Brain Functional Networks of Intergroup Empathy Bias: A Meta-Analysis of fMRI Studies

Authors: Sun Luwen, Zhou Yue, Jiang Zhongqing, Jiang Zhongqing

Date: 2025-03-21T00:00:00+00:00

Abstract

Intergroup Empathy Bias refers to the phenomenon of unequal empathy toward ingroup and outgroup members. Currently, the neural mechanisms underlying intergroup empathy bias, particularly the brain functional networks and neuromodulatory mechanisms behind it, remain unclear. This study employed activation likelihood estimation (ALE) to analyze brain coordinate data reported in 19 intergroup empathy bias studies, identifying two significant activation clusters located in the left anterior insula (lAI) and medial prefrontal cortex (mPFC), respectively. Furthermore, using Meta-Analytic Connectivity Modeling (MACM) and Neurosynth functional decoding, the study revealed that these two significant activation clusters not only play a critical role in intergroup empathy bias, but their associated brain functional networks also exhibit functional overlap with the central executive network (CEN). This finding not only deepens the theoretical foundation of the neuroscience of intergroup empathy, but also provides important neuroscientific evidence for the development of educational intervention strategies. Future research should focus on investigating the specific functional characteristics of these key brain regions and their neuromodulatory mechanisms, and strive to translate neuroscientific findings into practically effective empathy bias intervention protocols to promote intergroup harmony and reduce social prejudice and conflict.

Full Text

Abstract

Intergroup empathy bias refers to the phenomenon where individuals exhibit unequal empathy toward in-group and out-group members. Currently, the neural mechanisms underlying intergroup empathy bias, particularly the functional brain networks and neural regulatory mechanisms, remain poorly understood. This study utilized Activation Likelihood Estimation (ALE) to analyze brain

region coordinates reported in 19 studies on intergroup empathy bias, identifying two significant activation clusters located in the left anterior insula (lAI) and medial prefrontal cortex (mPFC). Further analysis using Meta-Analytic Connectivity Modeling (MACM) and Neurosynth functional decoding revealed that these clusters not only play crucial roles in intergroup empathy bias but also exhibit functional overlap with the central executive network (CEN). This discovery enhances the neuroscientific theoretical foundation of intergroup empathy while providing critical neural evidence for developing educational intervention strategies. Future research should focus on investigating the specific functional characteristics of these key brain regions and their neural regulatory mechanisms, with the aim of translating neuroscientific findings into practical interventions that reduce empathy bias, promote intergroup harmony, and mitigate social prejudice and conflict.

Keywords: intergroup empathy bias, ALE meta-analysis, MACM, Neurosynth

Intergroup empathy bias refers to the phenomenon where individuals exhibit unequal empathy toward in-group and out-group members (Cikara et al., 2011). Specifically, individuals tend to show greater empathy for in-group members while displaying reduced empathy—or even counter-empathic responses, such as deriving pleasure from out-group members' misfortunes or resentment toward their successes—toward out-group members (Cikara et al., 2014; Gutsell & Inzlicht, 2012; Vanman, 2016). This bias not only reveals the limitations of empathy as influenced by group identity but also provides a critical perspective for understanding the formation of social prejudice, escalation of group conflict, and barriers to cross-group cooperation. Moreover, research on intergroup empathy bias holds significant importance for elucidating the neural mechanisms of empathy, garnering widespread attention in both social psychology and neuroscience in recent years.

From an evolutionary psychology perspective, intergroup empathy bias has adaptive value, as empathy toward in-group members helps maintain group interests and enhance cohesion, thereby providing survival advantages (Efferson et al., 2008). However, despite these potential benefits, the negative consequences of this bias cannot be ignored. In healthcare settings, for instance, white physicians exhibit lower pain empathy toward minority patients (e.g., Black patients), resulting in inadequate pain management (Losin et al., 2020; Vaughn et al., 2019). In judicial contexts, jurors' empathy bias toward defendants of different racial backgrounds may lead to disparate verdicts (Mitchell et al., 2005). Critically, reducing intergroup empathy bias can help mitigate intergroup conflict and promote cross-group cooperation. Research demonstrates that individuals' empathy toward out-groups positively correlates with conflict-mitigating attitudes and support for cross-group cooperation (Caspar et al., 2023). Furthermore, administering oxytocin to enhance empathy among Israeli Jews toward Palestinians has shown promise in ameliorating the Israeli-Palestinian conflict (Influs et al., 2018; Shamay-Tsoory et al., 2013).

To suppress the negative effects of intergroup empathy bias and promote social

harmony and cooperation, researchers must investigate its underlying neural mechanisms. Neuroimaging techniques such as functional magnetic resonance imaging (fMRI) enable observation and analysis of differential neural activity during empathy tasks involving in-group versus out-group members. Moreover, researchers can objectively evaluate the effectiveness of various intervention methods by measuring the degree to which they reduce these neural activity differences, thereby optimizing intervention strategies.

Previous explorations of the neural mechanisms of intergroup empathy bias have primarily focused on racial groups as empathy targets (Berlingeri et al., 2016; Cheon et al., 2011; Fourie et al., 2017; Hein et al., 2016). For example, Xu et al. (2009) used fMRI to investigate neural activity during pain empathy for racial in-group and out-group members among White and Asian participants. Their neuroimaging results showed that painful expressions on in-group faces elicited stronger neural activity in the anterior cingulate cortex (ACC) and anterior insula (AI) compared to out-group faces. Other studies have employed alternative group categorizations; Hein et al. (2010) used sports team affiliations to define in-groups and out-groups, finding consistent results—images of in-group fans' suffering elicited stronger ACC and AI activation than those of rival team fans. Additionally, Ruckmann et al. (2015) used a minimal group paradigm to randomly assign participants to groups, revealing significant neural activity differences in the right fusiform gyrus, cerebellum, hippocampus, and amygdala across different group empathy conditions. Notably, racial categorization based on physical features such as skin color offers advantages in stability and reliability, resulting in a substantial body of research. However, systematic investigations into whether different group categorization methods influence brain activation patterns in intergroup empathy bias are currently lacking.

