

Reputational stakes shape partisans' reactions to material suffering

Charles A. Dorison^{a,1} & Nour Kteily^{a,b}

^aKellogg School of Management
Northwestern University
Evanston, IL 60208

^bHarvard Business School
Harvard University
Boston, MA 02163

¹Address correspondence to charles.dorison@kellogg.northwestern.edu.

Classification: Social Sciences, Psychological and Brain Sciences

Manuscript is currently under review.

Abstract

When groups compete, a proclivity to ‘merely’ favor ingroup members can turn into more insidious intergroup animosity. Among political partisans in the United States, for example, outparty hate has recently emerged as a stronger force than inparty love. In such contexts, the relative standing between groups can take on heightened psychological significance, with group members prioritizing staying ahead of the other side over other important considerations. Here, we examined how Democrats’ and Republicans’ sensitivity to the reputational standing of their group influenced their affective and behavioral responses to material harm. When considering a policy affecting society as a whole (themselves included), would partisans prefer that the policy fail and see the outgroup blamed or succeed and see the outgroup praised? Would partisans be less willing to donate to COVID-19 relief efforts if an outparty’s leader would get credit for the efforts? Five experiments (N=6698) revealed that Democrats and Republicans were more accepting of (and less willing to intervene to stop) widespread material harm if that harm bolstered the ingroup’s reputational standing relative to the outgroup. These patterns persisted (1) across both affective responses and (financially-incentivized) behavior; (2) both when reacting to past harm and evaluating opportunities to stop future harm; and (3) after ruling out non-reputational interpretations. Strikingly, incentives for reputational standing can blunt—and sometimes even eliminate—preferences for positive over negative material outcomes. Our work sheds light on the central role of group-based reputational standing in contributing to cycles of socially destructive intergroup animus.

Significance Statement

As competition between groups intensifies, relative group-based standing rises in significance. One key pathway through which groups reinforce their standing is by bolstering the ingroup's (or damaging the outgroup's) reputation. But just how far are group members willing to go to gain the upper hand? We examined this question in the context of partisan politics in the United States. Across five experiments, Democrats and Republicans were more accepting of (and less willing to intervene to stop) widespread material suffering when it improved their group's relative reputational standing. The present work highlights the critical role of group-based reputational incentives in reinforcing conflict by decreasing sensitivity to material suffering, a key outcome that participants in such conflicts typically aim to alleviate.

Our ability to affiliate and empathize with others is an important part of what makes us human. Humans cooperate within large, socially-constructed groups and extend care and compassion for the suffering of others within those groups (de Waal, 2008; Goetz, Keltner, & Simon-Thomas, 2010; Greene, 2013; Hare, 2017; Henrich & Muthukrishna, 2021; Kurzban, Burton-Chellew, & West, 2015). And yet, our proclivity toward cooperation and empathy is often parochial, limited to those within ingroup boundaries (Amodio & Cikara, 2021; Cikara, Bruneau, & Saxe, 2011; Esses, 2021; Hewstone, Rubin, & Willis, 2002; Richeson & Sommers, 2016). Indeed, in contexts marked by group competition, our ingroup love can be overwhelmed by outgroup hate, and we sometimes experience indifference or even joy at the suffering of outgroup others (Cikara, Bruneau, & Saxe, 2011; Cikara & Fiske, 2012; Combs, Powell, Schurtz, & Smith, 2009; Leach, Spears, Branscombe, & Doosje, 2003).

One striking modern example of such dynamics is political partisanship in the United States (Abramowitz & Webster, 2016; Davidai & Ongis, 2018; Finkel et al., 2021; Iyengar, Lelkes, Levendusky, & Malhotra, 2019; Iyengar, Sood, & Lelkes, 2012; Kalin & Sambanis, 2018). Not only has Democrats' and Republicans' warmth toward the outparty diminished, but outparty hate has emerged in recent years (and for the first time) as a stronger force than inparty love (Finkel et al., 2021). In turn, Democrats and Republicans increasingly view each other as evil and immoral (Druckman & Levendusky, 2019; Finkel et al., 2021).

This animus can be constructive when it fuels civic action: for example, voter turnout in the 2020 United States Presidential Election reached its highest point in over a century (Corasiniti & Rutenberg, 2021), likely reflecting the energizing effects of a strong commitment to electorally defeating outparty rivals. Further, anger can catalyze social movements, such as in

the large-scale protests for social justice in the wake of police killings of black men in the United States in the summer of 2020 (Kudesia, 2021).

