest mean difference = 0.06, all $p \geq .28$) (Figure 2).

For reaction time, we found a main effect of condition ($F_{4,224} = 4.892, p = .001$), but no main effect for group or a group × condition interaction. Pairwise comparisons for each condition, with an LSD correction, revealed increased reaction time for the sad condition compared to the fearful (mean difference = 40.42, $p = .001$) and shapes (mean difference = 43.56, $p < .001$) conditions and an increased reaction time for the anger, compared to the shape conditions (mean difference = 25.92, $p = .005$), but not between any other conditions. Results did not implicate a speed–accuracy trade-off (i.e., significant positive correlation between speed and accuracy) in the BPSD and control groups. Overall speed and accuracy were positively correlated for the non-BPSD group ($r_{18} = 0.55, p = .012$).

Neural Activity

The 3 (group: BPSD, non-BPSD, control) × 4 (condition: happy, sad, angry, fearful) ANOVA revealed a main effect of group in the left inferior frontal gyrus ($F_{2,57} = 7.82, p < .001$; see Table 2 for specific cluster size, statistical value, and location information). Contrasts of the average baseline-corrected neural activity $\beta$ values extracted from the whole of the active cluster, computed using SPSS v20, revealed significant mean differences between all groups. The control group had the highest change from baseline (mean = 0.16, SE = 0.03), followed by the non-BPSD group (mean = 0.07, SE = 0.03), followed by the BPSD group (mean = -0.05, SE = 0.03), whose activity decreased from baseline. There was also a main effect of condition ($F_{3,171} = 5.67, p < .001$), located in the bilateral superior frontal gyrus. Here, fear (mean = 0.10, SE = 0.04) and happy (mean = 0.05, SE = 0.03) were increased from baseline, though not significantly different from each other, whereas sad (mean = -0.07, SE = 0.04) and anger (mean = -0.17, SE = 0.04) decreased from baseline and were each significantly different from all other conditions.

Finally, a significant group × condition interaction revealed a single cluster of activation located in the left FG ($F_{6,171} = 3.95, p < .001$) (Figure 3). This cluster was located in the vicinity of the fusiform face area based on the work of Kanwisher et al. We created a region of interest based on a 3-mm sphere centered voxel $x = -35, y = -63, z = -10$, which was reported by Kanwisher et al. as the center of the left fusiform face area. $\beta$ values for baseline corrected neural activity levels were then extracted from this region of interest for the group × condition interaction to further explore the contributions of each individual condition and group to the interaction. The results of a 1-way ANOVA for each condition revealed significant group differences only for the angry condition ($F_{2,57} = 18.12, p < .001$) in this area of the FG. Post-hoc LSD contrasts revealed that the BPSD group had significantly lower activity levels within the left fronto-temporal cortex. This area is involved in the processing of facial expressions and has been implicated in emotion regulation and social cognition.
TABLE 2  Neural Activity Significant Cluster Information

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Peak Voxel</th>
<th>Average Stat Value</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole-brain analysis&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effect, group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left IFG</td>
<td>47</td>
<td>−22, 10, −15</td>
<td>582</td>
<td>9.86</td>
<td>.0003</td>
</tr>
<tr>
<td>Right SFG</td>
<td>9</td>
<td>32, 34, 30</td>
<td>926</td>
<td>6.80</td>
<td>.0004</td>
</tr>
<tr>
<td>Right SFG</td>
<td>10</td>
<td>41, 46, 18</td>
<td>2,367</td>
<td>6.49</td>
<td>.0004</td>
</tr>
<tr>
<td>Left SFG</td>
<td>10</td>
<td>−37, 43, 15</td>
<td>4,334</td>
<td>7.42</td>
<td>.0002</td>
</tr>
<tr>
<td>Group × condition interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left FG</td>
<td>37</td>
<td>−40, −65, −15</td>
<td>685</td>
<td>4.63</td>
<td>.0003</td>
</tr>
<tr>
<td>Face System ROI Mask&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effect, group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left FG</td>
<td>37</td>
<td>−34, −56, −13</td>
<td>347</td>
<td>8.67</td>
<td>.0006</td>
</tr>
<tr>
<td>Main effect, condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right STG</td>
<td>41</td>
<td>44, −38, 9</td>
<td>408</td>
<td>6.68</td>
<td>.0004</td>
</tr>
<tr>
<td>Group × condition interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left FG</td>
<td>37</td>
<td>−40, −65, −15</td>
<td>285</td>
<td>4.47</td>
<td>.0004</td>
</tr>
</tbody>
</table>

Note: BA = Brodmann area; IFG = inferior frontal gyrus; FG = fusiform gyrus; ROI = region of interest; SFG = superior frontal gyrus; Stat = statistical; STG = superior temporal gyrus.
<sup>a</sup>Whole map statistical threshold set at p < .001, 20-voxel extent.
<sup>b</sup>Whole map statistical threshold set at p < .001, no voxel-wise correction.

