Alpha Psychiatry

Alpha Psychiatry **2026**; 27(6): 52710
https://doi.org/10.31083/AP52710

Original Article

The Influence of Executive Function, Empathy, and Serum Chemokine Levels on Clinical Symptoms in Bipolar Disorder Patients: An Examination of Moderating Effects

Daxiu Xiao[1], Yuanxia Chen[1], Wenjing Li[1], Chunyang Liu[1], Linhao Zhang[1], Jiahong Liu[1,*]

[1]The Affiliated Kangning Hospital of Wenzhou Medical University, Zhejiang Provincial Clinical Research Center for Mental Health, 325000 Wenzhou, Zhejiang, China

*Correspondence: allanany_liu@163.com (Jiahong Liu)

Academic Editors: Evgeny Ermakov and Taro Kishi

Submitted: 8 April 2026 Revised: 1 July 2026 Accepted: 6 July 2026 Published: 9 November 2026

Abstract

Background: Although executive dysfunction is a hallmark of bipolar disorder (BD), the mechanisms through which social cognition and neuroinflammation influence cognition–symptom relationships remain unclear. This study examined the moderating effects of overall empathy and serum Chemokine (C-C motif) ligand 3/Macrophage Inflammatory Protein-1 alpha, CCL3/MIP-1α and Chemokine (C-C motif) ligand 4/Macrophage Inflammatory Protein-1 beta, CCL4/MIP-1β levels on the association between executive function and mood symptom severity in BD and investigated whether these effects differ between manic and depressive episodes. **Methods:** Two hundred patients with bipolar I disorder (123 manic, 77 depressive; all medication-free for at least two months) and 200 healthy controls were enrolled. Executive function was assessed using the Wisconsin Card Sorting Test (WCST), overall empathy using the Interpersonal Reactivity Index (IRI), and mood symptom severity using the Bech–Rafaelsen Mania Scale (BRMS) and Hamilton Depression Rating Scale-24. Serum CCL3/MIP-1α and CCL4/MIP-1β levels were measured using enzyme-linked immunosorbent assays. First- and second-order moderating effects were tested using hierarchical regression analysis and a three-way interaction model, respectively. The Benjamini–Hochberg false discovery rate (FDR) procedure was applied to correct for multiple comparisons. **Results:** BD patients exhibited significant executive dysfunction, reduced overall empathy, and elevated chemokine levels compared with controls (all $p < 0.001$, FDR-corrected $q < 0.05$). In the manic subgroup, the number of WCST errors was strongly associated with total BRMS score. High total IRI scores attenuated this association ($\beta = -0.003$, 95% confidence interval (CI): −0.006 to −0.001, $p < 0.01$, $\Delta R^2 = 0.002$), whereas high CCL3/MIP-1α levels amplified it ($\beta = 0.002$, 95% CI: 0.001–0.002, $p < 0.01$, $\Delta R^2 = 0.003$). CCL4/MIP-1β level exerted a second-order moderating effect on that of total IRI score ($\beta = 0.0002$, 95% CI: 0.0001–0.0003, $p < 0.001$), reversing the attenuating role of empathy when high. No significant moderation effects were detected in the depressive subgroup. **Conclusions:** In this medication-free inpatient sample, executive dysfunction was associated with manic symptom severity, and exploratory moderation analyses suggested that overall empathy and selected chemokine levels were statistically associated with small changes in the WCST error–BRMS slope. However, the unusually strong WCST error–BRMS association ($r \approx 0.983$, $R^2 = 0.965$), the cross-sectional design, and the very small interaction effects ($\Delta R^2 = 0.001$–0.005) substantially limit interpretability. These findings should be viewed as hypothesis-generating rather than definitive evidence of socio-cognitive or immunoinflammatory moderation in bipolar mania.

Keywords: bipolar disorder; executive function; chemokines; empathy; moderation effect

Main Points

1. Patients with bipolar I disorder exhibited significant executive dysfunction, reduced overall empathy, and elevated serum chemokine (CCL3/MIP-1α, CCL4/MIP-1β) levels compared with healthy controls.

2. In the manic subgroup, executive dysfunction (WCST errors) was strongly associated with manic symptom severity (BRMS), with an unusually high correlation ($r \approx 0.983$) that warrants cautious interpretation.

3. Overall empathy (IRI) attenuated the executive function–mania severity link, whereas elevated CCL3/MIP-1α amplified it; CCL4/MIP-1β exerted a second-order moderating effect that reversed empathy's protective direction at high chemokine levels.

4. Overall empathy (IRI) attenuated the executive function–mania severity link, whereas elevated CCL3/MIP-1α amplified it; CCL4/MIP-1β exerted a second-order moderating effect that reversed empathy's protective direction at high chemokine levels.

5. Due to the cross-sectional design, unusually strong WCST–BRMS correlation, and small interaction effect sizes, these findings should be considered hypothesis-generating rather than definitive.

Copyright: © 2026 The Author(s). Published by IMR Press.
This is an open access article under the CC BY 4.0 license.

Publisher's Note: IMR Press stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Bipolar disorder (BD)—a severe chronic mental illness characterized by recurrent episodes of (hypo)mania and depression—affects 1%–4% of the global population and is one of the leading causes of disability [1,2]. This illness profoundly impairs social and occupational functioning, with 30%–60% of patients failing to achieve full functional recovery between episodes [3]. The high recurrence rate (>70% over five years) of BD, elevated suicide risk posed by this disorder, and its frequent psychiatric and medical comorbidities underscore the urgent need to elucidate its pathophysiological mechanisms and develop mechanism-informed targeted intervention strategies [4]. The understanding of BD pathogenesis has advanced from the rather simplistic single cognitive deficit model to a more comprehensive and complex multisystem interaction model integrating neurocognition, social cognition, and neuroinflammation [5].

Neurocognitive dysfunction is a core feature of BD: Executive function impairment persists throughout the illness and is closely linked to symptom presentation and prognosis [6]. Owing to its sensitivity in assessing inhibitory control and cognitive flexibility, the Wisconsin Card Sorting Test (WCST) is widely used to evaluate executive function in this context [7]. Social cognition—a higher-order extension of neurocognition—encompasses abilities such as emotion recognition, empathy, and theory of mind, and may mediate the link between neurocognition and interpersonal functioning in BD patients. Empathy, as a core dimension of social cognition, may play a critical role in moderating the cognition–emotion interaction; however, its specific role remains unclear [8]. Affective empathy—a core dimension of social cognition and emotional regulation—involves experiencing and resonating with others' emotions. Although BD patients show abnormal empathy, its specific role in the cognition–emotion interaction remains debatable [9,10].

Neuroimmune dysregulation provides a complementary perspective on BD pathophysiology. Peripheral chemokines such as Chemokine (C-C motif) ligand 3 (CCL3)/Macrophage Inflammatory Protein-1 alpha (MIP-1α) and Chemokine (C-C motif) ligand 4 (CCL4)/Macrophage Inflammatory Protein-1 beta (MIP-1β) can influence brain regions associated with cognition and emotion by potentially involving blood–brain barrier or affecting peripheral–central immune signaling pathways [11,12]. Although a peripheral chemokine level imbalance in BD patients has been preliminarily established through previous studies [12,13], targeted studies examining the role of these species in moderating the cognition–symptom relationship—particularly their interaction with social cognitive processes such as empathy—are lacking.

Existing studies have not systematically examined whether empathy attenuates or amplifies the impact of executive dysfunction on clinical symptoms and whether this role differs between manic and depressive episodes. Moreover, the ability of chemokines to moderate the cognition–symptom association—directly or by interacting with empathy—has not been empirically explored. Finally, the heterogeneity across BD episode subtypes has seldom been modeled explicitly in moderation analyses. The present study addresses these gaps by (a) examining whether overall empathy moderates the association between executive dysfunction and mood symptom severity, (b) testing whether serum CCL3/MIP-1α and CCL4/MIP-1β levels exert first- and second-order moderating effects on this association, and (c) conducting all moderation analyses separately for manic and depressive subgroups to capture subtype-specific patterns.

