

# Motive attribution asymmetry for love vs. hate drives intractable conflict

Adam Waytz<sup>a,1</sup>, Liane L. Young<sup>b</sup>, and Jeremy Ginges<sup>c,d</sup>

<sup>a</sup>Department of Management and Organizations, Kellogg School of Management at Northwestern University, Evanston, IL 60208; <sup>b</sup>Department of Psychology, Boston College, Chestnut Hill, MA 02467; <sup>c</sup>School of Psychological Sciences, University of Melbourne, Parkville, VIC 3010, Australia; and <sup>d</sup>Department of Psychology, The New School of Social Research, New York, NY 10011

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved September 22, 2014 (received for review July 25, 2014)

**Five studies across cultures involving 661 American Democrats and Republicans, 995 Israelis, and 1,266 Palestinians provide previously unidentified evidence of a fundamental bias, what we term the “motive attribution asymmetry,” driving seemingly intractable human conflict. These studies show that in political and ethnoreligious intergroup conflict, adversaries tend to attribute their own group’s aggression to ingroup love more than outgroup hate and to attribute their outgroup’s aggression to outgroup hate more than ingroup love. Study 1 demonstrates that American Democrats and Republicans attribute their own party’s involvement in conflict to ingroup love more than outgroup hate but attribute the opposing party’s involvement to outgroup hate more than ingroup love. Studies 2 and 3 demonstrate this biased attributional pattern for Israelis and Palestinians evaluating their own group and the opposing group’s involvement in the current regional conflict. Study 4 demonstrates in an Israeli population that this bias increases beliefs and intentions associated with conflict intractability toward Palestinians. Finally, study 5 demonstrates, in the context of American political conflict, that offering Democrats and Republicans financial incentives for accuracy in evaluating the opposing party can mitigate this bias and its consequences. Although people find it difficult to explain their adversaries’ actions in terms of love and affiliation, we suggest that recognizing this attributional bias and how to reduce it can contribute to reducing human conflict on a global scale.**

intergroup conflict | ingroup love | outgroup hate | attribution | cognitive bias

**M**any human conflicts appear extraordinarily difficult to resolve even when outsiders can see the contours of a rational resolution. Ideological opponents risk the health of their economies and their planet because they are unable to make political compromises. Ethnic and religious groups across the world engage in mass acts of violence, rejecting solutions of mutual benefit that involve sharing power, land, or religious sites. Why are so many conflicts so intractable when people on both sides could gain from a compromise?

Approaches to this question typically center on the complexity and profundity of the issues involved and each side’s construal of the conflict (1, 2). An additional, deeper barrier to conflict resolution may be simple pessimism toward compromise. If adversaries believe inflexibility on the other side renders mutual compromise impossible, they will be unlikely to adopt seemingly rational strategies for conciliation (3). In other words, the perception of conflict intractability may be an independent cause of a stalemate. Here, we identify a fundamental cognitive bias that contributes to the belief in conflict intractability, and may therefore contribute to conflict spirals (4–6).

We predict and observe that people often attribute their outgroup’s actions during conflict to hatred. This attribution occurs despite evidence that intergroup aggression stems at least as much, if not more, from ingroup love (7–13). We hypothesized that whereas people can clearly identify ingroup love as the source of their own group’s engagement in conflict (14), when attempting to explain their outgroup’s actions, people focus on dislike (15) and “why they hate us” (16). Expanding on prior research, we propose a previously untested hypothesis: People will attribute ingroup

engagement in conflict to love more than hate, but they will attribute outgroup engagement in conflict to hate more than love. We term this pattern the “motive attribution asymmetry” and document this bias across five studies among American Democrats and Republicans as well as Palestinians and Israelis, demonstrating consequences of this bias that further fuel intergroup conflict. We use the term “bias” to mean response tendency (rather than error); in this case, a tendency to attribute love vs. hate to one’s ingroup to a greater degree than to one’s outgroup and to attribute hate vs. love to one’s outgroup to a greater degree than to one’s ingroup.

These studies provide a number of contributions:

- i) This research offers an account of why people fail to recognize or implement the solutions that research on conflict resolution has established to be effective.
- ii) This research examines specific attributions of love vs. hate rather than dispositional vs. situational attributions more broadly.
- iii) This research compares these attributions toward one’s outgroup with the outgroup’s attributions toward itself.
- iv) The studies themselves contain numerous previously unexamined features, investigating this phenomenon in diverse linguistic and political contexts, examining different types of real intergroup conflicts (one intensely ideological yet predominantly nonviolent and the other ethnoreligious and chronically violent), characterizing previously unexamined consequences of this attributional bias, and establishing a manipulation aimed at mitigating this bias and its effects.
- v) This research reveals the disconnect between the motives pertaining to love and hate that actually appear to drive intergroup conflict (7–13) and people’s own intuitive assessment of these motives.

## Significance

**Political conflict between American Democrats and Republicans and ethnoreligious conflict between Israelis and Palestinians seem intractable, despite the availability of reasonable compromise solutions in both cases. This research demonstrates a fundamental cognitive bias driving such conflict intractability: Adversaries attribute their ingroup’s actions to ingroup love more than outgroup hate and attribute their outgroup’s actions to outgroup hate more than ingroup love. This biased attributional pattern increases beliefs and intentions associated with conflict intractability, including unwillingness to negotiate and unwillingness to vote for compromise solutions. In addition, offering financial incentives for accuracy in evaluating one’s outgroup mitigates this biased attributional pattern and its consequences. Understanding this bias and how to alleviate it can contribute to conflict resolution on a global scale.**

Author contributions: A.W., L.L.Y., and J.G. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed. Email: a-waytz@kellogg.northwestern.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1414146111/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1414146111/-DCSupplemental).

