The Neuroscience of Intergroup Emotion

Marius C. Vollberg and Mina Cikara

Harvard University

William James Hall, 33 Kirkland Street, Cambridge, MA, 02138

Corresponding author: mcikara@fas.harvard.edu
What happens to our emotions and in our brains when we experience the world through the lens of our group memberships rather than as individuals? Here we review recent advances in social and affective neuroscience that have identified potential input variables and processing mechanisms underlying one widely studied emotion in intergroup contexts: empathy. There is a well-documented in-group bias in empathy but the mental processes that generate it are poorly understood. Drawing from recent insights in memory research, we suggest that episodic simulation—the ability to imagine events—is an underexplored candidate process that is likely to be involved in shaping emotional experience in intergroup settings.
For more than a century, the idea that social groups shape human behavior has been a pillar of social psychology [1–3]. Numerous experiments have documented the consequences of our propensity to see ourselves not just as individuals but also in terms of our group memberships [4,5]. Here we focus on how group membership shapes emotion. According to Intergroup Emotion Theory, self-categorization into groups (see Self-Categorization Theory [6]) changes emotional appraisal of group-relevant events [7]. In other words, self-categorization can cause one’s emotional repertoire to reflect the priorities and interests of the group instead of the individual.

Based on the observation that groups influence emotion and following advances in neuroimaging, there has been growing interest in the neural mechanisms underlying intergroup emotion. Much of the recent social and affective neuroscience research in this area has examined how group membership shapes our responses to others’ suffering: empathy. Specifically, mounting evidence indicates that people are less likely to empathize with others when they are socially distant, for example when they belong to different racial or national groups [8]. This intergroup empathy bias is reflected in a variety of physiological indices, which have offered unique insights into when and why people empathize less with out-group members.

**Intergroup empathy**

Empathy is an umbrella term that encompasses cognitive and affective components of how we react to others’ experiences and emotions [9]. The cognitive component refers to understanding what the target is feeling in the absence of any concomitant affect; this component is typically associated with a distributed set of brain regions including medial prefrontal cortex (mPFC), temporoparietal junction (TPJ), temporal pole, and precuneus. The affective component, on the other hand, refers to an observer’s emotion in reaction to someone else’s emotional state, where the valence of the observer’s emotion can be congruent (empathic) or incongruent (counter-empathic) with that of the experiencer. Brain regions linked to the congruent affective component typically include anterior cingulate cortex (ACC), anterior insula (AI), inferior parietal lobe (IPL), premotor cortex, amygdala, and sensorimotor
cortex [10,11]. Although these two components have been consistently linked to distinct neural substrates, open questions remain as to how this should be interpreted. The AI and ACC, for example, are two areas reliably associated with both the first-hand experience of pain and empathy for another person’s pain. However, those areas are also involved in a range of functions of no specific relevance to empathy, including valuation and the processing of salient changes in the environment [12,13]. Therefore, some have argued that these areas relate to empathy for pain due to their general function of encoding salient information related to threat [14,15]. Although further research is required to establish the exact relationship between these regions and the psychological processes they support, there is little doubt that our capacity to share and understand each other’s emotions is central to successful interactions and relationships [11,16].

Empathy researchers are further unified in the observation that social groups affect neural and behavioral responses to others’ emotions [17**]. This observation pertains to empathy for pain and misfortunes but also to perceivers’ motor resonance with in-group and out-group targets, which is thought to be directly related to empathy [18]. Specifically, several studies have documented reduced motor resonance with out-group relative to in-group targets [19–21]. For example, watching in-group members as opposed to out-group members receive an injection resulted in increased event-related desynchronization of beta rhythms in sensorimotor cortex, which the authors interpreted as greater resonance with in-group pain [22*].

This pattern of results replicates across numerous functional magnetic resonance imaging (fMRI) studies [23–26]. For example, participants in an fMRI study [25] showed increased blood-oxygen-level dependent (BOLD) activity in the ACC when watching members of their racial in-group (Caucasian or Chinese) relative to the out-group being pricked by a needle. Similarly, another fMRI study [24] revealed increased activity in the left AI when participants observed preferred versus rival sports fans receive painful stimulation. Notably, AI activity predicted participants’ willingness to absorb some pain on behalf of fellow fans, consistent with the hypothesized role of empathy in altruism [27]. (Intergroup empathy bias among soccer fans was not limited to reduced empathy for out-group members; out-group pain was also associated with activity in the ventral striatum, a brain region that has
been previously associated with schadenfreude, or pleasure in response to out-group members’ suffering [26]. Although biased empathy for pain has been linked to several brain areas, the ACC and AI bias pattern in particular has replicated across cultures including Chinese [28], Australian [29], and European [30] populations [for review see 31**].

