Peace Psychology Book Series
Series Editor: Daniel J. Christie

Eran Halperin
Keren Sharvit *Editors*

The Social Psychology of Intractable Conflicts

Celebrating the Legacy of Daniel Bar-Tal, Volume I

Springer

Chapter 11
Putting Neuroscience to Work for Peace

Emile Bruneau

How can seemingly mild-mannered people be induced to murder their neighbors? What motivates someone to strap a bomb to his body? Why are some human conflicts so seemingly resistant to resolution? The immense human tragedy of ethno-political violence, war and genocide has deeply motivated social psychology, spawning some of the most powerful demonstrations of human susceptibility to violence and conflict, and launching large research efforts to better understand the psychological underpinnings of intergroup conflict. At the forefront of this effort has been the social psychologist Daniel Bar-Tal. Writing from the context of the Arab–Israeli conflict, Bar-Tal has been an academic force in the effort to better understand the psychological processes that grip communities in conflict. Bar-Tal offers a comprehensive description of the array of biases that are at play particularly in conflicts characterized by cycles of violence and failed peace efforts, which include "hot" emotional biases (anger, fear, and (lack of) empathy) (Cikara et al., 2011b, 2013; Halperin et al. 2008; Tam et al. 2007), and a suite of "cold" high-level cognitive biases (devaluation of out-group compromises, uncritical acceptance of belief-confirming evidence) (Porat et al. 2015; Ross and Ward 1994, 1996). The integrated model articulated by Bar-Tal and colleagues suggests that these emotional and cognitive biases are bound together with social factors by an "ethos of conflict," which provides biases with a scaffold and helps to freeze them in place (Bar-Tal 2007; Bar-Tal and Halperin 2011, see also Cohrs et al.; Jost et al.; Oren vol. 2; Sharvit vol. 2; Nahhas et al. vol. 2).

In all, this model provides us with a rich descriptive tapestry of the psychological landscape faced by groups in conflict. Where do we go from here? How do we go from theoretical richness to mechanistic understanding? And how can we proceed from describing and demonstrating these psychological barriers, to effectively dismantling them?

E. Bruneau (✉)
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology,
Cambridge, MA 02139, USA
e-mail: ebruneau@mit.edu

E. Halperin, K. Sharvit (eds.), *The Social Psychology of Intractable Conflicts,*
Peace Psychology Book Series 27, DOI 10.1007/978-3-319-17861-5_11

I will argue in this chapter that one path lies through examining psychological biases at their cognitive roots, by looking "under the hood" directly at neural activity using functional neuroimaging techniques. In particular, I will present some ways in which neuroimaging technologies—functional magnetic resonance imaging (fMRI), in particular—can carry the torch of this research effort forward. In three sections, I will offer some examples of how neuroimaging can (1) help characterize intergroup biases, (2) expand our theoretical understanding of the psychological processes driving intergroup conflict, and (3) aid practical evaluations of conflict resolution efforts.

Neuroimaging to Identify Neural Correlates of "Hot" and "Cold" Psychological Biases

A primary goal of most social cognitive neuroscience studies is to identify the cognitive mechanisms involved in psychological processes: to establish the cognitive underpinnings of bias. In functional neuroimaging studies, the association between psychology and cognition is generally inferred through the demonstration of localized brain activity. This inference can be better justified if the brain region activated by a psychological paradigm is cognitively well characterized. Therefore, a first approach to examining the neurocognitive basis of intergroup conflict is to design experiments with hypotheses tailored to specific brain regions where the connection between functional activity and cognitive process is well established (Fig. 11.1).

By way of illustration, the fusiform face area (FFA) is a small patch of cortex on the underside of brain above the cerebellum that is selectively active in response to faces (Kanwisher et al. 1997). If a psychological prediction can be molded to this cognitive reality—for example, that groups in protracted conflict process enemy faces similarly to how they process animal faces—then this specific brain region may be able to be used to ask mechanistic questions of a specific psychological process (e.g., is dehumanization encoded in lower-level perceptual processing?). This approach has the benefit of linking a psychological paradigm to a strong neuroscience scholarship, which can provide practical and theoretical benefits: the FFA is anatomically indistinct from the surrounding cortex, but it can be localized within each individual after performing a task in the scanner known to activate the FFA (e.g., looking at a validated battery of race-neutral faces versus objects). A "functional localization" makes the inferred connection between a cognitive mechanism and a psychological process less problematic (Saxe et al. 2006a). This example serves as a hypothetical—current imaging techniques and paradigms may not be sensitive enough to detect the FFA responses to human versus animal faces (though the technology and the analysis techniques are always improving in the young field of cognitive neuroscience). However, this overall approach has been used success-