The neural mechanisms of intergroup empathy bias can also be examined through the lens of empathy' s three separable yet interconnected psychological components: affective empathy (Decety, 2011), cognitive empathy (Morelli et al., 2014), and empathic concern (Decety, 2015; Rameson et al., 2012; Zaki & Ochsner, 2012). Affective empathy refers to emotional contagion, where an individual' s emotions are influenced by others' emotions. Its bias manifests neuroimaging as significantly reduced ACC and AI activation during out-group empathy conditions (Cikara et al., 2011; Cikara & Van Bavel, 2014; Jackson et al., 2006; Han, 2018; Sheng et al., 2014). Cognitive empathy involves understanding others' perspectives and emotional states, with its bias reflected in diminished mPFC (Mitchell, 2009) and temporoparietal junction (TPJ) activation during out-group empathy. Empathic concern describes the psychological state of feeling sympathy and exhibiting altruistic behavioral tendencies (Marsh, 2018; Rameson et al., 2012). Drwecki et al. (2011) found that white participants showed greater empathic concern and provided higher levels of pain treatment to injured white individuals compared to Black individuals. This process involves neural activity in the ventral tegmental area (VTA), subgenual anterior cingulate cortex (sACC), and striatum (FeldmanHall et al.,

2015; Genevsky & Knutson, 2015). Research indicates that using cognitive reappraisal strategies during cognitive empathy can influence emotion inference and top-down regulation of affective empathy responses (Thompson et al., 2019). Both affective and cognitive empathy can trigger empathic concern, thereby promoting prosocial behavior (Marsh, 2018).

Existing research has primarily focused on key brain regions underlying empathy bias for specific in-group/out-group categorizations and single empathy components, lacking comprehensive analysis and comparison across studies, particularly regarding regions of significant concordance across studies. Consequently, the reliability of current findings remains insufficiently validated. For instance, research on pain empathy suggests that cognitive and affective processing share at least partially overlapping brain networks, with the AI likely serving as a key hub (Cox et al., 2012; Gu et al., 2013). However, whether the AI represents a region of significant concordance in intergroup empathy bias and whether other consistent regions exist require further investigation. Therefore, it is necessary to employ Activation Likelihood Estimation (ALE) to identify consistent brain regions underlying intergroup empathy bias and elucidate the core neural substrates of this phenomenon.

While consistent brain regions provide key localization for intergroup empathy bias, this perspective represents a simplified approach that restricts related functions to individual brain regions and focuses interpretation on specific differentially activated areas. This neglects the fact that the human brain comprises a complex network system of functionally and structurally interacting neural organizations (Bullmore & Sporns, 2009). Specifically, brain function depends on extensive interactions among multiple brain regions (Varela et al., 2001), with dynamic interactions across these networks being crucial for generating intergroup empathy. Therefore, comprehensive interpretation of brain region functions should be based on functional connectivity among co-activated regions during task states, achievable through Meta-Analytic Connectivity Modeling (MACM) supplemented by Neurosynth functional decoding.

To identify significant consistent brain regions in intergroup empathy bias and clarify their functional significance, this study proposes the following approach: First, use ALE to identify consistent brain regions involved in intergroup empathy bias and conduct subgroup analyses on potential moderating variables, including empathy type (pain, emotion), group categorization basis (race, other), and empathy task (implicit, explicit). Second, employ MACM to identify brain regions functionally connected to the consistent regions. Finally, use Neurosynth —an online meta-analytic tool containing over 14,000 functional neuroimaging studies—to functionally decode the brain network activated during intergroup empathy tasks and determine its functional characteristics. Integrating multiple independent studies through neuroimaging meta-analysis can identify consistent brain regions showing differential activation during in-group versus out-group empathy tasks, providing a more systematic theoretical framework for future neuroimaging research. Additionally, meta-analytic results can inform neuro-

scientifically grounded regulatory interventions by identifying key brain regions and functional networks to optimize intervention strategies for more effectively reducing intergroup empathy bias and promoting social harmony and cooperation.

2.1 Literature Search and Inclusion Criteria

We conducted a systematic search of fMRI studies on intergroup empathy bias in PubMed, Web of Science, ProQuest, CNKI, Wanfang, and VIP databases through January 31, 2024. Chinese database searches used four keyword dimensions: (1) neuroimaging terms (fMRI, brain imaging), (2) intergroup terms (group, intergroup, in-group, out-group), (3) empathy terms (empathy, sympathy, compassion), and (4) bias terms (prejudice, bias). English database searches employed equivalent terms: (1) fMRI, functional MRI, brain imaging; (2) group, intergroup, ingroup, outgroup; (3) empathy, sympathy, compassion; (4) prejudice, bias. Boolean operators connected keywords, yielding the Chinese search string: "(fMRI OR 脑成像) AND (群体 OR 群际 OR 内群体 OR 外群体) AND (共情 OR 移情 OR 同情) AND (偏见 OR 偏差)" and the English string: "(fMRI OR (functional MRI) OR (brain imaging)) AND (group OR intergroup OR ingroup OR outgroup) AND (empathy OR sympathy OR compassion) AND (prejudice OR bias)." Forward and backward citation searches of identified studies ensured comprehensive inclusion. This search strategy yielded 61 articles.

Inclusion and exclusion criteria were: (1) Study type: Empirical studies only, excluding reviews, meta-analyses, and conference abstracts; (2) Participants: Healthy adults aged 18-60, excluding patients with brain damage or psychological disorders; (3) Methods: fMRI studies only, excluding EEG, MEG, DTI, and structural MRI studies; (4) Analysis: Whole-brain analyses only, excluding region-of-interest analyses; (5) Results: Studies reporting coordinates in standardized Talairach or Montreal Neurological Institute (MNI) space for significant activation differences between in-group and out-group conditions. The screening process and results are illustrated in [Figure 1: see original paper]. Ultimately, 19 studies comprising 583 participants and 141 coordinates were included in the ALE meta-analysis.