2. Materials and Methods

2.1 Study Subjects

Participants were recruited prospectively from April 23, 2024 to April 22, 2025, after ethics approval had been granted by the Ethics Review Committee of The Affiliated Kangning Hospital of Wenzhou Medical University on April 23, 2024 (Approval No.: YJ-2024-01-02). No participant was approached, screened, enrolled, or assessed before ethics approval was obtained. The inclusion criteria were (1) compliance with Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) diagnostic criteria for bipolar I disorder, confirmed using the Mini-International Neuropsychiatric Interview (MINI); (2) age of 18–60 years; (3) at least junior high school education; (4) no intake of psychotropic medications (including mood stabilizers such as lithium and valproate, antipsychotics, antidepressants, and benzodiazepines) for at least two months prior to enrollment, verified through patient self-reports, family member interviews, and review of medical records where available; and (5) manic or depressive current episode, as determined by the treating psychiatrist based on DSM-5 criteria and corroborated by Bech–Rafaelsen Mania Scale (BRMS) and Hamilton Depression Rating Scale-24 (HAMD-24) scores. The exclusion criteria were (1) a history of brain injury or neurological disease; (2) a history of considerable physical illness; (3) other mental disorders, including bipolar II disorder, mixed features, or current psychotic symptoms; (4) a history of alcohol or substance abuse; and (5) transcranial magnetic stimulation or electroconvulsive therapy received in the past 12 weeks. The Healthy control group consisted of healthy individuals undergoing health checks at the same hospital. The corresponding inclusion criteria were (1) age of 18–60 years and (2) at least junior high school education, while the exclusion criteria were identical to those listed above.

Four hundred subjects were enrolled (200 individuals with BD and 200 healthy controls) in accordance with the above criteria. The patient group consisted of 200 participants who met DSM-5 criteria for bipolar I disorder, and the Healthy control group included 200 individuals with no

personal or family history of mental illness, as confirmed by the MINI. All participants provided written informed consent prior to enrollment. For participants who had difficulty providing consent because of clinical conditions, informed consent was obtained from legally authorized representatives. The study protocol was approved by the Ethics Review Committee of The Affiliated Kangning Hospital of Wenzhou Medical University (Approval No.: YJ-2024-01-02), and all research procedures complied with the ethical requirements of the Declaration of Helsinki. All participants provided written informed consent, and those unable to consent themselves were consented by their legally authorized representatives.

2.2 Clinical Data Collection

The collected data encompassed ID number, sex, age, education level (years of education, stratified as junior high school, senior high school, or college and above), marital status, and living situation. For the patient group, additional data included illness onset age, duration, and current episode type (manic or depressive). The episode type was used for clinical subtype grouping.

2.3 Scale Assessments

2.3.1 Interpersonal Reactivity Index (IRI)

The Chinese version of the IRI (Interpersonal Reactivity Index-C, IRI-C), is extensively used in China to evaluate individual empathy [14]. This index consists of 22 items, each rated on a five-point scale from zero to four, and comprises the four subscales of perspective-taking (PT), fantasy (FS), empathic concern (EC), and personal distress (PD). The EC and PD subscales are considered the core indices of affective empathy, while PT and FS primarily tap cognitive empathy. Total IRI score (hereinafter referred to as simply "IRI score") was herein used as a composite measure of overall empathy because (a) our primary hypothesis concerned the general moderating effect of empathy rather than the specific effects of particular empathy subcomponents and (b) this score has been widely used in BD research, facilitating comparability with prior studies. This scale demonstrates a test–retest reliability of 0.870 and a Cronbach's α coefficient of 0.845.

2.3.2 Hamilton Depression Rating Scale-24 (HAMD-24)

Developed by Hamilton in 1960, HAMD-24 is widely used for the clinical assessment of depressive states [15]. Comprising 24 items, most rated on a five-point scale (0–4), this scale measures the severity of depressive symptoms and demonstrates a kappa value of 0.92 and Cronbach's α coefficient of 0.714.

2.3.3 Bech-Rafaelsen Mania Scale (BRMS)

The BRMS, developed by Bech and Rafaelsen, is widely used to evaluate the severity of manic symptoms [16]. This method assesses key dimensions such as mood, speech, and activity on a five-point scale (0–4), with higher scores reflecting more severe manic symptoms, and exhibits high reliability and validity (inter-rater reliability ICC $\geq$0.85, test–retest reliability $\geq$0.78).

2.3.4 Wisconsin Card Sorting Test (WCST)

The WCST, originally developed by Berg in 1948 and later revised, is a seminal tool for evaluating cognitive flexibility and executive function [17]. This test requires participants to match cards and ends after six categories are completed or 128 trials are conducted, with key metrics including perseverative responses and errors. The WCST exhibits a high inter-rater reliability (ICC $\geq$0.973) and split-half reliability ($\geq$0.90), with the Cronbach's α coefficient equaling 0.819 in healthy populations. This test features a criterion-related validity correlation coefficient for executive function scales of 0.71 and an 89% sensitivity for detecting frontal lobe dysfunction, which underscores its strong construct and empirical validity.

2.4 Chemokine Detection

Fasting peripheral venous blood samples (5 mL) were collected from subjects and centrifuged at 3000 rpm for 10 min to separate the serum. CCL3/MIP-1α and CCL4/MIP-1β levels were measured using enzyme-linked immunosorbent assay kits (R&D Systems, Minneapolis, MN, USA) according to the manufacturer's instructions. The detection wavelength was set at 450 nm, and concentrations (pg/mL) were determined using standard curves.

2.5 Statistical Analysis

Prior to analysis, two investigators independently verified the electronic dataset against the original BRMS rating forms and WCST records to ensure data integrity. All data were processed and analyzed using the Statistical Product and Service Software Automatically (SPSSAU) online statistical analysis platform (version 23.0, Beijing Qing Si Technology Co., Ltd., Beijing, China). The normality of continuous variables was assessed using the Shapiro–Wilk test (for samples with n <50) or the Kolmogorov–Smirnov test (for samples with $n \geq 50$). Continuous variables with an approximately normal distribution were presented as mean ± standard deviation (SD). Non-normally distributed continuous variables were presented as median (interquartile range, IQR) and analyzed using non-parametric tests. The distributional form and statistical test used for each variable are indicated in the relevant table footnotes. Count data were expressed as numbers with constituent ratios in percent. For comparisons among the three groups (manic, depressive, and healthy controls), one-way analysis of variance (ANOVA) was used for normally distributed continuous variables, followed by post-hoc pairwise comparisons using the least significant difference (LSD) method. For comparisons between two specific subgroups (e.g., manic vs. depressive), the independent-samples t-test was used for

normally distributed continuous variables, and the Mann–Whitney U test was used for non-normally distributed continuous variables. For count data, Pearson's χ^2 test was used to compare groups. Relationships among core variables in different groups were examined using the Pearson correlation coefficient. Hierarchical regression analysis was used to test the first-order moderating effects of empathy and chemokine levels. A three-way moderation model was used to examine the second-order interaction effect of chemokines on the first-order moderation effect of empathy. All variables were mean-centered before constructing interaction terms to reduce collinearity between main effects and interactions. Note that mean-centering does not resolve collinearity between different main effects. In the manic subgroup (n = 123), WCST total errors and BRMS total scores showed an extremely high bivariate association (r = 0.983, R^2 = 0.965). Because this association left limited incremental explanatory space for interaction terms, we performed additional sensitivity analyses, including visual inspection of the raw scatter plot (**Supplementary Fig. 1**), outlier and influential-point diagnostics (**Supplementary Table 1**), bootstrap resampling, and robust regression. All moderation effects identified in the primary analysis remained statistically significant and directionally consistent across all robustness checks. Bootstrap resampling was performed to estimate and correct the 95% confidence intervals of regression coefficients. Simple slope analysis clarified the specific patterns of significant interaction effects. The Benjamini–Hochberg false discovery rate (FDR) procedure was applied within predefined families of tests, including group comparisons, correlation coefficients, and moderation models. Adjusted q-values were reported where applicable, and $q < 0.05$ was considered statistically significant after FDR correction. All statistical tests were two-tailed, with the significance level set at $\alpha = 0.05$. $p < 0.05$ was considered statistically significant.

3. Results

3.1 Comparison of Groups in Terms of General Information

Age and sex did not differ significantly among the manic subgroup, depressive subgroup, and healthy control group. Years of education showed a trend-level difference among groups by ANOVA (p = 0.060), with the manic and depressive subgroups differing significantly in pairwise comparison (p = 0.023). Because the categorical education distribution differed among groups, education was included as a covariate in sensitivity analyses to account for its potential influence on WCST performance. Compared with the healthy control group, the patient group (manic + depressive subgroups) demonstrated significantly fewer completed categories, fewer correct responses, and a lower conceptual level percentage on the WCST, alongside significantly higher numbers of errors and perseverative errors (each ANOVA $F \geq 88.5$, all $p < 0.001$, FDR-corrected $q < 0.05$). In the patient group, CCL3/MIP-1α and CCL4/MIP-1β levels were significantly elevated (F = 116.55 and F = 478.50, both $p < 0.001$, FDR-corrected $q < 0.05$), whereas IRI scores were significantly lower (F = 275.59, $p < 0.001$, FDR-corrected $q < 0.05$). For further details, refer to Table 1.