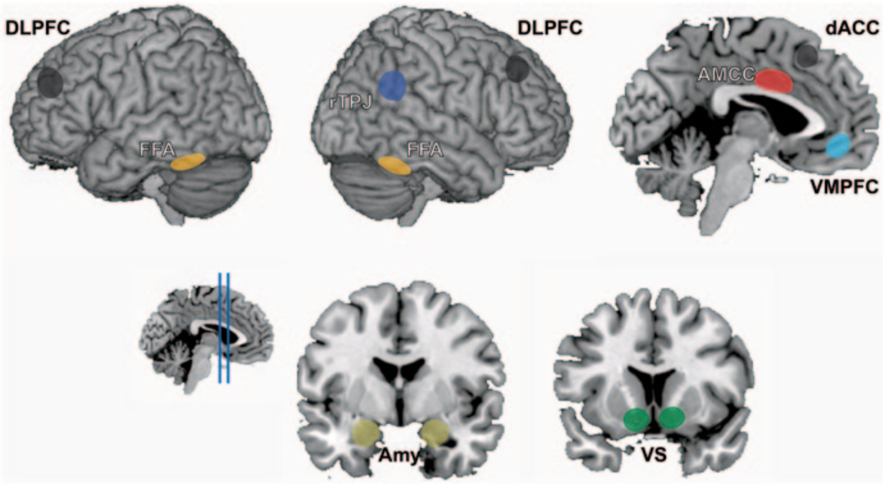

Fig. 11.1 A selection of brain regions potentially relevant to intergroup conflict. Activity in each labeled brain region is relatively well characterized to be selective for a specific cognitive function; included is a classic paper or recent review for each. *Thinking about other people's thoughts/intentions/desires* (*rTPJ* = right temporoparietal junction) (Saxe and Kanwisher 2003); *subjective value* (*VMPFC* = ventromedial prefrontal cortex) (Bartra et al. 2013); *cognitive control of emotion* (*DLPFC* = dorsolateral prefrontal cortex, *dACC* = dorsal anterior cingulate cortex) (Ochsner, Silvers and Buhle 2012); *face perception* (*FFA* = fusiform face area) (Kanwisher et al. 1997); *perceived threat* (*Amy* = amygdala) (LeDoux 2007); *pleasure/schadenfreude* (*VS* = ventral striatum) (Cikara et al. 2011a); *physical pain in self and others (i.e., empathy for physical pain)* (*AMCC* = anterior middle cingulate cortex) (Bernhardt and Singer 2012)

fully by a number of groups to examine responses to in-group and out-group faces in a different brain region: the amygdala (Golby et al. 2001; Hart et al. 2000; Lieberman et al. 2005).

The amygdala is an anatomically constrained subcortical brain region that is integral to fear conditioning in mammals, and extends to encompass a broad range of threat associated stimuli in humans (LeDoux 2007; Zald 2003). In one illustrative study (Cunningham et al. 2004), images of White and Black Americans were presented to White participants in the scanner either subliminally (i.e., faster than conscious perception), or supraliminally (i.e., slow enough to be consciously perceived). They found greater amygdala activity in response to out-group (Black) versus in-group (White) faces. Moreover, subliminal presentation of the faces resulted in a strong amygdala response, but when participants were consciously aware of the images, amygdala response was muted, and well-characterized regions of the brain in the dorsolateral prefrontal cortex (DLPFC) associated with cognitive control came online. These results suggest that the amygdala response is unconscious, and partially suppressed with conscious awareness through the DLPFC. This overall interpretation is bolstered by evidence that (1) the strength of amygdala activity was associated with implicit measures of bias (IAT d-score), and (2) that modulation of amygdala activity was associated with increased activity in DLPFC.

This study provides a potential cognitive description of the out-group fear response, and also illustrates two potentially useful features of neuroimaging measures. First, they can provide multidimensional measures of bias—i.e., amygdala response (unconscious fear) and the DLPFC response (conscious regulation of negative affect)—that may each be useful in evaluating de-biasing efforts. Second, they provide measures of bias that are unconscious. Both of these implications will be explored more fully in the following sections.