2.2 Literature Coding and Quality Assessment

We coded the following data from included studies: publication information (first author and year), sample size (total and male participants), age (mean and standard deviation), experimental control conditions, and coordinate space (Talairach or MNI). The first author performed coding on three occasions (August 5, 2023; August 31, 2023; January 31, 2024) according to inclusion criteria, with subsequent verification by co-authors. Results showed high consistency across the three coding rounds, with only minor discrepancies. Basic information for included studies is presented in . We assessed risk of bias for the 19 included studies using the Cochrane Risk of Bias tool across six domains (selec-

tion, implementation, measurement, incomplete data, reporting, and other bias) with seven items. All studies were rated as low risk of bias across all domains.

2.3 Activation Likelihood Estimation

This study employed Activation Likelihood Estimation (ALE), a method that calculates the probability of each voxel being included in activation regions corresponding to peak coordinates based on Gaussian probability density distribution models. Modeled activation (MA) maps are combined voxel-wise to derive cross-experiment activation likelihoods (Hu et al., 2015). Data processing used GingerALE software (version 3.0.2) (Eickhoff et al., 2009; Turkeltaub et al., 2012). The workflow involved: (1) converting MNI coordinates to Talairach space using built-in conversion functions, (2) performing meta-analysis on intergroup empathy bias coordinates with statistical threshold $p < 0.05$ using family-wise error (FWE) correction and cluster threshold $p < 0.001$ with 5,000 permutations (Eickhoff et al., 2017). Methodological recommendations suggest ALE meta-analysis requires at least 17–20 independent studies for adequate statistical power (Eickhoff et al., 2016; Müller et al., 2018). Our analysis included 19 studies, meeting minimum requirements.

2.4 Meta-Analytic Connectivity Modeling (MACM)

MACM operates on the principle that functionally related brain regions are more likely to be co-activated by the same task. By calculating co-occurrence probabilities across neuroimaging experiments, MACM reflects functional connectivity between brain regions (Hu et al., 2015). We used Sleuth software (version 3.0.4) with the BrainMap functional database (http://www.brainmap.org), which contains 3,406 articles, 111 paradigm classes, 76,016 subjects, 16,901 experiments, and 131,598 coordinates (Fox & Lancaster, 2002; Laird et al., 2009). Literature screening criteria were: experiment level—"context: normal mapping" and "activations: activation only" ; subject level— "diagnosis: normals," "age is less than 60," and "age is more than 18." Using ALE results, we input coordinates in "Locations" to retrieve relevant literature and corresponding brain region coordinates for ALE analysis of co-activation regions.

2.5 Neurosynth Tool

Neurosynth is an online meta-analytic database containing over 14,000 functional neuroimaging studies that pairs brain activation patterns and peak coordinates with relevant keywords (Yarkoni et al., 2011; https://neurosynth.org). We uploaded MACM-derived co-activation regions to the Neurosynth platform for functional decoding, yielding keywords associated with the uploaded regions. Each keyword receives a correlation score indicating its relationship strength with the co-activation regions, facilitating interpretation of brain region functional characteristics.

3.1 ALE Meta-Analysis Results

presents ALE meta-analysis results and detailed activation cluster information, while [Figure 2: see original paper] illustrates cluster distributions (visualized using Mango). Integrating coordinates from 19 studies (583 participants) revealed two significant activation clusters. The largest cluster was located near the left anterior insula (center coordinates: x = −32.7, y = 12.9, z = 3) with three peaks (peak 1: x = −28, y = 16, z = 2; peak 2: x = −38, y = 14, z = 8; peak 3: x = −32, y = 10, z = −6). This cluster comprised 55.3% claustrum, 34.2% insula, 5.3% lentiform nucleus, and 5.3% precentral gyrus. The second cluster was located in the medial prefrontal cortex (center coordinates: x = −2.1, y = 27.9, z = 45.3) with two peaks (peak 1: x = −2, y = 30, z = 44; peak 2: x = 0, y = 20, z = 52), comprising 52.9% medial frontal gyrus and 47.1% superior frontal gyrus.

3.2 MACM Analysis Results

We used activation clusters 1 and 2 center coordinates and spatial extents as region-of-interest parameters in MACM literature screening. This yielded 155 articles (196 experiments, 2,520 subjects, 3,364 coordinates) for cluster 1 and 9 articles (12 experiments, 132 subjects, 255 coordinates) for cluster 2. ALE meta-analysis identified two sets of co-activation regions (, [Figure 3: see original paper]). For cluster 1' s functional network, the largest cluster showed bilateral distribution (52.3% right hemisphere, 47.7% left hemisphere), primarily encompassing thalamus (20.8%), lentiform nucleus (17.9%), middle frontal gyrus (12.5%), precentral gyrus (12.5%), insula (12.1%), inferior frontal gyrus (10.6%), caudate nucleus (6.4%), claustrum (4.6%), and superior frontal gyrus (1.3%). For cluster 2' s functional network, the largest cluster was 58.1% left-lateralized and exclusively distributed in the anterior cingulate cortex.

3.3 Neurosynth Analysis Results

We performed functional decoding on both co-activation sets using Neurosynth (https://neurosynth.org), excluding anatomical terms like "anterior cingulate." Decoding revealed that cluster 1' s functional network was primarily associated with task processing and semantic processing, including keywords such as "task," "working memory," "load" for task processing, and "word," "language," "phonological" for semantic processing. Cluster 2' s network was associated with evaluation, reward, and psychopathology, including keywords like "noxious," "value," "reward," "money," "anticipation" for evaluation/reward, and "Post-Traumatic Stress Disorder (PTSD)" for psychopathology (, [Figure 4: see original paper]). Both networks additionally involved keywords such as "gain," "emotion," and "painful."

