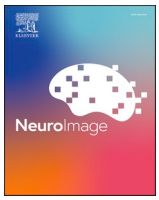
NeuroImage 327 (2026) 121738



Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg





Neural representations of emotional response inhibition reveal trait and state biomarkers in pediatric bipolar disorder

Jia Li [a], Rong Wang [a], Jianze Wu [a], Qian Xiao [d], Yuan Zhong [a,b,c,*]

[a] School of Psychology, Nanjing Normal University, Nanjing 210097, China
[b] Adolescent Education and Intelligence Support Lab of Nanjing Normal University, Laboratory of Philosophy and Social Sciences at Universities in Jiangsu Province, Nanjing 210097, China
[c] Huaian Hospital Affiliated to Nanjing Normal University, Huaian 223001, China
[d] Mental Health Centre of Xiangya Hospital, Central South University, Changsha 410008, China

ARTICLE INFO

Keywords:
Pediatric bipolar disorder(PBD)
Emotional response inhibition
Trait effects
State effects
Representational similarity analysis (RSA)
fMRI
Prefrontal cortex

ABSTRACT

Pediatric bipolar disorder (PBD) is characterized by disrupted cognitive control, particularly in response inhibition under emotional interference. However, the neural underpinnings of these deficits, particularly how these impairments vary across emotional valence and whether they reflect trait markers or state alterations, remain unclear. While traditional univariate fMRI analyses reveal broad activation differences, they lack sensitivity to fine-grained neural patterns. This study aims to examine the neural representations of emotional response inhibition in PBD under valence-dependent interference using representational similarity analysis(RSA). We included manic (n = 15) and euthymic (n = 18) PBD patients, along with matched healthy controls (n = 17). Participants completed an emotional Go/NoGo task with happy, sad, and neutral faces during fMRI. Six contrast conditions were modeled to assess trait- and state-related effects. Whole-brain searchlight RSA (8 mm radius) was used to identify regions showing group differences in neural representational patterns. Results showed that emotional response inhibition engaged distributed neural systems, with distinct patterns across valence conditions. Compared to controls, PBD patients exhibited trait-related representational differences during happy inhibition, sad inhibition, and sad-specific inhibition, involving regions such as the precentral gyrus, middle frontal gyrus, and inferior parietal lobule. Manic patients showed state-related reductions in neural representations during sad-specific inhibition within frontal areas compared to euthymic patients. These findings indicate that emotional response inhibition deficits in PBD arise from both trait- and state-dependent abnormalities in neural representations. The study highlights the value of multivariate fMRI in uncovering clinically relevant biomarkers and provides a novel framework for developing phase-specific interventions.

1. Introduction

Bipolar disorder (BD) is a severe mental illness characterized by recurrent episodes of depression and mania (Carvalho et al., 2020; Ferrari et al., 2016). Nearly 59 % of cases begin before age 20, a population referred to as pediatric bipolar disorder (PBD) (Van et al., 2019). Compared to adult-onset BD, PBD features more abrupt onset, frequent mood shifts, more mixed episodes, comorbidities, and suicide risk (Geller et al., 2000; Miller and Black, 2020). Deficits in cognitive control are considered a core endophenotype of BD (Bora et al., 2009; Stefanopoulou et al., 2009) and a risk factor for mood disorder onset and maintenance (Bortolato et al., 2015; Gotlib and Joormann, 2010). Cognitive control enables goal-directed behavior despite interference, relying on executive functions like inhibition and working memory. Notably, BD-related emotional dysregulation exacerbates emotional response inhibition deficits (Sadana et al., 2019).

Building on this background, prior work has generally suggested that BD is characterized by impaired response inhibition in emotional contexts; however, the literature remains heterogeneous. A central source of variability is that many studies have treated BD as a relatively homogeneous condition and have not simultaneously modeled illness phase and emotional valence as key moderators, thereby inflating between-study inconsistencies (Wessa et al., 2007; Wolkenstein et al., 2017; Förster et al., 2023; Pomarol-Clotet et al., 2015; Xiao et al., 2021). Overall, emotional inhibition paradigms have frequently reported heightened responses in emotion-generative regions (e.g., the amygdala

* Corresponding author at: School of Psychology, Nanjing Normal University, Nanjing 210097, China.
E-mail address: zhongyuan@njnu.edu.cn (Y. Zhong).

https://doi.org/10.1016/j.neuroimage.2026.121738
Received 20 September 2025; Received in revised form 17 January 2026; Accepted 19 January 2026
Available online 20 January 2026
1053-8119/© 2026 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

and broader limbic circuitry) alongside relatively reduced engagement of regulatory and cognitive control systems (e.g., prefrontal and frontoparietal regions), consistent with an imbalance between affective drive and top–down regulation(13,14). Nonetheless, other studies have observed hyperactivation in both affective and regulatory systems, suggesting that neural recruitment may differ as a function of clinical state and task demands (Wessa et al., 2007; Xiao et al., 2021). Phase-sensitive evidence further supports state dependence: depressive states are more strongly associated with limbic hyperreactivity, whereas manic states more prominently involve diminished prefrontal regulation (e.g., dorsolateral prefrontal cortex hypoactivation); at the circuit level, manic adults tend to show ventral circuit hyperactivity, while manic PBD may exhibit co-recruitment of ventral and dorsal systems, and euthymic PBD may show limbic predominance with attenuated dual-system engagement (Blumberg et al., 2003; Hummer et al., 2013; Passarotti et al., 2009; Pavuluri et al., 2008). Emotional valence also shapes inhibitory processing: adult BD is often characterized by a predominantly negative bias, whereas PBD appears to show inhibitory difficulties in both positive and negative contexts, implying a broader and more valence-sensitive pattern of impairment in youth (Wessa et al., 2007; Bauer et al., 2015; Seymour et al., 2015).

Importantly, these phase–valence distinctions may be especially consequential in PBD. Firstly, PBD occurs during a critical neurodevelopmental period when frontoparietal control networks and their coupling with limbic circuitry are still maturing (Arain et al., 2013; Gu et al., 2015), which increases the developmental sensitivity of emotional response inhibition mechanisms. Failure to stratify by phase and valence may more readily yield inconsistent findings. Secondly, previous resting-state fMRI studies in PBD have demonstrated state-specific alterations in higher-order circuits involved in emotion regulation (Lei et al., 2023; Wang et al., 2024). Thirdly, PBD may exhibit distinct neural mechanisms from adult BD, as evidenced by numerous studies (Wessa et al., 2007; Blumberg et al., 2003; Hummer et al., 2013; Passarotti et al., 2009; Pavuluri et al., 2008; Bauer et al., 2015; Seymour et al., 2015). Thus, it is essential to study PBD separately, rather than simply extrapolating conclusions from adult BD research. Lastly, from a clinical perspective, PBD is frequently associated with significant functional impairment and elevated long-term risk (Geller et al., 2000; Miller and Black, 2020). Therefore, distinguishing between relatively stable deficits and phase-linked abnormalities is crucial for early assessment, stratified intervention, and longitudinal management.

Prior neuroimaging studies of emotional response inhibition in BD have mainly used univariate activation analyses, which capture differences in response magnitude but obscure multivoxel pattern changes. As univariate measures compress distributed responses into a single value, they may miss disease-relevant effects when conditions show similar overall activation yet differ in spatial configuration, condition separability, or cross-network coordination. This limitation is especially important in PBD, which emerges during the maturation of frontoparietal control networks and their integration with limbic circuitry; thus, abnormalities may be expressed more as altered distributed representations and network coupling than as uniform regional increases or decreases. Given the stronger phase instability and broader valence sensitivity in PBD, a pattern-based approach is needed to test how illness phase and emotional valence jointly shape inhibitory processing. Representational similarity analysis (RSA) quantifies the similarity of multivoxel response patterns across conditions and can reveal fine-grained differences missed by univariate methods (Kriegeskorte, 2008; Haxby, 2012). RSA has also been informative in inhibition research (e.g., Tourette syndrome) (Rae et al., 2023). Moreover, basic emotions (e.g., happiness, sadness) show partly distinct neural representations that stabilize by late childhood (Camacho et al., 2023; Qin, 2023), supporting the examination of valence-related representational abnormalities in PBD. Accordingly, we use RSA to probe two forms of emotional response inhibition: traditional response inhibition under single-emotion interference, and emotion-specific inhibition, indexed by distinct suppression patterns that differentiate one emotion from others.

This study aims to examine the neural representations of emotional response inhibition in PBD under valence-dependent interference using RSA. Participants included individuals with manic-phase PBD, euthymic-phase PBD, and healthy controls, who performed an emotional Go/NoGo task during fMRI. To dissociate relatively stable disease-related abnormalities from phase-dependent alterations in a developmental population, we conceptually define a trait effect as neural abnormalities that are consistently present across illness phases in PBD, reflecting stable differences relative to healthy controls, and a state effect as abnormalities that vary as a function of illness phase, reflecting phase-specific alterations in representational organization or inhibitory processing between manic and euthymic states. Within this framework, we hypothesize that all PBD patients will exhibit shared inhibitory deficits under emotional interference, indicative of a cross-phase trait-related impairment, whereas finer-grained, valence-specific abnormalities in inhibitory representations will differ by illness phase, reflecting state-dependent effects. These abnormalities are expected to involve both cognitive regulatory regions and networks (e.g., the frontoparietal control network) and emotion-related regions (e.g., the limbic system), thereby providing representational-level evidence for the coexisting trait and state mechanisms underlying emotional response inhibition deficits in PBD.

2. Materials and methods

2.1. Participants

This cross-sectional study recruited adolescents with bipolar disorder meeting the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria from the Child and Adolescent Psychiatry Clinic of the Second Xiangya Hospital, Central South University, alongside demographically matched healthy controls (HCs) from local schools. All procedures were approved by the Ethics Committee of the Second Xiangya Hospital, Central South University(2013-S040) and complied with the Helsinki Declaration, with written informed consent obtained from participants and guardians. Participant characteristics are presented in Table 1.

PBD participants met the following inclusion criteria: 1) aged 10–18 years, 2) right-handed, 3) of Han Chinese ethnicity, 4) capable of cooperating with MRI procedures. DSM-5 diagnoses were established using a structured diagnostic interview (Section 2.2).