## Results and Discussion

Study 1 randomly assigned American Democrats and Republicans ( $n = 285$ ) to rate the motives for engaging in conflict, in terms of love and hate, of either their own party or the other party. A 2 (target: own party vs. other party)  $\times$  2 (motive: love vs. hate) mixed ANOVA revealed a target  $\times$  motive interaction [ $F(1, 283) = 83.57, P < 0.0001, \eta^2_p = 0.23$ ], with no main effects for condition ( $P = 0.97$ ) or motive ( $P = 0.36$ ). Follow-up  $t$  tests revealed that all simple effects in the condition  $\times$  motive interaction were significant ( $t > 5.63, P < 0.0001, d \geq 0.73$ ). People indicated that their own party was more motivated by love for their own party [mean ( $M$ ) = 4.47, standard deviation ( $SD$ ) = 1.31] than by hate for the other party ( $M = 3.63, SD = 1.17$ ) and that the opposing party was more motivated by hate ( $M = 4.56, SD = 1.10$ ) than love ( $M = 3.53, SD = 1.27$ ). In brief, these findings represent a fundamental discrepancy between how people view the motives of their own political party and the motives of the opposing political party (Fig. S1).

Studies 2 and 3 investigated whether this bias would replicate across cultures and in the context of a protracted violent conflict: the Israeli-Palestinian dispute. These studies used a within-subjects rather than between-subjects design such that all participants answered questions both for their own group and for the other side. We asked representative samples of Israelis (study 2,  $n = 497$ ) and Palestinians living in the West Bank and Gaza (study 3,  $n = 1,266$ ) why many Israelis and Palestinians supported violence in the conflict across the border of Gaza and Israel. Scale items and forced-choice items asked participants the extent to which they saw aggression committed by Israelis and aggression committed by Palestinians as driven by love and by hate.

In study 2 with the Israeli population, a 2 (target: Israel vs. Palestine)  $\times$  2 (motive: love vs. hate) ANOVA on the scale items revealed no main effect for target ( $P = 0.28$ ) but a main effect for motive such that participants reported hate ( $M = 2.79, SD = 0.71$ ) as a more likely motive than love ( $M = 2.65, SD = 0.80$ ) for both populations [ $F(1, 399) = 7.64, P = 0.006, \eta^2_p = 0.02$ ]. This main effect was qualified by a target  $\times$  motive interaction [ $F(1, 399) = 392.56, P < 0.0001, \eta^2_p = 0.50$ ].

Follow-up  $t$  tests on the initial ANOVA revealed that all simple effects in the target  $\times$  motive interaction were significant ( $t \geq 11.69, P < 0.0001, d \geq 1.78$ ). Israelis reported they were motivated by love for Israelis ( $M = 3.10, SD = 1.04$ ) more than by hate for Palestinians ( $M = 2.30, SD = 1.00$ ) and believed that Palestinians were driven by hate for Israelis ( $M = 3.27, SD = 0.88$ ) more than by love for Palestinians ( $M = 2.23, SD = 1.09$ ). In addition, Israelis believed love was a stronger motive for Israelis than for Palestinians and hate was a stronger motive for Palestinians than for Israelis (Fig. 1). The forced-choice questions revealed the same pattern (SI Materials and Methods).

In study 3 with the Palestinian population, a 2 (target: Israel vs. Palestine)  $\times$  2 (motive: love vs. hate) ANOVA on the four scale items revealed a main effect for target, with higher ratings for Palestinians than for Israelis ( $M = 3.32, SD = 0.52$  vs.  $M = 3.11, SD = 0.52$ ) [ $F(1, 1,221) = 163.15, P < 0.0001, \eta^2_p = 0.12$ ], and a main effect for motive, with higher ratings for hate than for love ( $M = 3.38, SD = 0.50$  vs.  $M = 3.06, SD = 0.56$ ) [ $F(1, 1,221) = 329.29, P < 0.0001, \eta^2_p = 0.21$ ]. More importantly, these main effects were qualified by a target  $\times$  motive interaction [ $F(1, 1,221) = 374.31, P < 0.0001, \eta^2_p = 0.24$ ].

Follow-up  $t$  tests on the initial ANOVA revealed that all simple effects in the target  $\times$  motive interaction were significant ( $t > 2.19, P \leq 0.028, d \geq 0.17$ ). Palestinians indicated they were motivated by love ( $M = 3.34, SD = 0.69$ ) more than hate ( $M = 3.29, SD = 0.66$ ) and believed that Israelis were driven by hate ( $M = 3.45, SD = 0.57$ ) more than love ( $M = 2.77, SD = 0.82$ ). In addition, Palestinians believed love was a stronger motive for Palestinians than for Israelis and hate was a stronger motive for Israelis than for Palestinians (Fig. 2). The forced-choice questions revealed the same pattern (SI Materials and Methods).

From scores on the scale items and forced-choice items, the major qualitative difference between the Palestinian and Israeli populations appears to be that Palestinians state that they are

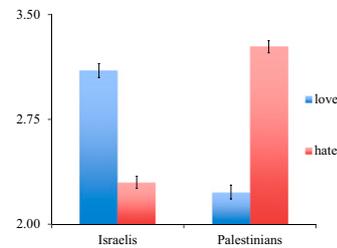


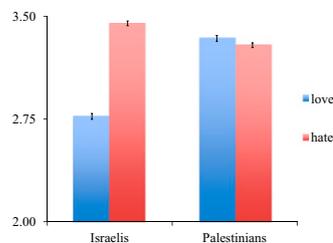
Fig. 1. Israelis' attributions to Israelis and Palestinians in study 2. Means and SEs are taken from the total sample rather than from specific tests that exclude subjects who responded "don't know/not applicable." Error bars represent SEs.

driven by hate almost as much as love (although ratings of these motives are significantly different). Israelis, on the other hand, much more strongly distinguish between these two motives, stating that they are driven far more by love than hate. A 2 (population: Israeli vs. Palestinian)  $\times$  2 (target: Israeli vs. Palestinian)  $\times$  2 (motive: love vs. hate) ANOVA combining data from studies 2 and 3 confirms this difference, revealing a three-way population  $\times$  target  $\times$  motive interaction [ $F(1, 1,620) = 920.90, P < 0.0001, \eta^2_p = 0.36$ ]. (This effect and all others for which the results of Box's test for equality of covariance matrices were significant across studies are unchanged compared with assuming sphericity.) It might be the case that in asymmetrical conflicts, for groups such as these Palestinians, who experience oppression and have relatively less power than Israelis, expressions of anger and hatred (in attempts to change the status quo) are more normative (17, 18). Nonetheless, these findings yet again represent a fundamental discrepancy between perceptions of one's ingroup's motives for conflict and perceptions of one's outgroup's motives for conflict.