Another approach to better understanding the neural processes involved in psychological biases is to generate research paradigms that are less constrained to the existing neuroscience literature, but that may reveal something new about neural processing. For example, high-level cognitive biases involved in the acceptance or rejection of ideological narratives may not map to well-characterized brain regions, but may instead provide insights into the role of specific brain regions in everyday cognition. In a study from our lab (Bruneau and Saxe 2010), and one of the few neuroimaging studies to examine processes in actual conflict group members, Arab and Israeli participants read short statements that were typical either of the Arab narrative about the Arab–Israeli conflict ("Israel is like a modern day Apartheid state…"), or the Israeli narrative about the conflict ("Palestinians could have a modern state next to Israel, but instead they have chosen violence…"). In a single region within the precuneus (PC), mean activity was higher for out-group versus in-group narratives, and the difference in response to out-group versus in-group narratives correlated with both explicit warmth felt toward the out-group versus the in-group ($r=0.64$, $p<0.001$), and Arab–Israeli IAT score ($r=0.69$, $p<0.001$). A follow-up study with American political partisans supports these findings: Democrats and Republicans asked to evaluate political arguments supporting out-group legislation (versus in-group legislation) show activity in a region of the PC similar to that seen in Arabs and Israelis (Bruneau, Coronel and Saxe, unpublished).

The results from these studies indicate that this region of the PC may be specifically sensitive to "motivated reasoning" about group ideological narratives. These studies flag this part of the PC as a brain region of interest that could help drive a high-level cognitive bias integrally involved in ideological conflict: the delegitimization of the other side's narrative (Bar-Tal and Halperin 2011). What cognitive process this activity represents (e.g., self-referential thinking, autobiographical recall, social cognition) is less clear, as all of these processes (and more) have previously been associated with PC activity. However, since this task generates such strong and localized neural activity, it may help in the process of teasing apart the functional subregions within the PC, which would then provide a better characterized brain region for future research.

In sum, one role of neuroimaging is to help characterize the neural mechanisms behind a psychological bias. These characterizations could help answer novel questions about bias (e.g., How low-level and perceptual is dehumanization?), identify multiple components of bias (e.g., emotional processing in amygdala and effortful control in DLPFC), or help map a brain region that may be involved in a particular bias (e.g., involvement of the PC in delegitimization).

Neuroimaging to Expand Theoretical Understanding of Psychological Phenomena

A second way in which neuroimaging can contribute to our understanding of intergroup conflict is by broadening and deepening our theoretical understanding of complex psychological biases. While some "hot" biases have rather specific psychological and physiological profiles (e.g., fear and anger), others are inherently ambiguous and multifaceted. Empathy, for example, refers to at least eight different phenomena (Batson 2009) that are likely driven by completely distinct cognitive mechanisms, that in turn motivate very different behavioral outcomes: while empathy defined as self-focused "personal distress" predicts avoidance of others in need, empathy defined as other-focused "empathic concern" predicts the opposite (Batson et al. 2002; FeldmanHall et al. 2015). Understanding which of these empathic responses someone is expressing is therefore important, and it may be possible to accurately and efficiently define and distinguish these processes with neuroimaging. For example, work from our lab and others show that empathic concern and personal distress may be neurally distinguishable (Bruneau et al. unpublished; Lamm et al. 2007).

"Cold" reasoning biases pose an even greater problem. Since high-level cognitive processes are generally opaque to introspection (Nisbett and Wilson 1977), people have a "bias blind spot," making them not just *unwilling*, but *unable* to assess (or even acknowledge) their own biases (Pronin et al. 2004; Pronin et al. 2002). Furthermore, high-level biases inherently describe complex cognitive processes, but are defined by single outcome measures, like reaction time or between-subject preferences. This poses a fundamental problem to fully characterizing the mechanisms of high-level biases, but this problem is potentially tractable with neuroimaging.

To date, few neuroimaging studies have examined high-level cognitive biases, much less in members of conflict groups. However, there is enough known about cognitive processing in relevant domains to allow us to envision the types of imaging studies that could bear fruit. Take, for example, reactive devaluation. Reactive devaluation describes the tendency to devalue, and therefore reject, a compromise proposal if it is delivered by someone affiliated with the opposing group; when the same proposal is delivered by an in-group affiliate, it is valued more highly and accepted more readily (Ross 1993). While the phenomenon is straightforward to measure behaviorally, the mechanism responsible for it is not: at least half a dozen mechanisms, all with potentially distinct cognitive processes, have been suggested to drive reactive devaluation (Ross 1993). For example, reactive devaluation could be generated from effortless heuristics ("The enemy is proposing it, so it must be bad, whatever it is"), or effortful cognitive processing ("That sounds good, but what are their *real* intentions?"). Therefore, the single behavioral outcome measure used to define reactive devaluation belies the heterogeneity of cognitive processes that could generate it.

