nature human behaviour

ARTICLES
https://doi.org/10.1038/s41562-019-0766-4

Inaccurate group meta-perceptions drive negative out-group attributions in competitive contexts

Jeffrey Lees[1]* and Mina Cikara[2]*

Across seven experiments and one survey ($n = 4{,}282$), people consistently overestimated out-group negativity towards the collective behaviour of their in-group. This negativity bias in group meta-perception was present across multiple competitive (but not cooperative) intergroup contexts and appears to be yoked to group psychology more generally; we observed negativity bias for estimation of out-group, anonymized-group and even fellow in-group members' perceptions. Importantly, in the context of US politics, greater inaccuracy was associated with increased belief that the out-group is motivated by purposeful obstructionism. However, an intervention that informed participants of the inaccuracy of their beliefs reduced negative out-group attributions, and was more effective for those whose group meta-perceptions were more inaccurate. In sum, we highlight a pernicious bias in social judgements of how we believe 'they' see 'our' behaviour, demonstrate how such inaccurate beliefs can exacerbate intergroup conflict and provide an avenue for reducing the negative effects of inaccuracy.

How we believe others perceive us—meta-perception—plays a critical role in how we interact with others[1–3]. In the context of intergroup interactions, these meta-perceptions may bring unpleasant, even harmful, evaluations to mind[4–8]. For example, when individuals believe they are being negatively stereotyped by an out-group member, they experience increased negative emotions and lower self-esteem[4], suffer increased anxiety[9], and subsequently exhibit greater intergroup bias[10].

Despite the important role played by beliefs about how 'they' see 'us' (and our actions)[11–14], past work has focused primarily on person-to-person interactions across group boundaries or on estimates of extremity of, and polarization in, out-group attitudes[15–18]. As an example of the latter, findings in the domain of values and attitudes indicate that group members overestimate the level of disagreement and polarization between groups (though note that these constitute first-order judgements, or 'how I see *X*')[16–18]. Evidence from the intergroup literature, more broadly, suggests that group labels exacerbate inaccuracy in social judgements because they activate stereotypes that cause people to adjust their judgements away from their initial, more accurate, anchors[12].

There is, in complement, a growing literature in the domain of second-order, intergroup meta-judgements (or 'what I think they think about us') which reveals that people tend to have overly negative and inaccurate judgements of out-group motives toward the in-group[11,14]. This foundational work on the effects of meta-perceptions in intergroup contexts raises two important questions: (1) are these meta-perceptions accurate? And (2) what happens when these judgements are made in response to collective action—when people consider how 'they' see 'our' (not my) behaviour?

Here, we tackle a particular form of intergroup inaccuracy by examining group meta-perceptions (GMPs)—how we believe our group's collective actions will be perceived by the out-group. In our view, GMPs represent an intergroup context-activated distortion of second-order judgements. This makes GMPs (1) distinct from first-order judgements and (2) unique in that they should be sensitive to functional relations between groups (that is, whether groups are cooperative, competitive and so on) but relatively invariant to the focal event/act/behaviour or the groups in question.

It is likely that GMPs serve an important role in determining the course of group-on-group interaction, because they allow us to make predictions about whether an out-group will be supportive or hostile towards our own group's efforts at cooperation; therefore, GMPs should also drive emotions, strategy and policy preferences. For example, US President George W. Bush, in his address to a joint session of Congress on 20 September 2001, laid out in stark terms how he believed Al-Qaeda perceived the United States, and how these second-order judgements ought to compel US foreign policy[19]: "Americans are asking 'Why do they hate us?' They hate what they see right here in this chamber: a democratically elected government … They hate our freedoms: our freedom of religion, our freedom of speech, our freedom to vote and assemble and disagree with each other … We will direct every resource at our command—every means of diplomacy, every tool of intelligence, every instrument of law enforcement, every financial influence, and every necessary weapon of war—to the destruction and to the defeat of the global terror network … Every nation in every region now has a decision to make: Either you are with us or you are with the terrorists." President Bush used the belief that "they hate our freedoms" to motivate his call to war and his ultimatum to other countries that they are either 'with us' or 'with the terrorists'. However, many have noted that this belief that Al-Qaeda "hate our freedoms" wrongly diagnosed the motivations of Al-Qaeda and the complex socio-political forces that drove their perception of the United States[20,21]. Furthermore, this essentializing language served to dehumanize Muslims and drive support for the 'War on Terror' among the American public[22].

This example highlights how inaccurate, and overly negative, beliefs about how the out-group perceives the behaviour (and values) of one's own group can drive intractable intergroup conflict. When group leaders and other group members believe that the out-group will react with animosity and perceive one's group in a highly negative fashion, they are likely to support antagonistic intergroup actions over cooperative and reconciliatory behaviours.

[1]Harvard Business School, Harvard University, Boston, MA, USA. [2]Psychology Department, Harvard University, Cambridge, MA, USA.
*e-mail: jlees@g.harvard.edu; mcikara@fas.harvard.edu

For example, when people believe they are dehumanized by an out-group, they are more likely to dehumanize the out-group in return, which leads to increased support for war and out-group torture[7]. This dynamic can unfold in contexts as hostile as war between nations, but also in legislative compromise across political parties, competitive sports and interaction among organizations. Nonetheless, interventions that directly inform individuals of their inaccurate beliefs may be able to induce positive behavioural change[23,24].

To investigate the nature of GMPs, we constructed a set of scenarios involving group-level conflict. For Experiments 1, 3, 4 and 6 and Study 5, these scenarios pertained to the behaviour of US political parties in a legislative context. In Experiment 2 the scenarios pertained to group-level conflict between men and women in educational and workplace settings. All scenarios presented instances where one group was attempting to pass a law or change a policy in a manner that would potentially disadvantage the other group (for example, requiring a sitting governor of the opposing party to disclose their taxes), except for Experiment 3 in which the behaviour would potentially benefit the other group. Supplementary Experiment A is a direct replication of Experiment 4 with a convenience sample, and Supplementary Experiment B is an exploratory follow-up to Experiment 6.

Experiments 1–4 were designed to test for participant accuracy in GMPs. At the beginning of these experiments, participants were asked to identify their political affiliation (or gender identity in Experiment 2) and were then randomly assigned to whether the group taking action in the scenario was their in-group or out-group. Those who read about their in-group taking action were asked for their GMPs (for example, "How much do you believe an (out-group member) will dislike this action?"), whereas those who read about their out-group taking action against their in-group were asked for their actual perceptions (for example, "How much do you dislike this action?"). In Experiment 4 we also asked about 'in-group perceptions' (for example, "How much do you believe an (in-group) member will dislike the (out-group) action?"). Across all experiments, the comparision of the GMP and actual perception conditions across groups (that is, Democrats versus Republicans and men versus women) allowed for a direct test of participant accuracy.

When reading the scenarios, participants were asked, either as a meta-perception, actual perception or in-group perception, their perceived dislike of, opposition to and political/social unacceptability of the action being taken in the scenario, which they reported on sliding scales with labels at the end of the scales (for example, 1 = not opposed, 100 = extremely opposed). After the ratings all participants, across all experiments and Study 5, completed a comprehension check which asked them to identify the group "taking action" in the scenario. Any participants who failed this check were excluded from all analyses. Lastly, all participants were asked their age, gender and whether they had comments for the experimenters (except in Experiment 4, in which demographic questions were asked at the beginning of the experiment).

All materials, data and analysis code for all experiments and studies, and pre-registrations for Experiments 4 and 6, are available on the Open Science Framework (OSF: https://osf.io/zhysa/).

Results

Experiments 1–4 and Study 5 were analysed using mixed-effects beta-regressions, and Experiment 6 was analysed using linear mixed-effects regression. All tests are two-sided. In Experiment 6, homoscedasticity and normality of errors were assumed but not formally tested. Further details regarding the analyses can be found in the Methods section.

In Experiment 1 ($n = 408$), participants were randomly assigned to either the GMP condition ($n = 129$), actual-perception condition ($n = 143$) or an unlabelled and anonymized control group meta-perception condition ($n = 136$) in which participants were asked how 'Party B' would perceive the behaviour of 'Party A'. Within each condition, participants were randomly assigned to read one of five scenarios (we included multiple scenarios in each experiment and study to assess the robustness of our effects and modelled scenario as a random effect).