Note, however, that there are several findings which challenge the simple version of the ACC/AI-intergroup empathy bias story. For example, in one study, participants, who viewed images of same-race and other-race targets suffering in the aftermath of Hurricane Katrina, exhibited similar degrees of ACC and AI activation across both conditions. The region that was more active in this context for in-group relative to out-group suffering was mPFC—what the authors dubbed a signal of extraordinary empathy for the in-group—which predicted participants’ reports of willingness to donate money and time to the people in the images [32]. In another fMRI study, Arabs and Israelis exhibited equivalent cingulate and AI responses to stories of in-group and out-group pain, which conflicted with their self-reports of significantly greater compassion for in-group relative to out-group members [33]. In contrast to other studies, which rely on static images of pin-pricks or symbols indicating mild electric shocks, these studies expose participants to highly elaborated scenes, which may be driving these divergent findings (see next section). But by including both the self-report and the neuroimaging data, these studies have raised important questions for both psychologists and neuroscientists about the component processes that drive empathy and helping behavior.

While most of the research discussed here relies on imaging approaches at the level of entire brain regions, approaches looking at neurochemistry have also provided evidence in support of a neural machinery that facilitates intergroup empathy bias. Specifically, some have argued that evolution has co-opted oxytocin, a neuropeptide implicated in pair-bonding, to facilitate in-group cooperation and intergroup competition by upregulating in-group empathy [34]. In fact, mounting evidence indicates that intergroup bias in both empathic and counter-empathic emotions play a central role in many intergroup conflicts [35,36].

Given that the neural architecture that underlies empathy also reflects intergroup bias with such far-reaching consequences, it raises the question of how we might counteract these tendencies. While there is a vast body of literature on the reduction of prejudice and intergroup conflict [37], few studies have tried to decrease intergroup empathy bias [38–40],
and even fewer have incorporated corroborating measures of neural activity. In line with an emerging trend to formalize human social behavior with computational models [41,42], a recent study has addressed this exact gap using a reinforcement learning approach [43**]. In this study, confederates posed as members of an ethnic out-group (in-group in the control condition) and gave up resources to save participants from receiving electric shocks. Neural prediction errors in the AI in response to this helping behavior were positively associated with participants’ subsequent empathy for yet another out-group member. In other words, the more participants were positively surprised by the out-group member’s altruism (as indexed via activity in the AI), the more they empathized with another member from that group. In sum, research on intergroup empathy indicates that sharing and understanding the emotions of others is a central component of intergroup emotion with promising potential for bias reduction.

**Beyond appraisals: New work in episodic simulation**

Our emotions are demonstrably influenced by factors that lie beyond events themselves, such as group membership. But how exactly are these factors incorporated to give rise to said emotions? Many theoretical accounts emphasize the role of subjective appraisals, which refer to the way in which we evaluate a given event [44–48]; though notably, extant descriptions of the underlying mechanisms driving appraisals are incomplete and a systematic integration of neuroscientific findings is lacking [49].

Some of the efforts to address this gap have identified cognitive processes that plausibly contribute to appraisals but are not specific to emotions. One of those processes is the aforementioned detection of salient events, which is commonly thought to precede appraisal [50]. In the context of group identities, a recent fMRI study using multivoxel pattern analysis (MVPA) looked at the neural patterns associated with the classification of in- versus out-group targets across arbitrary and political group boundaries. Crucially, participants exhibited distinct signatures for in- versus out-group members (indifferent to whether they were categorized by arbitrary or political group membership) in two regions that both belong to a network associated with salience (and empathy) processing: ACC and AI [51*]. A second, general process that is well studied and related to empathic responses is the representation of
other’s mental states and affect [52]; if there is no representation of another’s emotions, affective reactions are unlikely to emerge as well. A third, far less studied process that might affect intergroup emotion is episodic simulation.

Constructive episodic simulation is the ability to imagine events, where “constructive” refers to the flexible use of episodic memories as building blocks for imagination [53]. The concept of episodic simulation is intricately linked to research on the so-called default mode network (DMN). The DMN is associated with internal modes of cognition, including perspective-taking, as well as prospection and memory retrieval [54]. Anatomically, this network consists of the ventromedial prefrontal cortex (vmPFC), posterior cingulate/retrosplenial cortex, IPL, lateral temporal cortex (LTC), dorsal medial prefrontal cortex, and the hippocampal formation (hippocampus, entorhinal cortex, and parahippocampus).