An examination of the neural basis of reactive devaluation could start with a number of candidate brain regions, including the DLPFC (involved in cognitive control), and the right temporoparietal junction (rTPJ; involved in "Theory of Mind"). Theory of Mind is the ability to think about other people's thoughts—the process of "getting inside their heads." While Theory of Mind tasks result in activity in a network of brain regions, only the rTPJ appears to be *selectively* sensitive to reasoning about other people's thoughts, desires, and intentions (Saxe and Kanwisher 2003; Saxe et al. 2006b; Scholz et al. 2009). Measuring neural activity in these two brain regions during a reactive devaluation task may differentiate between cognitive strategies: if two participants register equal amounts of reactive devaluation behaviorally (i.e., equally, and more strongly, endorse proposals attributed to the in-group versus the out-group), but subject 1 shows more rTPJ and DLPFC than subject 2, subject 1 may be reactively devaluing more through consideration of out-group intentions, while subject 2 is using a heuristic approach. Ultimately, understanding *how* someone is reactively devaluing, rather than by *how much*, may better inform interventions aimed at decreasing bias. Most importantly, this neuroimaging approach could be similarly applied to unpack the neural mechanisms behind an array of psychological biases that are potentially cognitively heterogeneous, like biased assimilation and the fundamental attribution error.

While the above applications of neuroimaging to intergroup conflict mostly helps extend or solidify our mechanistic understanding of bias, neuroimaging also has the potential to provide a quantitative measure of bias that is directly proximal to behavior and potentially immune to self-reporting biases. In the final section, I turn from the theoretical towards the more practical and explore how neuroimaging could provide prospective and retrospective quantitative measures of psychological biases that could be used to evaluate conflict resolution interventions.

Neuroimaging as a Tool to Evaluate De-biasing Efforts

While Daniel Bar-Tal and others offer comprehensive theoretical descriptions of intergroup bias, many of the specific biases lack quantitative measures. For example, delegitimization describes the tendency to negate the other side's perspective or narrative about a conflict, but this cognitive process has proven difficult to operationalize through psychological measures. Neuroimaging studies, like the one described above with Israelis and Arabs over ideological narratives, puts us on the path towards building a cognitive profile and quantitative measure of delegitimization. Even when quantitative psychological measures have been developed for other biases, these measurement techniques have their limitations, which functional neuroimaging may be able to surmount.

Currently, the most common way to assess intergroup bias is explicitly, with self-report measures. Explicit measures are simple and convenient, but pose well-known methodological challenges because participants are motivated to present themselves in a positive light (Greenwald and Banaji 1995), and therefore may be unwilling to report their honest views. For example, White Americans who express positive attitudes and behavioral intentions towards Black Americans nevertheless show impulsive avoidance of a Black confederate (Dovidio et al. 2002). In the context of protracted intergroup conflict, demand characteristics may actually work in the other direction—a group's norm about expressions of general out-group antipathy may drive partisans to report *more* prejudice than they actually feel. Either way, the explicit measure fails to capture reality.

An alternative approach to assessing intergroup prejudice, or affect more specifically, is through implicit measures that tap physiological changes (e.g., heart rate, blood pressure, and skin conductance) (Amodio et al. 2003; Olsson et al. 2005) or response latency (Dovidio et al. 2002; Dovidio et al.1997). The most widely used response latency measure is the implicit association test (IAT) (Greenwald et al.1998), which has been used to assess implicit bias towards groups, including those defined by race, gender, and political partisanship (Aberson et al.2004; Greenwald et al.2003; Knutson et al.2007b; Phelps et al. 2000). Implicit tests have been shown to be less susceptible to cognitive control: even when participants are aware that the test is being used to assess bias, the effect remains (Kim 2003). There are also limitations to standard implicit measures, however. First, the output is usually a single measure generalized to positivity or negativity, so multiple interacting processes could be confounded. Second, as a difference measure, the IAT does not distinguish between "in-group love" and "out-group hate." Third, what exactly the IAT measures is still debated, particularly since the IAT has been shown to be influenced by priming effects and training (Foroni and Mayr 2005; Kawakami et al. 2007).

Neuroimaging has the potential to sidestep the limitations of both explicit and implicit assessments, providing a measure that is both transparent to demand characteristics and multidimensional. One way to evaluate the utility of neuroimaging measures to de-biasing or conflict resolution efforts is to determine how well they predict relevant behaviors (e.g., willingness to negotiate rather than fight) relative to explicit and implicit measures. If the predictive validity of imaging measures is significantly better than explicit or implicit measures, or if it helps account for enough variance that is unexplained by explicit and implicit measures, then its use may be justified. As proof of principle, neuroimaging measures have been demonstrated to outperform behavioral tasks in predicting outcomes in a number of domains (Gabrieli et al. 2015), including forecasting reading skills in children (Hoeft et al. 2007), which would improve with training in dyslexic children (Hoeft et al. 2011), and even recidivism rates in incarcerated criminals (Aharoni et al. 2013).