We theorized that this motive attribution asymmetry may be associated with specific beliefs that undermine conflict resolution, including a belief in the unalterable intransigence of the outgroup (3). Thus, the primary aim of study 4 was to examine the implications of the tendency demonstrated in studies 1–3 for people to attribute the outgroup's motives for conflict to outgroup hate more than ingroup love. Hatred is a highly persistent and passionate dislike associated with a belief that the hated object is evil (19). Because hatred is associated with the belief that the other is bad in nature (i.e., is essentialized as evil), it leads to an inclination not to change the other but to separate from or destroy the other (19, 20). Thus, we predicted that attributions of outgroup aggression to hate for the outgroup rather than love for the aggressors' ingroup would be associated with beliefs and intentions associated with conflict intractability (e.g., unwillingness to negotiate, pessimism that a peace deal can be attained). To test this hypothesis, we recruited another representative sample of Israelis ( $n = 498$ ) and asked them to rate the extent to which they believed that Palestinian violence was motivated by love for Palestine and by hatred of Israel. We also asked them about beliefs and intentions associated with conflict intractability across three different versions of instructions that did not affect results (SI Materials and Methods).

Israelis again attributed Palestinian support for conflict to hate ( $M = 3.29, SD = 0.95$ ) more than love ( $M = 2.52, SD = 1.14$ ) [ $t(430) = 12.70, P < 0.0001, d = 1.65$ ]. The forced-choice item revealed the same pattern (SI Materials and Methods). Having determined that the general outgroup attribution bias replicates, we next turned to other variables of interest to examine the implications of this bias.

From the scale items, we computed a measure of attributional bias by subtracting scores for attribution of love from scores for attribution of hate to produce a difference score. We then standardized this difference score, standardized people's responses on the forced-choice love-vs.-hate item (love = 1, hate = 2), and averaged these two standardized items [ $r(392) = 0.36, P < 0.0001$ ] to produce a bias score. This score includes people without a difference score or forced-choice item response because of responding,



**Fig. 2.** Palestinians' attributions to Israelis and Palestinians in study 3. Means and SEs are taken from the total sample rather than from specific tests that exclude subjects who responded "don't know/not applicable." Error bars represent SEs.

"don't know/not applicable." Analyses excluding these people reveal comparable relationships between bias and the various implications measures. Higher bias scores mean a greater perception that Palestinians are motivated by hate vs. love (separate analyses of love attributions, hate attributions, and forced-choice attributions are provided in *SI Materials and Methods*). This measure of bias was correlated with reduced willingness to negotiate [ $r(453) = -0.23, P < 0.0001$ ], reduced perceptions of a win-win [ $r(409) = -0.21, P < 0.0001$ ], reduced optimism [ $r(463) = -0.10, P = 0.038$ ], reduced personal willingness to vote for a peace deal [ $r(471) = -0.15, P = 0.001$ ], reduced expectation that Palestinians will vote for a peace deal [ $r(471) = -0.11, P = 0.016$ ], reduced positive compromise outcome beliefs [ $r(471) = -0.26, P < 0.0001$ ], and increased essentialist beliefs about Palestinians [ $r(471) = 0.27, P < 0.0001$ ]. Unsurprisingly, bias was not correlated with beliefs about Israelis (expected Israeli voting for a peace deal and essentialist beliefs about Israelis:  $P > 0.10$ ). From these findings, it is clear that this biased perception of one's outgroup as driven by hate vs. love, a critical component of the motive attribution asymmetry, has significant consequences for beliefs and behavioral intentions.

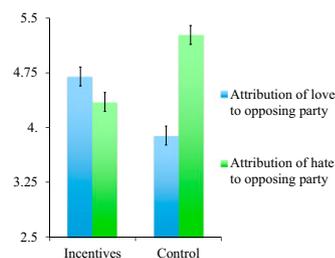
Study 5 introduced an experimental manipulation designed to reduce the motive attribution asymmetry and consequent perceived conflict intractability. American Democrats and Republicans ( $n = 331$ ) evaluated how much their own party is motivated by ingroup love (compassion and empathy toward their own party) and outgroup hate (dislike and animosity toward the opposing party). Then, participants were randomly assigned to an incentives condition that offered financial incentives for accurate evaluations of the outgroup or to a control condition that offered no incentives. All participants then evaluated how much the opposing party is motivated by ingroup love and outgroup hate on the same measures as for their own party. Then, all participants answered questions about their willingness to negotiate with the opposing party, optimism that a successful resolution could be reached, expectations for a win-win compromise, and essentialist beliefs toward the other party.

A 2 (condition: control vs. incentives)  $\times$  2 (target: own party vs. other party)  $\times$  2 (motive: love vs. hate) mixed ANOVA revealed a significant effect of target [ $F(1, 329) = 9.50, P = 0.002, \eta^2_P = 0.03$ ]; a significant motive  $\times$  condition interaction [ $F(1, 329) = 12.33, P = 0.001, \eta^2_P = 0.04$ ]; a significant target  $\times$  motive interaction [ $F(1, 329) = 54.34, P < 0.0001, \eta^2_P = 0.14$ ]; and, most importantly, a significant condition  $\times$  target  $\times$  motive interaction [ $F(1, 329) = 42.05, P = 0.001, \eta^2_P = 0.11$ ] (all other effects,  $P > 0.39$ ). Replicating previous studies, overall participants attributed more love ( $M = 4.66, SD = 1.49$ ) than hate ( $M = 4.09, SD = 1.54$ ) to their own party [ $t(330) = 4.62, P < 0.0001, d = 0.84$ ] and more hate ( $M = 4.81, SD = 1.70$ ) than love ( $M = 4.29, SD = 1.71$ ) to the opposing party [ $t(330) = 3.59, P < 0.0005, d = 0.68$ ] (attributions of love to ingroup vs. love to outgroup and hate to ingroup vs. hate to outgroup differ significantly as well:  $P \leq 0.0001$ ). More importantly, examining attributions to the outgroup by experimental condition reveals that whereas participants in the control condition indeed attribute more hate ( $M = 5.27, SD = 1.62$ ) than love ( $M = 3.89, SD = 1.65$ ) to the opposing party [ $t(166) = 7.56, P < 0.0001, d = 1.89$ ], in the incentives

condition, attribution of hate diminishes significantly compared with the control condition ( $M = 4.35, SD = 1.67$ ) [ $t(329) = 4.40, P < 0.0001, d = 0.49$ ] and attribution of love increases significantly compared with the control condition ( $M = 4.70, SD = 1.69$ ) [ $t(329) = 5.05, P < 0.0001, d = 0.56$ ] such that the attributional pattern is reversed (Fig. 3).