Based on disease status, PBD participants were further divided into two subgroups. To ensure subgroup classification reflected both syndromal status and symptom severity at the time of scanning, subgrouping was defined using DSM-5 episode criteria in combination with the Young Mania Rating Scale (YMRS) and the Mood and Feelings Questionnaire (MFQ). The manic episode subgroup ($n = 15$) met DSM-5 criteria for a (hypo)manic episode and was required to have a YMRS score > 26 and an MFQ score < 18 at the time of scanning. The euthymic subgroup ($n = 18$) was required to have maintained symptom remission for ≥ 1 month (defined as no longer meeting DSM-5 criteria for manic, hypomanic, depressive, or psychotic episodes) and to have YMRS score < 12 and MFQ score < 18 at the time of scanning.

Uniform exclusion criteria included: 1) full-scale IQ ≤ 80 on the Wechsler Intelligence Test, 2) comorbid psychiatric disorders (including autism spectrum disorder, schizophrenia spectrum disorders, Attention-Deficit/Hyperactivity Disorder, and Obsessive-Compulsive Disorder), 3) organic mental disorders, 4) history of substance abuse (within 2 months prior to enrollment), 5) pregnancy, 6) MRI contraindications (e.g., metal implants, claustrophobia).

In addition to demographic matching, healthy controls met the following criteria: 1) no personal history of psychiatric disorders, 2) no family history of psychiatric disorders in first-degree relatives, 3) no major neurological diseases or trauma. Their mental health status was confirmed via Structured Clinical Interview for DSM (SCID).

Table 1
Clinical and demographic characteristics of subjects with PBD and healthy individuals.

Characteristic	Manic PBD (n = 15)	Euthymic PBD (n = 18)	HCs (n = 17)	$F/t/\chi^2$	p	φ/η_p^2/Cohen d
Gender (male/female)	5/10	9/9	7/10	0.940	0.625	0.137
Age (years)	14.93 ± 1.94	15.17 ± 1.69	14.24 ± 1.56	1.353	0.268	0.054
Education (years)	8.33 ± 1.63	8.22 ± 1.77	7.59 ± 1.28	1.080	0.348	0.044
IQ	106.20 ± 11.56	109.56 ± 10.11	106.71 ± 6.33	0.618	0.543	0.026
YMRS scores	37.13 ± 4.65	10.17 ± 3.99	7.88 ± 3.30	261.922	<0.001	0.918
HAMD scores	5.93 ± 2.89	6.33 ± 1.91	5.24 ± 2.28	0.962	0.389	0.039
Onset age (years)	14.20 ± 2.04	13.76 ± 1.52		0.689	0.496	0.241
Illness duration (months)	15.13 ± 11.95	22.72 ± 18.89		−1.346	0.188	−0.471
Number of episodes	2.87 ± 1.36	5.00 ± 7.90		−1.030	0.311	−0.360
Handedness (left/right)	0/15	0/18	0/17			

A total of 60 participants were recruited for this study, consisting of 20 individuals each from three groups: manic-phase PBD, euthymic-phase PBD, and healthy controls. After rigorous imaging quality control (QC), which included the exclusion of participants based on criteria such as severe head motion (FD > 0.5 mm), significant artifacts, and missing data, the final sample included 15 manic-phase PBD participants, 18 euthymic-phase PBD participants, and 17 healthy controls.

2.2. Clinical assessments

All participants were assessed using the Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime Version (K-SADS-PL) (Kaufman et al., 1997) by two experienced psychiatrists (inter-rater reliability Kendall's W = 0.9). Independent interviews were conducted with both adolescents and their guardians to improve diagnostic accuracy and to capture current and lifetime psychopathology. Cognitive functioning was evaluated using the Wechsler Abbreviated Scale of Intelligence (WASI), with IQ > 80 required for inclusion. Symptom severity was assessed using the YMRS (Young et al., 1978) for manic symptoms and the MFQ (Wood et al., 1995) for depressive symptoms, while handedness was confirmed using the Edinburgh Handedness Inventory. The YMRS and MFQ were administered in close temporal proximity to MRI acquisition to ensure consistency between clinical state and neuroimaging measures. Comprehensive clinical data, including illness characteristics (age of onset, illness duration, episode frequency), and comorbidities, were then collected on the day of fMRI scanning to maintain temporal consistency with neuroimaging data. Detailed information can be found in Table 1.

2.3. Go/NoGo task paradigm

We employed a modified emotional Go/NoGo paradigm (adapted from Wessa et al.)(11) using a block design to assess emotional response inhibition (Fig. 1). The task comprised 8 experimental blocks (6 emotional conditions + 2 control conditions), each lasting 36 s and preceded by a 4 s instruction screen. Emotional stimuli consisted of standardized facial expressions depicting happiness, sadness, and neutrality, arranged in a full-factorial design to create 6 emotional pairing conditions (e.g., happy-go / neutral-nogo). Control conditions used neutral male/female faces (male-go / female-nogo or vice versa). Each block presented Go and NoGo stimuli in a 2:1 ratio, with 16 Go stimuli (0.5 s presentation) and 8 NoGo stimuli presented pseudo-randomly (ISI = 1 s). This ratio was chosen to establish a dominant response tendency, enhancing the sensitivity of inhibition control measurement and amplifying the inhibition difference under emotional conditions while reducing ceiling effects. The design followed classic emotional Go/NoGo paradigms (Wessa et al., 2007). A 36 s rest baseline (fixation cross) followed each block. Behavioral data (reaction time, accuracy) were recorded using E-Prime 2.0, synchronized with fMRI

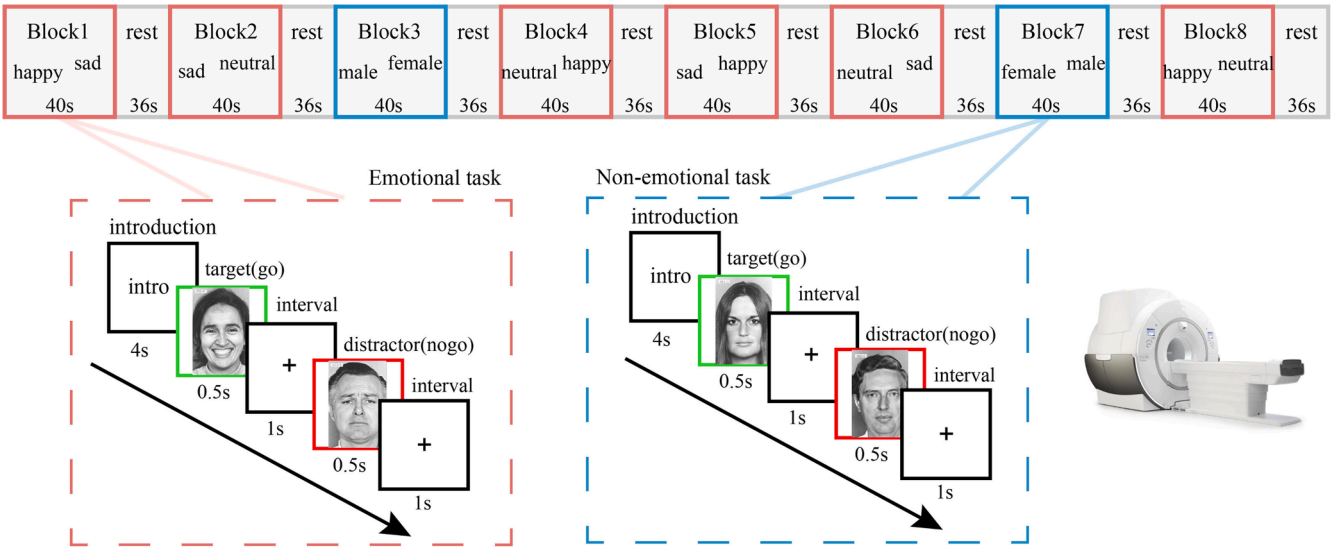


Fig. 1. Experimental procedure of the Emotional Go/NoGo task. The paradigm alternated between emotion-modulated (red outline) and neutral (blue outline) conditions across eight blocks (Block1-Block8). Each block commenced with a 40-second instructional introduction, followed by a 36-second active task period. Trial sequencing within task periods adhered to the following structure: a 0.5-second fixation cross ("+") preceded a 1-second cue presentation, which then transitioned to either (a) target stimuli (Go trials; 1-second duration) or (b) distractor stimuli (NoGo trials; 4-second duration). Emotional blocks employed facial stimuli with varied affective valence (happy/sad/neutral) and gender characteristics (female/male), whereas neutral blocks utilized matched stimuli devoid of emotional salience. Inter-block intervals incorporated standardized rest periods to allow for hemodynamic response normalization in fMRI protocols.

acquisition.

2.4. fMRI data acquisition

All imaging data were acquired using a 3.0-T Siemens Trio scanner. High-resolution T1-weighted structural images were obtained with a 3D magnetization-prepared rapid gradient-echo (MPRAGE) sequence (TR = 2300 ms, TE = 2.03 ms, TI = 900 ms, flip angle = 9°, FOV = 256 × 256 mm^2, matrix = 256 × 256, 1 mm slice thickness with no gap). Functional images were collected using a T2*-weighted gradient-echo planar imaging (EPI) sequence (TR = 2000 ms, TE = 30 ms, flip angle = 90°, FOV = 240 × 240 mm^2, matrix = 64 × 64, 30 axial slices with 4 mm thickness and 0.4 mm gap), with the first 5 vol discarded for magnetic field stabilization.

Quality control measures included: 1) foam padding to minimize head motion, 2) noise-reducing earplugs for hearing protection, and 3) stimulus presentation via E-Prime 3.0 (Psychology Software Tools, USA) projected onto a screen viewable through a head coil mirror.

2.5. fMRI data preprocessing

The neuroimaging data were preprocessed using fMRIPrep 23.1.3, a robust pipeline built on Nipype 1.8.6 (Gorgolewski et al., 2011).