Neuroimaging techniques may be particularly effective for interventions targeting "cold" reasoning biases that are more inaccessible to conscious introspection and subject to strong demand characteristics. For example, one proposed intervention involves inoculating people against reactive devaluation by educating them about the phenomenon itself. Bias inoculation has been shown to effectively attenuate or eliminate other biases, such as stereotype threat (Johns et al. 2005), but evaluating the effect of the intervention on a self-report measure like reactive devaluation is much more problematic than for a performance-based measure like stereotype threat. Therefore, even though preliminary data shows that participants taught about reactive devaluation decrease their expressions of this bias (Bruneau and Saxe unpublished; Halperin et al. unpublished), the demand characteristics are extremely high, making it difficult to distinguish between true belief change and mere social conformity. How can these two possibilities be distinguished?

Neuroimaging has already demonstrated an ability to distinguish between real change and inaccurate self-report—not with studies on reactive devaluation, but through studies on social conformity (Asch 1956; Cialdini and Trost 1998). In one fMRI conformity study (Zaki et al. 2011), participants were asked to rate the attractiveness of opposite gender pictures while undergoing fMRI. After rating the faces, participants were then told how their peers (supposedly) rated the same faces. At the end of the study, participants were then given the opportunity to revise their attractiveness ratings to the same pictures. Behaviorally, participants showed the classic social conformity response: attractiveness ratings increased or decreased slightly if their peers had rated the pictures as more or less attractive, respectively. But the neuroimaging data added critically to this picture by examining activity in the ventromedial prefrontal cortex (VMPFC), a brain region shown to provide "neural currency" for subjective value (Levy and Glimcher 2012) across a range of rewarding stimuli, from monetary (Knutson et al. 2007a) and gustatory (Plassmann et al. 2007), to aesthetic (Kirk et al. 2009) and social (Lin et al. 2012). In their study, Zaki et al. (2011) found that activity in this subjective value brain region increased between initial and final ratings if the peer group judged the face to be more attractive than the participant, and decreased between initial and final ratings if the peer group's attractiveness rating was lower than the individual's. In other words, conformity to peer ratings indeed was matched with changes in actual subjective neural value, suggesting that the participants *actually* saw the face as more or less attractive, according to peer ratings.

While social conformity is directly relevant to intergroup conflict (Paluck 2009), the relevance of these studies to conflict resolution goes beyond the specific instance of social conformity—it illustrates that neuroimaging could be used to measure subjective and implicit value that cuts through demand characteristics. For example, a current fMRI study in our lab aims to use the VMPFC value signal to measure how much someone reactively devalues out-group versus in-group compromise proposals, both before and after an inoculation intervention. This could sidestep demand characteristics entirely to provide a measure of cognitive bias immune to demand characteristics.

The above examples illustrate the potential of neuroimaging to retrospectively evaluate the effect of an intervention on biases that people may be unwilling or unable to report. Another exciting possibility is that neuroimaging may be deployed in intergroup conflict settings as a *prospective* tool, to determine which of a set of potential interventions may be most effective for a group, demographic, or individual.

Prospectively assessing interventions is the principle behind focus groups: if you want to know what intervention will best convince people to engage in a particular behavior—to use sunscreen, to avoid unprotected sex—you can convene a small group of people (a "focus group") and obtain their opinions about a number of different test interventions. The problem with this model is that people are notoriously limited in their ability to predict their own future behaviors (Nisbett and Wilson 1977), making focus groups imperfect predictors of future population-level behavior (Noar 2006). This could result from our tendency to discount the power of implicit processes in decision-making: we may be consciously aware of some aspects of a sunblock commercial (consequences of sunburns), but those may have less of an impact on our eventual behavior than the aspects of the commercial that appeal to unconscious processes (attractiveness of the spokesperson, normative views about the use of the product, and color and shape of the sunscreen bottle). But all of these components of the message are being processed, even if implicitly, and so could potentially be captured with neuroimaging.

Although neuroimaging has not yet been used as a forecasting tool in the context of intergroup conflict, a study by Emily Falk et al. (2011) demonstrates the utility of this approach for a large scale public health intervention. In their fMRI study, smokers with the intent to quit watched ads from three different antismoking campaigns, and then provided their assessment of which ad campaign they thought would be most effective, just like a traditional focus group. Participants were also scanned while viewing the ads to obtain a measure in the VMPFC of their subjective value of the ads. When the ads were actually aired, call volume to the 1-800-QUIT number that appeared at the bottom of each ad was predicted by smokers' neural responses, and not by their explicit self-report predictions.