To assess the impact of the motive attribution asymmetry on consequence measures, we computed a composite bias score by subtracting attributions of love to the outgroup from attributions of hate to the outgroup (separate analyses of love and hate attributions are provided in *SI Materials and Methods*). Significant correlations emerged such that bias reduced optimism [ $r(329) = -0.30, P < 0.0001$ ], bias reduced perceptions of a win-win [ $r(329) = -0.16, P = 0.004$ ], and bias increased essentialism [ $r(328) = 0.14, P = 0.014$ ]. These findings largely replicate those findings of study 4, although willingness to negotiate was only slightly negatively related to bias ( $r = -0.05, P = 0.38$ ). In addition, we computed a similar composite bias score for the ingroup by subtracting attributions of love to the ingroup from attributions of hate to the ingroup, and found this bias similarly correlated negatively with optimism [ $r(329) = -0.16, P = 0.003$ ] and perceptions of a win-win [ $r(329) = -0.10, P = 0.065$ ] and did not significantly affect willingness to negotiate ( $r = -0.04, P = 0.53$ ) or essentialist beliefs ( $r = -0.01, P = 0.85$ ). Although we did not have explicit predictions about the effect of ingroup attributions on consequence measures, these findings suggest that ingroup attributions have implications for conflict resolution as well.

It was also important to determine whether our manipulation affected these consequence measures so as to demonstrate that incentives for accuracy did not simply increase people's willingness to infer the other side's motivations momentarily (with no meaningful consequences for conflict intractability beliefs). Condition significantly influenced optimism such that incentives ( $M = 3.52, SD = 1.50$ ) increased optimism compared with the control condition ( $M = 3.14, SD = 1.64$ ) [ $t(329) = 2.17, P = 0.031, d = 0.24$ ] and incentives ( $M = 2.86, SD = 0.84$ ) reduced essentialist beliefs compared with the control condition ( $M = 3.08, SD = 0.93$ ) [ $t(328) = 2.20, P = 0.028, d = 0.24$ ]. Furthermore, bootstrapping mediation analysis (21) with condition coded as incentives = 1 and control = 0 (bias-corrected, 20,000 resamples) suggested that decreased biased attribution toward the outgroup (attributions of hate to the outgroup minus attributions of love to the outgroup) mediates the relationship between condition and optimism [95% confidence interval (CI) = 0.17–0.45; the path coefficients are illustrated in Fig. S2] and that attribution of hate to the outgroup and attribution of love to the outgroup separately appear to mediate this effect as well (95% CI = 0.13–0.40 and 95% CI = 0.02–0.21, respectively). An equivalent mediation analysis suggested that biased attribution toward the outgroup mediates the effect of condition on essentialist beliefs (95% CI = -0.13 to -0.0002; the path coefficients are illustrated in Fig. S2) and that attribution of hate to the outgroup separately appears to mediate this effect as well (95% CI = -0.20 to -0.06). In other words, incentives for accuracy increase optimism toward a mutual political agreement through simultaneously increasing attributions of love and decreasing attributions of hate toward the opposing political party. Incentives for accuracy also



**Fig. 3.** Comparison by condition of Democrats' and Republicans' attributions of love and hate to the opposing party in study 5. Error bars represent SEs.

increase perceptions that the opposing political party is malleable through reducing this biased attribution pattern as well, primarily through reducing attributions of hate toward the opposing political party. Condition did not significantly affect willingness to negotiate ( $P = 0.30$ ) or perceptions of a win-win ( $P = 0.17$ ), although participants in the incentives condition did indicate increased willingness to negotiate and increased expectations for a win-win solution.

Although the effect of monetary incentives is robust in this context, in that incentives not only mitigate but actually reverse the motive attribution asymmetry, we acknowledge that the strength of this particular intervention might vary for conflicts of a more violent and volatile nature, and we welcome future research to test its effects.

Five studies document the first evidence, to our knowledge, of a motive attributional asymmetry whereby adversaries attribute their own support for aggression more to love than to hate but attribute the other group's aggression more to hate than to love. By comparing attributions toward one's outgroup with the outgroup's attributions toward itself, we demonstrate for the first time, to our knowledge, that outgroup attributions of love vs. hate are biased, if not erroneous. Furthermore, this research illustrates important consequences of this attributional bias (e.g., willingness to negotiate, voting intentions, beliefs about compromise) and provides a manipulation of incentivizing accuracy to mitigate this attributional asymmetry, suggesting two mechanisms driving this asymmetry.

One potential mechanism concerns the cognitive availability of love vs. hate. Affiliative motives like love may be less observable in outgroup members than in ingroup members, especially during intergroup conflict. Expressions of love are more likely to occur in affiliative and familial settings than in the presence of ostensible competitors, whereas expressions of anger or aggression are more likely to emerge and capture attention in the presence of competitors (22, 23). Therefore, expressions of animosity (vs. expressions of love) may be more observable to people outside the group. This pattern of relative visibility between love and hate resembles the lack of awareness that drives other biases, such as the bias blind spot (24) and naive realism (1). Related to this misperception is the tendency for people to interpret others' actions as approach-motivated even when recognizing their own actions to be avoidance-motivated (because approach motives seem more salient) (25). When people evaluate their enemies, they apparently link enemy actions to approach motives (toward the ingroup) rather than avoidance motives to protect their own.