2.5.1. Anatomical data preprocessing

The T1w image underwent INU correction using N4BiasFieldCorrection (ANTs 2.5.0), followed by skull-stripping via antsBrainExtraction.sh (OASIS30ANTs template). Brain tissue segmentation (CSF/WM/GM) was performed using FSL's fast. Surface reconstruction used FreeSurfer's recon-all, with mask refinement via Mindboggle. Spatial normalization to MNI152NLin2009cAsym was conducted using antsRegistration (ANTs 2.5.0) with the ICBM 152 2009c template (TemplateFlow 23.1.0).

2.5.2. Functional data preprocessing

Each subject's BOLD data underwent standard preprocessing including motion correction, co-registration to structural images, and nuisance regression. Head motion parameters were estimated using FSL's mcflirt, followed by boundary-based registration to T1w images via FreeSurfer's bbregister. Confound regressors were computed including framewise displacement, DVARS, global signals, and CompCor components to account for physiological noise. Volumes exceeding motion thresholds (framewise displacement (FD) > 0.5 mm or DVARS (root mean square of the temporal derivative of the BOLD signal across voxels) > 1.5) were flagged as outliers. The preprocessed data were then resampled to standard surface (fsaverage) and volumetric spaces using FreeSurfer and ANTs tools respectively, incorporating all transformations in a single interpolation step.

2.6. Single-trial response estimation

Single-trial neural responses were estimated using the Least Squares Separate (LSS) approach (Mumford et al., 2012). For each trial, we constructed an independent General Linear Model (GLM) with the target trial as a separate regressor while grouping all other trials as another regressor. The model incorporated six motion parameters and FD as covariates, with additional detector regressors added for volumes where FD exceeded 0.3 mm. Stimulus onset was modeled as a delta function (duration = 500 ms) convolved with a canonical double-gamma hemodynamic response function. This procedure generated a beta statistical map for each trial, which was subsequently used for representational similarity analysis.

2.7. Representational similarity analysis

To investigate the neural representational mechanisms underlying emotional response inhibition in PBD, we conducted representational similarity analysis on our emotional Go/NoGo task data. Using a whole-brain searchlight RSA approach, we generated condition-specific neural representational maps for each participant, capturing the distributed patterns of brain activity associated with different emotional inhibition conditions. Fig. 2 shows the schematic diagram of the Searchlight-RSA approach.

2.7.1. Construction of model RDMs

To examine valence-specific representational patterns of emotional response inhibition, we constructed theoretical representational dissimilarity matrices (RDMs) (Kriegeskorte, 2008) based on our emotional Go/NoGo task design, comprising six distinct experimental conditions. Specifically, these experimental conditions can be grouped into two categories. The first category captures traditional emotion-related response inhibition; the second category focuses on contrasting response inhibition under different valence interferences, a key advantage of multivariate approaches. The correspondence between representational matrices and experimental conditions is as follows:

1) RDM1 (48 × 48): happy inhibition (happy-go vs. happy-nogo) - identifying regions specifically encoding inhibition of positive emotion, 2) RDM2 (48 × 48): sad inhibition (sad-go vs. sad-nogo) - marking regions specific to negative emotional inhibition, 3) RDM3 (48 × 48): emotional inhibition (happy-go/sad-go vs. happy-nogo/sad-nogo) - detecting emotion-general inhibitory regions, 4) RDM4 (48 × 48): happy-specific inhibition (happy-nogo vs. sad-nogo/neutral-nogo) - localizing regions exclusively sensitive to happy emotion inhibition, 5) RDM5 (48 × 48): sad-specific inhibition (sad-nogo vs. happy-nogo/neutral-nogo) - screening regions specific to sad emotion inhibition, 6) RDM6 (32 × 32): valence-specific inhibitory (happy-nogo vs. sad-nogo) - revealing neural patterns distinguishing emotional valence during inhibition.

Each RDM is a symmetric square matrix with zeros along the diagonal, while the off-diagonal entries represent trial-by-trial dissimilarities under specific emotion suppression conditions. All RDMs are binary matrices (0 = matched conditions, 1 = mismatched conditions), generated from a subset of 192 trials using a MATLAB script while strictly adhering to the block-wise pseudorandomized order of the experimental design. After construction, each matrix was verified for symmetry and zero-diagonal properties. The matrix dimensions vary depending on the number of trials in the compared conditions (e.g., 48 trials for within-emotion comparisons, 32 trials for between-emotion comparisons). The representational dissimilarity matrices are visualized in Fig. 2.

2.7.2. Searchlight analysis

This study employed whole-brain searchlight-RSA to investigate the neural representation patterns of emotion response inhibition. For each participant in standard MNI space, we constructed spherical searchlights with a 4-voxel radius (≈ 8 mm) centered on each voxel within the whole-brain gray matter mask (Kriegeskorte et al., 2006). The searchlight neighborhoods were generated using CoSMo MVPA toolbox's cosmo_spherical_neighborhood function, with each searchlight serving as a local region of interest (ROI) for subsequent analyses. The searchlight radius was selected to balance spatial resolution and signal-to-noise ratio (Oosterhof et al., 2016).

For each of the six theoretical model RDMs, we extracted condition-specific beta values within each searchlight to construct neural RDMs. Specifically, following each theoretical RDM's trial-pairing requirements (e.g., RDM1: happy Go vs. happy NoGo trials; RDM2: sad Go vs. sad NoGo trials), we computed pairwise Pearson correlation coefficients between corresponding trials and generated neural RDMs with elements defined as $1 - r$ (where r is the correlation coefficient). Each theoretical RDM had its specific trial combinations (e.g., RDM1: 48 trials, RDM6: 32 trials), ensuring precise dimensional matching between neural and theoretical RDMs.

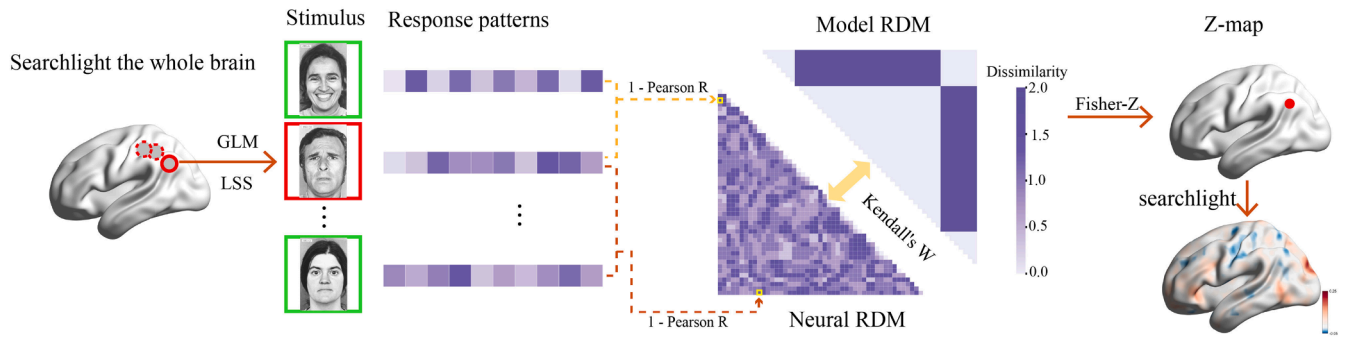


Fig. 2. Schematic diagram of whole-brain Searchlight-RSA. For each theoretical model RDM, we computed neural RDMs within spherical searchlights by calculating 1 - Pearson's r between trial-pairs, then correlated these with theoretical RDMs using Spearman's rank correlation. The resulting representational strength maps (Fisher z-transformed) identified brain regions where neural patterns aligned with each inhibition model.

We compared neural RDMs with theoretical RDMs using Spearman rank correlation. To reduce false positives, only upper triangular matrix data were used for correlation computation. The resulting correlation values were Fisher z-transformed and assigned to the central voxel of each ROI. This process was iterated across the entire brain, generating six whole-brain z-maps (representational strength maps) per participant corresponding to each theoretical model. This analytical pipeline systematically evaluated how well different brain regions' neural representation patterns aligned with theoretical predictions.

2.8. Statistical analysis

The searchlight-RSA analysis generated individual whole-brain z-maps representing neural pattern strength for each of the six emotional response inhibition conditions. These z-maps were subsequently used for group-level analyses. Given the relatively small sample size and non-normal distribution of our data, we performed nonparametric single-sample *t*-tests. A null distribution was generated through 5000 permutation tests, and statistical inference was performed using threshold-free cluster enhancement (TFCE) with whole-brain (analysis-mask) family-wise error (FWE) correction at $p < 0.05$ (i.e., controlling the maximal TFCE statistic across the search volume), without applying an additional arbitrary cluster-forming voxel-wise threshold. This approach eliminates the subjectivity inherent in traditional cluster-based threshold selection while maintaining sensitivity to detect both focal and distributed neural patterns. Brain regions surviving this rigorous correction threshold were considered to represent valid neural signatures of emotion-specific response inhibition functions. We selected TFCE because it provides robust FWE control within a permutation framework while avoiding dependence on an initial cluster-forming threshold and retaining sensitivity to both focal high-amplitude and spatially extended moderate-amplitude effects, compared with standard voxel-wise or conventional cluster-based FWE approaches.

To examine group differences among manic PBD, euthymic PBD, and healthy control participants, we conducted one-way ANOVAs using the PALM toolbox, implementing TFCE-based permutation testing. For each of the six emotional inhibition conditions, we created a union mask comprising all significant clusters from the initial searchlight analysis to constrain subsequent analyses to task-relevant regions. Permutation testing (5000 iterations) was performed under the null hypothesis of no between-group differences, with statistical significance determined by TFCE-based FWE-corrected inference at $p < 0.05$.

For brain regions showing significant ANOVA results, we conducted post-hoc pairwise comparisons (manic PBD vs. HCs, euthymic PBD vs. HCs, manic PBD vs. Euthymic PBD) within the same PALM framework using independent samples *t*-tests with 5000 permutations and TFCE correction. Multiple comparisons were controlled via TFCE-based FWE-corrected inference ($p < 0.05$).

Demographic characteristics and behavioral performance were analyzed using one-way ANOVA, chi-square tests, and independent samples *t*-tests as appropriate.