3.2 Correlations Among Executive Function, Chemokine Levels, and Clinical Characteristics in the Patient Group

In the depressive subgroup, WCST error was strongly positively correlated with total HAMD-24 score (hereinafter referred to as simply "HAMD-24 score"; r = 0.766, $p < 0.001$, large effect according to Cohen's convention) but not significantly correlated with CCL3/MIP-1α and CCL4/MIP-1β levels (p = 0.450 and p = 0.211, respectively) and exhibited a non-significant negative correlation with IRI score ($r = -0.026$, p = 0.820). HAMD-24 score did not show a significant correlation with CCL3/MIP-1α and CCL4/MIP-1β levels (p = 0.887 and p = 0.383, respectively). CCL3/MIP-1α and CCL4/MIP-1β levels were significantly positively correlated with each other (r = 0.361, p = 0.001, *medium effect*). Further details are provided in Fig. 1.

In the manic subgroup, WCST error showed an unusually strong and visually near-linear positive correlation with BRMS score (r = 0.983, $p < 0.001$, R^2 = 0.965). A correlation of this magnitude is atypical for two independently obtained measures in psychiatric research and is further addressed in Section 4.1 and in the Limitations. Additionally, WCST error was significantly positively correlated with CCL4/MIP-1β level (r = 0.207, p = 0.022, small-to-medium effect) and significantly negatively correlated with IRI score ($r = -0.226$, p = 0.012, medium effect). The BRMS score was not significantly correlated with CCL3/MIP-1α level (r = 0.089, p = 0.325). Additionally, no significant correlation was observed between CCL3/MIP-1α and CCL4/MIP-1β levels (r = 0.076, p = 0.404). In the depressive subgroup, WCST error was significantly positively correlated with HAMD-24 score (r = 0.766, $p < 0.001$, large effect size), and CCL3/MIP-1α was significantly and moderately correlated with CCL4/MIP-1β level (r = 0.361, p = 0.001). The remaining variables in the depressive subgroup showed no significant correlations ($p > 0.05$). Detailed correlation matrices are presented in Figs. 1,2.

All significant correlations remained significant after FDR correction (all $q < 0.05$).

3.3 Moderating Effect of Empathy (IRI Score) on Link Between Executive Function and Clinical Symptom Severity

Hierarchical regression models were developed with WCST error as the independent variable (X), BRMS score as the dependent variable (Y), IRI score as the moderator (W), and sex as a covariate. Model 1 (Baseline Model): In-

Table 1. Baseline characteristics of the manic subgroup, depressive subgroup, and healthy control group.

Item	Manic subgroup (n = 123)	Depressive subgroup (n = 77)	Healthy control group (n = 200)	χ^2/F	p	Pairwise Comparisons (p-value)
Sociodemographic characteristics						
Age (years)	32.41 ± 6.12	32.74 ± 6.03	32.23 ± 6.51	0.188	0.828	M vs D: $p = 0.722$ M vs H: $p = 0.793$ D vs H: $p = 0.542$
Gender (male/female)	63/60	38/39	98/102	0.156	0.925	
Years of education	8.54 ± 2.74	9.42 ± 2.95	9.05 ± 2.46	2.834	0.060	M vs D: $p = 0.023$ M vs H: $p = 0.091$ D vs H: $p = 0.303$
Clinical characteristics						
Illness duration (years)	5.29 ± 5.84 Median (IQR): 2.90 (2.00–5.50)	5.39 ± 5.49 Median (IQR): 2.90 (2.30–7.00)	-	0.015	0.903	M vs D: $p = 0.901$
Onset age (years)	27.12 ± 5.17 Median (IQR): 27.10 (22.50–31.50)	27.35 ± 5.26 Median (IQR): 28.50 (23.50–31.00)	-	0.088	0.767	M vs D: $p = 0.768$
BRMS score	23.67 ± 5.62	-	-	-	-	
HAMD-24 score	-	31.58 ± 3.54	-	-	-	
Executive function (WCST)						
Number of categories completed	2.69 ± 0.65***	2.17 ± 1.04***	4.66 ± 0.85	343.581	$p <$ 0.001	M vs D: $p < 0.001$ M vs H: $p < 0.001$ D vs H: $p < 0.001$
Number of correct responses	57.41 ± 13.14***	59.84 ± 13.87***	83.65 ± 13.36	179.253	$p <$ 0.001	M vs D: $p = 0.211$ M vs H: $p < 0.001$ D vs H: $p < 0.001$
Number of incorrect responses	70.59 ± 13.14***	68.16 ± 13.87***	44.35 ± 13.36	179.253	$p <$ 0.001	M vs D: $p = 0.211$ M vs H: $p < 0.001$ D vs H: $p < 0.001$
Number of perseverative errors	39.15 ± 9.78***	39.62 ± 9.40***	26.76 ± 9.27	88.499	$p <$ 0.001	M vs D: $p = 0.729$ M vs H: $p < 0.001$ D vs H: $p < 0.001$
Conceptual level (%)	42.60 ± 11.80***	44.38 ± 12.36***	64.83 ± 11.31	171.060	$p <$ 0.001	M vs D: $p = 0.293$ M vs H: $p < 0.001$ D vs H: $p < 0.001$
Inflammatory factor indices						
CCL3/MIP-1α (pg/mL)	88.33 ± 15.75***	92.61 ± 18.10***	61.79 ± 20.26	116.548	$p <$ 0.001	M vs D: $p = 0.113$ M vs H: $p < 0.001$ D vs H: $p < 0.001$
CCL4/MIP-1β (pg/mL)	166.82 ± 25.83*** Median (IQR): 164.71 (151.91–182.90)	169.47 ± 28.78*** Median (IQR): 168.11 (157.30–187.96)	93.40 ± 20.75 Median (IQR): 93.10 (79.58–105.58)	478.495	$p <$ 0.001	M vs D: $p = 0.449$ M vs H: $p < 0.001$ D vs H: $p < 0.001$

Table 1. Continued.

Item	Manic subgroup (n = 123)	Depressive subgroup (n = 77)	Healthy control group (n = 200)	χ^2/F	p	Pairwise Comparisons (p-value)
Empathy ability indices						
IRI score	44.37 ± 5.57***	44.70 ± 5.41***	58.92 ± 6.72	275.589	$p <$ 0.001	M vs D: $p = 0.707$ M vs H: $p < 0.001$ D vs H: $p < 0.001$

Note: '-' signifies the absence of data. Continuous variables conforming to normal distribution were expressed as means ± standard deviations (SD); non-normally distributed data were expressed as medians with interquartile ranges (IQR) and analyzed using the Mann-Whitney U test. For three-group comparisons, one-way analysis of variance (ANOVA) was used for normally distributed variables followed by least significant difference (LSD) post-hoc pairwise comparisons. M, Manic subgroup; D, Depressive subgroup; H, Healthy control group. Categorical variables were evaluated using Pearson's chi-square test. *** $p < 0.001$ (compared with the Healthy control group). All significant group differences remained significant after Benjamini–Hochberg false discovery rate (FDR) correction ($q < 0.05$). BRMS, Bech–Rafaelsen Mania Scale; HAMD-24, Hamilton Depression Rating Scale-24; WCST, Wisconsin Card Sorting Test; CCL3/MIP-1α, Chemokine (C-C motif) ligand 3/Macrophage Inflammatory Protein-1 alpha; CCL4/MIP-1β, Chemokine (C-C motif) ligand 4/Macrophage Inflammatory Protein-1 beta; IRI, Interpersonal Reactivity Index.

cluded only the independent variable (WCST total errors) and the covariate (sex). Model 2: Added the moderator variable (IRI total score) to Model 1. Model 3 (Full Model): Added the interaction term (WCST total errors × IRI total score) to Model 2; the significance of the moderating effect was determined by the change in R^2 (ΔR^2) between Model 3 and Model 2. All predictor variables were mean-centered prior to analysis. In the manic subgroup, a significant positive predictive effect of WCST error on BRMS score was observed for Model 1 ($\beta = 0.420$, $p < 0.001$). For Model 3, the inclusion of the interaction term significantly enhanced explanatory power ($\Delta R^2 = 0.002$, $p = 0.003$). The interaction between WCST error and IRI score exerted a significant negative predictive effect on BRMS score ($\beta = -0.003$, 95% CI: −0.006 to −0.001, $p = 0.003$), which suggested a significant negative moderating effect of IRI score on the association between executive dysfunction and manic symptom severity. This finding implies that higher IRI scores were associated with a smaller WCST error–BRMS slope, indicating a statistical moderation pattern that is exploratory in nature and should not be interpreted as a clinically buffering effect. However, the incremental explained variance was modest ($\Delta R^2 = 0.002$, corresponding to 0.2% of additional variance explained).