A second potential mechanism underlying this asymmetry concerns motivation: People may engage in motivated reasoning to see their ingroups as loving and their outgroups as hateful. This attributional pattern bears conceptual similarity to the ultimate attribution error, whereby people attribute positive (negative) ingroup behaviors to internal (external) factors and to attribute outgroup positive (negative) behaviors to external (internal) factors (26). Importantly, the present work differs from this prior research because it examines evaluations of the same behavior (involvement in conflict) in terms of solely internal factors (i.e. motivations for love vs. hate) rather than examining evaluations of different behaviors (i.e. positive behaviors and negative behaviors) in terms of internal vs. external factors, yet a similar motivational component of this bias might be present in this bias and in the motive attribution asymmetry. The present bias also bears similarity to people's motivational tendency to deny prosocial values (e.g., love) to outgroup members to dehumanize them and justify harm toward them (27).

We believe our results, including those results of study 5, are consistent with both of these mechanisms. Monetary incentives can reduce egocentrism through increasing effortful perspective-taking (28), as well as motivating people to detect others' emotions more accurately (29). Monetary incentives can thus increase effortful attention to the other side's true underlying mental states and can increase motivation to assess others accurately (30). Future research can examine the separate contributions of these two mechanisms.

Overall, these results suggest that the motive attribution asymmetry is a significant barrier to resolution of intergroup conflict. Independent of factors such as divergent ideologies, clashing religious beliefs, or external economic crises, the biased attribution

of an outgroup's actions to hate makes disputants perceive conflict to be more intractable. Just as the overestimation of group differences leads to conflict escalation (31), failing to recognize love as a shared motive between one's ingroup and outgroup likely exacerbates conflict as well. Furthermore, we speculate that this specific intergroup misperception is one that is particularly difficult to detect. Because constructs like love and loyalty hold such positive connotations, it is difficult to notice others' affiliative motives as sources of conflict, similar to the difficulty of noticing others' ingroup favoritism (rather than outgroup derogation) as a primary contributor to intergroup discrimination (32). Identifying this motive attribution asymmetry as a consequential source of conflict and identifying how to mitigate this attributional bias can contribute to reducing long-standing conflict.

## Materials and Methods

**Study 1: Democrat-Republican Conflict in American Politics.** Two hundred eighty-five United States residents (205 male, 68 female, and 12 unreported;  $M_{age} = 27.93$ ,  $SD = 9.17$ ) completed the study online through mTurk (Amazon Mechanical Turk) using Qualtrics software, in exchange for a small monetary reward. For this and all subsequent studies, participants were included in analyses if they completed all measures assessing attributions of love and hate (even if they did not complete all demographic measures). Institutional review board approval for all studies was obtained from Northwestern University, Boston College, and the New School of Social Research, and informed consent was obtained from participants across all studies.

Participants were randomly assigned to one of two conditions, the own-party or other-party condition. In both conditions, participants read, "Think about the political party that you belong to. Now, think about members of the opposing political party."

In the own-party condition, participants read, "When your party engages in conflict with the opposing party, how much is YOUR PARTY motivated by each of the following:" and then evaluated six items on a seven-point scale (1 = not at all, 7 = very much). Three items pertained to love ("empathy/compassion/kindness for people in your political party"), and three items pertained to hate ("dislike/indifference/hatred toward people in the opposing political party").

In the other-party condition, participants read, "When the opposing party engages in conflict with your party, how much is THE OPPOSING PARTY motivated by each of the following:" and evaluated six equivalent items. Three pertained to love (empathy/compassion/kindness toward the people in their political party), and three pertained to hate (dislike/indifference/hatred toward the people in your political party). We averaged the empathy, compassion, and kindness items across conditions to generate a love composite score ( $\alpha = 0.84$ ), and we averaged the dislike, indifference, and hatred items to generate a hate composite ( $\alpha = 0.59$ ). This slightly lower reliability for hate results from the distinction between the indifference item as a passive form of opposition and the dislike and hatred items as more active forms of opposition. Nonetheless, all significant results reported remain when the hate composite consists only of the dislike and hatred items or when the items are considered separately.

Participants then completed demographic measures, including two items of interest: one assessing political party ("What political party do you feel more closely aligned with?" with response options: Democratic and Republican; 19% of our sample indicated Republican affiliation) and one assessing political ideology (1 = extremely liberal, 7 = extremely conservative). Results did not differ meaningfully by political party (*SI Materials and Methods*).

**Study 2: Israelis' View of the Israeli-Palestinian Conflict.** Four hundred ninety-seven Israeli residents were recruited using computer-assisted telephone interviews, and reported personal and demographic information. This sample consisted of 229 males and 268 females, with the following age distribution (age was assessed categorically): 18–21 y (5.0%), 22–24 y (4.2%), 25–29 y (5.0%), 30–34 y (6.8%), 35–39 y (9.5%), 40–44 y (11.5%), 45–49 y (8.9%), 50–54 y (9.7%), 55–59 y (10.9%), 60–64 y (9.7%), 65+ y (17.5%), and 1.4% who did not respond. Our sample was primarily Jewish, with the following religion distribution: Secular Jewish (46.5%), Traditional Jewish (25.2%), National Orthodox (11.1%), Ultra Orthodox (9.7%), Muslim (2.6%), Druze (1.4%), Christian (1.2%), 0.8% who reported "other," and 1.6% who refused to answer.

All participants were given materials translated into Hebrew, stating: "As you know, there is often conflict between Israel and Palestinians living in Gaza. Palestinians have fired rockets into Israel, and Israelis have launched bombing attacks against Gaza. Many Palestinians and Israelis supported these respective attacks." Then, participants answered three questions about the motives of Israelis and three questions about the motives of Palestinians (order of Israel/Palestine counterbalanced). The first two questions pertaining to Israel were

scale items that read, "Why do you think many Israelis support bombing of Gaza during this conflict?" (response options: certainly yes, yes, no, certainly no, don't know/not applicable): (i) "Is it because of their love of Israel?" or (ii) "Is it because they hate Palestinians?" The third question pertaining to Israelis asked was presented in a forced-choice format, "What do you think is more important in causing Israelis to support bombing of Gaza: their love of Israel or their hatred of Palestinians?" (response options: their love of Israel, their hatred of Palestinians, don't know/not applicable).