The same animus, however, can have destructive effects. Polarization can erode democratic functioning by leading to institutional gridlock (Fiorina & Abrams, 2008; McCarty, Poole, & Rosenthal, 2016; Svobik, 2019). Affective polarization can also damage personal relationships (Druckman & Levendusky, 2019; Iyengar et al., 2019; Iyengar, Sood, & Lelkes, 2012; Iyengar & Westwood, 2015). Indeed, partisans sometimes go as far as blatantly dehumanizing members of other groups (Cassese, 2021; Kteily, Bruneau, Waytz, & Cotterill, 2015; Kteily, Hodson, & Bruneau, 2016; Martherus et al., 2021).

In many respects, then, the state of partisan conflict in the United States increasingly resembles sectarian conflicts such as those in the Middle East (Finkel et al., 2021). As we detail further below, when intergroup animus reaches such heights, the standing of one's group relative to the outgroup can, psychologically, take on existential significance. In turn, partisans may be willing to go to great lengths in order to harm the outgroup's standing or bolster the ingroup's.

In the present research, we focus on group members' willingness to tolerate material harm to the world to the extent that this damages the outgroup's reputation or burnishes the ingroup's. Importantly, whereas prior work not focused on reputation suggests that animus towards another group can lead us to sometimes feel happy when they experience (relatively minor) pain (Leach et al., 2003; Leach & Spears, 2008; Combs et al., 2009), we focus on how competition over group-based standing rooted in reputational considerations affects individuals' affective and behavioral responses to (relatively grave) widespread material harm borne by society as a whole (including the ingroup and its causes). When considering a policy affecting society as a whole (themselves included), would partisans prefer that the policy fail and see the

outgroup blamed or succeed and see the outgroup praised? Would partisans be less willing to donate to COVID-19 relief efforts if an outparty's leader would get credit for the efforts?

Theoretical background

Why might group members place so much weight on reputational standing? At the individual level, we are often highly sensitive to reputation, becoming (among other things) less likely to engage in unethical behavior (Bateson, Nettle, & Roberts, 2006) and more willing to enact costly punishment to enforce group norms (Jordan, Hoffman, Bloom, & Rand, 2016; Jordan & Rand, 2020) when our reputation is on the line. This sensitivity to reputation is functional insofar as our reputation helps determine our social status and affects others' willingness to interact and cooperate with us (Baumeister & Leary, 1995; Dorison, Umphres, & Lerner, 2021; Jordan, Hoffman, Nowak, & Rand, 2016; Lerner & Tetlock, 1999; Schlenker & Weigold, 1992; Tetlock, 2000, 2002).

At the group level, reputational standing may have similar functional value. When groups are engaged in political battle, actions that damage the outgroup's reputation or bolster the ingroup's can shape groups' odds of gaining power, for example, by influencing the sides' respective likelihood of winning elections. Relatedly, when groups are locked in a competition over moral claims, actions that make the outgroup look bad can tilt third-party moral—and material—support in favor of the ingroup (Noor, Shnabel, Halabi, & Nadler, 2012). Consider for example the conflict between Israelis and Palestinians, in which the two groups compete for third-party support over the legitimacy of their respective territorial claims (Shamir, 2007). In contexts like these, actions that tarnish the outgroup's reputation can have important strategic benefits for one's ingroup (e.g., increasing the odds of third parties intervening on one's behalf in the conflict; changing the narrative around moral righteousness; see also Saguy & Kteily, 2011).

In early 2021, for example, Israel was publicly criticized for not sending enough COVID-19 vaccines to Palestinians in the West Bank and Gaza (Kingsley, 2021; Rasgon, 2021). Israel's reputational loss could benefit Palestinians strategically by increasing the perceived legitimacy of Palestinians' grievances before the international community. From this perspective, and despite the material harm it imposes, outcomes that are materially bad for the ingroup (e.g., lacking life-saving vaccines) might nevertheless be experienced with (at least some modicum of) affective pleasure given the boost they provide in the reputational tug-of-war. Such reputational incentives might in turn make Palestinians less sensitive to the difference between good material outcomes for their group (which come with reputational benefits for Israel) and bad material outcomes for their group (which inflict reputational costs on the outgroup).¹

Of note, positivity toward material suffering that tips the reputational scales in favor of the ingroup over the outgroup could occur with or without conscious strategic calculation. That is, individuals might sometimes explicitly make the calculation that a specific material harm borne by the ingroup is worth the reputational benefit—perhaps particularly if the situation makes the reputational upside highly salient. For example, some Palestinians might consciously calculate that the strategic value of the reputational boost to Palestinians if Israel is seen to withhold vaccines from Palestinians outweighs the material cost of delayed vaccinations to Palestinians. But the affective enjoyment of strategically beneficial reputational harm to the outgroup—particularly in conflicts that have taken on sectarian qualities—might even become divorced from any conscious strategic deliberation. That is, when groups are locked in intense