We also obtained brain maps associated with the central executive network (CEN), also known as the executive control network (ECN), using Neurosynth.

Spatial overlay of this network with our intergroup empathy bias network is shown in [Figure 5: see original paper].

3.4 Exploratory Subgroup Analysis Results

For empathy type (pain vs. emotion), insufficient studies examined emotional empathy (only one experiment in one study), precluding subgroup analysis. After excluding emotional empathy tasks, we conducted separate subgroup analysis on pain empathy, which revealed two activation clusters from 19 studies (547 subjects): the largest cluster near left AI (center: x = −30.6, y = 12.9, z = 1.5) with two peaks (peak 1: x = −28, y = 16, z = 2; peak 2: x = −32, y = 10, z = −6), comprising 87.5% claustrum, 8.3% lentiform nucleus, and 4.2% insula; and a second cluster in mPFC (center: x = −1.9, y = 27.8, z = 45.5) with two peaks (peak 1: x = −2, y = 30, z = 44; peak 2: x = 0, y = 20, z = 52), comprising 50.8% medial frontal gyrus and 49.2% superior frontal gyrus.

For group categorization basis (race vs. other), race-based categorization (19 experiments, 434 subjects) yielded two clusters: the largest near left AI (center: x = −30, y = 13.2, z = 1.1) with three peaks (peak 1: x = −28, y = 16, z = 2; peak 2: x = −32, y = 10, z = −6; peak 3: x = −26, y = 16, z = −6), comprising 89.3% claustrum, 7.1% lentiform nucleus, and 3.6% insula; and a second cluster in mPFC (center: x = −2, y = 29, z = 44) with one peak (peak 1: x = −2, y = 30, z = 44), comprising 62.7% medial frontal gyrus and 37.3% superior frontal gyrus. However, non-race-based categorization (5 experiments, 130 subjects) revealed no significant co-activation regions.

For empathy task type (implicit vs. explicit), only one implicit study existed [Study 17, Sheng et al. (2014)], precluding subgroup analysis. After excluding this study, explicit empathy tasks (18 studies, 526 subjects) revealed one activation cluster (center: x = −32.7, y = 12.9, z = 3) with three peaks (peak 1: x = −28, y = 16, z = 2; peak 2: x = −38, y = 14, z = 8; peak 3: x = −32, y = 10, z = −6), comprising 55.3% claustrum, 34.2% insula, 5.3% lentiform nucleus, and 5.3% precentral gyrus.

Further exploratory analysis of explicit task subtypes (rating vs. judgment) showed that rating tasks (13 studies, 358 subjects) activated one cluster (center: x = −34.6, y = 11.7, z = 3.7) with three peaks (peak 1: x = −38, y = 14, z = 8; peak 2: x = −32, y = 10, z = 2; peak 3: x = −32, y = 10, z = −6), comprising 53.3% insula, 40% claustrum, and 6.7% precentral gyrus. Judgment tasks (7 studies, 204 subjects) activated one cluster (center: x = −1.8, y = 26.9, z = 44.2) with one peak (peak: x = −2, y = 28, z = 44), comprising 65.4% medial frontal gyrus, 32.7% superior frontal gyrus, and 1.9% cingulate cortex. Conjunction analysis revealed no significant clusters. Difference analysis showed no significant clusters for rating > judgment contrast, but judgment > rating contrast revealed activation in medial frontal gyrus, superior frontal gyrus, and cingulate cortex (center: x = −3, y = 22.5, z = 45.4).

This study screened 19 fMRI studies to systematically investigate consistent

brain regions and functional networks involved in intergroup empathy bias. ALE meta-analysis localized activation clusters primarily near the left AI and mPFC. MACM analysis identified brain regions co-activated with these clusters, constructing the functional network underlying intergroup empathy bias. Neurosynth functional decoding revealed the functional significance of this network and demonstrated spatial overlap with the central executive network. Finally, we conducted subgroup analyses based on group categorization (race vs. other), empathy type (pain vs. emotion), task type (implicit vs. explicit), and explicit task subtypes (rating vs. judgment).

4.1 Significant Consistent Brain Regions in Intergroup Empathy Bias

Our findings identified two significant consistent brain region clusters in intergroup empathy bias. Cluster 1 was located near the left anterior insula (lAI), where neural activity is significantly reduced when empathizing with out-group members (Cikara et al., 2011; Cikara & Van Bavel, 2014; Han, 2018; Jackson et al., 2006; Sheng et al., 2014; Xu et al., 2009). This cluster exhibited marked hemispheric lateralization: unlike the left AI, the right AI showed no significant neural activity changes during in-group versus out-group empathy. Previous research indicates that lAI primarily functions in emotion-related processing tasks (Raschle et al., 2018; Smith et al., 2017; Wang et al., 2020), whereas right AI plays a key role in attention regulation and switching between default mode and central executive networks (Perri et al., 2018; Sridharan et al., 2008; Wen et al., 2013). This lateralization likely reflects that cluster 1' s emotional processing is significantly modulated by group category, while right hemisphere attention regulation and network switching functions remain unaffected.

Cluster 2 was located in the medial prefrontal cortex (mPFC), where neural activity also decreases significantly when empathizing with out-group members (Masten et al., 2011; Mathur et al., 2010; Mitchell, 2009). The mPFC plays multiple roles in negative emotion regulation, involving both emotional evaluation and expression and emotion regulation through interactions with limbic structures (Banks et al., 2007; Ochsner et al., 2004; Phan et al., 2005). For example, mPFC modulates negative emotions elicited by observing or experiencing pain. Cheriyan and Sheets (2018) found increased mPFC neuronal activity in chronic pain mice, which modulates endogenous analgesia via the periaqueductal gray (PAG). Thus, during intergroup empathy, stronger negative emotions elicited by in-group members correspond to more pronounced regulatory effects and stronger activation in cluster 2.