2.9. Neural–behavioral and clinical correlation analyses

To investigate the behavioral and clinical relevance of the brain regions exhibiting representational abnormalities in patients, we performed correlation analyses between neural indices and behavioral/clinical measures. The neural indices were defined as the neural representational strength extracted specifically from the brain regions that showed significant differences between the patient group and healthy controls. These neural indices were derived from the specific emotion-related response inhibition conditions under which these neural representational abnormalities were identified. Subsequently, these indices were correlated with behavioral performance metrics from the emotional Go/NoGo task, including omission errors, false response errors, and reaction time. Additionally, clinical symptom severity was assessed using standardized instruments: the YMRS and the MFQ. These correlations aimed to examine whether individual differences in neural representation within the identified aberrant brain regions were associated with behavioral task performance and symptom severity, thereby linking neural and behavioral levels to better explain group abnormalities in the patient cohort.

3. Results

3.1. Behavioral results

Demographic and clinical characteristics are shown in Table 1 across manic PBD, euthymic PBD, and HC groups. Statistical analyses revealed no significant group differences in gender distribution, age, education years, or IQ scores, with all effect sizes($\eta_p^2 < 0.06$) indicating adequate matching. While depression scores (MFQ) showed no group differences, manic symptoms (YMRS) differed significantly, with manic PBD patients exhibiting markedly higher scores. Patient groups did not differ in illness characteristics including onset age, duration, or episode frequency, and all participants were confirmed right-handed.

Behavioral task performance for the manic PBD, euthymic PBD, and HC groups is summarized in Table 2. Although no significant between-group differences were detected in omission errors, false response errors, or reaction times across all conditions, several informative trends emerged.

For omission errors, the manic PBD group showed a noticeably higher tendency under emotional stimuli compared with the euthymic PBD and HC groups, whereas the euthymic PBD group was only slightly higher than HCs. Under the neutral and control conditions, omission errors were comparable between the manic PBD and euthymic PBD groups, but both appeared elevated relative to HCs. For false response errors, the manic PBD group consistently showed higher error rates than

Table 2
Task performances in an emotional Go/Nogo task in the manic PBD, euthymic PBD and control groups.

Performances	Manic PBD (n = 15)	Euthymic PBD (n = 18)	HCs (n = 17)	F	p	η_p^2
Omission errors (%)						
Emotional targets (Go)	5.41 ± 1.64	4.52 ± 2.00	4.27 ± 2.29	1.397	0.257	0.056
Neutral targets (Go)	3.51 ± 2.09	3.47 ± 1.94	3.28 ± 1.95	0.064	0.938	0.003
Control targets (Go)	1.82 ± 1.11	1.89 ± 1.07	1.64 ± 1.05	0.242	0.786	0.010
False response error (%)						
Emotional distractors (Nogo)	3.36 ± 1.17	3.07 ± 1.26	2.93 ± 1.52	0.428	0.654	0.018
Neutral distractors (Nogo)	2.67 ± 1.25	2.35 ± 1.44	2.14 ± 0.92	0.735	0.485	0.030
Control distractors (Nogo)	2.32 ± 0.95	2.40 ± 0.95	2.47 ± 1.34	0.071	0.932	0.003
Reaction time (ms)						
Emotional targets (Go)	589.47 ± 68.84	623.84 ± 100.70	604.64 ± 62.11	0.767	0.470	0.032
Neutral targets (Go)	592.44 ± 71.25	616.79 ± 94.70	592.27 ± 69.03	0.540	0.587	0.022
Control targets (Go)	562.61 ± 66.33	600.38 ± 90.37	555.72 ± 65.86	1.755	0.184	0.069

the euthymic PBD and HC groups in both emotional and neutral conditions, while the euthymic PBD group again showed a modest increase relative to HCs. Regarding reaction time, the manic PBD group exhibited shorter responses in the emotional condition than the other two groups. This pattern may tentatively suggest heightened approach responding or increased responsiveness to emotional stimuli in manic PBD, which could compromise inhibitory control when emotional cues are present. In contrast, under neutral and control conditions, reaction times in the manic PBD group were broadly similar to those of HCs. Across all conditions, the euthymic PBD group showed a general tendency toward longer reaction times than both the manic PBD and HC groups, potentially reflecting a more cautious response style or increased processing demands.

Overall, while behavioral differences did not reach statistical significance, the trend-level patterns are broadly consistent with the notion that manic PBD may be more susceptible to emotional interference, showing relatively reduced behavioral inhibition in the presence of emotional stimuli.

3.2. Searchlight analysis

Using Searchlight-RSA, we systematically examined whole-brain neural representation patterns in manic PBD, euthymic PBD, and HCs across six emotional response inhibition conditions. As an initial step, we localized condition-specific representational regions within each group under each emotional inhibition model($p_{FWE} < 0.05$). This data-driven delineation of multivariate representational profiles provided a spatial foundation for subsequent between-group statistical comparisons, enabling the identification of neural markers that reflect disorder-related deviations in emotional inhibitory processing. The detailed information is shown in Fig. 3 and Tables S1-S18.

Previous studies have shown that the regulation of each emotion involves unique and widely distributed cortical control networks (Camacho et al., 2023; Qin, 2023). In line with this, our results revealed that emotional response inhibition engaged a broad array of brain regions across the cerebral cortex, including the executive control network, emotional processing network, sensorimotor network, default mode network, visual processing network, and salience network. Each emotional inhibition condition elicited a distinct spatial pattern of neural representations, further supporting the notion that emotional regulation relies on distributed and emotion-specific cortical substrates. The specific cortical distribution of representational regions for each condition is detailed below.

In the happy inhibition condition, in which the theoretical RDM distinguished between happy-go and happy-nogo trials, healthy controls primarily activated the executive control network (right middle frontal gyrus [MFG], right superior frontal gyrus [SFG]) and visual processing network (right occipital lobe). In contrast, manic patients demonstrated predominant engagement of the salience network (left putamen) and visual network (right middle occipital gyrus [MOG], right cuneus). euthymic patients showed an intermediate activation pattern involving both the frontoparietal network (left inferior frontal gyrus [IFG], left SFG) and sensorimotor regions (left precentral gyrus). These spatial patterns of group-specific representational regions are visualized in Fig. 3A. As visually evident in Fig. 3A, the distribution of representational regions in the manic group was notably sparser compared to the other two groups, indicating a reduced number of cortical areas effectively representing happy inhibition. This reduction may reflect functional impairments in the brain regions that are typically responsible for this form of emotional inhibitory control in patients with manic-phase PBD.

When examining sad inhibition, HCs exhibited robust activation across multiple networks including the executive control network (bilateral MFG, right IFG), default mode network (right posterior cingulate), and temporal-parietal regions (bilateral middle temporal gyrus [MTG], right inferior parietal lobule [IPL]). Manic patients, however, displayed reduced activation primarily confined to the executive control network (bilateral MFG). euthymic patients demonstrated partial restoration of temporal-frontal connectivity (bilateral superior temporal gyrus [STG], left IFG). These network-level differences across groups are illustrated in Fig. 3B

The emotional inhibition condition revealed distinct network engagement profiles among groups. HCs maintained coordinated activation across executive control networks (bilateral MFG, left IFG), default mode networks (bilateral medial frontal gyrus), and visual processing systems (right lingual gyrus). Manic patients consistently shifted toward salience network dominance (right putamen, left insula), while euthymic patients showed intermediate reorganization with improved connectivity between frontal (bilateral MFG) and temporal (left MTG) regions. A full depiction of these group-specific engagement patterns can be found in Fig. 3C.

For happy-specific inhibition, which contrasts inhibitory responses to happy stimuli against those to other emotional valences (including sad and neutral), activation patterns were more focal yet distinct. HCs selectively recruited motor control regions (right precentral gyrus), whereas manic patients primarily engaged frontal-striatal pathways (left MFG, right putamen). euthymic patients presented transitional activation focused on prefrontal regulatory zones (left IFG), indicating emerging recovery of top-down control mechanisms. These activation differences are shown in Fig. 3D

In the sad-specific inhibition condition, HCs showed integrated activation of the limbic network (left amygdala, left hippocampus) and prefrontal regulatory regions (left IFG). Conversely, manic patients exhibited a pronounced shift toward subcortical structures within the salience network (bilateral putamen, left insula). euthymic patients displayed transitional patterns with partial restoration of fronto-limbic connectivity (bilateral MTG, right SFG), as illustrated in Fig. 3E. Compared to healthy controls, both patient groups showed fewer representational regions under this condition, along with weaker representational strength. Notably, manic patients exhibited not only a further reduction in the number of representational regions relative to