FG region compared to the controls (p < .001) and the non-BPSD group (p < .001) (Figure 3). There were nonsignificant trends for the 1-way ANOVAs for the sad (F<sub>2,57</sub> = 2.99, p = .058) and fear (F<sub>2,57</sub> = 2.80, p = .069) conditions.

To increase our power to find the effects of interest solely within the face processing system, we computed the above analysis using a bilateral mask of the amygdala, FG, and STS with a liberal overall threshold of p < .01, uncorrected. This ANOVA revealed a main effect of group in the left FG (F<sub>6,171</sub> ≥ 7.82, p < .001) (Table 2 provides specific cluster size and location information). Contrasts of baseline corrected neural activity β values extracted from this region revealed significant mean differences between the BPSD and both other groups but not between the control and non-BPSD group. The non-BPSD group had the highest change from baseline (mean = 0.36, SE = 0.06), followed by the control group (mean = 0.30, SE = 0.06), followed by the BPSD group (mean = −0.01, SE = 0.06). There was also a main effect of condition (F<sub>3,171</sub> ≥ 5.67, p < .001), located in the right STS. Here, angry (mean = 0.25, SE = 0.04), fearful (mean = 0.24, SE = 0.05), and sad (mean = 0.19, SE = 0.04) did not significantly differ from each other but were all 3 higher than happy (mean = 0.08, SE = 0.04). Finally, a significant group × condition interaction revealed a cluster of activation located in the left FG (F<sub>6,171</sub> ≥ 3.95, p < .001), overlapping with the interaction effect found at the whole-brain level. No effect was found in the right or left amygdala.

Relationships Among Demographic Variables, Task Performance, and Clinical Variables
To examine the potential effects of demographic variables, task performance, and clinical variables on the between-group differences found within the FG, the baseline corrected extracted activation β values, taken from the region of the FG in which the group × condition interaction, was significant (Figure 2), were used as the dependent variable in a series univariate (ANCOVAs). Each ANCOVA model had 2 independent variables, diagnostic group and the covariate of interest. Here we used the FG activation for only the anger condition because it was noted to have the largest difference between groups. This focused analysis also allowed us to reduce the likelihood of committing a type 1 error.

The demographic variables, age, sex, SES, and IQ did not have an independent effect on FG activation to anger (largest F<sub>1,56</sub> = 2.13, p = .15). In each model, the effect of diagnostic group remained significant even including the covariate in the model (all F<sub>2,56</sub> ≥ 15.29, all p < .001). For behavioral variables, task accuracy (F<sub>1,56</sub> = 0.68, p = .41) and reaction time (F<sub>1,56</sub> = 1.96, p = .17) for the anger condition did not have independent effects on FG activation. The effect of diagnostic group remained significant when controlling for task accuracy (F<sub>2,56</sub> = 16.85, p < .001) and reaction time (F<sub>2,56</sub> = 18.54, p < .001).

For clinical variables, we examined only the BPSD and non-BPSD groups. Lifetime history of comorbid anxiety disorders, ADHD, or disruptive behavior disorders did not have an...
independent effect on anger FG activation (all $F_{1,37} \leq 1.57$, all $p \geq .22$), and the effect of diagnostic group remained significant in each model (all $F_{1,37} \geq 31.00$, all $p < .001$). Finally, when examining medication effects, use of neither antidepressants, antipsychotics, mood stabilizers, nor stimulants (taking vs. not taking) had an independent effect on anger FG activation (all $F_{1,37} \leq 1.55$, all $p \geq .22$), and the effect of diagnostic group remained significant in each case (all $F_{1,37} \geq 27.48$, all $p < .001$).

Data Combination From Multiple Scan Sites
As described earlier, scanner SNR was collected monthly with a BIRN phantom and linked to each participant’s scan data. Following the same univariate ANCOVA procedure utilized to examine the effects of demographic, task performance, and clinical variables, we examined the effect of SNR on average FG activation across emotions. In this model, diagnostic group, scanner SNR, and scan site were entered as independent variables. Neither SNR nor scan site had an independent effect on average FG activation (all $F_{1,56} \leq 0.647$, all $p \geq .43$), and the effect of diagnostic group on average FG activation remained significant (all $F_{2,56} \geq 4.99$, all $p \leq .01$). For additional analyses relating to multisite data compilation, see supplemental materials.

DISCUSSION
Using a dynamic face processing task, we found evidence for decreased activation of the left FG in youth with BPSD relative to both a clinical group matched for demographic characteristics and comorbid psychopathology and a demographically matched healthy control group. This effect was present in all emotion conditions, but was most apparent for the anger condition.