¹ Note that we did not theoretically differentiate between damage to the outgroup's reputation and benefit to the ingroup's reputation (or between benefit to the outgroup's reputation and damage to the ingroup's). Rather, we focus on individuals' sense of how a given event shapes the relative balance between the two sides in a reputational tug-of-war. From that perspective, harm to the outgroup's reputation or benefit to the ingroup's both tilt the balance in favor of the ingroup.

conflict, individuals might simply come to intuitively enjoy reputational damage to the outgroup—even when the strategic implications are not explicitly salient or obvious. At a proximate level, individuals might simply enjoy ‘sticking it’ to the other side, even when that means taking on material harm.

Our theorizing thus far considers when and why individuals might be sensitive to reputational stakes in the context of widespread material suffering. But will individuals *actually* be motivated by reputation when material harm is at stake? Tolerating widespread material harm to others—especially those belonging to the ingroup or its causes—might be a bridge too far, no matter the implications for group standing. In fact, prior research suggests a reticence to make certain “taboo” tradeoffs, particularly when these involve loss of life. Instead, individuals treat lives as sacred (i.e., possessing transcendent significance that precludes comparison; Tetlock et al., 2000). Furthermore, and despite the fact that Democrats and Republicans routinely engage in outgroup derogation, they may be unwilling to impose material harm on the world (including themselves) in exchange for reputational gains viz. their counterparts. After all, Democrats and Republicans still share a common national bond, and willingness to accept material harm in the service of reputational gain might reasonably be viewed as “un-American.” From this perspective, partisans’ concern with avoiding widespread material suffering may overwhelm any reputational considerations.

Overview of studies

In the present work, we examine how competition over group-based standing in contexts characterized by deep intergroup animus shapes individuals’ affect and behavior. In Study 1, we provide an initial test of whether group-based reputational incentives shape partisans’ reactions when responding to materially good and bad outcomes, even in the context of immediate, large-

scale harm to society. In Studies 2-4, we investigate whether group-based reputational incentives can reduce affective differentiation between material outcomes that are good vs. bad for the world (including the ingroup/ingroup causes). Finally, in Study 5, we use an incentive-compatible design to test whether individuals are less willing to actively intervene to stop suffering when doing so would yield reputational gains for the outgroup. Taken together, the present work sheds light on the central role of group-based reputational standing in contributing to cycles of socially destructive intergroup animus.

Open science statement. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in all studies (Simmons, Nelson, & Simonsohn, 2012). Sample sizes were chosen based on power analyses from pilot studies. All studies were preregistered on aspredicted.org and all data, materials, pre-registrations, and analysis code can be found on [Researchbox.com](https://researchbox.com).²

Study 1

In Study 1 (N = 981 Democrats in July 2021), we investigate whether partisans take group-based reputation into account or not when responding to materially good and bad outcomes. Put another way, are partisans' affective reactions determined solely by what actually materially happens, or do partisans additionally factor in whether their ingroup's reputation is burnished or not in the process?

² Researchbox: https://researchbox.org/388&PEER_REVIEW_passcode=KDPXLZ

Study 1: https://aspredicted.org/483_6GN

Study 2a: https://aspredicted.org/YXM_MPC

Study 2b: https://aspredicted.org/RF5_XFS

Study 3a: https://aspredicted.org/NZ9_587

Study 3b: https://aspredicted.org/4NW_67L

Study 4: https://aspredicted.org/YL6_422

Study 5: https://aspredicted.org/1V5_CW5

Democrats were randomly assigned to one of four between-subjects experimental conditions. In all conditions, participants read an ostensibly real news story in which they learned that President Biden had made a correct private prediction about the success or failure of a new domestic economic policy. We held Biden's correct private prediction constant across conditions to ensure that any differences in affective responses across conditions did not come from differences in feeling confirmed in the belief about Biden being right vs. wrong. The story also made clear that while the policy was put into effect while President Biden was in office, the policy was developed and implemented independently of his private prediction, before he came into power.