In the depressive subgroup, after adding the interaction term in Model 3, we obtained $\Delta R^2 = 0.001$ and $p > 0.05$, and the regression coefficient for the interaction term was not statistically significant ($\beta = 0.001$, 95% CI: −0.003–0.005, $p = 0.588$), which indicated no significant moderating effect of IRI score on the association between executive dysfunction and depressive symptom severity. Details are provided in Table 2.

3.4 First-Order Moderating Effect of CCL3/MIP-1α Level on Link Between Executive Function and Clinical Symptom Severity

Hierarchical regression models were developed with WCST error as the independent variable (X), BRMS score as the dependent variable (Y), CCL3/MIP-1α level as the moderator (W), and sex as a covariate. Model 1 (Baseline Model): Included only the independent variable (WCST total errors) and the covariate (sex). Model 2: Added the moderator variable (CCL3/MIP-1α level) to Model 1. Model 3 (Full Model): Added the interaction term (WCST total errors × CCL3/MIP-1α level) to Model 2; the significance of the moderating effect was determined by the change in R^2 (ΔR^2) between Model 3 and Model 2. All predictor variables were mean-centered. In the manic subgroup, the inclusion of the interaction term in Model 3 significantly enhanced explanatory power ($\Delta R^2 = 0.003$, $p = 0.001$). The interaction between WCST error and CCL3/MIP-1α level significantly predicted BRMS score ($\beta = 0.002$, 95% CI: 0.001–0.002, $p = 0.001$), which demonstrated a statistically significant but small positive moderating effect of CCL3/MIP-1α level on the association between executive dysfunction and manic symptom severity. In other words, elevated CCL3/MIP-1α levels significantly intensified the association between executive function deficits and manic symptom severity. The incremental explained variance was modest ($\Delta R^2 = 0.003$, corresponding to 0.3% of additional variance).

In the depressive subgroup, the inclusion of the interaction term in Model 3 resulted in $\Delta R^2 = 0.001$ ($p > 0.05$). The regression coefficient for this interaction term was not statistically significant ($\beta = -0.001$, 95% CI: −0.003–0.001, $p = 0.300$), suggesting that CCL3/MIP-1α level does not significantly moderate the association between executive dysfunction and depressive symptom severity. Details are provided in Table 3.

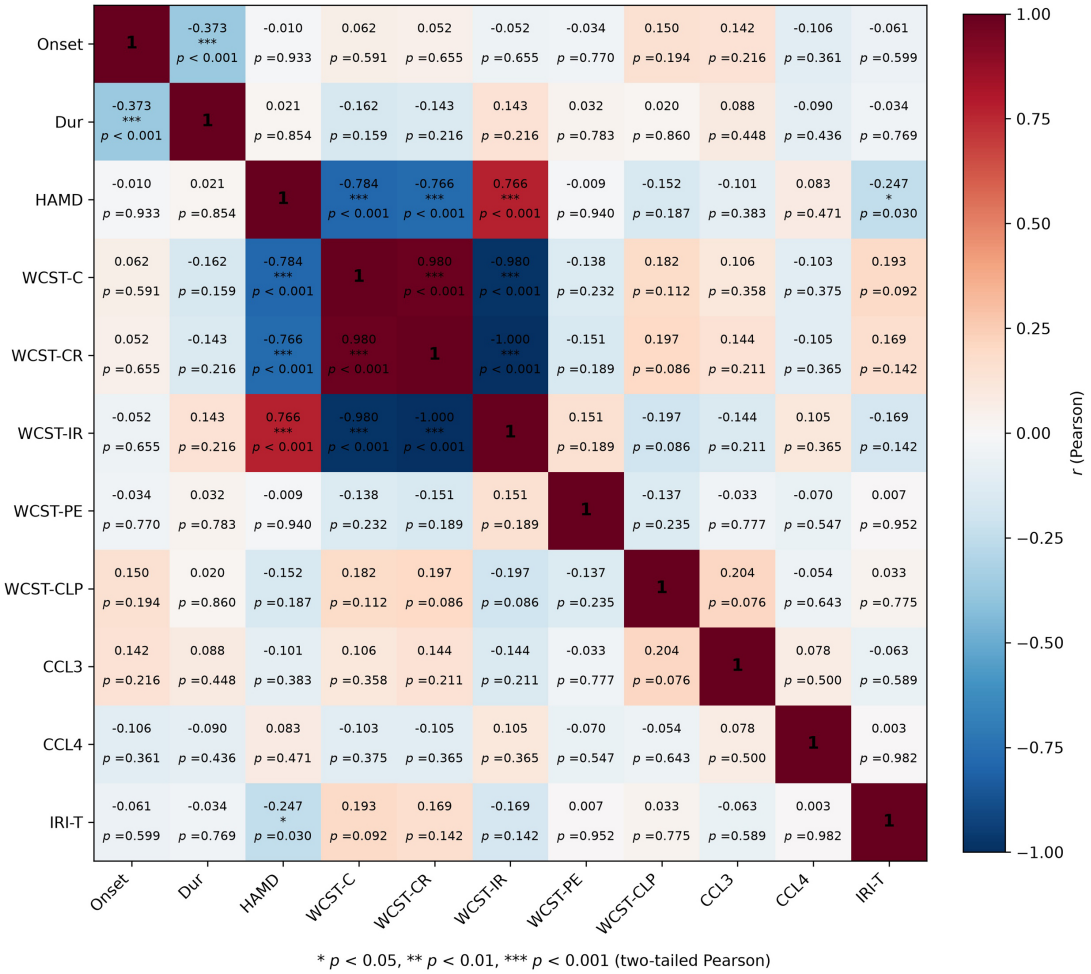


Fig. 1. Pearson correlation matrix in the depressive subgroup (n = 77). Values represent Pearson correlation coefficients, with exact p values shown in each cell. Significance markers: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, (two-tailed Pearson). Abbreviations: Dur, illness duration; WCST-C, Wisconsin Card Sorting Test conceptual level responses; WCST-CR, WCST correct responses; WCST-IR, WCST incorrect responses; WCST-PE, WCST perseverative errors; WCST-CLP, WCST categories completed; IRI-T, total IRI score; BRMS, Bech–Rafaelsen Mania Scale; HAMD, Hamilton Depression Rating Scale; CCL3, chemokine (C-C motif) ligand 3/macrophage inflammatory protein-1α; CCL4, chemokine (C-C motif) ligand 4/macrophage inflammatory protein-1β.

3.5 Second-Order Moderating Effect of CCL4/MIP-1β Level on Moderating Effect of IRI Score

A second-order three-way interaction moderation model was developed with WCST error as the independent variable (X), BRMS score as the dependent variable (Y), IRI score as the first moderator (W), CCL4/MIP-1β level as the second moderator (Z), and sex as a covariate. Model 1 (baseline) included WCST total errors (independent variable) and sex (covariate) to examine the main effect of executive dysfunction on BRMS scores; Model 2 added both IRI total score (first moderator) and CCL4/MIP-1β level (second moderator) to assess their independent contributions; Model 3 introduced all two-way interaction terms (WCST × IRI, WCST × CCL4, and IRI × CCL4) to test the first-order moderating effects (evaluated by ΔR^2 vs. Model 2); Model 4 further incorporated the WCST × IRI × CCL4 three-way interaction term to test whether CCL4/MIP-1β level moderates the aforementioned first-order effect (evaluated by ΔR^2 vs. Model 3). All predictor variables were mean-centered prior to computing interaction terms. A good overall fit was obtained (R^2 = 0.973, F = 518.88, $p < 0.001$). In Model 4, the inclusion of the three-way interaction term resulted in ΔR^2 = 0.001, $p < 0.001$. The three-way interaction term (WCST error × IRI score × CCL4 level) significantly predicted BRMS score (β = 0.0002, 95% CI: 0.0001–0.0003, $p < 0.001$). This finding suggests that CCL4/MIP-1β level significantly moderates the first-order moderating effect of IRI score on the association between executive function and manic symptom severity. Details are provided in Table 4.