While using a "neural focus group" to predict population level outcomes could be used to adjudicate between potential public service interventions aimed more at ethnic conflict (e.g., to ease anti-Roma bias in Europe), neural measures could also aid in individual forecasting, akin to the recent trend towards "personalized medicine." Although this has again not been implemented in the context of intergroup conflict resolution, individual forecasting using fMRI has shown promise in predicting the effectiveness of treatments for those suffering from post-traumatic stress disorder (PTSD). The treatment of choice for PTSD is cognitive behavioral therapy (CBT; as opposed to pharmacological interventions), however responsiveness to CBT is only approximately 50 %. Initial results from a neurodiagnostic study found that activity in the amygdala in response to threatening images in PTSD patients predicted positive outcomes to CBT treatment 6 months later (Bryant et al. 2008). Functional and structural MRI have provided similarly promising results for predicting treatment outcomes for pharmacological or CBT treatment across a number of other neuropsychiatric conditions, including depression (Fu et al. 2013; Pizzagalli 2010), schizophrenia (Kumari, Antonova, Fannon and Peters, 2010; Kumari et al. 2009), and social anxiety disorder (Doehrmann et al. 2013). This same tailored approach could theoretically be applied to conflict resolution programs, to determine which type of intervention might have the greatest likelihood of success for an individual (and which may backfire). Although a long way off, personal forecasting of conflict resolution programs is the type of innovation that functional neuroimaging could enable.

Conclusion

For the past 40 years, Daniel Bar-Tal and others have provided a detailed description of the psychological forces arrayed against conflict resolution efforts. The daunting task in front of us now is determining how to dismantle these psychological barriers to peace. One way forward is to build evidence-based interventions (see Hameiri and Halperin), and then evaluate using randomized control trials (RCTs). RCTs have been implemented previously in field assessments of full-scale social programs, such as inner-city antiviolence campaigns (Webster et al. 2013), antipoverty programs (Banerjee et al. 2013), and "edutainment" aimed at changing norms of intergroup violence and preventing genocide in Africa (Paluck 2009; 2010). Another way to enact evidence-based approaches is through examining the effect of specific, small-scale interventions, such as perspective-taking, emotion regulation, or paradoxical thinking in more controlled or laboratory settings (Bruneau and Saxe 2012; Halperin et al. 2013; Hameiri et al. 2014). While these approaches are rare (Paluck and Green 2009), the results cited above are promising, and have led to large-scale sviolence and antipoverty efforts.

A key difficulty with applying this approach to conflict resolution efforts is that many of the biases that we might wish to mitigate are (1) cognitively complex and operationally un- or ill-defined, (2) subject to demand characteristics (3), and/or inaccessible to introspection. Neuroimaging has the potential to aid in all of these limitations by allowing a broader and deeper understanding of the forces driving conflict, a means to measure them, and a tool by which to forecast their effects.

References

Aberson, C. L., Shoemaker, C., & Tomolillo, C. (2004). Implicit bias and contact: the role of interethnic friendships. *Journal of Social Psychology, 144,* 335–347.

Aharoni, E., Vincent, G. M., Harenski, C. L., Calhoun, V. D., Sinnott-Armstrong, W., Gazzaniga, M. S., et al. (2013). Neuroprediction of future rearrest. *Proceedings of the National Academy of Sciences, 110,* 6223–6228.

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology, 84,* 738–753.

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied, 70*(9), 1–70.

Banerjee, A. V., Duflo, E., Glennerster, R., & Kinnan, C. (2013). The miracle of microfinance? Evidence from a randomized evaluation. MIT Department of Economics Working Paper No. 13-09. http://ssrn.com/abstract=2250500 or http://dx.doi.org/10.2139/ssrn.2250500.

Bar-Tal, D. (2007). Sociopsychological foundations of intractable conflicts. *American Behavioral Scientist, 50,* 1430–1453.

Bar-Tal, D., & Halperin, E. (2011). Socio-psychological barriers to conflict resolution. In D. Bar-Tal (Ed.), *Intergroup conflicts and their resolution: A social psychological perspective*, 217–240.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage, 76,* 412–427.