We manipulated both the material outcome of the policy for society (good vs. bad) and the reputational stakes for Biden (high vs. low). To manipulate the material outcome, participants learned that a new report found either that the policy had reduced unemployment and efficiently allocated resources (good material outcome) or had no effect on unemployment and inefficiently allocated resources (bad material outcome). To manipulate the reputational outcome, participants learned either that President Biden had received public credit in the news media for his correct private prediction (high reputation) or had not received public credit for his correct private prediction (low reputation). Specifically, in the high reputation/good material outcome condition, participants read (1) that credit for the policy had been attributed to Biden; (2) that media reporting had emphasized Biden's support of the policy, and (3) that public opinion polling suggested that perceptions of Biden had improved as a result. In contrast, in the low reputation/good material outcome condition, participants read (1) that credit for the policy had been attributed to the independent advisory committee; (2) that media reporting had clarified Biden's lack of involvement, and (3) that public opinion polling suggested that perceptions of

Biden had remained unchanged as a result. In the bad material outcome conditions, participants read analogous text in which Biden was either praised for predicting the policy's struggles (high reputation) or not praised for predicting the policy's struggles (low reputation). Biden was not blamed for the policy in any condition.

The primary outcome variable was participants' affective responses to reading the news story. We measured affective reactions in two ways: with a single global evaluation and with a combined index of ten specific emotion items, including five positive items (e.g., happy, relieved, glad) and five negative items (e.g., upset, sad, distressed; see Method for details).

We had two main hypotheses for this study. First, we expected people to show a clear overall preference for good (vs. bad) material outcomes. That is, irrespective of whether or not Biden received praise, we expected individuals to generally show more positive affect when the policy was successful versus when it was unsuccessful.

Second, and more centrally, we hypothesized that Democrats would report more positive affective reactions when President Biden received (vs. did not receive) public credit for his private correct prediction. That is, we expected reputational considerations to shape individuals' affective responses to material outcomes: namely, that individuals' joy at the success of the policy would be greater when Biden was credited than when he was not, and that individuals' distress at the failure of the policy to be lower when Biden was credited for predicting its failure than when he was not.

We found support for both hypotheses. First, participants reported more positive affective reactions when the policy was a success compared to when it was a failure, and this was true for both the single general affect item ($M_{\text{bad material}} = 42.41$ vs. $M_{\text{good material}} = 67.92$, $t(977) = 21.18$, p

< .001, Cohen's $d = 1.35$) and the 10-item composite ($M_{\text{bad material}} = -1.18$ vs. $M_{\text{good material}} = 2.07$, $t(979) = 25.17$, $p < .001$, Cohen's $d = 1.61$).

Second, and more central to our theorizing, we found that participants simultaneously responded to reputational incentives, reporting more positive affective responses when Biden received a reputational boost for his correct private prediction than when he did not receive a reputational boost for his correct private prediction (single general affect item: $M_{\text{low reputation}} = 52.60$ vs. $M_{\text{high reputation}} = 58.02$, $t(977) = 3.76$, $p < .001$, Cohen's $d = 0.24$; 10-item composite: $M_{\text{low reputation}} = 0.14$ vs. $M_{\text{high reputation}} = 0.79$, $t(979) = 3.95$, $p < .001$, Cohen's $d = 0.25$). Of note, such effects held both when the material outcome was positive (single general affect item: $M_{\text{low reputation}} = 64.39$ vs. $M_{\text{high reputation}} = 71.35$, $t(494) = 4.12$, $p < .001$, Cohen's $d = 0.37$; 10-item composite: $M_{\text{low reputation}} = 1.62$ vs. $M_{\text{high reputation}} = 2.50$, $t(496) = 4.71$, $p < .001$, Cohen's $d = 0.42$) and when the material outcome was negative (single general affect item: $M_{\text{low reputation}} = 40.60$ vs. $M_{\text{high reputation}} = 44.21$, $t(481) = 2.14$, $p = .033$, Cohen's $d = 0.20$; 10-item composite: $M_{\text{low reputation}} = -1.38$ vs. $M_{\text{high reputation}} = -1.00$, $t(481) = 2.20$, $p = .029$, Cohen's $d = 0.20$).³

Study 1 demonstrated that Democrats attended to group-based reputational incentives both in the context of material gain and in the context of material suffering. One might have imagined that partisans would simply focus on the fact that policy was effective (or not). Instead, whether an ingroup leader (i.e., President Biden) received reputational credit shaped affective responses, regardless of whether the policy was a success or failure. Of note, we held worldview confirmation constant, by having Biden make a correct private prediction in all cases. We consider worldview confirmation again in Study 4.

³ In Study 1, as in all studies, we included individual difference measures to examine the extent to which the effect of reputation salience on tolerance for material harm might be enhanced vs. dampened by ingroup attachment and ingroup glorification (Roccas, Klar, & Liviatan, 2006). Across studies, we found some evidence, albeit inconsistent evidence, supporting this theorizing (data available on ResearchBox.org via the link above).