Exploratory subgroup analyses revealed that pain empathy bias and race-based empathy bias clusters remained consistent with overall activation clusters, and explicit task clusters aligned with cluster 1. Notably, subgroup analysis by explicit task type (rating vs. judgment) showed that rating and judgment tasks activated clusters spatially corresponding to cluster 1 (left AI) and cluster 2 (mPFC), respectively. This suggests task-specific neural activation patterns:

rating tasks, emphasizing subjective emotional experience and affective resonance, preferentially activated left AI, whereas judgment tasks, emphasizing cognitive evaluation of emotional categories, primarily activated mPFC. However, these subgroup analyses have limitations due to relatively small sample sizes in each subgroup, necessitating further validation in future research.

Although consistent brain regions identify key areas for intergroup empathy bias, functional interpretation requires analysis of their broader functional networks, which MACM and Neurosynth can provide.

4.2 Functional Brain Networks of Intergroup Empathy Bias

MACM results demonstrated significant functional connectivity between activation clusters 1 (left AI) and 2 (mPFC) and multiple other brain regions, suggesting these regions form a coordinated neural functional network that optimizes connectivity patterns for intergroup empathy bias processing. Neurosynth functional decoding systematically revealed three functional dimensions of this network:

First, keywords including "working memory,""task,""load,""phonological,""word," and "language"indicate cognitive control functions. These features closely resemble the central executive network (CEN), which regulates attention allocation, working memory maintenance, and cognitive flexibility through dorsolateral prefrontal cortex (dlPFC), ACC, and parietal cortex coordination (Niendam et al., 2012; Seeley et al., 2007). Spatial overlay of Neurosynth-derived CEN maps with our network ([Figure 5: see original paper]) revealed overlap in key nodes including dlPFC and ACC, suggesting the intergroup empathy bias network may suppress out-group empathy through attention resource modulation and negative memory retrieval enhancement.

Second, keywords "pain,""emotion,""PTSD,"and "affective"indicate emotional information processing. Anatomically, this network resembles the empathy affective network proposed by Lieberman et al. (2019), encompassing mPFC, mirror neuron systems (temporoparietal junction, posterior temporal cortex), ventral ACC, and amygdala. Functionally, it also resembles the orbito-affective network (ORA network) described by Ji et al. (2019), which implements emotional valence evaluation through orbitofrontal cortex, ventromedial PFC, and ACC. Thus, the intergroup empathy bias network may weaken out-group empathy by reducing emotion recognition efficiency (insufficient ventral ACC activation) and inhibiting emotional contagion (diminished mirror system function).

Third, keywords "gain," "demands," "reward," "value," "money," and "anticipation" indicate involvement in reward expectation and value evaluation. During intergroup empathy, individuals weigh empathy costs against expected rewards (Cameron, 2018). Empathy toward in-group members typically yields higher social rewards, as in-group members represent important social resources with whom interactions are more frequent, important, and beneficial (Correll & Park, 2005). The underlying neural mechanism involves mPFC' s role in re-

ward expectation and decision-making (Floresco et al., 2015; Jenni et al., 2017; Starkweather et al., 2018), encoding and integrating reward value information (Robbins & Dalley, 2017; Zhuo et al., 2023) and cooperating with the septal area in prosocial motivation formation (Masten et al., 2011; Morelli et al., 2014; Rameson et al., 2012). These mechanisms contribute to heightened empathy bias toward in-group members.

5 Limitations and Future Directions

First, all 19 included studies examined physical pain empathy, with only one study comparing physical pain and emotional empathy across group dimensions. This predominance of pain empathy studies may bias meta-analytic results toward converging on pain-specific brain regions. Moreover, both pain empathy and the emotional empathy included in this meta-analysis represent negative empathy, whereas positive empathy—understanding and sharing others' positive emotional states (Yue & Huang, 2016)—remains understudied. This imbalance likely stems from humans' processing advantage for negative emotional cues and their adaptive significance, creating methodological path dependence toward negative empathy paradigms like social exclusion and pain empathy. Given this limitation, interpretations of brain activation patterns in intergroup empathy bias should consider potential valence effects. Future research should investigate emotional and positive empathy neural mechanisms and develop relevant experimental paradigms to comprehensively understand general principles underlying intergroup empathy bias.

Second, ALE meta-analysis relies on published significant activation coordinates, potentially introducing publication bias. While traditional effect-size-based bias assessment is unsuitable for coordinate data, ALE employs stringent multiple comparison corrections including false discovery rate (FDR) (Laird et al., 2005) and family-wise error rate (FWE) (Eickhoff et al., 2012) to ensure voxel-wise stability across experiments (Hu et al., 2015). Although we used the latest ALE algorithm revision enhancing reliability, potential publication bias should be considered when interpreting results. Additionally, while MACM is widely used to identify functionally connected regions, these associations do not establish causality. Future research should employ targeted experimental tasks and techniques such as transcranial magnetic stimulation (TMS) to verify whether these brain regions directly participate in intergroup empathy bias.