Batson, C. (2009). These things called empathy: Eight related by distinct phenomena. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 3–15). Cambridge: MIT Press.
Batson, C., Chang, J., Orr, R., & Rowland, J. (2002). Empathy, attitudes, and action: Can feeling for a member of a stigmatized group motivate one to help the group? *Personality and Social Psychology Bulletin*, *28*, 1656.
Bernhardt, B. C., & Singer, T. (2012). The neural basis of empathy. *Annual Review of Neuroscience, 35,* 1–23.
Bruneau, E. G., & Saxe, R. (2010). Attitudes towards the outgroup are predicted by activity in the precuneus in Arabs and Israelis. *Neuroimage, 52,* 1704–1711.
Bruneau, E. G., & Saxe, R. (2012). The power of being heard: The benefits of 'perspective-giving' in the context of intergroup conflict. *Journal of experimental social psychology,* 48, 855–866.
Bryant, R., Felmingham, K., Kemp, A., Das, P., Hughes, G., Peduto, A., et al. (2008). Amygdala and ventral anterior cingulate activation predicts treatment response to cognitive behaviour therapy for post-traumatic stress disorder. *Psychological Medicine, 38,* 555–561.
Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, & L. Gardner (Eds), *The handbook of social psychology*, (pp. 151–192). New York: McGraw-Hill.
Cikara, M., Botvinick, M. M., & Fiske, S. T. (2011a). Us versus them: social identity shapes neural responses to intergroup competition and harm. *Psychologcial Science, 22,* 306–313.
Cikara, M., Bruneau, E. G., & Saxe, R. (2011b). Us and them. *Current Directions in Psychological Science, 20,* 149–153.
Cunningham, W. A., Johnson, M. K., Raye, C. L., Chris Gatenby, J., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science, 15,* 806–813.
Doehrmann, O., Ghosh, S. S., Polli, F. E., Reynolds, G. O., Horn, F., Keshavan, A., et al. (2013). Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging. *JAMA Psychiatry, 70,* 87–97.
Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: automatic and controlled processes. *Journal of Experimental Social Psychology, 33,* 510–540.
Dovidio, J., Kawakami, K., & Gaertner, S. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology, 82,* 62–68.
Falk, E. B., Berkman, E. T., Whalen, D., & Lieberman, M. D. (2011). Neural activity during health messaging predicts reductions in smoking above and beyond self-report. *Health Psychology, 30,* 177.
FeldmanHall, O., Dalgleish, T., Evans, D., & Mobbs, D. (2015). Empathic concern drives costly altruism. *NeuroImage, 105,* 347–356.
Foroni, F., & Mayr, U. (2005). The power of a story: new, automatic associations from a single reading of a short scenario. *Psychological Bulletin Review, 12,* 139–144.
Fu, C. H., Steiner, H., & Costafreda, S. G. (2013). Predictive neural biomarkers of clinical response in depression: A meta-analysis of functional and structural neuroimaging studies of pharmacological and psychological therapies. *Neurobiology of Disease, 52,* 75–83.
Gabrieli, J., Ghosh, S. S., & Whitfield-Gabrieli, S. (2015). Prediction as a Humanitarian and Pragmatic Contribution from Human Cognitive Neuroscience. *Neuron, 85,* 11–26.
Golby, A. J., Gabrieli, J. D., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *National Neuroscience, 4,* 845–850.
Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27.
Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1480.
Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197–216.
Halperin, E. (2013). Emotion, emotion regulation, and conflict resolution. *Emotion Review, 6,* 68–76.

Halperin, E., Bar-Tal, D., Nets-Zehngut, R., & Drori, E. (2008). Emotions in conflict: Correlates of fear and hope in the Israeli-Jewish society. *Peace and Conflict: Journal of Peace Psychology, 14,* 233–258.
Halperin, E., Pliskin, R., Saguy, T., Liberman, V., & Gross, J. J. (2013). Emotion Regulation and the Cultivation of Political Tolerance Searching for a New Track for Intervention. *Journal of Conflict Resolution, 58,* 1110–1138
Hameiri, B., Porat, R., Bar-Tal, D., Bieler, A., & Halperin, E. (2014). Paradoxical thinking as a new avenue of intervention to promote peace. *Proceedings of the National Academy of Sciences, 111,* 10996–11001.
Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *Neuroreport, 11,* 2351–2355.
Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. L., et al. (2007). Functional and morphometric brain dissociation between dyslexia and reading ability. *Proceedings of the National Academy of Sciences, 104,* 4234–4239.
Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., et al. (2011). Neural systems predicting long-term outcome in dyslexia. *Proceedings of the National Academy of Sciences, 108,* 361–366.
Johns, M., Schmader, T., & Martens, A. (2005). Knowing Is Half the Battle Teaching Stereotype Threat as a Means of Improving Women's Math Performance. *Psychological Science, 16,* 175–179.
Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302–4311.
Kawakami, K., Phills, C. E., Steele, J. R., & Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder:Improving implicit racial attitudes and interracial interactions through approach behaviors. *Journal of Personality and Social Psychology, 92,* 957–971.
Kim, D. Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly, 66,* 83–96.
Kirk, U., Skov, M., Hulme, O., Christensen, M. S., & Zeki, S. (2009). Modulation of aesthetic value by semantic context: An fMRI study. *Neuroimage, 44,* 1125–1132.
Knutson, B., Rick, S., Wimmer, G., Prelec, D., & Loewenstein, G. (2007a). Neural predictors of purchases. *Neuron, 53,* 147–156.
Knutson, K. M., Mah, L., Manly, C. F., & Grafman, J. (2007b). Neural correlates of automatic beliefs about gender and race. *Human Brain Mapping, 28,* 915–930.
Kumari, V., Peters, E. R., Fannon, D., Antonova, E., Premkumar, P., Anikumar, A. P., Williams, S.C.R., Kuipers, E. (2009). Dorsolateral prefrontal cortex activity predicts responsiveness to cognitive-behavioral therapy in schizophrenia. *Biological Psychiatry, 66,* 594–602.
Kumari, V., Antonova, E., Fannon, D., & Peters, E. R. (2010). Beyond dopamine: functional MRI predictors of responsiveness to cognitive behaviour therapy for psychosis. *Frontiers in behavioral neuroscience, 4*.
Lamm, C., Batson, C., & Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience, 19*(1), 42–58.
LeDoux, J. (2007). The amygdala. *Current Biology, 17,* R868–R874.
Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current opinion in neurobiology, 22,* 1027–1038.
Lieberman, M. D., Hariri, A., Jarcho, J. M., Eisenberger, N. I., & Bookheimer, S. Y. (2005). An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *National Neuroscience, 8,* 720–722.
Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive And Affective Neuroscience, 7,* 274–281.
Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231.
Noar, S. M. (2006). A 10-year retrospective of research in health mass media campaigns: Where do we go from here? *Journal of Health Communication, 11,* 21–42.