While participants factored in reputation both when the material outcome was good and when it was bad, it is notable that they were still happier in the low reputation/good material outcome condition (Means = 64.39 and 1.62) than in the high reputation/bad material outcome condition (Means = 44.21 and -1.00). That being said, a shortcoming of the present design is that it does not allow for a clean comparison of the relative effects of material and reputational outcomes. Specifically, while material outcomes ranged from positive to negative, reputational outcomes ranged only from positive to absent (rather than negative).

For this reason, in the following studies (Studies 2-4), we more cleanly investigate the relative effects of material and reputational outcomes by explicitly pairing bad reputational outcomes with good material outcomes, and vice versa (i.e., good reputational outcomes with bad material ones). This design allows us to not only directly compare their relative contributions to affective responses, but also to examine a novel implication of our framework: whether reputational stakes can blunt affective differentiation between good and bad material outcomes. That is, does pairing good material outcomes with bad reputational ones (and vice versa) reduce the baseline differentiation between good and bad material outcomes? We test this hypothesis in Studies 2-4.

Study 2

In Study 2, we build on Study 1 by again examining whether partisans remain sensitive to reputational considerations, including in situations involving large amounts of material harm. Critically, as described above, we also begin to assess whether reputational incentives can blunt partisans' affective differentiation between positive vs. negative material outcomes by pairing good material outcomes with bad reputational outcomes (and vice versa). Of note, although Study 1 happened to focus on Democrats, our theorizing applies to both Democrats and

Republicans. In Study 2 (August 2020: Study 2a = 991 Republicans and Study 2b = 1,023 Democrats), we broaden our investigation to include both Republicans and Democrats.

Participants were randomly assigned to one of four between-subjects experimental conditions. In all conditions, participants read an ostensibly real news story in which a public official made a pessimistic prediction, forecasting an ensuing economic collapse due to either failure to re-open businesses quickly enough from COVID-19 related shutdowns (for Republicans) or failure to pass a COVID-19 stimulus bill (for Democrats). Specifically, participants read that the public figure warned that, if the policy they were advocating for was not implemented, homelessness rates would rise by 30% in the next month and jobless claims would remain above one million per week.

As in Study 1, we manipulated the salience of reputational outcomes (high vs. low). While in Study 1 we operationalized reputation salience through whether President Biden received public credit (for making a correct private prediction), in Study 2 we manipulated reputational salience by who made the prediction: either a partisan leader (high reputation salience) or an economist (low reputation salience; no information about the economist's political leaning was provided or implied). For Republicans, the partisan leader was (then) Senate Majority Leader Mitch McConnell (a Republican). For Democrats, the partisan leader was House Majority Leader Nancy Pelosi (a Democrat). To ensure the reputational stakes were salient in the high reputation salience condition, we included language from an outgroup member implying that the ingroup figure's prediction would be wrong. In particular (depending on condition), McConnell's and Pelosi's concerns were dismissed by an outgroup Senator (Democrat Kristin Gillibrand or Republican Lindsey Graham, respectively) as merely "alarmist politics."

We again crossed reputation salience with another between-subjects manipulation: the valence of the material outcome that ultimately transpired (negative vs. positive economic news). In the negative economic news condition, participants read that a new report on the U.S. economy commissioned by the Federal Reserve confirmed that homelessness rates had increased by 30% and that the report cited data from the Labor Department indicating that jobless claims had persisted above one million per week. In the positive economic news condition, participants learned about the same report from the Federal Reserve (and data from the Labor Department), but that homelessness rates had remained stable and that jobless claims had fallen below one million per week.

Note that the valence of the material outcome also determined whether the public official's prediction was confirmed (reputational gain for the ingroup, in the high reputation salience condition) or contradicted (reputational loss for the ingroup, in the high reputation salience condition). Whereas the eventual material outcome cast no reputational shadow for participants in the low reputation salience condition (i.e., in which the prediction was made by an economist), this was not true for participants in the high reputation salience condition: a positive material outcome would imply relief from predicted suffering but hinder the reputational standing of their partisan leader (and, by extension, the ingroup); a negative material outcome would imply the reverse.

The primary outcome variable was participants' affective reactions to reading the news story. We measured affective reactions through a composite of the same five positive and five negative emotion items from Study 1 (see Method for details).

We had two hypotheses for this study. First, we expected people to again show a clear overall preference for good (vs. bad) material outcomes for society. We expected this to be very

clearly true in the case where the predictor was an economist and no reputation was at stake. Second, and of more central interest, was whether individuals would remain sensitive to reputation in a context with dramatic material consequences. We predicted that they would, and that this would manifest in an interaction in which for partisan predictors, the preference for good vs. bad material outcomes would be blunted because the good vs. bad material outcomes were in opposition to bad vs. good reputational outcomes for an ingroup leader.