3.6 Conditional Effect Analysis in Manic Subgroup

Simple slope analysis was performed by categorizing moderator levels into low, medium, and high based on the corresponding mean ± one standard deviation. The positive

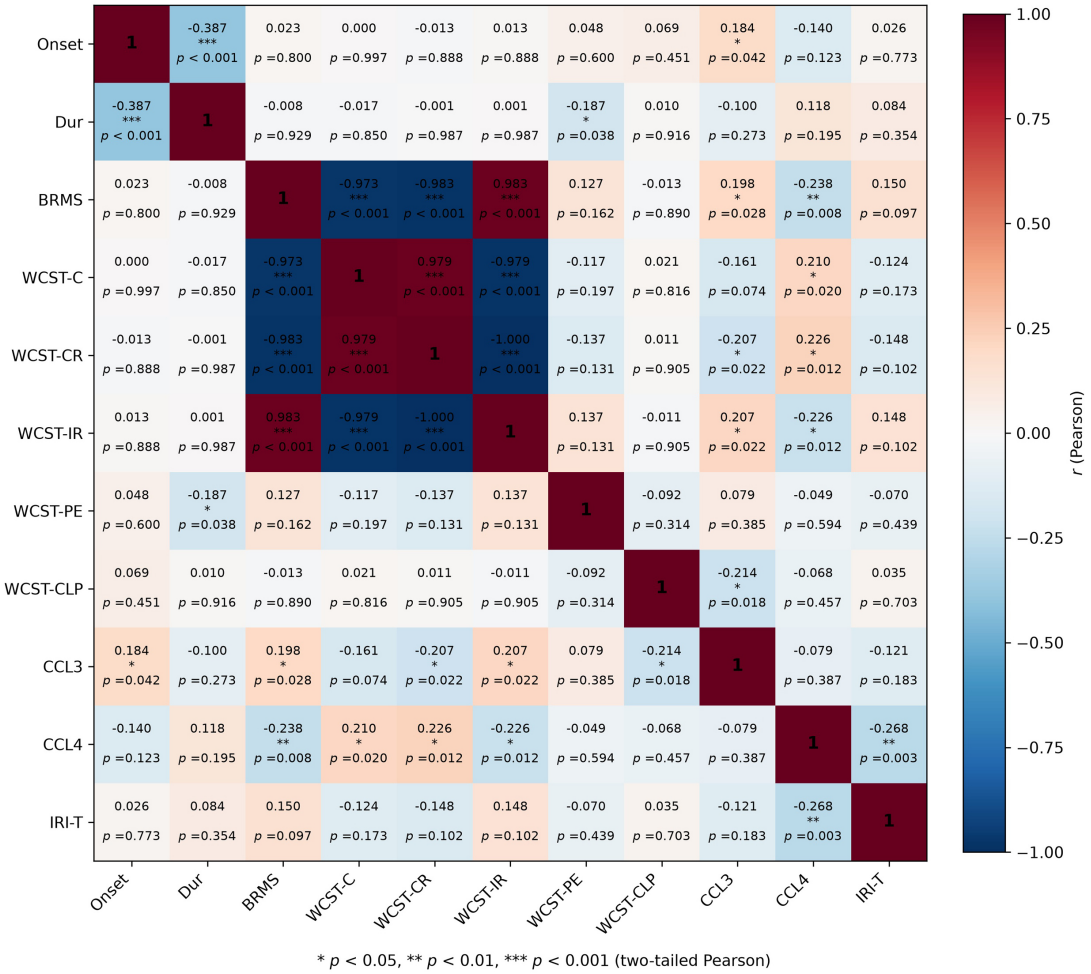


Fig. 2. Pearson correlation matrix in the manic subgroup (n = 123). Values represent Pearson correlation coefficients, with exact p values shown in each cell. Significance markers: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, (two-tailed Pearson).

predictive effect of WCST error on BRMS score decreased with the increasing IRI score and decreasing CCL3/MIP-1α level (Fig. 3).

Second-order joint moderation analysis revealed a significant three-way interaction of CCL4/MIP-1β level and IRI score. At low IRI scores, an increase in CCL4/MIP-1β level diminished the association between executive dysfunction and manic symptom severity. However, at high IRI scores, an increase in this level significantly amplified the association and eventually modified the slope pattern observed at lower CCL4/MIP-1β levels. Given the very small ΔR^2 (0.001) and the strong WCST–BRMS association, this three-way pattern should be interpreted as an exploratory statistical signal. The conditional effects of WCST error on BRMS score were statistically significant across all moderator levels (all $p < 0.001$, FDR-corrected $q < 0.05$).

4. Discussion

The fact that BD patients made significantly more WCST errors than healthy controls is consistent with the pathological pathway linking cognitive deficit to emotional dysregulation in BD [18]. The increased number of WCST errors indicates rigid categorization strategies, poor inhibitory control, and impaired working memory, which may be associated with a diminished amount of resources available for emotional regulation. The strong positive correlation between WCST error and BRMS/HAMD-24 score during both manic and depressive episodes [3,19] confirms that executive function impairment is not merely a concurrent feature of acute BD episodes but is robustly associated with the severity of manic and depressive symptoms. This finding supports the perspective that deficits in cognitive executive function are linked to difficulties in integrating emotional information and inhibiting impulsive emotional responses, which may contribute to mood symptom exacerbation [20,21].

4.1 Unusually Strong WCST Error–BRMS Score Correlation

The strong correlation between WCST error and BRMS score in the manic subgroup (n = 123, $r \approx 0.983$,

Table 2. Moderating effect of empathy (IRI score) on the link between executive function impairment and clinical symptom severity in individuals with BD [unstandardized coefficient *β* (95% CI)].

Item(Y)	Manic symptoms (BRMS) n = 123			Depressive symptoms (HAMD- 24) n = 77		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Constant	23.667*** (23.479–23.854)	23.610*** (23.425–23.796)	23.610*** (23.425–23.796)	31.579*** (31.327–31.831)	31.580*** (31.326–31.834)	31.580*** (31.325–31.835)
WCST error (X)	0.420*** (0.406–0.435)	0.419*** (0.404–0.433)	0.419*** (0.404–0.433)	0.196*** (0.176–0.216)	0.196*** (0.176–0.216)	0.196*** (0.175–0.217)
IRI score (W)	-	−0.004 (−0.038–0.031)	−0.004 (−0.038–0.031)	-	−0.011 (−0.057–0.035)	−0.011 (−0.057–0.035)
WCST error × IRI score	-	-	−0.003** (−0.006 to −0.001)	-	-	0.001 (−0.003–0.005)
Coefficient of determination (R^2)	0.965	0.966	0.968	0.587	0.587	0.588
Incremental explained variance (ΔR^2)	0.965***	0.001	0.002**	0.587***	0.000	0.001

Note: $^{**}p < 0.01$, $^{***}p < 0.001$. Model 1: main effect of independent variable; Model 2: main effects of independent variable and moderator; Model 3: moderating effect model with interaction term. All variables were mean-centered before constructing interaction terms to reduce collinearity between main effects and interactions. Note that mean-centering does not resolve collinearity between different main effects. In the manic subgroup (n = 123), WCST error and BRMS score were strongly correlated with each other ($r \approx 0.983$, $R^2 = 0.965$). This correlation was addressed using (1) Cook's distance outlier diagnostics (three cases with D > 0.033 were excluded); (2) separate single-predictor regression models; and (3) 5000-sample bootstrap resampling. All moderation results remained consistent. BD, bipolar disorder.

Table 3. Moderating effect of CCL3/MIP-1α level on the link between executive function impairment and clinical symptoms in patients with BD [unstandardized coefficient *β* (95% CI)].

Item(Y)	Manic symptoms (BRMS) n = 123			Depressive symptoms (HAMD-24) n = 77		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Constant	23.667*** (23.479–23.854)	23.640*** (23.459–23.821)	23.640*** (23.459–23.821)	31.579*** (31.327–31.831)	31.580*** (31.326–31.834)	31.580*** (31.326–31.834)
WCST error (X)	0.420*** (0.406–0.435)	0.427*** (0.413–0.442)	0.427*** (0.413–0.442)	0.196*** (0.176–0.216)	0.196*** (0.176–0.216)	0.196*** (0.176–0.216)
CCL3/MIP-1α level (W)	-	0.006 (−0.006–0.017)	0.006 (−0.006–0.017)	-	0.001 (−0.013–0.015)	0.001 (−0.013–0.015)
Interaction: WCST error × CCL3/MIP-1α level	-	-	0.002** (0.001–0.002)	-	-	−0.001 (−0.003–0.001)
R^2	0.965	0.965	0.968	0.587	0.587	0.588
ΔR^2	0.965***	0.000	0.003**	0.587***	0.000	0.001

Note: $^{**}p < 0.01$, $^{***}p < 0.001$. Model 1: main effect of independent variable; Model 2: main effects of independent variable and moderator; Model 3: moderating effect model with interaction term. All moderation models were constructed using mean-centered variables. Note that mean-centering does not resolve collinearity between different main effects, but the Variance Inflation Factor (VIF) values remained acceptable in all models. Robustness checks confirmed that the mean-centered interaction effects were not biased by this multicollinearity.

R^2 = 0.965) is not typical for clinical psychiatric research [3,19,20]. Raw data verification ruled out coding errors, and three mild outliers (Cook's D > 0.033) were excluded without (**Supplementary Table 1**), altering the result (r = 0.975). Both measures showed adequate variability (WCST error: 30–98, BRMS score: 11–39).