Examining all areas of the brain, we found that it was only the FG (“fusiform face area”) region that showed significant interaction between groups and conditions. This region is widely noted as a region involved not only in face processing but in key components of social communication such as eye contact, facial identity recognition, and mouth movement. Although this region has rarely been studied directly in pediatric BPSD, inferences can be drawn from widely replicated results of hypoactivation and social impairment in the autism literature. Dalton et al., for example, found that visual fixation upon the eyes of a face was positively correlated with FG activation, whereas Kleinhans et al. found that amygdala to FG functional connectivity was negatively correlated with social impairments as diagnosed through clinical interview. Future studies may therefore show that similar impairments in social functioning may be linked to FG hypoactivation in pediatric bipolar...
disorder. Indeed, in the current study we found FG hypoactivation in our BPSD group, but not in our group matched for comorbid diagnosis, whose fusiform activation did not differ from that of a healthy control group. This study provides preliminary evidence that dysfunction in emotion processing occurs not only in emotional regions of the brain, which might be expected, given the affective challenges of this mood disorder, but also in basic visual processing of faces. In addition, we found a main effect of group in the inferior frontal gyrus, which was highest for the control group, possibly indicating increased regulation of affect in response to emotionally evocative stimuli.

Although participants were asked to perform a task that did not require explicit attention to emotional faces, we found that FG activation varied according to the type of emotional distracter in the BPSD group. Specifically, the differences in activation between the BPSD group and the non-BPSD group were largest for the angry face condition and smaller, albeit still statistically significant, for the sad face condition. Not only is anger an emotion that is often present in the clinical symptoms of pediatric BPSD,2,48 but previous studies have pointed to deficits in anger perception in BPSD. In a pediatric BPSD sample, Guyer et al.8 reported the most errors in identifying angry faces compared to other emotions, and Rich et al.5 reported that neutral faces were perceived as more “hostile” by youth with BPSD compared to control participants. Neuroimaging studies also have found increased activation in the amygdala and prefrontal cortex during passive viewing of angry faces in comparison to control participants.49 Our results indicate decreased visual processing of angry faces in participants with BPSD. This may underlie social impairments often seen in these youth, who may not effectively detect or may misinterpret the facial cues present in angry friends or family members.

Many previous studies indicate that there is an overall right greater than left laterality in FG face-selective activity.14,50-52; however, our results were localized to the left hemisphere. Our findings were mostly driven by the angry face condition. Previous research aimed at emotional brain asymmetry has indicated lateralized brain activity in the right hemisphere for withdrawal emotions (e.g., sadness) and lateralized brain activity in the left hemisphere for approach emotions (e.g., anger).53 For example, in a study of electroencephalography (EEG) asymmetry, Harmon-Jones54 found that in adults who self-reported high trait anger, greater relative left frontal cortical activity was observed in response to anger-invoking pictures, but not to other emotional pictures. Although measures of trait anger were not collected in this sample, the FG asymmetry results observed may be related to differences in emotional perception of this approach emotion between groups.

Although the findings of this study are promising, some limitations must be noted. Nearly all clinical participants in this study were medicated, and, although covariate analyses indicate that medication use did not affect the specific findings of our study, effects of medication on face perception systems are not well understood. However, previous studies have argued that medication is more likely to have a normalizing effect in clinical neuroimaging studies,40 so we would not expect medication to introduce spurious differences. Second, limitations of the current work may be related to the fMRI study design. In contrast to previous studies of emotional face processing in pediatric bipolar disorder,5,55 our analyses did not find differences in amygdala functioning when viewing emotional faces. This could be related to the task-irrelevant emotional nature of our design (requiring no emotional judgment). This could also be related to the lack of task conditions requiring differential focus on the emotional faces. Indeed, other research that has employed this technique has found modulation of the amygdala due to varying attentional demands placed on the stimulus56. Finally, the block design of our study, combined with the differing accuracy between subject groups on emotion and baseline conditions, can be seen as a limitation to our interpretation of study findings. The block design of the study did not allow us to analyze only correct trials, which were significantly fewer in the BPSD and non-BPSD groups. As previous studies have emphasized,46,57 one must investigate the nature of the behavior involved in the task to disambiguate an effect of interest from other neural processes related to error rate (e.g., conflict monitoring,58 executive function59). In addition, the possible speed–accuracy trade-off observed in the non-BPSD group points to possible behavioral differences among groups that is not easily noted with a block design. Thus, we encourage our
results to be interpreted as an early-phase investigation requiring further replication using an event-related task design. Future studies should attempt the analysis of only correct trials when performance differences arise, which is a confound that is often unavoidable when investigating young patient groups. We note, however, that in our study, behavioral differences were only found between the patient and control groups and did not exist when comparing both patient groups. Thus any possible confound caused by behavioral performance cannot account for the observed differences between the BPSD and non-BPSD children with regard to FG activity to sad and angry faces. Further mitigating the concern that performance confounds may be responsible for any of the results that we observed, the non-BPSD and healthy control groups did not differ from each other with regard to FG activity to anger despite the fact that the non-BPSD children were significantly less accurate than the controls during this task condition.