Finally, this study identified significant consistent brain regions in intergroup empathy bias, revealing differential activation in left AI and mPFC when empathizing with in-group versus out-group members. Based on these functional characteristics, future research can develop targeted intervention training. For left AI, a key region for emotional processing that shows reduced activity during out-group empathy, training could enhance out-group emotional perception through facial expression evaluation and simulated emotional experience tasks (e.g., virtual scenario immersion). For mPFC, involved in cognitive control and decision-making, cognitive reappraisal training and emotion regulation strate-

gies (e.g., mindfulness meditation) could improve cognitive processing efficiency for out-group emotional states. For individuals showing extreme empathy bias and very low out-group empathy, TMS could be applied to target left AI and mPFC activity while engaging in out-group empathy tasks. Additionally, our novel finding of overlap between the intergroup empathy bias network and CEN suggests top-down regulatory mechanisms play important roles. Attention training could help allocate attentional resources more equitably, while guiding individuals to retrieve positive memories about out-group members may promote more active out-group empathy. Since intergroup empathy bias emerges early in development—children aged 3–12 already show stronger in-group empathy (Tompkins et al., 2023)—early intervention during childhood, when neural plasticity is greater, may be particularly effective. Such interventions require coordinated efforts from families, schools, and society to cultivate empathy, educate children about different cultural groups, facilitate intergroup cooperation, and promote positive media representations of intergroup interactions, with continuous evaluation and optimization of intervention effectiveness. Multi-level approaches may effectively reduce early intergroup empathy bias, promote social harmony, bridge social divisions, and create more inclusive and diverse societies.

Acknowledgments

We sincerely thank Professor Hu Chuanpeng from the School of Psychology at Nanjing Normal University and Professor Zhang Yali from the School of Education at Hebei Normal University for their constructive suggestions during manuscript revision. We also thank the reviewers and editorial board for their valuable feedback.

*Indicates studies included in the meta-analysis.

References

Hu, C., Di, X., Li, J., Sui, J., & Peng, K. (2015). Meta-analysis of neuroimaging data. *Advances in Psychological Science*, 23(7), 1118–1129.

Yue, T., & Huang, X. (2016). Positive empathy in cognitive neuroscience research. *Advances in Psychological Science*, 24(3), 402–409.

Zhuo, L., Zeng, X., Wu, B., Niu, R., Yu, P., & Wang, W. (2023). The role of the medial prefrontal cortex-nucleus accumbens circuit in decision-making impulsivity: Evidence from animal models. *Acta Psychologica Sinica*, 55(4), 556–571.

Azevedo, R. T., Macaluso, E., Avenanti, A., Santangelo, V., Cazzato, V., & Aglioti, S. M. (2013). Their pain is not our pain: Brain and autonomic correlates of empathic resonance with the pain of same and different race individuals. Human Brain Mapping*, 34(12), 3168–3181.

Banks, S., Eddy, K., Angstadt, M., Nathan, P., & Phan, K. (2007). Amygdala-

frontal connectivity during emotion regulation. *Social Cognitive and Affective Neuroscience*, 2(4), 303–312.

Berlingeri, M., Gallucci, M., Danelli, L., Forgiarini, M., Sberna, M., & Paulesu, E. (2016). Guess who' s coming to dinner: Brain signatures of racially biased and politically correct behaviors. Neuroscience*, 332, 231–241.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3), 186–198.

Cameron, C. D. (2018). Motivating empathy: Three methodological recommendations for mapping empathy. *Social and Personality Psychology Compass*, 12(11), e12418.

Cao, Y., Contreras-Huerta, L. S., McFadyen, J., & Cunnington, R. (2015). Racial bias in neural response to others' pain is reduced with other-race contact. Cortex*, 70, 68–78.

Cao, Y., Yusri, N. M., Powell, T., & Cunnington, R. (2019). Neural and behavioral markers of observed pain of older adults. Neuropsychologia*, 131, 84–90.

Caspar, E. A., Pech, G. P., Gishoma, D., & Kanazayire, C. (2023). On the impact of the genocide on the intergroup empathy bias between former perpetrators, survivors, and their children in Rwanda. *American Psychologist*, 78(7), 825–841.

Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and them: Intergroup failures of empathy. *Current Directions in Psychological Science*, 20(3), 149–153.

Cikara, M., Bruneau, E., Van Bavel, J. J., & Saxe, R. (2014). Their pain gives us pleasure: How intergroup dynamics shape empathic failures and counter-empathic responses. *Journal of Experimental Social Psychology*, 55, 110–125.

Cikara, M., & Van Bavel, J. J. (2014). The neuroscience of intergroup relations: An integrative review. *Perspectives on Psychological Science*, 9(3), 245–274.

Cheriyan, J., & Sheets, P. (2018). Altered excitability and local connectivity of mPFC-PAG neurons in a mouse model of neuropathic pain. *The Journal of Neuroscience*, 38(21), 4829–4839.

Cheon, B. K., Im, D. M., Harada, T., Kim, J. S., Mathur, V. A., Scimeca, J. M., ···& Chiao, J. Y. (2011). Cultural influences on neural basis of intergroup empathy. NeuroImage*, 57(2), 642–650.

Contreras-Huerta, L. S., Baker, K. S., Reynolds, K. J., Batalha, L., & Cunnington, R. (2013). Racial bias in neural empathic responses to pain. PloS One*, 8(12), e84001.

Correll, J., & Park, B. (2005). A model of the ingroup as a social resource. *Personality and Social Psychology Review*, 9(4), 341–359.

Cox, C., Uddin, L., Martino, A., Castellanos, F., Milham, M., & Kelly, C. (2012). The balance between feeling and knowing: Affective and cognitive empathy are reflected in the brain' s intrinsic functional dynamics. *Social Cognitive and Affective Neuroscience*, 7(6), 727–737.

Decety, J. (2011). Dissecting the neural mechanisms mediating empathy. *Emotion Review*, 3(1), 92–108.

Decety, J. (2015). The neural pathways, development and functions of empathy. *Current Opinion in Behavioral Sciences*, 3, 1–6.

Drwecki, B. B., Moore, C. F., Ward, S. E., & Prkachin, K. M. (2011). Reducing racial disparities in pain treatment: The role of empathy and perspective-taking. *Pain*, 152(5), 1001–1006.

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage*, 59(3), 2349–2361.

Efferson, C., Lalive, R., & Fehr, E. (2008). The coevolution of cultural groups and ingroup favoritism. *Science*, 321(5897), 1844–1849.