The first two questions pertaining to Palestinians were scale items that read, "Why do you think many Palestinians support firing rockets from Gaza into Israel during this conflict?" (response options: certainly yes, yes, no, certainly no, don't know/not applicable): (i) "Is it because of their love of Palestine?" or (ii) "Is it because they hate Israelis?" The third question pertaining to Palestinians was presented in a forced-choice format, "What do you think is more important in causing Palestinians to support firing rockets from Gaza into Israel: their love of Palestinians or their hatred of Israel?" (response options: their love of Palestine, their hatred of Israelis, don't know/not applicable).

Don't know/not applicable responses represented 13.3% or less of responses for each question and were excluded from analysis. For the remaining analyses, we coded responses on the scale items as certainly yes = 4, yes = 3, no = 2, and certainly no = 1.

**Study 3: Palestinians' View of the Israeli-Palestinian Conflict.** One thousand two hundred sixty-six Palestinians (628 male, 638 female, and 5 unreported;  $M_{\text{age}} = 37.85$ ,  $SD = 13.43$ ; 98.7% Muslim) living in the West Bank and Gaza were recruited to participate in face-to-face interviews. This sample was 98.7% Muslim (1.3% Christian) (data collection procedures are described in *SI Materials and Methods*).

Participants were asked the same questions, in the same format as with the Israeli sample, as well as additional questions on personal and demographic information. Responses to the target measures of attribution were coded accordingly, with don't know/not applicable responses representing 4.3% or less of responses for each question.

**Study 4: Implications of Bias.** Four hundred ninety-eight Israeli residents were recruited for a telephone interview as in study 2 and reported demographic information. This sample consisted of 241 males and 257 females, with the following age distribution (age was again assessed categorically): 18–21 y (4.0%), 22–24 y (3.8%), 25–29 y (2.8%), 30–34 y (6.8%), 35–39 y (9.2%), 40–44 y (10.8%), 45–49 y (7.8%), 50–54 y (9.6%), 55–59 y (11.4%), 60–64 y (10.6%), 65+ y (21.7%), and 1.2% who refused to say. Our sample was primarily Jewish, with the following religion distribution: Secular Jewish (46.2%), Traditional Jewish (27.3%), National Orthodox (11.6%), Ultra Orthodox (8.6%), Muslim (3.2%), Christian (1.4%), 0.8% who reported other, and 0.8% who refused to answer.

Participants were randomly assigned to one of three conditions to assess the robustness of bias to different experimental inductions. These inductions and their effects are described in *SI Materials and Methods*; because they did not affect the phenomenon of interest (the attribution of love vs. hate to the outgroup), we collapsed over conditions for analyses. Similar to studies 2 and 3, all participants were told, "We would now like to ask you about how a typical Palestinian thinks" and answered three questions about attributions of Palestinian love vs. hate worded similar to study 2. The first two questions pertaining to Palestine were scale items that read, "Why do you think the typical Palestinian supports firing rockets from Gaza into Israel during this conflict?" (response options: certainly yes = 4, yes = 3, no = 2, certainly no = 1, don't know/not applicable): (i) "Is it because of love of Palestine?" or (ii) "Is it because of hate for Israelis?" The third question pertaining to Palestine was presented in a forced-choice format, "What is a more likely motivation behind the average Palestinian's support of firing rockets from Gaza into Israel: their love of Palestinians or their hatred of Israel?" (response options: their love of Palestine = 1, their hatred of Israelis = 2, don't know/not applicable). The don't know/not applicable responses represented 12.4% or less of responses for each question, did not differ significantly across conditions, and were excluded from analysis.

After these initial questions, participants answered a series of additional questions (listed in order below) about the relationship between bias in attribution of hate vs. love to Palestinians to other consequential beliefs and intentions.

A first question assessed willingness to negotiate and read, "Israelis and Palestinians are hoping to begin negotiations on issues such as the Palestinian right of return, the status of Jerusalem, final borders, as well as the makeup of a future Palestinian state. Would you be willing/unwilling to negotiate these issues?" (response options: not willing at all = 1, somewhat not willing = 2, somewhat willing = 3, very willing = 4, don't know/not applicable = excluded from analysis) (adapted from ref. 17).

A second question assessed perceptions of a win-win (33) and read, "As you know, delegations of both Israel and Palestine have been working to negotiate a peace agreement between the two sides. What is the likely outcome of such an agreement?" (response options: the agreement will be good for Israel only = 0, the agreement will be good for Palestine only = 0, the agreement will be bad for both Israel and Palestine = 0, the agreement will be good for both Israel and Palestine = 1, don't know/not applicable = excluded from analysis). This question assessed the belief that a compromise would constitute a mutually beneficial outcome for Palestine and Israel vs. any other outcome.

A third question assessed optimism, asking, "Are you optimistic or pessimistic that a successful peace agreement between Israel and Palestine can be implemented?" (response options: optimistic = 2, pessimistic = 1, don't know/not applicable = excluded from analysis).

Next, participants responded to three questions assessing response to negotiation. They read:

Suppose Israeli and Palestinian negotiators reached a peace deal. Under this deal, there would be two states living side by side: a Palestinian state in the West Bank and Gaza alongside the Jewish State of Israel. Both the government of Israel and the government of Palestine would apologize for the wrongs they have done, and recognize the legitimacy of the other state. Before this deal is signed, both Israel and the Palestinians decide to hold a referendum to let their people decide. If majorities of both populations vote for the deal, it will be signed and implemented. If either Israelis or Palestinians reject the deal, it will not be signed.

Then, they were asked, "Do you think you would vote for this deal or against the deal if the referendum was held tomorrow?" (response options: for = 2, against = 1), "What percentage of Israelis do you think would vote for the deal?" (response options: between 0% and 20% = 1, between 21% and 40% = 2, between 41% and 60% = 3, between 61% and 80% = 4, between 81% and 100% = 5), and "What percentage of Palestinians do you think would vote for the deal?" (response options: between 0% and 20% = 1, between 21% and 40% = 2, between 41% and 60% = 3, between 61% and 80% = 4, between 81% and 100% = 5).