Results of this study suggest that FG deficits in facial processing are specific to pediatric BPSD rather than a common functional brain abnormality observed in general child psychopathology. Thus, it is possible that this patient group could benefit from therapeutic training designed to enhance FG functioning to facilitate positive social interaction. Similar interventions have been attempted, with mixed success, in studies of autism spectrum disorders through eye contact and facial expression recognition training, as well as in psychosocial treatments for pediatric BPSD. Before such treatment development work can commence, however, future studies should consider the use of eye-tracking methodology during fMRI scanning to determine whether the observed effect is due to visual avoidance of emotional faces or is simply an effect of inefficiency in the brain’s face-processing circuitry in pediatric BPSD. In addition, studies enrolling young children at risk for BPSD (e.g., unaffected siblings) could examine longitudinally whether FG hypoaivation is a biological precursor of this disorder or the result of repeated difficulties in social interaction throughout the course of development.

REFERENCES


SUPPLEMENT 1

Data Combination From Multiple Scan Sites

We completed a whole-brain 3 (site: Pittsburgh, Cleveland, Cincinnati) x 4 (condition: happy, sad, angry, fearful) analysis of variance (ANOVA). Here, we examined the main effect of scan site to identify any area of the brain that may have been differentially affected by varying magnetic resonance imaging (MRI) scanners. Lowering the threshold for the whole-brain map to a liberal, uncorrected p < .05 threshold, to be able to detect any possible effect, many areas of the brain showed activation patterns related to scan site (F2,57 = 3.16, p < .05, uncorrected). However, none of these areas overlapped with the left fusiform region found for the group x condition interaction (Figure S1). There were no statistically significant areas of activation at the whole-brain corrected FDR(q) < .05 threshold.1 Only 3 clusters (minimum size, 300 voxels) were significantly activated at a p < .001 threshold (left middle temporal gyrus BA 21, left middle frontal gyrus BA 10, and left cerebellum).

Given the small sample sizes at each site, we did not have adequate power to test the site x group x condition interaction effect. However, to determine whether any 1 site was driving the observed results, we conducted a series of analyses in which each site was removed, 1 at a time, and the primary models described in the main text (3 [group: bipolar, nonbipolar, control] x 4 [condition: happy, sad, angry, fearful]) were rerun using our face-processing system region of interest mask (amygdala, fusiform gyrus [FG], superior temporal gyrus [STG]). We examined the group x condition interaction in the fusiform gyrus at a liberal, whole-map p < .05, uncorrected threshold to allow us to detect similar effects as the original models with less power. In each pairing of sites, we found a group x condition interaction in the left FG that overlapped with the left FG region found in the original analysis that included all 3 sites (Figure S2). For the Cleveland and Cincinnati pairing, the FG region was 791 1 mm x 1 mm x 1 voxels (peak: x = −40, y = −56, z = −16). For the Cleveland and Pittsburgh pairing, the FG region was 1,146 voxels (peak: x = −37, y = −60, z = −15). For the Cincinnati and Pittsburgh pairing, the FG region was 1,187 voxels (peak: x = −40, y = −65, z = −15).

SUPPLEMENTAL REFERENCES

FIGURE S1  Main effect of site, $F(2,57) = 3.16, p < .05$ (uncorrected). Note: Results of a statistically lenient ($p < .05$ uncorrected), $3$ (site: Pittsburgh, Cleveland, Cincinnati) $\times$ $5$ (condition: happy, sad, anger, fear, shapes) analysis of variance revealed a main effect of site (shown in red). Note that this effect was not present in the fusiform gyrus in which the primary results of the current study were localized (shown in blue). Furthermore, regression analyses revealed that signal-to-noise ratio (SNR), collected through phantom scanning, did not contribute to the effects observed in the fusiform gyrus (FG) area identified in blue.
FIGURE S2  2 (site) × 4 (condition) fusiform gyrus interactions. Note: Separate site × condition analyses of variance, each leaving out a single scan site, indicated that, in each case, results of a significant interaction in the fusiform gyrus (FG; pictured in red) overlapped with the overall site × condition interaction found in all sites combined (pictured in blue).

Cincinnati & Cleveland
\[ F(6,108)=2.18, \ p < .05 \]

Cincinnati & Pittsburgh
\[ F(6,99)=2.19, \ p < .05 \]

Cleveland & Pittsburgh
\[ F(6,126)=2.17, \ p < .05 \]

\( x=-34, \ y=-57, \ z=15 \)