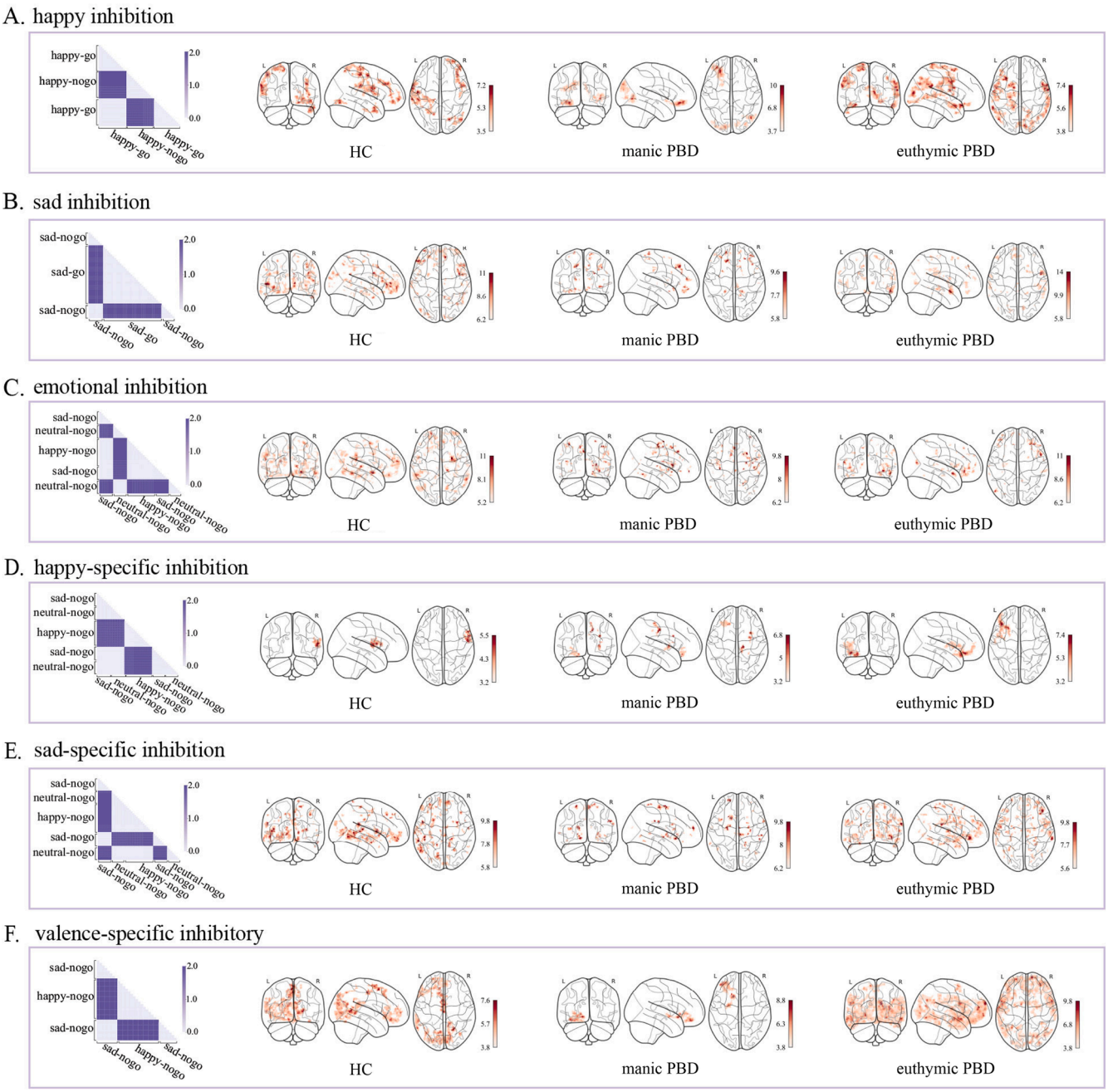


Fig. 3. Searchlight RSA reveals neural representations of emotional inhibition in three groups. A, Representational Dissimilarity Matrix (RDM) for the happy inhibition(happy-go vs. happy-nogo) condition, alongside whole-brain neural representation patterns across three subject groups. The color bar indicates representational strength, with higher values reflecting more distinct neural encoding of happy inhibition. B-F, Corresponding RDMs and neural representation patterns for: (B) sad inhibition (sad-go vs. sad-nogo), (C) emotional inhibition (happy-go/sad-go vs. happy-nogo/sad-nogo), (D) happy-specific inhibition (happy-nogo vs. sad-nogo/neutral-nogo), (E) sad-specific inhibition (sad-nogo vs. happy-nogo/neutral-nogo), and (F) valence-specific inhibition (happy-nogo vs. sad-nogo). All panels share the same interpretation of representational strength as in (A).

HCs, but also a significantly sparser distribution compared to euthymic patients, suggesting more pronounced disruption during the manic state.

During valence-specific inhibitory tasks, HCs demonstrated widespread network involvement including the executive control network (left MFG, left IFG, left medial frontal gyrus, left SFG), visual processing network (bilateral occipital lobe, bilateral cuneus, right fusiform gyrus, left lingual gyrus), and parietal-temporal regions (left parietal lobe, left temporal lobe). Manic patients showed substantially reduced activation limited to frontal regions (left frontal lobe, left MFG), while euthymic patients exhibited partial network recovery spanning bilateral frontal (bilateral frontal lobe, left SFG), temporal-parietal (bilateral temporal lobe, bilateral parietal lobe, right angular gyrus), and subcortical networks (right putamen, right lentiform nucleus). A summary of these group-level activation profiles is provided in Fig. 3F.

Under each condition, the spatial distribution of representational brain regions in the patient group was sparser compared to that in the healthy control group. Although the degree of reduction varied across conditions, with more pronounced differences observed in certain contrasts such as sad-specific inhibition and more subtle discrepancies in others, the overall pattern consistently indicated attenuated representational engagement in PBD. This suggests that key functional brain areas in patients with bipolar disorder exhibited reduced multivariate

encoding fidelity during emotional inhibitory processing, potentially reflecting impaired stimulus differentiation and regulatory control. These findings provide neurobiological support for the presence of distributed emotional response inhibition deficits in PBD. For condition-specific group differences, see post hoc comparison results.

3.3. Between–group analysis

The following results constitute the primary neuroimaging findings of this study and directly test our hypotheses. The neural differences between PBD patients and HCs are depicted in Fig. 4 and Tables S19.

3.3.1. Trait effect: Group differences between PBD and HCs

One of our core hypotheses was that PBD patients, regardless of illness phase, would exhibit inhibitory control deficits under emotional interference. To test this hypothesis, we first compared neural representational strength between the patient groups and HCs across the six inhibitory conditions in this study. PBD patients showed significant differences in neural representations relative to healthy controls during happy inhibition, sad inhibition, and sad-specific inhibition conditions, demonstrating trait-related effects specific to the disorder.

Specifically, during happy inhibition, both manic and euthymic PBD groups demonstrated weaker representation strength in the left precentral gyrus (manic PBD: $x = -62.5$, $y = 7.5$, $z = 29.5$, $t_{(31)} = 3.193$, $p_{FWE} = 0.017$; euthymic PBD: $x = -64.5$, $y = 1.5$, $z = 23.5$, $t_{(31)} = 5.667$, $p_{FWE} = 0.001$) compared to HCs.

For sad inhibition, patients exhibited reduced activation in the left MFG (manic PBD: $x = -34.5$, $y = 59.5$, $z = -4.5$, $t_{(31)} = 4.622$, $p_{FWE} = 0.004$; euthymic PBD: $x = -28.5$, $y = 41.5$, $z = -12.5$, $t_{(31)} = 3.676$, $p_{FWE} = 0.021$), with manic patients additionally showing deficits in the left occipital lobe ($x = -26.5$, $y = -88.5$, $z = 29.5$, $t_{(31)} = 4.396$, $p_{FWE} = 0.001$), indicating broader disruptions in top-down regulatory networks during negative emotion inhibition.

During sad-specific inhibition, both patient groups displayed weaker representations in the left IFG (manic PBD: $x = -50.5$, $y = 37.5$, $z = 1.5$, $t_{(31)} = 4.701$, $p_{FWE} = 0.001$; euthymic PBD: $x = -46.5$, $y = 35.5$, $z = 5.5$, $t_{(31)} = 4.387$, $p_{FWE} = 0.002$) and left MTG (manic PBD: $x = -66.5$, $y = -18.5$, $z = -6.5$, $t_{(31)} = 4.833$, $p_{FWE} = 0.001$; euthymic PBD: $x = -68.5$, $y = -20.5$, $z = -10.5$, $t_{(31)} = 3.662$, $p_{FWE} = 0.013$), while manic patients further exhibited deficits in the left IPL ($x = -60.5$, $y = -40.5$, $z = 25.5$, $t_{(31)} = 4.550$, $p_{FWE} = 0.005$). Overall, patient–control differences were mainly localized to frontal, frontoparietal, and temporal regions