We again found support for both hypotheses. First, when the predictor was an economist (i.e., the low reputation salience condition), participants reported a more positive overall response for positive material outcomes, and this effect was quite substantial in size, both when collapsing across both the Republican and Democrat sub-studies ($M_{\text{bad}} = -3.34$ vs. $M_{\text{good}} = 0.32$, $t(893) = 26.45$, $p < .001$, Cohen's $d = 1.72$) and when considering each of the Republican ($M_{\text{bad}} = -2.87$ vs. $M_{\text{good}} = 0.84$, $t(435) = 18.62$, $p < .001$, Cohen's $d = 1.75$) and Democrat ($M_{\text{bad}} = -3.37$ vs. $M_{\text{good}} = -0.15$, $t(451) = 19.75$, $p < .001$, Cohen's $d = 1.76$) sub-studies individually.

Second, and supporting our key theoretical prediction, we found that this preference for good over bad material outcomes was significantly attenuated when the predictor was a partisan (i.e., in the high reputation salience condition). Results are depicted in Figure 1. Collapsing across the two sub-studies yielded robust evidence for a significant interaction ($b = 0.81$, $se = 0.19$, $t = 4.24$, $p < .001$). In Study 2a (i.e., among Republicans), we found significant evidence to support this hypothesis ($b = 1.18$, $se = 0.26$, $t = 4.49$, $p < .001$) and in Study 2b (i.e., among Democrats) we found marginally significant evidence ($b = 0.44$, $se = 0.26$, $t = 1.71$, $p = .088$). Thus, reputational considerations blunted individuals' preferences for good over bad material outcomes for the world. That is, partisans let reputational considerations influence their (degree of) preference for good over bad material outcomes. Rather than simply focusing on whether or

not the policy had wrought economic destruction on society as a whole, their affective responses factored in how these respective material outcomes made their group look.

Of note, although we predicted a significant interaction effect, we also theorized that the effect size of the interaction would be relatively small in this specific experimental context. After all, the difference between material outcomes here was grave, representing the presence or absence of economic devastation amid a global pandemic. Although we thought it would be impressive if, even within this context, individuals' affective responses were shaped by reputation, we expected that we would still observe a more substantial effect of material outcome, with individuals overall preferring good over bad material outcomes in both the partisan and economist conditions. As depicted in Figure 1, this was in fact the case: although affective differentiation was smaller when reputation was on the line (vs. when it was not), individuals felt more positive about good versus bad material outcomes in both conditions. Collapsing across Studies 2a and 2b, the effect of material outcome was reduced by approximately 20% in high reputation salience condition relative to the low reputation salience condition (i.e., from $d = 1.72$ to $d = 1.42$), suggesting that despite a substantial effect of reputational considerations, the relative size of the effect of material outcomes was larger.



Figure 1. In the context of domestic economic policy, Republicans’ and Democrats’ preferences for good vs. bad material outcomes were significantly blunted when the good vs. bad material outcomes were yoked, respectively, to bad vs. good reputational outcomes for an ingroup leader (Study 2). Error bars represent 1 SE and colored dots represent raw data.

Study 3

Study 2 provided clear evidence that reputational considerations can blunt individuals’ preferences for good over bad material outcomes for the world. Still, at least in that specific experimental context, partisans’ affective reactions were driven more strongly by material outcomes than by reputational considerations.

However, this needn’t always be the case. Rather, individuals’ relative sensitivity to material and reputational outcomes is likely to depend on the size of material and reputational

stakes in the context in question. In Study 2, the material stakes were great, and the reputational stakes were substantially more modest. If the reputational stakes were higher (or the material stakes were lower), then the size of the effect of reputation on partisans' affective responses might have been more comparable to that of material outcomes. That is, in some contexts, individuals might even be just as happy with outcomes involving material harm (but reputational benefit) as they are with outcomes involving material benefit (but reputational harm). Study 3 (April 2021: Study 3a = 581 Republicans and Study 3b = 869 Democrats) explores these possibilities by examining the relative effects of material and reputational outcomes in a new context: national security.