Across all scenarios, participants in the GMP condition substantially overestimated the negative perceptions of out-group participants (that is, out-group members in the actual-perception condition) on our three measures: action dislike (unstandardized log-odds regression coefficient (b) = 1.51, 95% confidence interval (CI) = [1.19,1.83], odds ratio (OR) = 4.53, Z-score (z) = 9.27, $P < 0.001$), opposition to the action ($b = 1.40$, 95% CI = [1.09,1.72], OR = 4.08, $z = 8.78$, $P < 0.001$) and political unacceptability of the action ($b = 1.36$, 95% CI = [1.04,1.67], OR = 3.89, $z = 8.46$, $P < 0.001$). Similarly, participants in the control meta-perception condition overestimated the negative perceptions of those in the actual-perception condition: dislike ($b = 1.32$, 95% CI = [1.02,1.62], OR = 3.74, $z = 8.55$, $P < 0.001$), opposition ($b = 1.22$, 95% CI = [0.93,1.52], OR = 3.40, $z = 8.15$, $P < 0.001$) and political unacceptability ($b = 1.13$, 95% CI = [0.83,1.42], OR = 3.08, $z = 7.45$, $P < 0.001$). Pairwise post hoc tests indicated no statistically significant difference between responses in the control meta-perception condition versus the GMP condition: dislike ($b = -0.19$, 95% CI = [−0.54,0.15], OR = 0.83 $t(402) = -1.30$, $P = 0.40$), opposition ($b = -0.18$, 95% CI = [−0.52,0.16], OR = 0.83, $t(402) = -1.24$, $P = 0.43$) and political unacceptability ($b = -0.23$, 95% CI = [−0.58,0.11], OR = 0.79, $t(401) = -1.58$, $P = 0.26$). We also examined the main effect of accuracy by party, modelled as a categorical fixed effect with two groups: "Democrat Accuracy"—Democrats in the GMP and control conditions compared to Republicans in the actual-perception condition—and "Republican Accuracy"—Republicans in the GMP and control conditions compared to Democrats in the actual-perception condition (see Methods for model details). This approach allowed the main effect to appropriately contrast meta/control versus actual perceptions (the baseline in the analyses) across parties, rather than within each party. Indeed, there was no statistically significant main effect of party accuracy: dislike ($b = -0.04$, 95% CI = [−0.29,0.21], OR = 0.96, $z = -0.32$, $P = 0.75$), opposition ($b = -0.00$, 95% CI = [−0.25,0.24], OR = 1.00, $z = -0.03$, $P = 0.98$) and political unacceptability ($b = -0.02$, 95% CI = [−0.27,0.23], OR = 0.98, $z = -0.18$, $P = 0.85$). Finally, pairwise post hoc tests found no statistically significant differences when examining whether Democrats and Republicans differed in their actual perceptions of the scenarios: dislike ($b = 0.00$, 95% CI = [−0.61,0.61], OR = 1.00, $t(400) = 0.02$, $P = 1.00$), opposition ($b = 0.15$, 95% CI = [−0.45,0.74], OR = 1.16, $t(400) = 0.72$, $P = 0.98$) and political unacceptability ($b = 0.11$, 95% CI = [−0.49,0.71], OR = 1.11, $t(399) = 0.51$, $P = 1.00$). See Fig. 1 for a visualization of the raw data by condition.

As predicted, GMPs in Experiment 1 were more negative than participants' actual perceptions of the out-group's behaviour. This was true even when we removed party labels. Thus, merely invoking the political intergroup context was sufficient to engender inaccuracy, supporting our proposition that GMPs are an intergroup, context-activated distortion, invariant to the groups in question. Furthermore, we found no credible evidence that this effect was moderated by participants' party membership. This suggests that Democrats and Republicans were equally pessimistic, and therefore inaccurate, in judging how members of the other party perceived the collective behaviour of their own party.

To further examine the generalizability of our findings, Experiment 2 ($n = 286$) used a design similar to that of Experiment 1 but in the context of gender relations. There were two changes from the design of Experiment 1. First, participants were assigned to one of three scenarios regarding group-level gender conflict (for example, integrating a single-gender school choir), rather than five scenarios regarding political conflict. Second, we did not include an

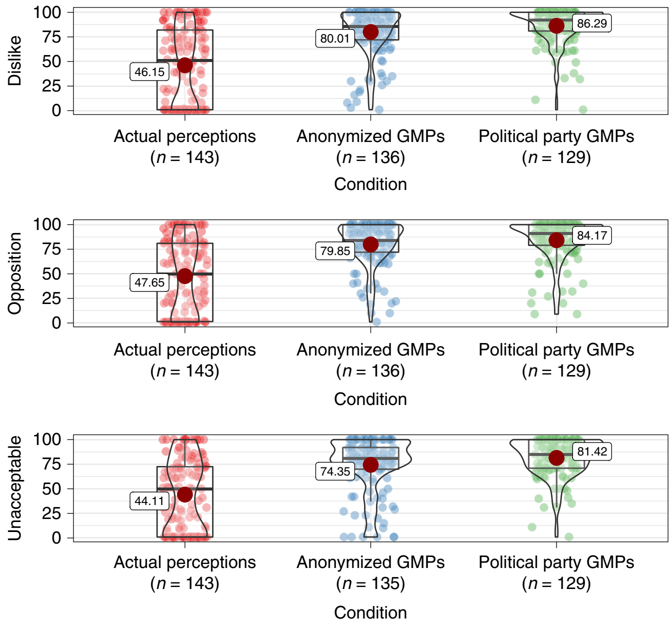

Fig. 1 | Raw data from Experiment 1 by condition and dependent variable. In this experiment, $n = 408$ (collected via Mechanical Turk). In the two GMP conditions, participants reported how much they thought their out-group, or an anonymized political party (control), would dislike, oppose and find unacceptable the in-group's/other party's action in the scenario. Solid red dots and corresponding numbers are sample means, the boxplot centre lines are sample medians. Participants in the political party GMP condition overestimated the negative perceptions of out-group participants in the actual-perception condition on action dislike ($b = 1.51$, 95% CI = [1.19,1.83], OR = 4.53, $z = 9.27$, $P < 0.001$), opposition to the action ($b = 1.40$, 95% CI = [1.09,1.72], OR = 4.08, $z = 8.78$, $P < 0.001$) and political unacceptability of the action ($b = 1.36$, 95% CI = [1.04,1.67], OR = 3.89, $z = 8.46$, $P < 0.001$). Participants in the control meta-perception condition overestimated the negative perceptions of those in the actual-perception condition on dislike ($b = 1.32$, 95% CI = [1.02,1.62], OR = 3.74, $z = 8.55$, $P < 0.001$), opposition ($b = 1.22$, 95% CI = [0.93,1.52], OR = 3.40, $z = 8.15$, $P < 0.001$) and political unacceptability ($b = 1.13$, 95% CI = [0.83,1.42], OR = 3.08, $z = 7.45$, $P < 0.001$). Pairwise post hoc tests indicate no statistically significant difference between responses in the control meta-perception versus GMP condition on dislike ($b = -0.19$, 95% CI = [−0.54,0.15], OR = 0.83 $t(402) = -1.30$, $P = 0.40$), opposition ($b = -0.18$, 95% CI = [−0.52,0.16], OR = 0.83, $t(402) = -1.24$, $P = 0.43$) and political unacceptability ($b = -0.23$, 95% CI = [−0.58,0.11], OR = 0.79, $t(401) = -1.58$, $P = 0.26$). These results provide evidence of overly pessimistic GMPs.

anonymized-group control condition. As with Experiment 1, participants were randomly assigned to the GMP condition ($n = 128$) or actual-perception condition ($n = 158$), read only one scenario and responded to items regarding perceived dislike of, opposition to and social unacceptability of the action in the scenario.