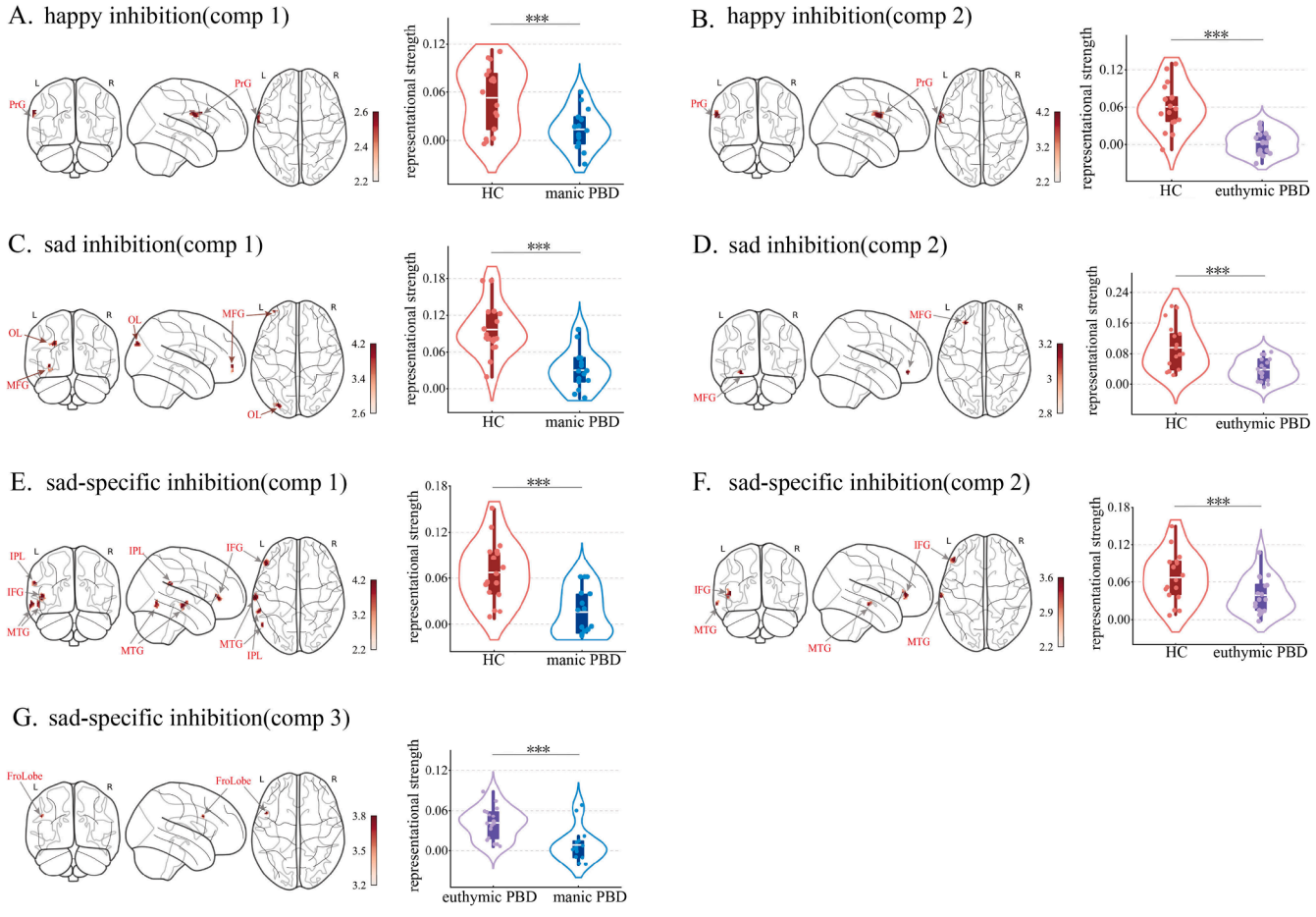


Fig. 4. Aberrant neural representations of emotional inhibition in PBD patients. A, Neural representation differences between manic-phase PBD patients and healthy controls during happy inhibition. The color bar represents t-values from between-group comparisons, with higher absolute values indicating more significant differences. The right panel displays descriptive statistics of representational strength (the correlation between observed neural activity patterns and theoretical similarity matrices of task conditions) in the identified differential brain regions, illustrating group-wise distinctions. B, Neural representation differences between euthymic PBD and HCs during happy inhibition. C, Neural representation differences between manic PBD and HCs during sad inhibition. D, Neural representation differences between euthymic PBD and HCs during sad inhibition. E, Neural representation differences between manic PBD and HCs during sad-specific inhibition. F, Neural representation differences between euthymic PBD and HCs during sad-specific inhibition. G, Neural representation differences between euthymic PBD and manic-phase PBD during sad-specific inhibition. Labels comp1 and comp2 refer to the primary and secondary comparison conditions in the representational similarity analysis respectively. ***p < 0.001. **Abbreviations:** PrG, precentral gyrus; MFG, middle frontal gyrus; OL, occipital lobe. IPL, inferior parietal lobule; IFG, inferior frontal gyrus; MTG, middle temporal gyrus; Frolobe, frontal lobe.

implicated in cognitive control and top-down regulation, providing strong support for the hypothesized inhibitory deficit under emotional interference. In contrast, no significant group differences were detected for emotional inhibition, happy-specific inhibition, or valence-specific inhibition after FWE correction (Table S19).

3.3.2. State effect: Differences between manic and euthymic PBD

Our second core hypothesis was that PBD patients in different illness phases would show phase-dependent abnormalities during inhibition of responses to specific emotional valence. To test this hypothesis, we further compared manic and euthymic patients across the inhibitory conditions. Patients in the manic phase exhibited weaker neural representations during sad-specific inhibition compared to euthymic patients ($x = -48.5$, $y = 13.5$, $z = 29.5$, $t_{(31)} = 3.885$, $p_{FWE} = 0.027$), reflecting a state effect that differentiates between clinical phases. Specifically, the differential brain regions were localized in the frontal lobe, particularly the MFG, suggesting that PBD patients during manic episodes exhibit further functional impairment in top-down regulatory networks. This finding supports the presence of a phase-related, valence-specific abnormality, with the critical effect expressed during sad-specific inhibition. Similarly, apart from this sad-specific inhibition contrast, no other between-phase differences survived FWE correction (Table S19).

3.4. Brain–behavior and symptom correlates

Correlation analyses revealed two significant associations, both confined to the manic PBD group. Specifically, mean neural representational strength during the sad inhibition condition was positively correlated with false response errors on the emotional Go/NoGo task ($r = 0.517$, $p = 0.048$), indicating that individuals showing stronger neural representations during sad inhibition also tended to commit more commission errors. In addition, sad-inhibition neural representations showed a positive association with depressive symptom severity as indexed by MFQ scores ($r = 0.505$, $p = 0.055$). While this effect was marginal, its direction was consistent with the behavioral correlation. No other correlations between neural representational indices (across the six inhibition conditions) and behavioral measures (omission errors, false response errors, reaction time) or symptom scales (YMRS, MFQ) reached significance in the manic PBD, euthymic PBD, or HC groups.

4. Discussion

This study is, to our knowledge, the first to apply multivariate representational similarity analysis (RSA) to characterize emotional response inhibition in pediatric bipolar disorder (PBD). By probing condition-specific representational structure during an emotional Go/NoGo task, we identified (i) cross-phase representational abnormalities in both manic and euthymic PBD relative to healthy controls (HCs) during happy inhibition, sad inhibition, and sad-specific inhibition, and (ii) a phase-dependent difference between manic and euthymic PBD selectively during sad-specific inhibition, expressed in prefrontal cortex. These effects were primarily located in regions supporting top–down control and stimulus–response mapping (frontoparietal and lateral prefrontal systems) as well as regions contributing to perceptual and motor components of inhibition (occipitotemporal and precentral cortices). Collectively, our findings suggest coexisting trait-like and state-related alterations in the representational organization of emotional response inhibition in PBD, extending prior univariate evidence of emotion–control imbalance in bipolar disorder (Wessa et al., 2007; Wolkenstein et al., 2017; Förster et al., 2023; Pomarol-Clotet et al., 2015; Xiao et al., 2021). Importantly, RSA tests how neural patterns differentiate task conditions, not merely whether overall response magnitude differs (Kriegeskorte, 2008; Haxby, 2012). In developmental psychopathology, such representational-level perturbations may be especially informative because frontoparietal control networks and their coupling with affective and perceptual systems are still maturing during adolescence (Arain et al., 2013; Gu et al., 2015). Thus, representational differences may emerge even when overt behavior shows only trend-level change.

4.1. Hypothesis testing: selective support for trait vs. state effects

Our a priori framework distinguished trait effects (abnormalities shared across manic and euthymic PBD relative to HCs) from state effects (differences between manic and euthymic PBD). The results supported this framework in a contrast-specific manner: trait-like abnormalities were observed in happy inhibition, sad inhibition, and sad-specific inhibition, whereas a state effect emerged only for sad-specific inhibition. Notably, no significant group differences survived whole-brain correction for emotional inhibition, happy-specific inhibition, or valence-specific inhibition. The null finding for emotional inhibition is compatible with a representational interpretation because this contrast collapses across valences (happy and sad) and may therefore be less sensitive when abnormalities are valence-dependent (Kriegeskorte, 2008; Haxby, 2012). We do not over-interpret the remaining non-significant contrasts, and we treat the observed pattern of selectivity as the primary empirical result of the hypothesis test.

4.2. Trait effect I: Happy inhibition and reduced precentral representational fidelity—action gating under positive interference

During happy inhibition, that is, response inhibition under happy emotional interference, operationalized in our RSA model as the contrast between happy-go and happy-nogo trials, both manic and euthymic PBD showed weaker representational strength in the left precentral gyrus relative to HCs. The precentral cortex is central to action planning and motor implementation, and in Go/NoGo contexts it contributes to forming a dominant response tendency and to gating action execution when a stop signal is required (Nakayama et al., 2016). Therefore, reduced representational correspondence to the happy-inhibition model may reflect diminished pattern-level differentiation between Go and NoGo action states under positive emotional interference, that is, weaker encoding of “execute vs withhold” mappings in motor-related circuitry.

This trait-like pattern is consistent with the idea that PBD may show broad emotion-related inhibition vulnerability that persists even outside acute episodes, potentially reflecting a stable alteration in the integration between emotional cues and response gating systems (Wessa et al., 2007; Wolkenstein et al., 2017; Förster et al., 2023; Pomarol-Clotet et al., 2015; Xiao et al., 2021). Importantly, our interpretation is framed at the representational level: it is not simply that precentral cortex is “less engaged,” but that its multivoxel patterns are less consistent with the inhibitory representational structure required to distinguish happy-go from happy-nogo trials.

4.3. Trait effect II: Sad inhibition and prefrontal (MFG) representational weakening—rule-based control under negative interference

For sad inhibition, both patient groups showed reduced representational strength in the left middle frontal gyrus (MFG), with additional occipital involvement in manic PBD. The MFG is frequently implicated in cognitive control and voluntary regulation processes, including the suppression of negative affect (Phan et al., 2005; Elliott et al., 2000). In the RSA framework, weaker MFG representational correspondence suggests that the prefrontal system encodes the “sad-go vs sad-nogo” distinction with reduced fidelity, consistent with compromised rule maintenance or conflict resolution under negative emotional interference.

The additional occipital representational alteration observed in manic PBD may reflect perturbations in perceptual differentiation of emotional face information or disrupted top–down modulation of early visual processing during inhibition (Förster et al., 2023). Rather than

interpreting occipital findings as unequivocal "compensation," a more conservative account is that altered occipital representational structure may index atypical visual–attentional encoding of sad cues that interacts with downstream control demands. This interpretation aligns with evidence that visual cortex representations can become sparser under negative stimulus processing and may relate to emotion regulation dynamics (Yin et al., 2020), but the functional direction (adaptive vs maladaptive) likely depends on broader network interactions and should be tested directly in future work.

4.4. Trait effect III: Sad-specific inhibition and weakened IFG–MTG representations—impaired binding between inhibitory rules and sad meaning

A key contribution of the present study is the identification of trait-like abnormalities during sad-specific inhibition, which isolates representational structure unique to sad NoGo trials relative to other NoGo valences—an analytic advantage afforded by multivariate RSA (Kriegeskorte, 2008; Haxby, 2012). In both PBD groups, representational strength was reduced in left inferior frontal gyrus (IFG) and left middle temporal gyrus (MTG), with additional left inferior parietal lobule (IPL) impairment in manic PBD.