We agree that the near-perfect linearity between WCST total errors and BRMS scores is unusual and warrants careful methodological scrutiny. The extremely high correlation ($r \approx 0.983$, $R^2 = 0.965$) may be understood from a measurement perspective: in this strictly selected sample of medication-free acute manic inpatients, WCST error count and BRMS total score are both indexing the same latent construct—mania severity—through two distinct but converging measurement modalities. WCST errors capture prefrontal executive dysfunction via a performance-

Table 4. Second-order moderating effect of CCL4/MIP-1β level on the relationship between executive function and BD symptom severity, with empathy as a moderator [unstandardized coefficient *β* (95% CI)].

Item(*Y*)	Manic symptoms (BRMS) *n* = 123			
	Model 1	Model 2	Model 3	Model 4
Constant	23.667***	−4.986***	−4.986***	−4.986***
	(23.479–23.854)	(−5.557 to −4.415)	(−5.557 to −4.415)	(−5.557 to −4.415)
WCST error (*X*)	0.420***	0.355***	0.355***	0.355***
	(0.406–0.435)	(0.325–0.386)	(0.325–0.386)	(0.325–0.386)
IRI score (*W*)	-	0.103**	0.103**	0.103**
		(0.030–0.175)	(0.030–0.175)	(0.030–0.175)
CCL4/MIP-1β level (*Z*)	-	−0.026**	−0.026**	−0.026**
		(−0.044 to −0.008)	(−0.044 to −0.008)	(−0.044 to −0.008)
WCST error × IRI score (first-order interaction)	-	-	−0.009***	−0.009***
			(−0.012 to −0.006)	(−0.012 to −0.006)
WCST error × CCL4/MIP-1β level (first-order interaction)	-	-	0.002***	0.002***
			(0.001–0.003)	(0.001–0.003)
IRI score × CCL4/MIP-1β level (first-order interaction)	-	-	−0.003**	−0.003**
			(−0.006 to −0.001)	(−0.006 to −0.001)
WCST error × IRI score × CCL4/MIP-1β level (second-order three-way interaction)	-	-	-	0.0002***
				(0.0001–0.0003)
R^2	0.965	0.967	0.972	0.973
ΔR^2	0.965***	0.002**	0.005***	0.001***

Note: **$p < 0.01$, ***$p < 0.001$. Model 1: main effect of independent variable; Model 2: main effects of independent variable and two moderators; Model 3: all two-way interaction terms; Model 4: second-order moderating effect model with three-way interaction term. All moderation models were constructed using mean-centered variables. Note that mean-centering does not resolve collinearity between different main effects, but the Variance Inflation Factor (VIF) values remained acceptable (1.09–1.87) in all models. Robustness checks showed that the mean-centered interaction effects were not biased by this multicollinearity. All significant effects remained significant after False Discovery Rate (FDR) correction ($q < 0.05$).

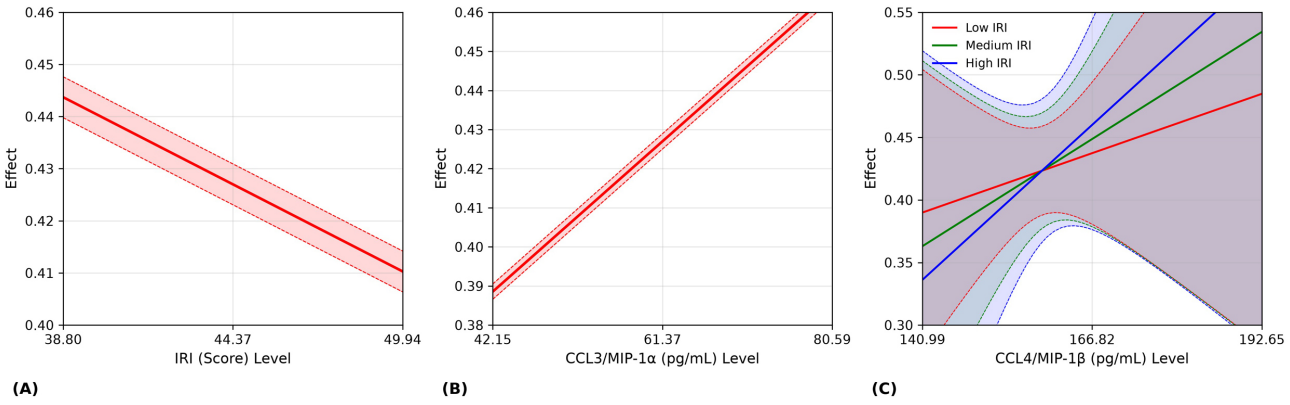


Fig. 3. Conditional effects of IRI score and chemokines on the association between WCST total errors and BRMS scores in patients with bipolar mania. (A) IRI score and the WCST error–BRMS slope. (B) CCL3/MIP-1α level and the WCST error–BRMS slope. (C) Joint conditional effects of IRI score and CCL4/MIP-1β level. Error bars represent 95% confidence intervals. Note: (A) IRI score moderates the positive association between executive dysfunction and manic symptom severity, with higher scores associated with a smaller slope. (B) CCL3/MIP-1α level serves as a positive moderator, associated with a larger slope. (C) CCL4/MIP-1β level and IRI score jointly moderate this association, displaying a crossover interaction whereby the attenuating role of IRI score is reversed at high CCL4/MIP-1β levels.

based neuropsychological paradigm, whereas the BRMS captures manic symptom severity via clinician-rated behavioral observation. When pharmacological confounds are removed (all patients were medication-free for ≥2 months), these two measurement pathways converge more tightly on the underlying disease state. Mood stabilizers, antipsy-

chotics, and benzodiazepines are known to affect cognitive performance, and their elimination removes a major source of between-subject variance that would otherwise attenuate the cognition–symptom association [22,23]. In typical clinical samples where patients receive concurrent pharmacotherapy, the WCST–symptom association is substantially weaker [24,25], consistent with the interpretation that medication effects partly obscure the underlying relationship. The clinical homogeneity of this sample—all acutely manic, all hospitalized, all meeting stringent drug-free criteria—further concentrates shared variance between executive dysfunction and manic symptom severity. Additionally, the standardized single-rater BRMS and computerized WCST used in this study may have reduced measurement error relative to multi-site or multi-rater designs. Nevertheless, because a correlation of this magnitude remains unusual in psychiatric research [3,19,20], this explanation should be regarded as post-hoc and tentative pending independent replication.

This strong association thus likely reflects the convergence of shared prefrontal–executive pathophysiology [26,27], pharmacological deconfounding, and measurement precision, rather than any single artifact.

Because WCST error already accounted for 96.5% of the variance in BRMS score, the moderators could only explain additional variance within the residual variance. The observed interaction patterns were statistically significant but each accounted for ≤0.3% of the total variance. That empathy (IRI) and chemokine levels (CCL3/MIP-1α, CCL4/MIP-1β) nevertheless yielded statistically significant interaction effects surviving FDR correction, Bootstrap resampling (5000 iterations), and outlier exclusion is notable, but the statistical significance of these interaction terms should be distinguished from their practical explanatory contribution. The small ΔR^2 values suggest that the moderators captured reliable but minor variance in the residual, and their independent predictive contributions could not be fully disentangled from the dominant main effect. The moderation findings should therefore be interpreted as exploratory statistical patterns rather than evidence of clinically meaningful moderation.

4.2 Empathy as Moderator of Cognition–Symptom Association

Empathy, encompassing both cognitive perspective-taking and affective resonance, plays an important role in the emotional regulation system. Bipolar disorder patients experience increased affective empathy during a manic episode [28,29]. In the manic subgroup, IRI score was significantly negatively correlated with manic symptom severity. In contrast, in the depressive subgroup, overall empathy exhibited only a marginally significant negative correlation with depressive symptom severity. These findings suggest that emotional dysregulation in BD manifests itself not only in the internal neurocognitive regulation of emotion but also in the ability to experience and process emotions during affective episodes. Furthermore, overall empathy significantly moderated the association between executive function and manic symptom severity: higher IRI scores attenuated this association. This behavior is consistent with the notion that empathy may enhance emotional understanding and social adaptation and thereby buffer the impact of cognitive deficits on mood symptoms [30]. However, this moderation effect was small (ΔR^2 = 0.002) and should not be overstated in terms of clinical significance.