Results indicated a statistically significant condition (actual versus meta-perception) by gender–accuracy interaction (that is, a fixed effect similar to party accuracy in Experiment 1, contrasting accuracy across gender rather than within gender), indicating that one gender had less accurate GMPs than the other: dislike ($b = 0.78$, 95% CI = [0.22,1.34], OR = 2.18, $z = 2.73$, $P = 0.006$), opposition ($b = 0.74$, 95% CI = [0.18,1.30], OR = 2.09, $z = 2.59$, $P = 0.010$) and social unacceptability ($b = 0.65$, 95% CI = [0.09,1.21], OR = 1.92, $z = 2.27$, $P = 0.023$). Pairwise post hoc tests revealed that female participants had highly negative and inaccurate GMPs, replicating Experiment 1: dislike ($b = -1.13$, 95% CI = [−1.66,−0.59], OR = 0.32, $t(280) = -5.42$, $P < 0.001$), opposition ($b = -1.07$, 95% CI = [−1.60,−0.54], OR = 0.34, $t(280) = -5.22$, $P < 0.001$) and social unacceptability ($b = -1.02$, 95% CI = [−1.56,−0.49], OR = 0.36, $t(280) = -4.93$, $P < 0.001$). However, male participants' GMPs were not significantly different from the actual perceptions of female participants: dislike ($b = -0.35$, 95% CI = [−0.86,0.17], OR = 0.71, $t(280) = -1.74$, $P = 0.30$), opposition ($b = -0.33$, 95% CI = [−0.85,0.20], OR = 0.72, $t(280) = -1.69$, $P = 0.33$) and social unacceptability ($b = -0.37$, 95% CI = [−0.89,0.14], OR = 0.69, $t(280) = -1.87$, $P = 0.24$). This interaction was driven by gender differences in actual perceptions. Pairwise post hoc tests indicated that male and female participants' GMPs were not significantly different across dislike ($b = 0.29$, 95% CI = [−0.25,0.82], OR = 1.33, $t(280) = 1.39$, $P = 0.51$), opposition ($b = 0.19$, 95% CI = [−0.34,0.73], OR = 1.21, $t(280) = 0.91$, $P = 0.80$) and social unacceptability ($b = 0.52$, 95% CI = [−0.02,1.07], OR = 1.69, $t(280) = 2.49$, $P = 0.063$). However, women's (relative to men's) actual perceptions of the behaviours were significantly more negative across dislike ($b = 1.07$, 95% CI = [0.56,1.58], OR = 2.91, $t(280) = 5.39$, $P < 0.001$), opposition ($b = 0.93$, 95% CI = [0.42,1.43], OR = 2.53, $t(280) = 4.75$, $P < 0.001$) and social unacceptability ($b = 1.18$, 95% CI = [0.66,1.69], OR = 3.24, $t(280) = 5.91$, $P < 0.001$).

Thus, while we found no credible evidence that men's group meta-perceptions about the degree to which women would be upset were inaccurate, women's GMPs were inaccurate and overly negative, replicating the results from Experiment 1 in the domain of gender. It is important to reiterate, however, that the men's 'accuracy' result was driven by differences in male and female participants' actual perceptions. In other words, men's GMPs were closer to women's actual perceptions because women reported being more upset about the policy changes than did men. This pattern is likely the result of real-world power differences between the genders: men may be marginally less impacted and therefore less upset by disadvantageous policies in the contexts featured in our scenarios. More generally, Experiments 1 and 2 demonstrated GMP inaccuracy but only insofar as it pertained to the out-group in competitive or zero-sum contexts. To examine whether GMPs reflect a negativity bias or a valence-independent extremity bias, Experiment 3 contrasted GMPs versus actual perceptions in response to cooperative rather than competitive behaviours.

Experiment 3 ($n = 499$) used the same design as the GMP and actual-perception conditions from Experiment 1. While the scenarios pertained to the same political content, the nature of the behaviours was inverted such that the groups were taking cooperative actions, which either benefited the other group or disadvantaged the group taking the action. For example, instead of trying to make equal a partisan redistricting board controlled by the other party, in Experiment 3 the party taking action was trying to make equal a partisan redistricting board controlled by their own party. Participants in the GMP ($n = 233$) and actual-perception ($n = 266$) conditions were asked for their positive perceptions (for example, 1 = not supportive, 100 = extremely supportive), rather than negative perceptions. Otherwise the procedure was the same as in Experiment 1, including the between-subjects random assignment to both condition and scenario.

In contrast to Experiments 1 and 2, Experiment 3 found no credible evidence for GMP inaccuracy in cooperative contexts across the measures support ($b = -0.02$, 95% CI = [−0.25,0.21], OR = 0.98, $z = -0.20$, $P = 0.84$), liking ($b = 0.12$, 95% CI = [−0.11,0.35], OR = 1.13, $z = 1.02$, $P = 0.31$) or political acceptability ($b = -0.05$, 95% CI = [−0.28,0.18], OR = 0.95, $z = -0.42$, $P = 0.67$). There was a main effect (but never an interaction) of party accuracy for support ($b = 0.44$, 95% CI = [0.20,0.67], OR = 1.55, $z = 3.69$, $P < 0.001$), liking ($b = 0.48$, 95% CI = [0.25,0.71], OR = 1.61, $z = 4.05$, $P < 0.001$) and political acceptability ($b = 0.52$, 95% CI = [0.29,0.75], OR = 1.69, $z = 4.46$, $P < 0.001$), such that Democrats' positive reactions were slightly higher than those of Republicans. GMPs for both parties accurately tracked this mean-level difference. The findings from Experiment 3 parallel other work demonstrating that dyadic meta-perceptions are more accurate when two people are cooperative, but less so when competing[25]. Broadly, Experiment 3 also provides evidence that GMP inaccuracy represents specifically a negativity bias in competitive contexts, rather than an extremity bias in how we believe the out-group will react to the in-group's actions in general.

Experiments 1, 2 and 3 are limited in several notable ways. First, they all use convenience samples (that is, Mechanical Turk workers) and, as such, do not represent general population GMPs and actual perceptions. Second, the previous experiments do not tell us whether people are inaccurate specifically about how the out-group sees the in-group's behaviour or, more generally, how any group sees any other group's behaviour. Experiment 4, a pre-registered (see OSF: https://osf.io/atck5) extension of Experiment 1, used a nationally representative sample and included an in-group perception condition to address these limitations.

Experiment 4 ($n=536$) featured the same scenarios from Experiment 1. Participants were randomly assigned, between subjects, to the actual-perception condition ($n=170$), GMP condition ($n=206$)—both of which were the same as Experiment 1—or a new condition called the in-group perception condition ($n=160$). Participants in the in-group perception condition read the same scenarios as those in the actual-perception condition, but instead of being asked for their individual perceptions they were asked how they believed 'another (in-group member)' would perceive the scenarios. In contrast to Experiment 1, participants read and responded to all five scenarios (a repeated-measures factor, modelled as a random effect for participant).

Experiment 4 revealed statistically significant differences among all three conditions on all three outcome measures (see Fig. 2 for raw data distributions). Actual perceptions were lower than in-group perceptions for opposition ($b=-0.26$, 95% CI = [−0.43,−0.09], OR = 0.77, $z=-2.93$, $P=0.003$), unacceptability ($b=-0.25$, 95% CI = [−0.43,−0.07], OR = 0.78, $z=-2.72$, $P=0.007$) and dislike ($b=-0.34$, 95% CI = [−0.52,−0.17], OR = 0.71, $z=-3.93$, $P<0.001$). GMPs were higher than in-group perceptions for opposition ($b=0.51$, 95% CI = [0.35,0.68], OR = 1.67, $z=6.10$, $P<0.001$), unacceptability ($b=0.43$, 95% CI = [0.25,0.60], OR = 1.53, $z=4.87$, $P<0.001$) and dislike ($b=0.41$, 95% CI = [0.24,0.57], OR = 1.50, $z=4.83$, $P<0.001$). The pairwise post hoc contrasts between actual perceptions and GMPs were also significant for opposition ($b=-0.77$, 95% CI = [−0.97,−0.58], OR = 0.46, $t(2669)=-9.27$, $P<0.001$), unacceptability ($b=-0.67$, 95% CI = [−0.87,−0.47], OR = 0.51, $t(2669)=-7.83$, $P<0.001$) and dislike ($b=-0.75$, 95% CI = [−0.95,−0.56], OR = 0.47, $t(2669)=-9.04$, $P<0.001$), directly replicating the main finding of inaccurate GMPs from Experiment 1 but this time in a nationally representative sample. We also performed a direct replication of Experiment 4 using a convenience sample (again Mechanical Turk workers) and found practically identical results (see Supplementary Experiment A).

Critically, the differences between in-group perceptions and GMPs indicate that our inaccuracy findings for Experiments 1 and 2 cannot be explained entirely by the difference in referents across the actual-perception judgements ("how would you feel?") versus GMP ("how would an out-group member feel?") judgements. In Experiment 4 the in-group judgement also uses a group-level referent ("how would an in-group member feel about the out-group's action?"), but is still significantly less negative than the GMP judgements.