Next, participants answered four questions regarding compromise outcome beliefs. They read, "This next question explores how you think Palestinians may respond to things done by Israel. Imagine that Israel unilaterally declared and implemented a freeze on settlement building and dismantled illegal settlement outposts. How do you think Palestinians would respond?" Then, they were asked two questions regarding potentially positive outcome beliefs: "Do you think that Palestinians would see that Israel was serious about taking steps towards peace and would end their support for the violent factions such as Hamas?" and "Do you think that Palestinians would see that Israel was serious about taking steps towards peace and would push their political leaders to make concessions—perhaps even recognition of Israel?" (yes = 2, no = 1). They then answered two questions regarding negative outcome beliefs: "Do you think these actions by Israel would be in vain because the settlements are not core to the conflict?" and "Do you think that Palestinians would see these [actions] as a triumph and use the victory to gather momentum for more attacks against Israel?" (yes = 2, no = 1).

Finally, participants answered questions about beliefs regarding essentialism of each group "As much as I hate to admit it, you can't teach an old dog new tricks—Palestinians/Israelis can't really change their basic characteristics"; "Palestinians/Israelis can at times seem to behave differently, but the important parts of who they are can't really be changed"; "Palestinians/Israelis who are characterized by extreme and violent behavior will never change because these characteristics are deeply rooted within them"; and "Palestinians/Israelis have basic moral values and beliefs that can't be changed significantly." All statements were then followed by, Do you strongly agree (4), somewhat agree (3), somewhat disagree (2), or strongly disagree (1)? These questions were adapted from previous work (3) and reflect beliefs regarding the fixed vs. malleable nature of both groups.

Willingness to negotiate, perceptions of a win-win, optimism, personal likelihood of voting for a peace deal, expected Israeli voting for a peace deal, and expected Palestinian voting for a peace deal were treated as single items. A composite score was generated for positive compromise outcome beliefs ( $\alpha = 0.69$ ) by averaging the two positive outcome beliefs and two negative outcome beliefs (reverse-scored). Composite scores were also generated for essentialist beliefs about Palestinians ( $\alpha = 0.85$ ) and Israelis ( $\alpha = 0.64$ ) by averaging the four items pertaining to each group.

Of chief importance for this study was to examine whether the biased perception of Palestinians as driven by hate vs. love correlates with any of these measures. We therefore computed a composite bias score as described above. Because condition did not significantly affect our individual measures

that constitute this bias score, we collapsed over the sample to examine correlations between bias and the other measures.

**Study 5: Incentivizing Accuracy.** Because studies 1–4 have demonstrated that the motive attribution asymmetry is robust across populations and has significant consequences, study 5 attempts to reduce this bias and its associated consequences. In particular, this study tests whether introducing monetary incentives for accuracy reduces the tendency of American Democrats and Republicans to attribute more hate than love to the opposing party (as was found in study 1), and whether reducing this bias can improve consequential perceptions of the other party. Given that paying people has been shown to both increase and diminish empathic accuracy (29, 34), this study also provides a test of the efficacy of monetary incentives for improving interpersonal perceptions.

Three hundred thirty-one US residents (223 male, 106 female, and 2 unreported;  $M_{\text{age}} = 29.43$ ,  $SD = 9.80$ ) completed the study online as in study 1. Participants were included in analyses if they completed all attributional measures as in study 1 (even if they did not complete all consequence measures described below, leading to differing degrees of freedom in some analyses).

All participants completed a series of demographic measures, including a question asking, “What political party do you feel most closely aligned with?” with the response options Democratic Party, Republican Party, and moderate/neither/somewhere in between. We excluded participants who selected this last category (they were redirected to complete a different study and did not complete the measures described below) so that we could examine participants for whom political conflict was particularly salient; similar to study 1, only 20% of our sample indicated Republican affiliation. Participants also answered a similar question assessing political ideology as in study 1 (1 = very liberal, 7 = very conservative). Overall, results did not differ meaningfully by political party (*SI Materials and Methods*).

Next, all participants were asked, “When your party engages in conflict with the opposing party, how much is YOUR PARTY motivated by each of the following?” and then evaluated four items on a seven-point scale (1 = not at all, 7 = very much). Two items pertained to love (“empathy for people in your political party” and “compassion for people in your political party”), and two items pertained to hate (“dislike toward the people in the opposing political party” and “animosity toward the people in the opposing political party”).

Then, participants were randomly assigned to one of two conditions, the control condition or the incentives condition. In both conditions, participants were told their responses would be compared with those responses of the opposing

party, but in the incentives condition we also told participants that they could earn an additional bonus of \$12 for accurately gauging the motivations of the opposing political party (*SI Materials and Methods*). All participants then evaluated four items on the same seven-point scale as for their own party. Two pertained to love (“empathy/compassion for people in their political party”), and two pertained to hate (“dislike/animosity toward the people in your political party”).

We averaged the empathy and compassion items to generate a love composite score for ingroup and outgroup [ $r(329) = 0.82$ ,  $P < 0.0001$ ;  $r(329) = 0.80$ ,  $P < 0.0001$ ], and we averaged the dislike and animosity items to generate a hate composite for ingroup and outgroup [ $r(329) = 0.80$ ,  $P < 0.0001$ ;  $r(329) = 0.78$ ,  $P < 0.0001$ ].

We also measured potential consequences of the motive attribution asymmetry, similar to study 4. As a measure of willingness to negotiate, participants responded to the following question, “This year, members from different political parties will begin negotiations on a number of issues, such as the implementation of the Affordable Care Act; how to handle the federal budget, including making cuts; and a number of social issues, such as immigration, abortion, and gay marriage. How unwilling/willing would you be to negotiate with a member of the opposing party on these issues?” (1 = completely unwilling, 7 = completely willing). To measure perceptions of a win-win, participants were asked, “Democrats and Republicans will be forced to reach resolutions on the issues mentioned above. What is the likely outcome of such resolutions?” (response options: the resolutions will be good for your party only = 0, the resolutions will be good for the other party only = 0, the resolutions will be bad for your party and the opposing party = 0, the resolutions will be good for both your party and the opposing party = 1). As a measure of optimism, participants answered, “How pessimistic/optimistic are you that successful resolutions on the issues mentioned above can be implemented?” (1 = very pessimistic, 7 = very optimistic). Finally, we asked the same four questions about essentialism as in study 4 worded toward the opposing party (and excluding the word “violent”) and averaged these items (including participants who did not answer all items) ( $\alpha = 0.79$ ). At the study’s conclusion, we paid the \$12 bonus to one participant in the incentives condition.