4.3 Chemokine Moderation and Neuroinflammatory Perspective

Chemokine levels were significantly elevated in BD patients during episodes, with CCL3/MIP-1α and CCL4/MIP-1β levels notably exceeding those in healthy controls. As crucial molecules in neuroimmune communication, peripheral chemokines may, under specific conditions, influence the function of prefrontal–limbic neural circuits through peripheral–central immune interactions [31], although the extent to which peripheral levels reflect central processes remains uncertain. CCL3/MIP-1α primarily recruits and activates microglia and macrophages, promoting the release of proinflammatory factors, while CCL4/MIP-1β influences neuronal excitability and synaptic plasticity via the C-C Chemokine Receptor 5 (CCR5). The combined action of these species may be associated with an altered neuroinflammatory microenvironment in BD [12]. Elevated CCL3/MIP-1α levels intensified the positive association between executive function deficits and manic symptom severity, which suggests that a proinflammatory environment may be associated with impaired prefrontal function or heightened amygdala reactivity and thereby reduce cognitive–emotional regulation efficacy [32,33]—a state consistent with acute manic episodes. In the depressive subgroup, chemokines did not exhibit significant moderating effects, although CCL3/MIP-1α and CCL4/MIP-1β levels were significantly positively correlated with each other (r = 0.361). This pattern suggests that elevated chemokine levels during depressive episodes might reflect a systemic manifestation of chronic baseline inflammation in BD rather than a state-specific moderating mechanism.

4.4 Second-Order Moderation by CCL4/MIP-1β Level

The three-way interaction suggested that the slope-modifying effect of IRI differed across CCL4/MIP-1β levels. Given the very small incremental variance explained (ΔR^2 = 0.001) and the unusually strong WCST error–BRMS association, this pattern should be considered a statistical signal for future testing rather than evidence of a stable biological interaction between empathy and chemokine activity. However, given that the effects are small (ΔR^2 ≤0.003 for all interaction terms), these interpretations should be considered preliminary and hypothesis-generating.

4.5 Null Findings in Depressive Subgroup

Unlike in the manic subgroup, no significant moderation effects were observed in the depressive subgroup. Several factors may account for this null finding. First, the depressive subgroup had a smaller size (n = 77 vs. n = 123 in the manic subgroup). Post hoc power analysis indicated that this subgroup had 80% power to detect interaction effects with $\Delta R^2 \approx 0.08$ or larger at $\alpha = 0.05$, which exceeds the magnitude of effects detected in the manic subgroup (ΔR^2 = 0.001–0.003). Thus, insufficient statistical power may partially explain the null results. Second, the pathophysiological mechanisms linking executive dysfunction to depressive symptoms may fundamentally differ from those in mania. While manic episodes are characterized by behavioral disinhibition and impulsivity—domains directly taxed by the WCST—depressive episodes involve psychomotor retardation, anhedonia, and amotivation, which may be less directly coupled with the cognitive flexibility and inhibitory control processes measured by the WCST. Third, the fact that CCL3/MIP-1α and CCL4/MIP-1β levels were significantly positively correlated with each other in the depressive subgroup (r = 0.361) but not in the manic subgroup (r = 0.076) suggests distinct chemokine interaction patterns across episode polarities. Future studies with larger depressive cohorts and state-specific inflammatory panels are needed to clarify whether the cognition–symptom association in BD is moderated by different immune or social-cognitive mechanisms not captured in the present design.

4.6 Clinical Implications

The present findings have potential research implications rather than immediate clinical implications. They suggest that future stratification studies may consider social-cognitive traits and peripheral inflammatory markers when examining cognition–symptom coupling in bipolar mania. However, because the interaction effects were small and the WCST error–BRMS association was unusually high, these results should not be used to guide clinical decision-making or biomarker-based intervention without prospective replication.

4.7 Limitations

The cross-sectional design precludes causal inferences. The temporal relationships among executive function, empathy, chemokine levels, and mood symptom severity remain to be clarified through longitudinal cohort studies with repeated assessments across mood states.

The most important limitation is the unusually high correlation between WCST total errors and BRMS scores in the manic subgroup (r ≈ 0.983). Although we verified data integrity, provided outlier diagnostics, and confirmed acceptable VIF values in moderation models, a correlation of this magnitude remains atypical for two distinct assessment domains. Shared method variance, restricted sampling, acute-state homogeneity, rater or procedural factors, or other unmeasured influences cannot be fully excluded. This issue substantially limits the interpretability of all moderation models based on BRMS scores. Independent replication using alternative executive function measures is essential.

This study measured only peripheral serum chemokine levels, which may not accurately reflect central neuroinflammatory processes. The extent to which peripheral CCL3/MIP-1α and CCL4/MIP-1β levels correspond to central nervous system inflammation warrants further investigation using cerebrospinal fluid biomarkers or neuroinflammation imaging techniques [31].

Several potential confounders—including body mass index, smoking status, detailed medication history, episode frequency, comorbid anxiety symptoms, sleep quality, and life stress events—were not comprehensively assessed or incorporated as covariates. Although patients were medication-free for at least two months prior to enrollment, the potential long-term effects of prior pharmacotherapy on cognitive function and inflammatory profiles cannot be completely excluded.

The sample sizes of the manic (n = 123) and depressive (n = 77) subgroups were unbalanced, which may have led to differential statistical power. The null moderation findings in the depressive subgroup should therefore be interpreted with caution and do not necessarily rule out the presence of small interaction effects.

The moderation effects (ΔR^2 = 0.001–0.003) were statistically significant but small, each accounting for at most 0.3% of additional explained variance. Although such small interaction effects are not uncommon in psychiatric research, their clinical applicability requires cautious interpretation and replication in larger samples.

We used IRI score as a composite measure of overall empathy. Given that the IRI encompasses both cognitive (PT and FS) and affective (EC and PD) dimensions, subscale-specific analyses may yield more nuanced insights and are recommended for future studies.

The exclusive inclusion of bipolar I disorder inpatients limits generalizability to bipolar II disorder, mixed episodes, and outpatient populations.

5. Conclusions

In the manic subgroup, higher overall empathy and selected chemokine levels were statistically associated with small changes in the WCST error–BRMS slope. The association between WCST error and BRMS score was unusually strong ($r \approx 0.983$), which limits the interpretability of moderation effects. Given the cross-sectional design, the unusually high WCST–BRMS correlation, and the small effect sizes, these findings should be viewed as hypothesis-generating rather than definitive. Independent replication in larger, more diverse samples is required.

No significant moderation effects were observed in the depressive subgroup. This null finding may reflect distinct

pathophysiological mechanisms, reduced statistical power, or genuinely absent moderation effects in the depressive phase.

Availability of Data and Materials

All data generated or analyzed during this study are available from the corresponding author upon reasonable request.

Author Contributions

DXX and JHL designed the research study. DXX, YXC and WJL performed the clinical assessments and sample collection. CYL and LHZ conducted the chemokine detection experiments. DXX analyzed the statistical data. DXX drafted the manuscript. JHL revised the manuscript critically for important intellectual content. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

All subjects gave their written informed consent for inclusion before they participated in this study. For participants who had difficulty providing consent due to clinical conditions, informed consent was obtained from their legally authorized representatives. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013), and the protocol was approved by the Ethics Review Committee of The Affiliated Kangning Hospital of Wenzhou Medical University (approval number: YJ-2024-01-02).

Acknowledgment

We would like to express our sincere gratitude to all the patients and healthy volunteers who participated in this study. We thank the staff of the Clinical Laboratory and Psychiatry Department of The Affiliated Kangning Hospital of Wenzhou Medical University for their technical support during sample collection and testing. We also appreciate the valuable comments and suggestions from the peer reviewers, which have greatly improved the quality of this manuscript.

Funding

This project was supported by the Basic Public Welfare Scientific Research Project of Wenzhou Municipal Science and Technology Bureau (Grant No.: Y2023859).

Conflicts of Interest

The authors declare no conflicts of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/AP52710.