Study 5 ($n=212$) tested whether inaccurate GMPs are consequential, by examining the relationship between GMPs and negative motive attributions towards the out-group. In this study, participants completed the GMP condition from Experiment 1. They then reported how much they agreed with the statement "(Out-group members) are purposefully obstructing the process surrounding the (specific scenario topic)" (1–100 slider scale, 'Strongly Disagree' to 'Strongly Agree'). Analyses indicated a significant positive linear association between the belief that the out-group is obstructionist and negative GMPs of dislike ($b=2.12$, 95% CI = [1.40,2.84], OR = 8.34, $z=5.76$, $P<0.001$), opposition ($b=1.95$, 95% CI = [1.19,2.70], OR = 7.00, $z=5.06$, $P<0.001$) and political unacceptability ($b=1.66$, 95% CI = [0.96,2.35], OR = 5.24, $z=4.69$, $P<0.001$). There was no significant main effect of party identification on dislike ($b=-0.04$, 95% CI = [−0.37,0.30], OR = 0.96, $z=-0.22$, $P=0.83$), opposition ($b=-0.11$, 95% CI = [−0.44,0.23], OR = 0.90, $z=-061$, $P=0.55$) or political unacceptability ($b=-0.08$, 95% CI = [−0.42,0.26], OR = 0.92, $z=-0.47$, $P=0.64$). Thus, the more negative (and therefore inaccurate) participants' GMPs, the more likely they were to believe that the out-group is motivated by obstructionism. See Fig. 3 for visualization of raw data and Pearson correlations.

Experiment 6 ($n=1,122$) sought to reduce the perception that the out-group is motivated by obstructionism by using a

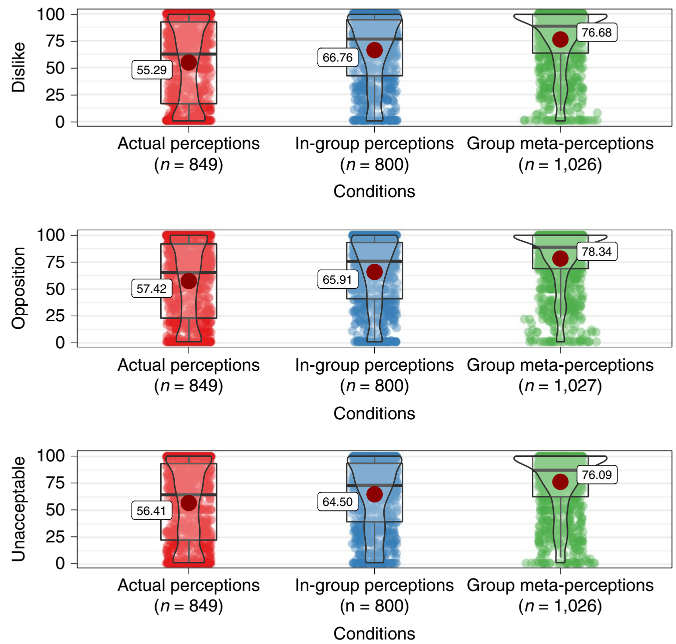

Fig. 2 | Raw data from Experiment 4 by condition and dependent variable. Sample sizes listed in figures are the number of judgements (across five repeated measures). Total $n=538$ (nationally representative sample collected via Qualtrics survey panels). By condition: actual perceptions, $n=170$; in-group perception, $n=160$; GMPs, $n=206$. Solid red dots and corresponding numbers are sample means, and the boxplot centre lines are sample medians. Actual perceptions were lower than in-group perceptions for opposition ($b=-0.26$, 95% CI = [−0.43,−0.09], OR = 0.77, $z=-2.93$, $P=0.003$), unacceptability ($b=-0.25$, 95% CI = [−0.43,−0.07], OR = 0.78, $z=-2.72$, $P=0.007$) and dislike ($b=-0.34$, 95% CI = [−0.52,−0.17], OR = 0.71, $z=-3.93$, $P<0.001$). GMPs were higher than in-group perceptions for opposition ($b=0.51$, 95% CI = [0.35,0.68], OR = 1.67, $z=6.10$, $P<0.001$), unacceptability ($b=0.43$, 95% CI = [0.25,0.60], OR = 1.53, $z=4.87$, $P<0.001$) and dislike ($b=0.41$, 95% CI = [0.24,0.57], OR = 1.50, $z=4.83$, $P<0.001$). The pairwise post hoc contrasts between actual perceptions and GMPs were also significant for opposition ($b=-0.77$, 95% CI = [−0.97,−0.58], OR = 0.46, $t(2,669)=-9.27$, $P<0.001$), unacceptability ($b=-0.67$, 95% CI = [−0.87,−0.47], OR = 0.51, $t(2,669)=-7.83$, $P<0.001$) and dislike ($b=-0.75$, 95% CI = [−0.95,−0.56], OR = 0.47, $t(2,669)=-9.04$, $P<0.001$). These results provide evidence of overly pessimistic GMPs and overly pessimistic judgements of the in-group's reactions.

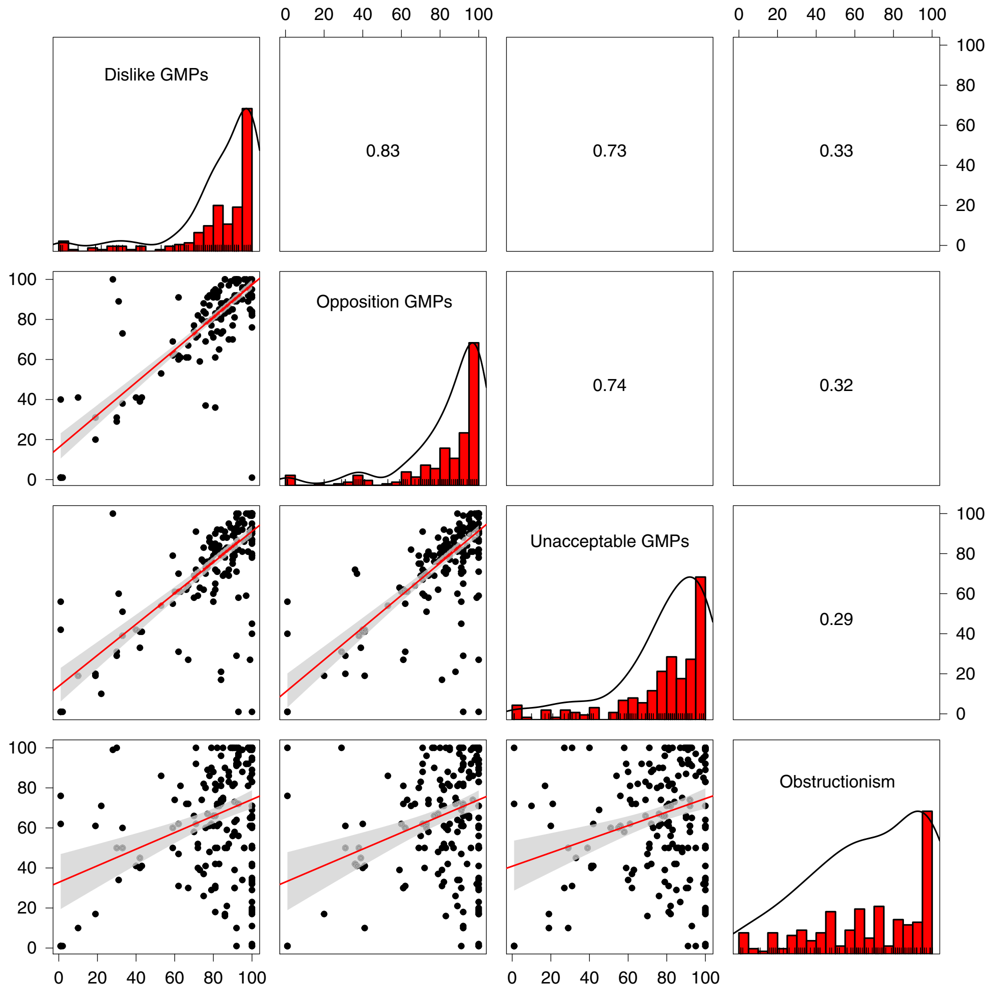

Fig. 3 | Distributions, Pearson correlations and scatterplots for the three GMP ratings and beliefs about out-group obstructionism in Study 5. Sample size, $n=212$ (collected via Mechanical Turk). Scatterplot lines are linear regression lines, and shaded area around lines are 95% CIs. Correlations: dislike–opposition ($r=0.83$, 95% CI = [0.79,0.87], $t(208)=21.73$, $P<0.001$), dislike–unacceptable ($r=0.73$, 95% CI = [0.66,0.79], $t(210)=15.50$, $P<0.001$), dislike–obstructionism ($r=0.33$, 95% CI = [0.20,0.45], $t(210)=5.08$, $P<0.001$), unacceptable–opposition ($r=0.74$, 95% CI = [0.68,0.80], $t(208)=16.02$, $P<0.001$), unacceptable–obstructionism ($r=0.29$, 95% CI = [0.16,0.40], $t(210)=4.32$, $P<0.001$) and obstructionism–opposition ($r=0.32$, 95% CI = [0.19,0.43], $t(208)=4.80$, $P<0.001$). These data indicate a positive linear association between pessimistic GMPs and the belief that the out-group is purposefully obstructionist.