Mechanistically, the IFG is a core hub for inhibitory control and is also implicated in emotion regulation, making it a plausible convergence point for emotional stopping computations (Aron, 2007; Mitchell, 2011). Reduced IFG representational strength therefore suggests that PBD may allocate fewer effective representational resources to encoding the inhibitory rule structure specifically required to suppress responses to sadness-related cues, consistent with prior reports of IFG abnormalities in BD emotion inhibition contexts (Förster et al., 2023; Xia et al., 2024). The MTG contributes to emotional stimulus processing, higher-level visual/semantic integration, and can participate in control-related representations when emotion meaning must be linked to task rules (Sabatinelli et al., 2011; Bhaijiwala et al., 2014). Accordingly, weakened MTG representational structure may reflect impaired encoding of sadness-related facial meaning or diminished coupling between semantic–affective representations and inhibitory rule representations. This offers a concrete representational mechanism: the system may struggle to bind "sadness meaning" to a distinct NoGo rule in a way that is separable from other emotional NoGo rules.

In manic PBD, the additional reduction in IPL representational strength may indicate compromised attentional allocation and stimulus–response mapping updates that support the resolution of dominant response tendencies under competing emotional distractors (Seeley et al., 2007; Seghier, 2013). While univariate studies sometimes report IPL hyperactivation in BD (Förster et al., 2023; Xiao et al., 2021), our representational findings support an "inefficiency" account: degraded representational fidelity can coexist with increased amplitude recruitment, such that greater activation may reflect effortful but less precise processing rather than improved function.

4.5. State effect: Manic-specific reduction in prefrontal representations during sad-specific inhibition—candidate phase marker

Only sad-specific inhibition differentiated manic from euthymic PBD, with manic patients showing weaker representational strength in prefrontal cortex (notably MFG). This pattern indicates that the manic state may be associated with further reductions in the precision of prefrontal representations required for fine-grained suppression of sadness-related responses, beyond the cross-phase trait vulnerability.

This finding is consistent with phase-sensitive models proposing that mood states modulate top–down regulatory dynamics in advanced control networks (Blumberg et al., 2003; Hummer et al., 2013; Passarotti et al., 2009; Pavuluri et al., 2008). It also converges with our prior work showing state-dependent changes in top-down regulatory connections (triple network → limbic network) and functional dynamics of higher-order networks (Lei et al., 2023; Wang et al., 2024). From this perspective, sad-specific inhibition may be a particularly sensitive computational probe of phase-related dysregulation because it requires stable maintenance of a valence-specific inhibitory rule boundary. Therefore, the manic-associated prefrontal representational weakening observed here should be viewed as a candidate state-related biomarker that warrants validation in larger and longitudinal designs.

4.6. Integrating trait and state effects: A mechanistic account of representational selectivity

Building on the condition-specific findings above, we propose a mechanistic account of why this pattern of selective representational disturbances delineates a trait–state architecture in PBD. The core point is that, in our RSA framework, representational strength indexes how reliably local multivoxel patterns instantiate the representational geometry specified by each model RDM (Kriegeskorte, 2008; Haxby, 2012). Accordingly, effects that were present in both manic and euthymic PBD relative to HCs can be interpreted as trait-like because they indicate a cross-phase reduction in the fidelity with which key inhibition computations are represented under emotional interference, even when syndromal symptoms have remitted. Importantly, these trait-like effects converge on computations that require maintaining a stable separation between "execute" and "withhold" response states within an emotional context (happy inhibition; sad inhibition) and, for sadness in particular, linking a NoGo rule to sadness-related stimulus meaning in a way that remains separable from other NoGo valences (sad-specific inhibition). In other words, the trait pattern is most parsimoniously understood as a persistent weakness in representational precision for emotion-laden inhibitory rules, expressed across multiple nodes of the emotion–control interface.

In contrast, the state effect was computationally specific: manic PBD showed an additional reduction in prefrontal representational strength only during sad-specific inhibition. This contrast places the strongest demand on maintaining a valence-unique inhibitory boundary within the NoGo class, because it requires representing sad NoGo trials as distinct from other emotional NoGo trials, rather than merely separating Go from NoGo. The selective emergence of a manic–euthymic difference in this contrast therefore suggests that manic episodes preferentially compromise prefrontal representations supporting context-sensitive rule maintenance and fine-grained negative emotion suppression, consistent with phase-sensitive models that emphasize reduced prefrontal regulatory efficacy in mania and state-dependent alterations in higher-order regulatory dynamics (Blumberg et al., 2003; Hummer et al., 2013; Passarotti et al., 2009; Pavuluri et al., 2008; Lei et al., 2023; Wang et al., 2024). Importantly, this account yields testable predictions: candidate trait markers should show longitudinal stability across mood phases, whereas the candidate state marker should covary with symptom fluctuations and normalize with remission in within-subject designs.

4.7. Behavioral null effects and brain–behavior/symptom links: representational alterations may precede overt performance deficits

Although behavioral group differences did not reach significance, trend-level patterns (e.g., higher error tendencies and faster responses in manic PBD under emotional conditions; slower responses in euthymic PBD) were broadly consistent with heightened emotional interference and altered response strategies. Behavioral null effects in adolescent Go/NoGo paradigms can reflect ceiling effects, limited sensitivity of summary error measures, and compensatory strategy use. In this context, multivariate representational indices may detect latent disruptions in condition separability that are not yet expressed as robust behavioral impairment (Kriegeskorte, 2008; Haxby, 2012).

Correlation analyses yielded associations confined to manic PBD, where stronger sad-inhibition representational strength correlated with more false responses and showed a marginal positive association with

depressive symptom severity (MFQ). These findings should be interpreted cautiously given sample size and multiple testing. One plausible account is that in mania, greater representational expression of sad inhibition might reflect inefficient or noisy recruitment of control representations, which does not translate to better stopping performance and may co-occur with greater symptom burden. Future work should test whether these correlations replicate and whether they reflect maladaptive over-engagement, failure of compensation, or altered coupling between representational precision and behavioral control.

4.8. A representational degradation–compensation framework to reconcile hypo- and hyper-findings

Across conditions, we observed reduced representational strength in distributed regions supporting emotional inhibition. This pattern motivates a "representational degradation" account: when neural systems fail to maintain distinct patterns for Go vs NoGo under emotional interference (i.e., reduced representational separability), information processing becomes less efficient and may require greater effort elsewhere in the network to sustain performance. This framework can reconcile prior reports of prefrontal hyperactivation in PBD during emotional inhibition (Xiao et al., 2021) with our representational reductions: hyperactivation may reflect compensatory effort in the face of degraded representational quality rather than improved inhibitory function. Over time, such compensatory demands might contribute to subjective fatigue during emotion regulation and attentional strain reported in BD (Brown et al., 2012). Importantly, the persistence of representational reductions during euthymia suggests a potential endophenotypic feature, whereas additional manic-specific prefrontal degradation during sad-specific inhibition suggests a phase-sensitive vulnerability.

4.9. Limitations and future directions

Several limitations warrant consideration. First, the sample size is modest and the design is cross-sectional; definitive trait–state inferences require longitudinal within-subject assessments across mood phases. Second, the block design supports stable emotional context but limits characterization of rapid trialwise dynamics. Third, the emotion set was limited to happy, sad, and neutral faces; extending to high-arousal negative emotions (e.g., anger, fear) may clarify valence–arousal specificity. Fourth, our binary model RDMs may not fully capture continuous representational dimensions; future studies should test richer models incorporating graded valence/arousal or behavioral similarity. Fifth, due to the variability in medication usage among participants, we were unable to include medication history as a covariate in the analysis. Given that pharmacological treatments can modulate brain activity, this introduces a potential limitation in interpreting the observed neural representations. Future research should consider accounting for medication effects in analyzing trait–state relationships. Finally, replication and predictive validation (e.g., phase classification, symptom prediction) in larger cohorts will be essential to establish the clinical utility of sad-specific inhibition representations as candidate biomarkers.

5. Conclusions

In summary, RSA revealed that emotional response inhibition deficits in PBD are characterized by both cross-phase trait-like and manic-specific state-related alterations in representational organization. Trait-like abnormalities were observed in precentral (happy inhibition), prefrontal and occipital (sad inhibition), and IFG–MTG (sad-specific inhibition) systems, whereas a manic-associated reduction in prefrontal representational strength emerged selectively during sad-specific inhibition. These findings support the view that PBD involves disrupted representational precision for emotion-laden inhibitory rules, and they motivate future longitudinal and mechanistic work to test phase-sensitive interventions targeting valence-specific inhibitory control.

Ethics statement

All procedures in this study were approved by the Ethics Committee of the Second Xiangya Hospital, Central South University(2013-S040). All procedures performed in this research involving human participants comply with the ethical standards of national and institutional research committees, as well as the ethical standards of the Declaration of Helsinki (1964) and its subsequent amendments. Written informed consent was obtained from all participants and their guardians.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

CRediT authorship contribution statement

Jia Li: Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation. **Rong Wang:** Writing – review & editing, Supervision, Investigation, Funding acquisition. **Jianze Wu:** Software, Methodology, Funding acquisition, Formal analysis. **Qian Xiao:** Investigation, Data curation. **Yuan Zhong:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

We gratefully acknowledge Chuanji Gao, Liang Shi and Yun Wu for their invaluable advice on fMRI data preprocessing and analysis. We are also grateful to Yun Yang for her expert advice on the literature review and experimental design. Lastly, we thank all members of the Yuan Zhong Lab for their insightful discussions and support of this work. This research was funded by the National Natural Science Foundation of China (No. 32571269, 81871344), National Key R&D Program of China (Grant No 2023YFC3341300), Natural Science Foundation of Jiangsu Province (No. BK20241881), and "333 Talent Project" of Jiangsu Province.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2026.121738.

References

Arain, M., Haque, M., Johal, L., Mathur, P., Nel, W., Rais, A., et al., 2013. Maturation of the adolescent brain. Neuropsychiatr. Dis. Treat. 9, 449–461.
Aron, A.R., 2007. The neural basis of inhibition in cognitive control. The neuroscientist 13 (3), 214–228.
Bauer, I.E., Frazier, T.W., Meyer, T.D., Youngstrom, E., Zunta-Soares, G.B., Soares, J.C., 2015. Affective processing in pediatric bipolar disorder and offspring of bipolar parents. J. Child Adolesc. Psychopharmacol. 25 (9), 684–690.
Bhaijiwala, M., Chevrier, A., Schachar, R., 2014. Withholding and canceling a response in ADHD adolescents. Brain Behav. 4 (5), 602–614.
Blumberg, H.P., Leung, H.C., Skudlarski, P., Lacadie, C.M., Fredericks, C.A., Harris, B.C., et al., 2003. A functional Magnetic resonance imaging study of Bipolar disorder: state- and trait-related dysfunction in ventral prefrontal cortices. Arch. Gen. Psychiatry 60 (6), 601–609.
Bora, E., Yucel, M., Pantelis, C., 2009. Cognitive endophenotypes of bipolar disorder: a meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. J. Affect. Disord. 113 (1–2), 1–20.
Bortolato, B., Carvalho, A.F., McIntyre, R.S., 2015. Cognitive dysfunction in major depressive disorder: a novel therapeutic target? Curr. Neuropharmacol. 13 (999), 1.