Eickhoff, S. B., Laird, A. R., Fox, P. M., Lancaster, J. L., & Fox, P. T. (2017). Implementation errors in the GingerALE software: Description and recommendations. *Human Brain Mapping*, 38(1), 7–11.

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping*, 30(9), 2907–2926.

Eickhoff, S. B., Nichols, T. E., Laird, A. R., Hoffstaedter, F., Amunts, K., Fox, P. T., ···& Eickhoff, C. R. (2016). Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. *NeuroImage*, 137, 70–85.

FeldmanHall, O., Dalgleish, T., Evans, D., & Mobbs, D. (2015). Empathic concern drives costly altruism. *NeuroImage*, 105, 347–356.

Floresco, S. B. (2015). The nucleus accumbens: An interface between cognition, emotion, and action. *Annual Review of Psychology*, 66, 25–52.

Fourie, M. M., Stein, D. J., Solms, M., Gobodo-Madikizela, P., & Decety, J. (2017). Empathy and moral emotions in post-apartheid South Africa: An fMRI investigation. Social Cognitive and Affective Neuroscience*, 12(6), 881–892.

Fox, G. R., Sobhani, M., & Aziz-Zadeh, L. (2013). Witnessing hateful people in pain modulates brain activity in regions associated with physical pain and reward. Frontiers in Psychology*, 4, 772.

Fox, P. T., & Lancaster, J. L. (2002). Mapping context and content: The BrainMap model. *Nature Reviews Neuroscience*, 3(4), 319–321.

Genevsky, A., & Knutson, B. (2015). Neural affective mechanisms predict market-level microlending. *Psychological Science*, 26(9), 1411–1422.

Gu, X., Hof, P. R., Friston, K. J., & Fan, J. (2013). Anterior insular cortex and emotional awareness. *Journal of Comparative Neurology*, 521(15), 3371–3388.

Gutsell, J., & Inzlicht, M. (2012). Intergroup differences in the sharing of emotive states: Neural evidence of an empathy gap. *Social Cognitive and Affective Neuroscience*, 7(5), 596–603.

Han, S. (2018). Neurocognitive basis of racial ingroup bias in empathy. *Trends in Cognitive Sciences*, 22(5), 400–421.

Hein, G., Engelmann, J. B., Vollberg, M. C., & Tobler, P. N. (2016). How learning shapes the empathic brain. Proceedings of the National Academy of Sciences*, 113(1), 80–85.

Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. Neuron*, 68(1), 149–160.

Influs, M., Pratt, M., Masalha, S., Zagoory-Sharon, O., & Feldman, R. (2018). A social neuroscience approach to conflict resolution: Dialogue intervention to Israeli and Palestinian youth impacts oxytocin and empathy. *Social Neuroscience*, 14(4), 378–389.

Jackson, P. L., Brunet, E., Meltzoff, A. N., & Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia*, 44(5), 752–761.

Jenni, N. L., Larkin, J. D., & Floresco, S. B. (2017). Prefrontal dopamine D1 and D2 receptors regulate dissociable aspects of decision making via distinct ventral striatal and amygdalar circuits. *The Journal of Neuroscience*, 37(26), 6200–6213.

Ji, J. L., Spronk, M., Kulkarni, K., Repovš, G., Anticevic, A., & Cole, M. W. (2019). Mapping the human brain' s cortical-subcortical functional network organization. *NeuroImage*, 185, 35–57.

Jie, J., Fan, M., Yang, Y., Luo, P., Wang, Y., Li, J., ···& Zheng, X. (2022). Establishing a counter-empathy processing model: Evidence from functional magnetic resonance imaging. Social Cognitive and Affective Neuroscience*, 17(3), 273–289.

Laird, A. R., Eickhoff, S. B., Li, K., Robin, D. A., Glahn, D. C., & Fox, P. T. (2009). Investigating the functional heterogeneity of the default mode network using coordinate-based meta-analytic modeling. *Journal of Neuroscience*, 29(46), 14496–14505.

Laird, A. R., Fox, P. M., Price, C. J., Glahn, D. C., Uecker, A. M., Lancaster, J. L., ···& Fox, P. T. (2005). ALE meta-analysis: Controlling the false discovery rate and performing statistical contrasts. *Human Brain Mapping*, 25(1), 155–164.

Li, X., Liu, Y., Luo, S., Wu, B., Wu, X., & Han, S. (2015). Mortality salience enhances racial in-group bias in empathic neural responses to others' suffering. NeuroImage*, 118, 376–385.

Lieberman, M. D., Straccia, M. A., Meyer, M. L., Du, M., & Tan, K. M. (2019). Social, self, (situational), and affective processes in medial prefrontal cortex (MPFC): Causal, multivariate, and reverse inference evidence. *Neuroscience & Biobehavioral Reviews*, 99, 311–328.

Losin, E. A. R., Woo, C. W., Medina, N. A., Andrews-Hanna, J. R., Eisenbarth, H., & Wager, T. D. (2020). Neural and sociocultural mediators of ethnic differences in pain. *Nature Human Behaviour*, 4(5), 517–530.

Luo, S., Li, B., Ma, Y., Zhang, W., Rao, Y., & Han, S. (2015). Oxytocin receptor gene and racial ingroup bias in empathy-related brain activity. NeuroImage*, 110, 22–31.

Marsh, A. A. (2018). The neuroscience of empathy. *Current Opinion in Behavioral Sciences*, 19, 110–115.

Masten, C. L., Morelli, S. A., & Eisenberger, N. I. (2011). An fMRI investigation of empathy for 'social pain' and subsequent prosocial behavior. *NeuroImage*, 55(1), 381–388.

Mathur, V. A., Harada, T., Lipke, T., & Chiao, J. Y. (2010). Neural basis of extraordinary empathy and altruistic motivation. NeuroImage*, 51(4), 1468–1475.

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1309–1316.