pre-registered intervention (see OSF: https://osf.io/jhnsb). Building on the design of Study 5, after participants provided their three GMP ratings in response to one of the five political scenarios, they were randomly assigned, between-subjects, to one of three conditions before reporting their perceived out-group obstructionism: the control ($n=396$), "truth intervention" ($n=358$) or "hypocrisy prevention intervention" ($n=368$). In the control condition, participants were simply reminded of the GMP ratings they had provided on the previous page (that is, no new information). In the truth intervention, participants were provided with the information from the control condition plus the true value for their out-group's actual perceptions (the mean of the representative sample responses from Experiment 4) for that same scenario. This allowed participants to see the (in-)accuracy of their GMPs. Recall that in Experiment 4 we also found that participants inaccurately believed their in-group would react less negatively than their out-group to the same behaviour. Therefore, in the hypocrisy prevention intervention, participants received all the information in the truth intervention while also receiving the exact true values for their in-group's actual perceptions (also drawn from Experiment 4) for the same scenario. As such, the hypocrisy intervention additionally prevented participants from anchoring on an inaccurate belief that the in-group's negativity would still be lower than the out-group's in the same scenario. This allowed us to test whether there was an added benefit to highlighting participants' (in-)accuracy regarding the extent to which their in- and out-groups were similar in their actual perceptions.

As hypothesized, participants who were assigned to the truth intervention condition had lower ratings of out-group obstructionism than did the control group ($b=-4.08$, 95% CI = [−7.67,−0.48], $\beta=-0.155$, $t(1114)=-2.22$, $P=0.027$). Those assigned to the hypocrisy prevention intervention also had lower obstructionism ratings relative to control ($b=-4.64$, 95% CI = [−8.22,−1.08], $\beta=-0.177$, $t(1114)=-2.55$, $P=0.011$). However, post hoc pairwise comparisons indicated no statistically significant difference in obstructionism between the hypocrisy prevention and truth interventions ($b=-0.57$, 95% CI = [−4.96,3.82], $t(1115)=-0.304$, $P=0.95$), suggesting that the hypocrisy prevention intervention provided no additional benefit over the truth intervention. There was also a main effect of party identification on obstructionism, with Democrats rating Republicans as higher on obstructionism that Republicans

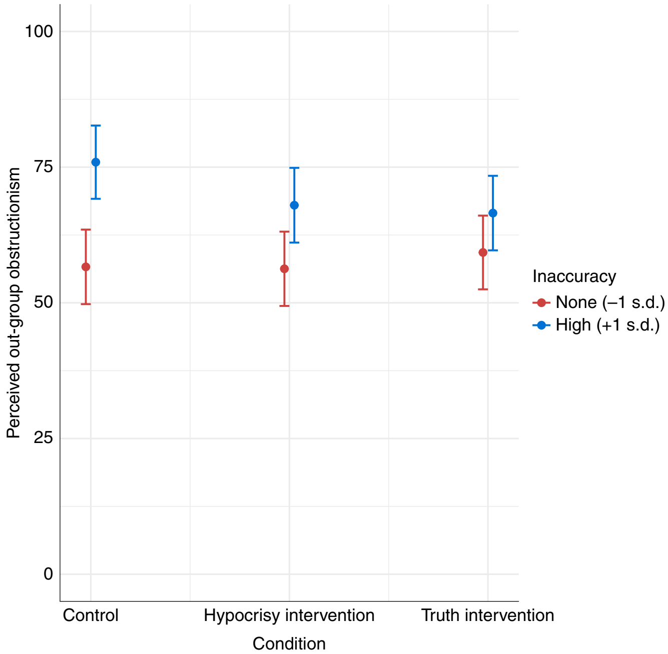

Fig. 4 | Effect of condition on obstructionism, by accuracy, in Experiment 6. Sample size, $n = 1{,}122$ (collected via Mechanical Turk). By condition: control, $n = 396$; hypocrisy intervention, $n = 368$; truth intervention, $n = 358$. GMP inaccuracy moderated the effectiveness of the hypocrisy prevention intervention ($b = -0.17$, 95% CI = [−0.33,−0.01], $\beta = -0.144$, $t(1{,}112) = -2.09$, $P = 0.037$) and truth intervention ($b = -0.27$, 95% CI = [−0.43,−0.12], $\beta = -0.23$, $t(1{,}113) = -3.39$, $P < 0.001$) at reducing obstructionism. In other words, the interventions were more effective at reducing obstructionism for participants whose GMPs were relatively more inaccurate and negative. Here inaccuracy is plotted at 1 s.d. above and below the mean inaccuracy ($M = 22$, s.d. = 22). An s.d value of −1 equals an inaccuracy of zero, meaning that the participant was on average perfectly accurate in their GMPs; an s.d. value of +1 equals an inaccuracy of 44, meaning that the participant on average overestimated out-group negativity by 44 points (on a 100-point scale). Bars are 95% CIs.

rated Democrats ($b = -3.84$, 95% CI = [−6.88,−0.79], $\beta = -0.146$, $t(1114) = -2.47$, $P = 0.014$); however, further analysis indicated no statistically significant party by condition interaction for either the truth intervention ($b = 4.44$, 95% CI[−2.97,11.88], $t(1113) = 1.17$, $P = 0.24$) or hypocrisy intervention ($b = 0.83$, 95% CI[−6.59,8.26], $t(1112) = 0.22$, $P = 0.83$). In other words, the interventions were not more effective at reducing negative motive attributions among one party relative to the other.

Further analysis revealed statistically significant interactions of condition on GMP inaccuracy (operationalized as the mean difference between participants' GMPs and the true values, such that higher values are considered more inaccurate and negative). We found that GMP inaccuracy moderated the effectiveness of the hypocrisy prevention intervention ($b = -0.17$, 95% CI = [−0.33,−0.01], $\beta = -0.144$, $t(1112) = -2.09$, $P = 0.037$) and truth intervention ($b = -0.27$, 95% CI = [−0.43,−0.12], $\beta = -0.23$, $t(1113) = -3.39$, $P < 0.001$), relative to control. In other words, the interventions were more effective at reducing obstructionism for participants whose GMPs were relatively less accurate and more negative. There was also a linear association between inaccuracy and perceived obstructionism ($b = 0.44$, 95% CI = [0.32,0.56], $\beta = 0.37$, $t(1114) = 7.33$, $P < 0.001$), replicating the finding from Study 5. See Fig. 4 for visualization of the effect of the interventions at one standard deviation above and below the mean of accuracy (see Supplementary Fig. 4 for raw data distributions). As an exploratory measure, we followed up with participants 1 week after they had completed Experiment 6, to determine whether the effect of the intervention persisted over time. We had a 73% response rate, but found no credible evidence for a continued effect of the intervention on a rating of general out-group obstructionism (see Supplementary Experiment B).

The results of Experiment 6 provided support for the hypothesis that negative motivational attributions towards the out-group, such as obstructionism, were driven in part by inaccurate beliefs regarding how negatively the out-group perceived the collective behaviour of one's in-group. They also suggest that simply providing individuals with concrete information regarding their inaccurate, and overly negative, GMPs can help reduce downstream negative attributions towards the out-group. However, we found no credible evidence that the hypocrisy prevention intervention provided additional benefit above the truth intervention, which suggests that participants were not anchoring on inaccurate beliefs about how the in-group would react to the same behaviour. Given the central role played by motive attributions in intergroup relations[26,27], our findings highlight a potential avenue for future attempts at reducing intergroup hostility and conflict, and an avenue for further understanding the antecedents of negative and inaccurate motive attributions[9,12].