**ACKNOWLEDGMENTS.** This study was supported by Northwestern University, Boston College, and the Dispute Resolution Research Center at Kellogg School of Management. Studies with Israelis and Palestinians were funded by the National Science Foundation (Grant SES-0962080) and the Office of Naval Research (Grant MINERVA-ONR N000141310054).

- Robinson RJ, Keltner D, Ward A, Ross L (1995) Actual versus assumed differences in construal: “Naive realism” in intergroup perception and conflict. *J Pers Soc Psychol* 68(3):404–417.
- Ginges J, Atran S, Medin D, Shikaki K (2007) Sacred bounds on rational resolution of violent political conflict. *Proc Natl Acad Sci USA* 104(18):7357–7360.
- Halperin E, Russell AG, Trzesniewski KH, Gross JJ, Dweck CS (2011) Promoting the Middle East peace process by changing beliefs about group malleability. *Science* 333(6050):1767–1769.
- Jervis R (1976) *Perception and Misperception in International Politics* (Princeton Univ Press, Princeton).
- Acemoglu D, Wolitzky A (2014) Cycles of distrust: An economic model. *Am Econ Rev* 104(4):1350–1367.
- Haushofer J, Biletzki A, Kanwisher N (2010) Both sides retaliate in the Israeli-Palestinian conflict. *Proc Natl Acad Sci USA* 107(42):17927–17932.
- De Dreu CK, et al. (2010) The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science* 328(5984):1408–1411.
- Choi J-K, Bowles S (2007) The coevolution of parochial altruism and war. *Science* 318(5850):636–640.
- Atran S (2010) *Talking to the Enemy: Faith, Brotherhood, and the (Un)Making of Terrorists* (HarperCollins, New York).
- Buttelmann D, Böhm R (2014) The ontogeny of the motivation that underlies ingroup bias. *Psychol Sci* 25(4):921–927.
- Epley N (2014) *Mindwise: How We Understand What Others Think, Believe, Feel, and Want* (Random House, New York).
- Brewer MB (1999) The psychology of prejudice: In-group love and out-group hate? *J Soc Issues* 55(3):429–444.
- Halevy N, Bornstein G, Sagiv L (2008) “In-group love” and “out-group hate” as motives for individual participation in intergroup conflict: A new game paradigm. *Psychol Sci* 19(4):405–411.
- Leidner B, Castano E (2012) Morality shifting in the context of intergroup violence. *Eur J Soc Psychol* 42(1):82–91.
- Frey FE, Tropp LR (2006) Being seen as individuals versus as group members: Extending research on metaperception to intergroup contexts. *Pers Soc Psychol Rev* 10(3):265–280.
- Zakaria F (2001) The politics of rage: Why do they hate us? *Newsweek* 138(16):22–25.
- Kteily N, Saguy T, Sidanius J, Taylor DM (2013) Negotiating power: Agenda ordering and the willingness to negotiate in asymmetric intergroup conflicts. *J Pers Soc Psychol* 105(6):978–995.
- Saguy T, Kteily N (2011) Inside the opponent’s head: Perceived losses in group position predict accuracy in metaperceptions between groups. *Psychol Sci* 22(7):951–958.
- Halperin E (2008) Group-based hatred in intractable conflict in Israel. *J Conflict Resolut* 52:713–736.
- Royzman EB, McCauley C, Rozin P (1999) From Plato to Putnam: Four ways to think about hate. *The Psychology of Hate*, ed Sternberg RJ (American Psychological Association, Washington, DC), pp 3–36.
- Preacher KJ, Hayes AF (2008) Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods* 40(3):879–891.
- Ackerman JM, et al. (2006) They all look the same to me (unless they’re angry): From out-group homogeneity to out-group heterogeneity. *Psychol Sci* 17(10):836–840.
- Dunham Y (2011) An angry = outgroup effect. *J Exp Soc Psychol* 47(3):668–671.
- Pronin E, Gilovich T, Ross L (2004) Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychol Rev* 111(3):781–799.
- Miller DT, Nelson LD (2002) Seeing approach motivation in the avoidance behavior of others: Implications for an understanding of pluralistic ignorance. *J Pers Soc Psychol* 83(5):1066–1075.
- Pettigrew TF (1979) The ultimate attribution error: Extending Allport’s cognitive analysis of prejudice. *Pers Soc Psychol Bull* 5(4):461–476.
- Struch N, Schwartz SH (1989) Intergroup aggression: Its predictors and distinctness from in-group bias. *J Pers Soc Psychol* 56(3):364–373.
- Epley N, Keysar B, Van Boven L, Gilovich T (2004) Perspective taking as egocentric anchoring and adjustment. *J Pers Soc Psychol* 87(3):327–339.
- Klein KJ, Hodges SD (2001) Gender differences, motivation, and empathic accuracy: When it pays to understand. *Pers Soc Psychol Bull* 27(6):720–730.
- Prior M, Sood G, Khanna K (2013) The impact of accuracy incentives on partisan bias in reports of economic perceptions. Working paper. Available at [www.princeton.edu/~mprior/PSKPartisanBias.current.pdf](http://www.princeton.edu/~mprior/PSKPartisanBias.current.pdf). Accessed July 24, 2014.
- Keltner D, Robinson RJ (1997) Defending the status quo: Power and bias in social conflict. *Pers Soc Psychol Bull* 23(10):1066–1077.
- Greenwald AG, Pettigrew TF (2014) With malice toward none and charity for some: Ingroup favoritism enables discrimination. *Am Psychol* 69(7):669–684.
- Malhotra D, Ginges J (2010) Preferring balanced vs. advantageous peace agreements: A study of Israeli attitudes towards a two state solution. *Judgment and Decision Making* 5(6):420–427.
- Ma-Kellams C, Blascovich J (2013) The ironic effect of financial incentive on empathic accuracy. *J Exp Soc Psychol* 49(1):65–71.