A recent study revisited the DMN by looking at its structural organization within (rather than across) individuals [55]. At this level of analysis, the DMN fractionates into two parallel networks A (including posterior IPL, lateral temporal cortex, ventromedial PFC, retrosplenial/ventral posteromedial cortex, and parahippocampal cortex) and B (including TPJ, lateral temporal cortex, an inferior region of ventromedial PFC, a dorsal region of anteromedial PFC, and posterior cingulate cortex). This fractionation speaks to the DMN’s intricate relationship with episodic simulation because separate investigations that averaged across individuals recently associated the construction of imagined events with activity in regions associated with both networks A and B: right IPL, right posterior cingulate cortex, right ventral precuneus, and left anterior hippocampus [56]. Not only are these correlates consistent with the traditional association of the hippocampus with memory and spatial navigation [57] but they also suggests substantial overlap with both interdigitated networks of the DMN.

These findings are intriguing because the DMN has also been argued to reflect the exceptionally social nature of our species as it supports perspective taking and social cognition more broadly [58–60]. Indeed, recent investigations suggest social and memory processing to be deeply intertwined and related to the DMN: connectivity between some of its core regions (TPJ and mPFC) during rest following a social knowledge task was shown to be positively associated with retrieval of that social information [61]. We believe that this is an
important avenue for intergroup neuroscience research, since episodic simulation might be
differentially engaged when simulating events related to in- versus out-group members. There
is no neuroscientific study to date that has specifically addressed this question. However,
some studies investigating intergroup empathy bias are consistent with involvement of parts
of the hippocampal formation in empathy for the in- but not for the out-group [24], or for
loved ones compared to strangers [62].

Meanwhile, behavioral studies have already begun to suggest a link between episodic
simulation and helping behavior. One study [63**] discovered that imagining or remembering
events of helping others increased participants’ willingness to help a present person in need;
this effect was driven by the vividness of the (re-)constructed event. Going beyond this
general effect, another study applied this approach to prosocial intentions in an intergroup
setting and found that the specificity of an imagined helping episode fully mediated the
elimination of a willingness to help gap between in-group and out-group targets [64].
Importantly for the current focus, recent evidence suggests that, at least regarding willingness
to help, it is indeed the vividness of the simulation that matters regardless of whether we
imagine ourselves or someone else as the agent within it [65]. A related series of experiments
indicates that considering moral scenarios with one’s eyes closed increases simulation,
thereby intensifying emotional reactions to those scenarios and altering subsequent judgments
and behaviors [66].

In trying to explain the link between emotionally relevant events and emotional
responses—especially in intergroup settings—concomitant processes such as episodic
simulation could constitute much more than an epiphenomenal peculiarity. Instead, we
suggest episodic simulation is a candidate mechanism that alters the relative salience and
subsequent impact of different aspects of an event on emotional experience and behavior.
Whereas perspective-taking describes the attribution of mental states to an individual,
episodic simulation constitutes a distinct process that generates the setting in which the
individual and their mental states are embedded. This separate pathway might provide
additional explanatory power in accounting for a range of phenomena. The identifiable victim
effect which describes the preferential allocation of resources to identifiable as opposed to
anonymous victims [67,68], for example, could partially result from altered episodic
simulation as a function of identifiability. Importantly, such effects would be expected to manifest similarly in intergroup contexts.

Conclusion
The last couple of decades have seen substantial advances in our understanding of affect, intergroup dynamics, and neuroscience. Here, we have focused on intergroup empathy to highlight ways in which neuroscience has informed key questions concerning the function, specificity, and interdependence of brain regions involved in intergroup emotion; in particular, we have pointed out salience and appraisal processing as related but not specific to intergroup emotion. The reinterpretation of memory-related brain circuitries as central to social cognition constitutes yet another example of how cognitive neuroscience findings may serve to reveal additional component processes. Specifically, we suggest the role of episodic simulation in intergroup emotion to be a promising avenue for future research in intergroup emotions and affective neuroscience more generally.
Acknowledgments

This work is funded by NSF BCS-1551559 awarded to MC.
References

** Review focusing on social psychological theory and functional relations between groups in social cognitive neuroscience. Provides an overview of relevant neuroscience research and open questions for interdisciplinary investigations comining intergroup relations and cognitive neuroscience
* Electroencephalographic investigation comparing activity in sensorimotor cortex in response to painful and painless stimulation administered to ethnic ingroup versus outgroup hands. Results are consistent with group-based bias, but only for painful stimulation.


** Comprehensive review of EEG and fMRI findings documenting neural substrates of racial ingroup bias in empathy across a range of brain regions, time-windows and ethnic groups.


** FMRI study in which participants received help from an in-group or an out-group member. Reinforcement learning models applied to neuroimaging data revealed subsequent increases in empathy for the out-group are associated with prediction errors in the anterior insula cortex during the intervention.


* FMRI study using MVPA to identify for the first time specific neural activation patterns associated with the classification of individuals as in-group or out-group members across political and arbitrary group boundaries.


** Behavioral study documenting for the first time the role of episodic simulation in prosocial cognitions and behavior. Imagining a helping episode significantly increased participants’ willingness to help compared to control conditions that do not involve helping or episodic simulation.