Discussion

Across seven experiments and one survey, we found that group meta-perceptions were consistently inaccurate and negatively biased across a variety of competitive intergroup contexts, scenarios and participant samples. Theoretically, our findings of negative and inaccurate GMPs across multiple intergroup domains—even in the absence of group labels as in the control condition of Experiment 1—parallel research on the interindividual–intergroup discontinuity effect, which demonstrates that intergroup interactions are more hostile and competitive then interindividual interactions[28,29]. Importantly, the interindividual–intergroup discontinuity effect is observed in both actual behaviour and expectations of behaviour, in that people expect future intergroup interactions to be more hostile than interpersonal interaction[30]. If people assume that intergroup interactions are going to be more hostile, this may partially explain why GMPs are overly negative and associated with negative motive attributions, although it does not explain why GMPs are so inaccurate. Similarly, while recent evidence suggests that perceptions of political party polarization in the United States have become more negative and inaccurate over the past four decades[18,31], this does not explain inaccurate GMPs in the domain of gender, why there is no evidence for GMP inaccuracy in cooperative political contexts and why there is no evidence that inaccurate GMPs vary across the scenario content or party of the perceiver.

Several limitations in these experiments highlight fruitful avenues for future research. One assumption embedded in these studies is that actual perceptions represent ground truth. An alternative source of GMP inaccuracy may be actual perceivers downplaying their reactions to these events. For example, in Experiment 2, men might have been under-reporting their dissatisfaction with losing resources, which would make women's GMPs look more inaccurate than they are. Furthermore, the use of random-probability sampling would be superior to the quota-matching methods we used in Experiment 4 for estimating the true population 'actual perceptions' of our scenarios. Second, we did not measure confidence in participants' own judgements, which should be related to GMP (in-)accuracy as it is in other meta-perception research[32]. Third, we found no statistically significant effect of our intervention on negative motive attributions 1 week after it was administered, though we hasten to note that we specifically designed our intervention to minimize the likelihood that our results were driven by demand effects. Furthermore, the attrition rate of participants meant that our follow-up measurement 1 week later was probably underpowered.

Future research should vary the strength and nature of any such interventions to better understand which qualities provide greater (if any) benefit over time.

Conceptually, future research ought to examine the relationship between GMPs and other second-order judgements in intergroup contexts. Here we operationalized GMPs as judgements regarding out-group members' reactions to collective in-group behaviours, but GMPs can be measured along many features, including attitude[33] and trait[34] attributions (that is, 'how they see us'), dehumanization[7] (that is, 'how human they think we are'), judgements of intent[35] and even group emotions[36]. Understanding how GMPs across these judgements relate to, and are distinct from, one another will be critical in building theory around the dynamics of, and outcomes associated with, GMPs in intergroup contexts. Lastly, future work should also seek to take advantage of current events as they are unfolding to see how inaccuracies in GMP are shaped during real-world events related to issues with which people are very familiar.

Our findings highlight a consistent, pernicious inaccuracy in social perception, along with how these inaccurate perceptions relate to negative attributions towards out-groups. More broadly, inaccurate and overly negative GMPs exist across multiple competitive intergroup contexts, and we find no evidence that they differ across the political spectrum. This suggests that there may be many domains of intergroup interaction where inaccurate GMPs could potentially diminish the likelihood of cooperation and, instead, exacerbate the possibility of conflict. However, our findings also highlight a straightforward manner in which simply informing individuals of their inaccurate beliefs can reduce these negative attributions.

Methods

All studies were approved by Harvard University's Institutional Review Board, and all participants gave their informed consent before participating. All participants, except those in Experiment 4, were collected on Amazon's Mechanical Turk platform (Mturk) and were located in the United States. Participants in Experiment 4 were collected through Qualtric Survey Panels, and the sample was quota-matched to US census data distributions of the following variables in the general population: age, gender, ethnicity, education and income (see Supplementary information for demographic breakdown and quotas). All surveys were administered via the Qualtrics survey platform.

Participants. Samples from Experiments 1, 3, 4 and 6 and Study 5 consist of self-identified Republicans and Democrats, and the sample of Experiment 2 consists of self-identified men and women. Experiment 1 ($n=408$) and Experiment 2 ($n=286$) had sample sizes of 170 per condition determined a priori, with the goal of attaining 144 per condition after exclusion of participants who failed comprehension checks (see Exclusions below). An a priori power analysis indicated that 144 per condition was necessary to detect a small effect size of $f=0.15$ with 80% power within a three-condition, between-subjects analysis of variance framework. Expecting to observe a reduced effect size in Experiment 3 ($n=499$) relative to Experiment 1, we increased the sample size to a target of 275 per condition and collected 675 in the hope of reaching 550 after exclusions. We did not conduct a formal power analysis for Experiment 3. Experiment 4 had a pre-registered sample size of $n=500$ (selected via a priori power analysis to detect standardized $b=0.20$ with 80% power; see pre-registration for details); Qualtrics purposefully oversampled to ensure a minimum of 500 quality responses (hence final $n=536$). For Study 5 ($n=212$) we selected an a priori sample size of $n=300$, with the goal of attaining approximately $n=250$ after exclusions, the sample size at which small correlations stabilize[37]. Experiment 6 ($n=1,122$) had a pre-registered sample size of $n=1,510$, in the hope of obtaining 1,260 after exclusions (selected via a priori power analysis to detect standardized $b=0.20$ with 80% power; see pre-registration for details).

Exclusions. In Experiment 1 we removed 12 responses due to three separate participants taking the study multiple times (all their responses were removed). A further 89 participants failed the comprehension check and one participant was excluded for not completing the dependent variable ratings, leaving a final $n=408$ (mean age (M_{age}) $=35.2$, 239 women). In Experiment 2 we removed two responses due to one participant completing the study twice, another response due to a participant not providing their gender identity and 56 participants who failed the comprehension check, leaving a final $n=286$ ($M_{age}=36.2$, 156 women). In Experiment 3, 165 participants failed the comprehension check and 12 responses were removed due to duplicate Internet protocol addresses, leaving a final $n=499$ ($M_{age}=35.1$, 29 women). In Experiment 4, 364 participants failed the comprehension check and the Qualtrics manager continued collecting data until 536 participants (273 women; age brackets: 165 in ages 18–34, 189 in ages 35–54, 182 in ages 55+) who met our demographic quotas completed the study. In Study 5, 86 participants failed the comprehension check and two were removed for not completing the dependent variable rating, leaving a final $n=212$ ($M_{age}=35.89$, 120 women). In Experiment 6, 349 participants failed the comprehension check and 26 responses were removed due to duplicate Mturk identity or Internet protocol addresses, leaving a final $n=1,122$ ($M_{age}=35.1$, 642 women). We did not weigh Mturk samples by political party or gender, because we were interested in in-group versus out-group dynamics and not the difference between, for example, Democrats and Republicans. In Experiment 4 we quota-matched to a 50/50 split of Democrats and Republicans. Self-identified Independents were allowed to complete all studies (except Experiment 4), but were excluded from all analyses a priori.

Compensation. Experiments 1 and 3 and Study 5 paid US$0.10 and were advertised as taking 1 min. Experiment 4 was advertised as taking 9 min (per se, 4 min, but it was bundled with a separate 5-min study which always followed Experiment 4), and participants were paid a preset amount of credit via Qualtrics Panel's internal payment system. Experiment 6 paid US$0.15 and was advertised as taking 60–90 s.

Procedure. We randomly assigned participants to condition and scenario (in Experiment 4, scenario order) across all the experiments and studies. Across scenarios, we also randomized the order of the dependent variable items (for example, dislike, opposition). All randomization was facilitated through Qualtrics' randomization functions. The surveys were programmed to pipe the appropriate out-group and/or in-group labels into the scenarios and dependent variables ratings based on the participants' self-reported group affiliation. All dependent variables across all studies appeared as sliding scales with end-labels and tick-marks, but no visible numbers (except for the ratings in Experiment 6, in which a numeric value (1–100) appeared next to the slider when participants provided a response). Across all experiments and studies, except Experiment 4, excluded participants received full compensation.