Brown, M.R.G., Lebel, R.M., Dolcos, F., Wilman, A.H., Silverstone, P.H., Pazderka, H., et al., 2012. Effects of emotional context on impulse control. Neuroimage 63 (1), 434–446.
Camacho, M.C., Nielsen, A.N., Balser, D., Furtado, E., Steinberger, D.C., Fruchtman, L., et al., 2023. Large-scale encoding of emotion concepts becomes increasingly similar between individuals from childhood to adolescence. Nat. Neurosci. 26 (7), 1256–1266.
Carvalho, A.F., Firth, J., Vieta, E., 2020. Bipolar disorder. N. Engl. J. Med. 383 (1), 58–66.
Elliott, R., Rubinsztein, J.S., Sahakian, B.J., Dolan, R.J., 2000. Selective attention to emotional stimuli in a verbal go/no-go task: an fMRI study. Neuroreport 11 (8), 1739–1744.
Förster, K., Maliske, L.Z., Schurz, M., Henneberg, P.M., Dannlowski, U., Kanske, P., 2023. How do bipolar disease states affect positive and negative emotion processing? Insights from a meta-analysis on the neural fingerprints of emotional processing. Bipolar. Disord. 25 (7), 540–553.
Ferrari, A.J., Stockings, E., Khoo, J., Erskine, H.E., Degenhardt, L., Vos, T., et al., 2016. The prevalence and burden of bipolar disorder: findings from the Global Burden of Disease Study 2013. Bipolar. Disord. 18 (5), 440–450.
Geller, B., Bolhofner, K., Craney, J.L., Williams, M., DelBello, M.P., Gundersen, K., 2000. Psychosocial functioning in a prepubertal and early adolescent bipolar disorder phenotype. J. Am. Acad. Child Adolesc. Psychiatry 39 (12), 1543–1548.
Gorgolewski, K., Burns, C.D., Madison, C., Clark, D., Halchenko, Y.O., Waskom, M.L., et al., 2011. Nipype: a flexible, lightweight and extensible neuroimaging Data processing framework in Python. Front Neuroinformatics 5, 13.
Gotlib, I.H., Joormann, J., 2010. Cognition and depression: current status and future directions. Annu Rev. Clin. Psychol. 6, 285–312.
Gu, S., Satterthwaite, T.D., Medaglia, J.D., Yang, M., Gur, R.E., Gur, R.C., et al., 2015. Emergence of system roles in normative neurodevelopment. Proc. Natl. Acad. Sci. U S A. 112 (44), 13681–13686.
Haxby, J.V., 2012. Multivariate pattern analysis of fMRI: the early beginnings. Neuroimage 62 (2), 852–855.
Hummer, T.A., Hulvershorn, L.A., Karne, H.S., Gunn, A.D., Wang, Y., Anand, A., 2013. Emotional response inhibition in Bipolar disorder: a functional magnetic resonance imaging study of trait- and State-related abnormalities. Biol. Psychiatry 73 (2), 136–143.
Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al., 1997. 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 36 (7), 980–988.
Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. Proc. Natl. Acad. Sci. 103 (10), 3863–3868.
Kriegeskorte, N., 2008. Representational similarity analysis – connecting the branches of systems neuroscience. Front. Syst. Neurosci. 2, 249.
Lei, W., Xiao, Q., Wang, C., Cai, Z., Lu, G., Su, L., et al., 2023. The disruption of functional connectome gradient revealing networks imbalance in pediatric bipolar disorder. J. Psychiatr. Res. 164, 72–79.
Miller, J.N., Black, D.W., 2020. Bipolar Disorder and suicide: a review. Curr. Psychiatry Rep. 22 (2), 6.
Mitchell, D.G.V., 2011. The nexus between decision making and emotion regulation: a review of convergent neurocognitive substrates. Behav. Brain Res. 217 (1), 215–231.
Mumford, J.A., Turner, B.O., Ashby, F.G., Poldrack, R.A., 2012. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. Neuroimage 59 (3), 2636–2643.
Nakayama, Y., Yamagata, T., Hoshi, E., 2016. Rostrocaudal functional gradient among the pre-dorsal premotor cortex, dorsal premotor cortex and primary motor cortex in goal-directed motor behaviour. In: Munoz, D. (Ed.), Eur J Neurosci, Eur J Neurosci, 43, pp. 1569–1589.
Oosterhof, N.N., Connolly, A.C., Haxby, J.V., 2016. CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. Front Neuroinformatics 10, 27.
Passarotti, A.M., Sweeney, J.A., Pavuluri, M.N., 2009. Differential engagement of cognitive and affective neural systems in pediatric bipolar disorder and attention deficit hyperactivity disorder. J. Int. Neuropsychol. Soc. 16 (1), 106–117.
Pavuluri, M.N., O'Connor, M.M., Harral, E.M., Sweeney, J.A, 2008. An fMRI study of the interface between affective and cognitive neural circuitry in pediatric bipolar disorder. Psychiatry Res. NeuroimAging 162 (3), 244–255.
Phan, K.L., Fitzgerald, D.A., Nathan, P.J., Moore, G.J., Uhde, T.W., Tancer, M.E., 2005. Neural substrates for voluntary suppression of negative affect: a functional magnetic resonance imaging study. Biol. Psychiatry 57 (3), 210–219.
Pomarol-Clotet, E., Alonso-Lana, S., Moro, N., Sarró, S., Bonnin, M.C., Goikolea, J.M., et al., 2015. Brain functional changes across the different phases of bipolar disorder. Br. J. Psychiatry 206 (2), 136–144.
Qin, S., 2023. Emotion representations in context: maturation and convergence pathways. Trends. Cogn. Sci. 27 (10), 883–885.
Rae, C.L., Raykov, P., Ambridge, E.M., Colling, L.J., Gould van Praag, C.D., Bouyagoub, S., et al., 2023. Elevated representational similarity of voluntary action and inhibition in Tourette syndrome. Brain Commun. 5 (5), fcad224.
Sabatinelli, D., Fortune, E.E., Li, Q., Siddiqui, A., Krafft, C., Oliver, W.T., et al., 2011. Emotional perception: meta-analyses of face and natural scene processing. Neuroimage 54 (3), 2524–2533.
Sadana, D., Gupta, R.K., Jain, S., Kumaran, S.S., G.S, R., Thennarasu, K., et al., 2019. Neurocognitive profile of patients with Bipolar Affective Disorder in the euthymic phase. Asian J Psychiatry 44, 121–126.
Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., et al., 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. J. Neurosci. 27 (9), 2349–2356.
Seghier, M.L., 2013. The Angular gyrus: multiple functions and multiple subdivisions. The Neuroscientist 19 (1), 43–61.
Seymour, K.E., Kim, K.L., Cushman, G.K., Puzia, M.E., Weissman, A.B., Galvan, T., et al., 2015. Affective processing bias in youth with primary bipolar disorder or primary attention-deficit/hyperactivity disorder. Eur. Child Adolesc. Psychiatry 24 (11), 1349–1359.
Stefanopoulou, E., Manoharan, A., Landau, S., Geddes, J.R., Goodwin, G., Frangou, S., 2009. Cognitive functioning in patients with affective disorders and schizophrenia: a meta-analysis. Int. Rev. Psychiatry 21 (4), 336–356.
Van, M.A., Moreira, A., Youngstrom, E., 2019. Updated meta-analysis of epidemiologic studies of Pediatric Bipolar disorder. J. Clin. Psychiatry 80 (3), 21938.
Wang, R., Wang, C., Zhang, G., Mundinano, I.C., Zheng, G., Xiao, Q., et al., 2024. Causal mechanisms of quadruple networks in pediatric bipolar disorder. Psychol. Med. 54 (16), 4713–4724.
Wessa, M., Houenou, J., Paillère-Martinot, M.L., Berthoz, S., Artiges, E., Leboyer, M., et al., 2007. Fronto-striatal overactivation in euthymic bipolar patients during an emotional go/NoGo task. Am. J. Psychiatry 164 (4), 638–646.
Wolkenstein, L., Kanske, P., Bailer, J., Wessa, M., Hautzinger, M., Joormann, J., 2017. Impaired cognitive control over emotional material in euthymic bipolar disorder. J. Affect. Disord. 214, 108–114.
Wood, A., Kroll, L., Moore, A., Harrington, R., 1995. Properties of the mood and feelings questionnaire in adolescent psychiatric outpatients: a research note. J. Child Psychol. Psychiatry 36 (2), 327–334.
Xia, Y., Wang, X., Sheng, J., Hua, L., Dai, Z., Sun, H., et al., 2024. Response inhibition related neural oscillatory patterns show reliable early identification of bipolar from unipolar depression in a go/No-go task. J. Affect. Disord. 351, 414–424.
Xiao, Q., Wu, Z., Jiao, Q., Zhong, Y., Zhang, Y., Lu, G., 2021. Children with euthymic bipolar disorder during an emotional go/nogo task: insights into the neural circuits of cognitive-emotional regulation. J. Affect. Disord. 282, 669–676.
Yin, S., Bo, K., Liu, Y., Thigpen, N., Keil, A., Ding, M., 2020. Fear conditioning prompts sparser representations of conditioned threat in primary visual cortex. Soc. Cogn. Affect. Neurosci. 15 (9), 950–964.
Young, R.C., Biggs, J.T., Ziegler, V.E., Meyer, D.A., 1978. A rating scale for mania: reliability, validity and sensitivity. Br. J. Psychiatry 133 (5), 429–435.