In Study 3a, Republicans were randomly assigned to one of four between-subjects experimental conditions. In all conditions, participants read an ostensibly real news story in which they learned that a leading public figure had predicted that Iran's cyber capabilities were stagnating (i.e., a predicted positive material outcome for Americans, given the intergroup animosity between the United States and Iran). Specifically, they learned that, despite some internal differences in perspective among foreign policy analysts, a public figure had staked much of his foreign policy credibility on his view that Iran's cyber capabilities were stagnating. As in Studies 1-2, we manipulated both the salience of reputational outcomes (high vs. low) and the valence of the material outcome (negative vs. positive news about Iran's actual cyber capabilities). To manipulate reputation salience, we mimicked the design of Study 2 by manipulating the partisan identity of the predictor—in this case, either Democratic President Biden (high reputation salience) or CIA Director Burns (low reputation salience; Burns' partisan group membership was not mentioned). To manipulate material outcome, participants learned that a new, independent report either confirmed (good material outcome) or contradicted (bad

material outcome) the prediction regarding Iran’s stagnating capabilities. Specifically, participants learned that a report relying on a systematic assessment by U.S. intelligence sources and American operatives had found either that Iran’s cyber capabilities were either less (good material outcome) or more (bad material outcome) sophisticated than previously believed. Thus, relative reputational gain for Republicans (i.e., Biden looking bad) co-occurred with material harm for the country (i.e., Iran’s cyber capabilities growing more sophisticated than anticipated) and vice-versa. The primary outcome variable was participants’ affective reactions to reading the news story, which we measured in two ways: with a single global evaluation and with a combined index of ten specific emotion items (five positive and five negative, see Method for details).

While now in the context of national security rather than domestic economic policy, the design parallels Study 2 in which participants read a news story in which a public figure makes a public prediction with reputational consequences. One other design feature merits note. In Study 3a specifically, we sought to provide a stringent test of our theorizing about the role of reputation by further raising the countervailing material stakes. In particular, we intentionally amplified participants’ perceptions of Biden’s (or CIA Director Burns’) involvement in national security. That is, when participants learned that the prediction was inaccurate, they further learned that this raised question marks about general foreign policy competence, increasing the national security risk to the United States. Consequently, the more desirable reputational outcome for Republicans (i.e., Biden looking bad) involved even greater material downsides: not only were Iran’s cyber capabilities more sophisticated, but the U.S. faced the additional material cost of reduced competence in handling related threats from China and Russia. The reverse was also true (i.e., when the prediction was accurate, it reflected well on general competence and ability to

handle related threats). With these heightened material stakes, any effects of reputation would be still more impressive.

As in Study 2, we first examined the effect of material outcomes in the low reputation conditions. As predicted, participants in the conditions involving low reputation salience responded much more positively to good vs. bad material outcomes, for both the single general affect item ($M_{\text{bad material}} = 33.74$ vs. $M_{\text{good material}} = 64.77$, $t(280) = 13.17$, $p < .001$, Cohen's $d = 1.55$) and the 10-item composite ($M_{\text{bad material}} = -2.15$ vs. $M_{\text{good material}} = 1.79$, $t(284) = 16.55$, $p < .001$, Cohen's $d = 1.95$).

As in Study 2, we theorized that this preference for the positive over negative material outcomes would be significantly attenuated when the positive outcome helped an outparty partisan leader's (in this case, President Biden's) reputation. As depicted in Figure 2, this prediction was confirmed: we found significant evidence that reputational gains blunted affective responses to bad material outcomes for both the single general affect item ($b = 10.01$, $se = 3.55$, $t = 2.82$, $p = .0049$) and for the combined index ($b = 0.97$, $se = 0.34$, $t = 2.86$, $p = .0044$). Of note, it represented an approximately 30-40% reduction in the gap in affective responses between low vs. high threat to the United States from Iran (Single item: $d_{\text{low reputation salience}} = 1.55$ vs. $d_{\text{high reputation salience}} = 0.93$; Composite: $d_{\text{low reputation salience}} = 1.95$ vs. $d_{\text{high reputation salience}} = 1.43$).