Analyses. We analysed Experiments 1–4 and Study 5 using mixed-effects beta-regressions (glmmTMB[38] package, v.0.2.3) in R (v.3.6.1), and Experiment 6 using linear mixed-effects modelling (lmerTest[39] R package, v.3.1-0). All post hoc tests used the Tukey method for P-value adjustment and were conducted with the emmeans[40] R package (v.1.4). We used beta-regressions for Experiments 1–4 and Study 5 due to the highly skewed GMP response data, and transformed the data for beta-regressions using established formulae[41]. As a robustness check we performed all non-pre-registered beta-regression analyses (Experiments 1, 2 and 3 and Study 5) using linear mixed-effects modelling via the lmerTest R package: none of our results changed meaningfully. For Experiments 1, 3 and 4, Study 5 and the main effects of Experiment 6, we report the results from models that include only the main effects because there were never any significant interactions among the fixed effects; furthermore, the saturated models including fixed effects and the corresponding interactions did not improve model fits. Results for Experiment 2 are from the saturated models and, while we report the interaction of accuracy on condition in Experiment 6, we never find an interaction with party identification and do not report those saturated models. Across Experiments 1–4 we regressed the relevant dependent variable rating (dislike, opposition, political/social unacceptability) onto fixed effects for condition and the relevant group variable ('party accuracy' in Experiments 1, 3 and 4, 'gender accuracy' in Experiment 2), a random effect with random intercepts for scenario (along with a random effect with random intercepts for participant in Experiment 4, due to the repeated measures), and in Experiment 2 an interaction term for the condition by group interaction. In Study 5 we regressed obstructionism onto each GMP item separately, including a fixed effect for party and a random effect with random intercepts for scenario. In Experiment 6 we regressed obstructionism onto condition, including a fixed effect for party and a random effect with random intercepts for scenario, then replaced the fixed effect for party with the interaction of accuracy with condition. All tests were two-sided. Data analyses were not performed blind to the conditions of the experiments and studies. Figures were created using the R packages ggstatsplot[42] (v.0.0.12), sjPlot[43] (v.2.7.0) and psych[44] (v.1.8.12).

Experiments 4 and 6 were pre-registered. Experiment 4 was pre-registered on 26 February 2019 and can be found at https://osf.io/atck5. Experiment 6 was pre-registered on 19 March 2019 and can be found at https://osf.io/jhnsb. No analyses deviate from the pre-registrations.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data that supported the findings of this study are publicly available in CSV format on the Open Science Framework: https://osf.io/zhysa/.

Code availability

All analyses reported in this study used the statistical software R (v.3.6.1). All R files are publicly available on the Open Science Framework: https://osf.io/zhysa/.

Received: 9 August 2018; Accepted: 1 October 2019;
Published online: 11 November 2019

References

1. Carlson, E. N. Meta-accuracy and relationship quality: weighing the costs and benefits of knowing what people really think about you. *J. Personal. Soc. Psychol.* **111**, 250–264 (2016).
2. Carlson, E. N., Vazire, S. & Furr, R. M. Meta-insight: do people really know how others see them? *J. Personal. Soc. Psychol.* **101**, 831–846 (2011).
3. Vazire, S. & Carlson, E. N. Others sometimes know us better than we know ourselves. *Curr. Dir. Psychol. Sci.* **20**, 104–108 (2011).
4. Vorauer, J. D., Main, K. J. & O'Connell, G. B. How do individuals expect to be viewed by members of lower status groups? Content and implications of meta-stereotypes. *J. Personal. Soc. Psychol.* **75**, 21 (1998).
5. Vorauer, J. D., Hunter, A. J., Main, K. J. & Roy, S. A. Meta-stereotype activation: evidence from indirect measures for specific evaluative concerns experienced by members of dominant groups in intergroup interaction. *J. Personal. Soc. Psychol.* **78**, 690–707 (2000).
6. Frey, F. E. & Tropp, L. R. Being seen as individuals versus as group members: extending research on metaperception to intergroup contexts. *Personal. Soc. Psychol. Rev.* **10**, 265–280 (2006).
7. Kteily, N., Hodson, G. & Bruneau, E. They see us as less than human: metadehumanization predicts intergroup conflict via reciprocal dehumanization. *J. Personal. Soc. Psychol.* **110**, 343–370 (2016).
8. Sigelman, L. & Tuch, S. A. Metastereotypes: blacks' perceptions of whites' stereotypes of blacks. *Public Opin. Q.* **61**, 87 (1997).
9. Finchilescu, G. Intergroup anxiety in interracial interaction: the role of prejudice and metastereotypes. *J. Soc. Issues* **66**, 334–351 (2010).
10. Klein, O. & Azzi, A. E. The strategic confirmation of meta-stereotypes: how group members attempt to tailor an out-group's representation of themselves. *Br. J. Soc. Psychol.* **40**, 279–293 (2001).
11. Waytz, A., Young, L. L. & Ginges, J. Motive attribution asymmetry for love vs. hate drives intractable conflict. *Proc. Natl Acad. Sci. USA* **111**, 15687–15692 (2014).
12. Lau, T., Morewedge, C. K. & Cikara, M. Overcorrection for social-categorization information moderates impact bias in affective forecasting. *Psychol. Sci.* **27**, 1340–1351 (2016).
13. Goldstein, N. J., Vezich, I. S. & Shapiro, J. R. Perceived perspective taking: when others walk in our shoes. *J. Personal. Soc. Psychol.* **106**, 941–960 (2014).
14. Saguy, T. & Kteily, N. Inside the opponent's head: perceived losses in group position predict accuracy in metaperceptions between groups. *Psychol. Sci.* **22**, 951–958 (2011).
15. Robinson, R. J., Keltner, D., Ward, A. & Ross, L. Actual versus assumed differences in construal: 'Naive realism' in intergroup perception and conflict. *J. Personal. Soc. Psychol.* **68**, 404–417 (1995).
16. Chambers, J. R. & Melnyk, D. Why do I hate thee? Conflict misperceptions and intergroup mistrust. *Personal. Soc. Psychol. Bull.* **32**, 1295–1311 (2006).
17. Chambers, J. R., Baron, R. S. & Inman, M. L. Misperceptions in intergroup conflict. *Psychol. Sci.* **17**, 38–45 (2006).
18. Westfall, J., Van Boven, L., Chambers, J. R. & Judd, C. M. Perceiving political polarization in the United States: party identity strength and attitude extremity exacerbate the perceived partisan divide. *Perspect. Psychol. Sci.* **10**, 145–158 (2015).
19. Bush, G. W. President Bush's address to a joint session of Congress and the nation. *The Washington Post.* http://www.washingtonpost.com/wp-srv/nation/specials/attacked/transcripts/bushaddress_092001.html (20 September 2001).
20. Sunstein, C. R. Why they hate us: The role of social dynamics. *Harvard J. Law Public Policy* **25**, 429–440 (2002).
21. Zakaria, F. The politics of rage: why do they hate us? *Newsweek.* https://www.newsweek.com/politics-rage-why-do-they-hate-us-154345 (14 October 2001).
22. Merskin, D. The construction of Arabs as enemies: post-September 11 discourse of George W. Bush. *Mass Commun. Soc.* **7**, 157–175 (2004).
23. Rogers, T. & Feller, A. Reducing student absences at scale by targeting parents' misbeliefs. *Nat. Hum. Behav.* **2**, 335–342 (2018).
24. Nyhan, B. & Reifler, J. The roles of information deficits and identity threat in the prevalence of misperceptions. *J. Elect. Public Opin. Parties* **29**, 1–23 (2018).
25. Eisenkraft, N., Elfenbein, H. A. & Kopelman, S. We know who likes us, but not who competes against us:dyadic meta-accuracy among work colleagues. *Psychol. Sci.* **28**, 233–241 (2017).
26. Reeder, G. D., Vonk, R., Ronk, M. J., Ham, J. & Lawrence, M. Dispositional attribution: multiple inferences about motive-related traits. *J. Personal. Soc. Psychol.* **86**, 530–544 (2004).
27. Miller, D. T. & Nelson, L. D. Seeing approach motivation in the avoidance behavior of others: implications for an understanding of pluralistic ignorance. *J. Personal. Soc. Psychol.* **83**, 1066–1075 (2002).
28. Insko, C. A., Schopler, J., Hoyle, R. H., Dardis, G. J. & Graetz, K. A. Individual-group discontinuity as a function of fear and greed. *J. Personal. Soc. Psychol.* **58**, 68–79 (1990).
29. Wildschut, T., Pinter, B., Vevea, J. L., Insko, C. A. & Schopler, J. Beyond the group mind: a quantitative review of the interindividual-intergroup discontinuity effect. *Psychol. Bull.* **129**, 698–722 (2003).
30. Pemberton, M. B., Insko, C. A. & Schopler, J. Memory for and experience of differential competitive behavior of individuals and groups. *J. Personal. Soc. Psychol.* **71**, 14 (1996).
31. Enders, A. M. & Armaly, M. T. The differential effects of actual and perceived polarization. *Polit. Behav.* **41**, 815–839 (2018).
32. Carlson, E. N., Furr, R. M. & Vazire, S. Do we know the first impressions we make? Evidence for idiographic meta-accuracy and calibration of first impressions. *Soc. Psychol. Personal. Sci.* **1**, 94–98 (2010).
33. Stern, C. & Kleiman, T. Know thy outgroup: promoting accurate judgments of political attitude differences through a conflict mindset. *Soc. Psychol. Personal. Sci.* **6**, 950–958 (2015).
34. Stroessner, S. J. & Dweck, C. S. in *Social Perception: From Individuals to Groups* (eds Stroessner, S. J. & Sherman, J. W.) 177–196 (Psychology Press, 2015).
35. Ames, D. & Fiske, S. Perceived intent motivates people to magnify observed harms. *Proc. Natl Acad. Sci. USA* **112**, 3599–3605 (2015).
36. Goldenberg, A., Saguy, T. & Halperin, E. How group-based emotions are shaped by collective emotions: evidence for emotional transfer and emotional burden. *J. Personal. Soc. Psychol.* **107**, 581–596 (2014).
37. Schönbrodt, F. D. & Perugini, M. At what sample size do correlations stabilize? *J. Res. Personal.* **47**, 609–612 (2013).
38. Brooks, M. et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R J.* **9**, 378–400 (2017).
39. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: Tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).
40. Lenth, R. emmeans: Estimated marginal means, aka least-squares means. R package version 1.4. https://CRAN.R-project.org/package=emmeans (2019).
41. Smithson, M. & Verkuilen, J. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychol. Methods* **11**, 54–71 (2006).
42. Patil, I. & Powell, C. ggstatsplot: "ggplot2" based plots with statistical details. R package version 0.0.12. https://CRAN.R-project.org/package=ggstatsplot (2018).
43. Lüdecke, D. sjPlot: Data visualization for statistics in social science. R package version 2.7.0. https://CRAN.R-project.org/package=sjPlot (2019).
44. Revelle, W. psych: Procedures for psychological, psychometric, and personality research. R Package version 1.8.12. https://CRAN.R-project.org/package=psych (2018).