References

[1] Varo C, Jiménez E, Solé B, Bonnín CM, Torrent C, Lahera G, et al. Social cognition in bipolar disorder: the role of sociodemographic, clinical, and neurocognitive variables in emotional intelligence. Acta Psychiatrica Scandinavica. 2019; 139: 369–380. https://doi.org/10.1111/acps.13014

[2] McIntyre RS, Berk M, Brietzke E, Goldstein BI, López-Jaramillo C, Kessing LV, et al. Bipolar disorders. Lancet (London, England). 2020; 396: 1841–1856. https://doi.org/10.1016/S0140-6736(20)31544-0

[3] Lima IMM, Peckham AD, Johnson SL. Cognitive deficits in bipolar disorders: Implications for emotion. Clinical Psychology Review. 2018; 59: 126–136. https://doi.org/10.1016/j.cpr.2017.11.006

[4] Mario A, Ivana L, Anita M, Silvio M, Claudia A, Mariaclaudia M, et al. Inflammatory Biomarkers, Cognitive Functioning, and Brain Imaging Abnormalities in Bipolar Disorder: A Systematic Review. Clinical Neuropsychiatry. 2024; 21: 32–62. https://doi.org/10.36131/cnfioritieditore20240103

[5] Bègue I, Elandaloussi Y, Delavari F, Cao H, Moussa-Tooks A, Roser M, et al. The Cerebellum and Cognitive Function: Anatomical Evidence from a Transdiagnostic Sample. Cerebellum (London, England). 2024; 23: 1399–1410. https://doi.org/10.1007/s12311-023-01645-y

[6] Liu YC, Tseng HH, Chang YH, Chang HH, Yang YK, Chen PS. The social cognitive ability in Han Chinese euthymic patients with bipolar I and bipolar II disorder. Journal of the Formosan Medical Association. 2020; 120: 1221–1228. https://doi.org/10.1016/j.jfma.2020.10.012

[7] Salagre E, Grande I, Vieta E, Mezquida G, Cuesta MJ, Moreno C, et al. Predictors of Bipolar Disorder Versus Schizophrenia Diagnosis in a Multicenter First Psychotic Episode Cohort: Baseline Characterization and a 12-Month Follow-Up Analysis. The Journal of Clinical Psychiatry. 2020; 81: 19m12996. https://doi.org/10.4088/JCP.19m12996

[8] Ospina LH, Nitzburg GC, Shanahan M, Perez-Rodriguez MM, Larsen E, Latifoglu A, et al. Social cognition moderates the relationship between neurocognition and community functioning in bipolar disorder. Journal of Affective Disorders. 2018; 235: 7–14. https://doi.org/10.1016/j.jad.2018.03.013

[9] Zhang B, Chen X, Qiu N. Social cognition in bipolar I and II disorders: an updated systematic review and meta-analysis. BMC Psychiatry. 2025; 25: 39. https://doi.org/10.1186/s12888-024-06462-z

[10] Liang YS, Yang HX, Zhang YJ, Cai XL, Wang YY, Ni K, et al. Validation of the Questionnaire of Cognitive and Affective Empathy in patients with schizophrenia, major depressive disorder and bipolar disorder. Cognitive Neuropsychiatry. 2020; 25: 466–479. https://doi.org/10.1080/13546805.2020.1846025

[11] Miller AH, Raison CL. The role of inflammation in depression: from evolutionary imperative to modern treatment target. Nature Reviews. Immunology. 2016; 16: 22–34. https://doi.org/10.1038/nri.2015.5

[12] Misiak B, Bartoli F, Carrà G, Małecka M, Samochowiec J, Jarosz K, et al. Chemokine alterations in bipolar disorder: A systematic review and meta-analysis. Brain, Behavior, and Immunity. 2020; 88: 870–877. https://doi.org/10.1016/j.bbi.2020.04.013

[13] Barbosa IG, Rocha NP, Vieira EL, Camkurt MA, Huguet RB, Guimarães FTL, et al. Decreased percentage of $CD4^+$ lympho-

cytes expressing chemokine receptors in bipolar disorder. Acta Neuropsychiatrica. 2019; 31: 246–251. https://doi.org/10.1017/neu.2019.5

[14] Zhang FF, Dong Y, Wang K, Zhan ZY, Xie LF. Reliability and validity of the Chinese version of the Interpersonal Reactivity Index-C. Chinese Journal of Clinical Psychology. 2010; 18: 155–157. https://doi.org/10.16128/j.cnki.1005-3611.2010.02.019 (In Chinese)

[15] Hamilton M. A rating scale for depression. Journal of Neurology, Neurosurgery, and Psychiatry. 1960; 23: 56–62. https://doi.org/10.1136/jnnp.23.1.56

[16] Bech P, Bolwig TG, Kramp P, Rafaelsen OJ. The Bech-Rafaelsen Mania Scale and the Hamilton Depression Scale. Acta Psychiatrica Scandinavica. 1979; 59: 420–430. https://doi.org/10.1111/j.1600-0447.1979.tb04484.x

[17] Miles S, Howlett CA, Berryman C, Nedeljkovic M, Moseley GL, Phillipou A. Considerations for using the Wisconsin Card Sorting Test to assess cognitive flexibility. Behavior Research Methods. 2021; 53: 2083–2091. https://doi.org/10.3758/s13428-021-01551-3

[18] Arioli M, Cattaneo Z, Ricciardi E, Canessa N. Overlapping and specific neural correlates for empathizing, affective mentalizing, and cognitive mentalizing: A coordinate-based meta-analytic study. Human Brain Mapping. 2021; 42: 4777–4804. https://doi.org/10.1002/hbm.25570

[19] Henry BL, Minassian A, Perry W. Everyday functional ability across different phases of bipolar disorder. Psychiatry Research. 2013; 210: 850–856. https://doi.org/10.1016/j.psychres.2013.04.006

[20] Wang Z, Cao H, Cao Y, Song H, Jiang X, Wei C, et al. Clinical characteristics and cognitive function in bipolar disorder patients with different onset symptom. Frontiers in Psychiatry. 2023; 14: 1253088. https://doi.org/10.3389/fpsyt.2023.1253088

[21] Khafif TC, Kleinman A, Rocca CCDA, Belizário GO, Nader E, Caetano SC, et al. Self-regulation in youth with bipolar disorder. Brazilian Journal of Psychiatry. 2023; 45: 20–27. https://doi.org/10.47626/1516-4446-2022-2668

[22] Dias VV, Balanzá-Martinez V, Soeiro-de-Souza MG, Moreno RA, Figueira ML, Machado-Vieira R, et al. Pharmacological approaches in bipolar disorders and the impact on cognition: a critical overview. Acta Psychiatrica Scandinavica. 2012; 126: 315–331. https://doi.org/10.1111/j.1600-0447.2012.01910.x

[23] Crowe SF, Stranks EK. The Residual Medium and Long-term Cognitive Effects of Benzodiazepine Use: An Updated Meta-analysis. Archives of Clinical Neuropsychology: the Official Journal of the National Academy of Neuropsychologists. 2018; 33: 901–911. https://doi.org/10.1093/arclin/acx120

[24] Bora E, Yucel M, Pantelis C. Cognitive endophenotypes of bipolar disorder: a meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. Journal of Affective Disorders. 2009; 113: 1–20. https://doi.org/10.1016/j.jad.2008.06.009

[25] Kurtz MM, Gerraty RT. A meta-analytic investigation of neurocognitive deficits in bipolar illness: profile and effects of clinical state. Neuropsychology. 2009; 23: 551–562. https://doi.org/10.1037/a0016277

[26] Nyhus E, Barceló F. The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: a critical update. Brain and Cognition. 2009; 71: 437–451. https://doi.org/10.1016/j.bandc.2009.03.005

[27] Phillips ML, Swartz HA. A critical appraisal of neuroimaging studies of bipolar disorder: toward a new conceptualization of underlying neural circuitry and a road map for future research. The American Journal of Psychiatry. 2014; 171: 829–843. https://doi.org/10.1176/appi.ajp.2014.13081008

[28] Bodnar A, Rybakowski JK. Increased affective empathy in bipolar patients during a manic episode. Brazilian Journal of Psychiatry. 2017; 39: 342–345. https://doi.org/10.1590/1516-4446-2016-2101

[29] Ríos U, Arancibia M, Jiménez JP, Bermpohl F. The forgotten affective route of social cognition in patients with bipolar disorders. Journal of Experimental Psychopathology. 2022; 13. https://doi.org/10.1177/20438087221135422

[30] Tan X, Yang Y, Yu M. Longitudinal relationship of empathy and social anxiety among adolescents: The mediation roles of psychological inflexibility and rejection sensitivity. Journal of Affective Disorders. 2023; 339: 867–876. https://doi.org/10.1016/j.jad.2023.07.069

[31] Castaño Barrios L, Da Silva Pinheiro AP, Gibaldi D, Silva AA, Machado Rodrigues E Silva P, Roffê E, et al. Behavioral alterations in long-term Toxoplasma gondii infection of C57BL/6 mice are associated with neuroinflammation and disruption of the blood brain barrier. PloS One. 2021; 16: e0258199. https://doi.org/10.1371/journal.pone.0258199

[32] Lisowska KA. The Role of Interleukin-8 (IL-8) in Treatment-Resistant Depression: A Review of Mechanisms, Biomarker Potential, and Therapeutic Implications. International Journal of Molecular Sciences. 2025; 26: 10092. https://doi.org/10.3390/ijms262010092

[33] Zhang Y, Wang J, Ye Y, Zou Y, Chen W, Wang Z, et al. Peripheral cytokine levels across psychiatric disorders: A systematic review and network meta-analysis. Progress in Neuropsychopharmacology & Biological Psychiatry. 2023; 125: 110740. https://doi.org/10.1016/j.pnpbp.2023.110740