Mitchell, T., Haw, R., Pfeifer, J., & Meissner, C. (2005). Racial bias in mock juror decision-making: A meta-analytic review of defendant treatment. *Law and Human Behavior*, 29(6), 621–637.

Morelli, S. A., Rameson, L. T., & Lieberman, M. D. (2014). The neural components of empathy: Predicting daily prosocial behavior. *Social Cognitive and Affective Neuroscience*, 9(1), 39–47.

Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., ···& Eickhoff, S. B. (2018). Ten simple rules for neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews*, 84, 151–161.

Niendam, T., Laird, A., Ray, K., Dean, Y., Glahn, D., & Carter, C. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, 12(2), 241–268.

Ochsner, K., Ray, R., Cooper, J., Robertson, E., Chopra, S., Gabrieli, J., & Gross, J. (2004). For better or for worse: Neural systems supporting the cognitive down- and up-regulation of negative emotion. *NeuroImage*, 23(2), 483–499.

Perri, R. L., Berchicci, M., Bianco, V., Spinelli, D., & Di Russo, F. (2018). Brain waves from an "isolated" cortex: Contribution of the anterior insula to cognitive functions. *Brain Structure and Function*, 223(3), 1343–1355.

Phan, L., Fitzgerald, D., Nathan, P., Moore, G., Uhde, T., & Tancer, M. (2005). Neural substrates for voluntary suppression of negative affect: A functional magnetic resonance imaging study. *Biological Psychiatry*, 57(3), 210–219.

Pluta, A., Mazurek, J., Wojciechowski, J., Wolak, T., Soral, W., & Bilewicz, M. (2023). Exposure to hate speech deteriorates neurocognitive mechanisms of the ability to understand others' pain. Scientific Reports*, 13(1), 4127.

Rameson, L. T., Morelli, S. A., & Lieberman, M. D. (2012). The neural correlates of empathy: Experience, automaticity, and prosocial behavior. *Journal of Cognitive Neuroscience*, 24(1), 235–245.

Raschle, N. M., Menks, W. M., Fehlbaum, L. V., Steppan, M., Smaragdi, A., Gonzalez-Madruga, K., ···& Stadler, C. (2018). Callous-unemotional traits and brain structure: Sex-specific effects in anterior insula of typically-developing youths. *NeuroImage: Clinical*, 17, 856–864.

Robbins, T. W., & Dalley, J. W. (2017). Impulsivity, risky choice, and impulse control disorders: Animal models. In J-C. Dreher & L. Tremblay (Eds.), *Decision Neuroscience* (pp. 81–93). Academic Press.

Ruckmann, J., Bodden, M., Jansen, A., Kircher, T., Dodel, R., & Rief, W. (2015). How pain empathy depends on ingroup/outgroup decisions: A functional magnet resonance imaging study. Psychiatry Research: NeuroImaging*, 234(1), 57–65.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., ···& Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27(9), 2349–2356.

Shamay-Tsoory, S., Abu-Akel, A., Palgi, S., Sulieman, R., Fischer-Shofty, M., Levkovitz, Y., & Decety, J. (2013). Giving peace a chance: Oxytocin increases empathy to pain in the context of the Israeli–Palestinian conflict. *Psychoneuroendocrinology*, 38(11), 3139–3144.

Sheng, F., Liu, Q., Li, H., Fang, F., & Han, S. (2014). Task modulations of racial bias in neural responses to others' suffering. NeuroImage*, 88, 263–270.

Smith, R., Lane, R. D., Alkozei, A., Bao, J., Smith, C., Sanova, A., ···& Killgore, W. D. (2017). Maintaining the feelings of others in working memory is associated

with activation of the left anterior insula and left frontal-parietal control network. *Social Cognitive and Affective Neuroscience*, 12(5), 848–860.

Sridharan, D., Levitin, D. J., & Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences*, 105(34), 12569–12574.

Starkweather, C. K., Gershman, S. J., & Uchida, N. (2018). The medial prefrontal cortex shapes dopamine reward prediction errors under state uncertainty. *Neuron*, 98(3), 616–629.

Tompkins, R., Vasquez, K., Gerdin, E., Dunham, Y., & Liberman, Z. (2023). Expectations of intergroup empathy bias emerge by early childhood. *Journal of Experimental Psychology: General*, 153(11), 2700–2714.

Thompson, N. M., Uusberg, A., Gross, J. J., & Chakrabarti, B. (2019). Empathy and emotion regulation: An integrative account. *Emotion and Cognition*, 247, 273–304.

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Human Brain Mapping*, 33(1), 1–13.

Vanman, E. (2016). The role of empathy in intergroup relations. *Current Opinion in Psychology*, 11, 59–63.

Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2(4), 229–239.

Vaughn, I. A., Terry, E. L., Bartley, E. J., Schaefer, N., & Fillingim, R. B. (2019). Racial-ethnic differences in osteoarthritis pain and disability: A meta-analysis. *The Journal of Pain*, 20(6), 629–644.

Wang, C., Wu, B., Liu, Y., Wu, X., & Han, S. (2015). Challenging emotional prejudice by changing self-concept: Priming independent self-construal reduces racial in-group bias in neural responses to other' s pain. Social Cognitive and Affective Neuroscience*, 10(9), 1195–1201.

Wang, Y., Zou, Q., Ao, Y., Liu, Y., Ouyang, Y., Wang, X., ···& Chen, H. (2020). Frequency-dependent circuits anchored in the dorsal and ventral left anterior insula. *Scientific Reports*, 10(1), 16394.

Wen, X., Liu, Y., Yao, L., & Ding, M. (2013). Top-down regulation of default mode activity in spatial visual attention. *Journal of Neuroscience*, 33(15), 6444–6453.

Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. Journal of Neuroscience*, 29(26), 8525–8529.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8(8), 665–670.

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience*, 15(5), 675–680.

Note: Figure translations are in progress. See original paper for figures.