Acknowledgements

Work on this project by M.C. was supported by a National Science Foundation Award (no. BCS-1551559). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We thank members of the Harvard Intergroup Neuroscience Lab, Sidanius Lab and attendees at the 2018 East Coast Doctoral Conference for their helpful comments, Z. Ingbretsen and N. Hunt for help with data collection and I. Zahn and S. Worthington for statistical assistance.

Author contributions

J.L. and M.C. designed all experiments and wrote the manuscript. J.L. completed data collection and analysis under the supervision of M.C.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41562-019-0766-4.

Correspondence and requests for materials should be addressed to J.L. or M.C.

Peer review information Primary handling editor: Aisha Bradshaw

nature research

Corresponding author(s): Jeffrey Lees & Mina Cikara

Last updated by author(s): Sep 26, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Amazon's Mechanical Turk platform was used to recruit participants in all experiments/studies except Experiment 4, which utilized Qualtrics Survey Panels, and the Qualtrics survey platform was used to facilitate all experiments/studies.
Data analysis	The statistical software package R (v. 3.6.1) was used to conduct all analyses.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data files (csv format) and analysis code (R files) can be found on the Open Science Foundation's website via the following link: https://osf.io/zhysa/

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	All experiments/studies are quantitative, cross-sectional, online survey-based studies. Experiments 1, 2, 3, 6 and Sup. Exp B utilize between-subjects designs. Study 5 is correctional and involves no manipulations, and Experiment 4 and Supp. Exp A utilize within-between subjects designs.
Research sample	All participants (expect those in Experiment 4) were collected via Amazon's Mechanical Turk platform, which allows for a more representative sample compared to university students. All participants had to be 18 years of age or older, residing in the United States, and have a 95% HIT rate or higher on Mturk. Experiment 4 was collecting using a general population Qualtrics survey panel. Experiment 1 (N = 408), Mean Age = 35,Male/Female = 169/239, Democrat/Republican = 271/137 Experiment 2 (N = 286), Mean Age = 36,Male/Female = 130/156 Experiment 3 (N = 499), Mean Age = 35,Male/Female = 206/293, Democrat/Republican = 328/117 Experiment 4 (N = 536), Age Brackets = 18-34 (165), 35-54 (189), 55+ (182) Male/Female = 263/273, Democrat/Republican = 269/267 Study 5 (N = 212), Mean Age = 36, Male/Female = 92/120, Democrat/Republican = 132/80 Experiment 6 (N = 1122), Mean Age = 35, Male/Female = 480/642, Democrat/Republican = 704/418 Supp. Experiment A (N = 397), Mean Age = 35, Male/Female = 198/199, Democrat/Republican = 260/137 Supp. Experiment B (N = 822), Male/Female = 343/479, Democrats/Republicans = 529/293
Sampling strategy	All experiments/studies aimed to achieve a minimum power of 0.80, and all a priori sample size determinations were made with the knowledge that a small plurality of participants would fail the comprehension check. See methods section for exact sample size and power calculations. All experiments/studies (except Experiment 4, see below) were focused solely on Democrats & Republicans (except in Experiment 2, which focused on men and women). However our surveys were open to the general population, and participants who self-identified as neither Democrat/Republican (or men/women in Experiment 2) were allowed to complete the survey and were paid. As such there was no targeted sampling used (e.g. specific survey panels), nor was our sampling weighted by party/gender. Experiment 4 utilized Qualtrics survey panels to obtain a nationally representative sample of American participants. Participants were quota matched to census distributions along the following variables: age, gender, ethnicity, education, and income. The sample also had a quota for a 50/50 split of Democrats and Republicans as to not imbalance the study design. See supplemental materials for the exact quotas and achieved sample along the five variables above. Participants who did not meet the demographic quotas were not allowed to proceed with the study.
Data collection	Data collection for all experiments/studies (except Experiment 4) was conducted online via Amazon's Mechanical Turk. Experiment 4 was conducted via Qualtrics Survey Panels. Experimenters were blinded to group allocation during data collection for all experiments/studies.
Timing	Experiment 1 was collected in November 2017 Experiment 2 was collected in December 2017 Experiment 3 was collected in December 2018 Experiment 4 was collected in March 2019 Study 5 was collected in May 2018 Experiment 6 was collected in March 2019 Supp. Experiment A was collected in December 2018 Supp. Experiment B was collected in April 2019
Data exclusions	All participants who identified as neither Republican or Democrat (or neither male/female in Experiment 2) were excluded a priori. Experiment 1: 248 Independents/other Experiment 2: 1 participant of unspecified gender Experiment 3: 300 Independents/other Experiment 4: No exclusion of participants who completed the experiment Study 5: 143 Independents/other Experiment 6: 572 Independents/other Supp. Experiment A: 108 Independents/other Supp. Experiment B: 64 whose demographic data at T2 did not match that from T1 Participants were also excluded from all analyses if they failed the comprehension check, took the study multiple times (i.e. duplicate IP addresses or Mturk IDs), or failed to respond to the relevant dependent variable (pairwise deletion). See methods section for exact breakdown of said exclusions by experiment/study.
Non-participation	Across all experiments/studies several participants began, but never completed, the online surveys. Incomplete responses were not included in any analyses. Number of Incomplete Responses (% Completion rate): Experiment 1: 28 (96%) Experiment 2: 9 (97%) Experiment 3: 40 (96%)

Experiment 4: 113 (97%, based on participants who finished or were excluded because of lack of qualification)
Study 5: 13 (97%)
Experiment 6: 81 (96%)
Supp. Experiment A: 36 (95%)
Supp. Experiment B: 5 (99%)

Randomization

All randomization was facilitated by the survey platform, Qualtrics. As such experimenters were unaware of which participants were allocated to which experimental condition. Randomization was not associated with any features of the participant (e.g. political party of gender).

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology
☒	☐	Animals and other organisms
☐	☒	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics

See above

Recruitment

See above

Ethics oversight

Harvard University's Institutional Review Board

Note that full information on the approval of the study protocol must also be provided in the manuscript.