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences, 125,* E1–E24.

Olsson, A., Ebert, J. P., Banaji, M. R., & Phelps, E. A. (2005). The role of social groups in the persistence of learned fear. *Science, 309*(5735), 785–787.

Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: a field experiment in Rwanda. *Journal of Personality and Social Psychology, 96,* 574–587.

Paluck, E. L. (2010). Is it better not to talk? Group polarization, extended contact, and perspective taking in eastern Democratic Republic of Congo. *Personality and Social Psychology Bulletin, 36,* 1170.

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: what works? A review and assessment of research and practice. *Annual Review of Psychology, 60,* 339–367.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., et al. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12,* 729–738.

Pizzagalli, D. A. (2010). Frontocingulate dysfunction in depression: toward biomarkers of treatment response. *Neuropsychopharmacology, 36,* 183–206.

Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of Neuroscience, 27,* 9984–9988.

Porat, R., Halperin, E., & Bar-Tal, D. (2015). The effect of sociopsychological barriers on the processing of new information about peace opportunities. *Journal of Conflict Resolution,* 59, 93–119.

Pronin, E., Lin, D., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin, 28,* 369.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: divergent perceptions of bias in self versus others. *Psychological Review, 111,* 781–799.

Ross, L. (1993). *Reactive devaluation in negotiation and conflict resolution*: Stanford Center on Conflict and Negotiation, Stanford University.

Ross, L., & Ward, A. (1994). Psychological barriers to dispute resolution. *Advances in Experimental Social Psychology, 27,* 255–255.

Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In T. Brown, E. S. Reed, & E. Turiel (Eds.), *Values and knowledge*. (pp. 103–135). Hillsdale: Erlbaum.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people The role of the temporoparietal junction in 'theory of mind'. *Neuroimage, 19,* 1835–1842.

Saxe, R., Brett, M., & Kanwisher, N. (2006a). Divide and conquer: A defense of functional localizers. *Neuroimage, 30,* 1088–1096; discussion 1097–1089.

Saxe, R., Moran, J., Scholz, J., & Gabrieli, J. (2006b). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. *Social Cognitive And Affective Neuroscience, 1,* 229.

Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E. N., & Saxe, R. (2009). Distinct regions of right temporo-parietal junction are selective for theory of mind and exogenous attention. *PLoS ONE, 4,* e4869.

Tam, T., Hewstone, M., Cairns, E., Tausch, N., Maio, G., & Kenworthy, J. (2007). The impact of intergroup emotions on forgiveness in Northern Ireland. *Group Processes & Intergroup Relations, 10,* 119.

Webster, D. W., Whitehill, J. M., Vernick, J. S., & Curriero, F. C. (2013). Effects of Baltimore's Safe Streets Program on gun violence: A replication of Chicago's CeaseFire Program. *Journal of Urban Health, 90,* 27–40.

Zaki, J., Schirmer, J., & Mitchell, J. P. (2011). Social influence modulates the neural computation of value. *Psychological Science, 22,* 894–900.

Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research Reviews, 4,* 88–123.