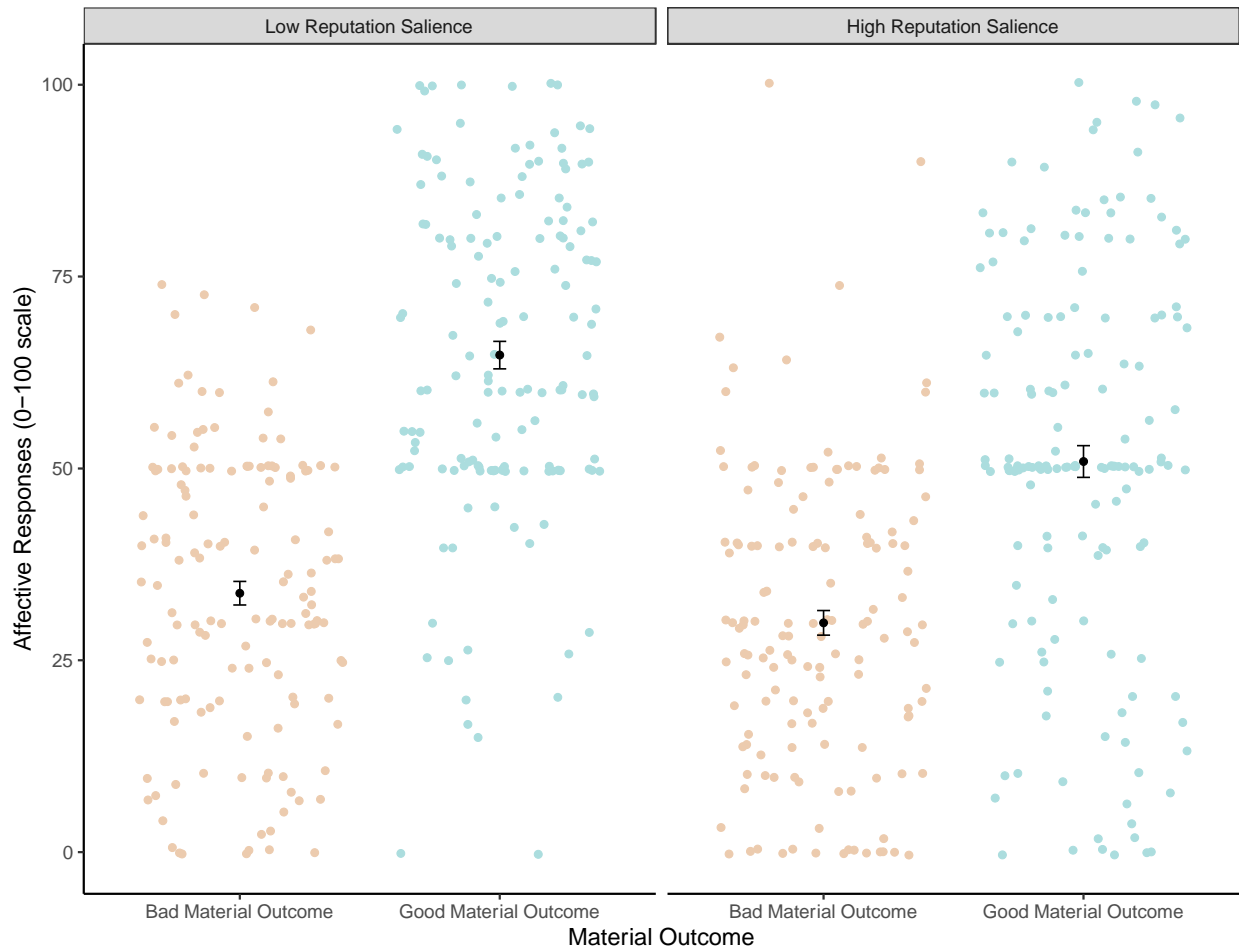


Figure 2. In the context of national security, Republicans’ preferences for good vs. bad material outcomes were blunted when the good vs. bad material outcomes were yoked, respectively, to bad vs. good reputational outcomes for an outgroup leader (i.e., President Biden) (Study 3a). Error bars represent 1 SE and colored dots represent raw data.

Study 3b was a parallel study with Democrats. Democrats also read an ostensibly real news story in which a public figure made a prediction regarding Iran’s cyber capabilities. Participants were randomly assigned to one of six between-subjects experimental conditions: four primary conditions that closely followed the conditions from Study 3a and two supplemental conditions (described later).

In the four primary conditions, Democrats were randomly assigned in a 2 (reputation salience: high vs. low) x 2 (material outcome: positive vs. negative) fully between-subjects

design. The predictors were again either President Biden (here, an ingroup member) or CIA Director Burns and the prediction regarding Iran's cyber capabilities was either contradicted or confirmed. In contrast to Study 3a, in Study 3b the public figure predicted that Iran's cyber capabilities were *growing more sophisticated* (rather than stagnating). Thus, reputational gains for Democrats in the high reputation salience condition (i.e., Biden looking good) again co-occurred with material harm (i.e., Iran's growing cyber capabilities) and vice versa. The primary outcome variable was again participants' affective reactions to reading the news story, which we measured in the same two ways as in Study 3a: with a single global evaluation and with a combined index of ten specific emotion items (five positive and five negative, see Method for details).

We again began by examining the effect of material outcomes in the low reputation conditions. As predicted, participants in the low reputation salience conditions reported a more positive affective response when the pessimistic prediction was contradicted (i.e., good material outcome: Iran's cyber capabilities not growing more sophisticated) compared to when the prediction was confirmed (i.e., bad material outcome: Iran's cyber capabilities growing more sophisticated). This was true both for responses regarding general global affect ($M_{\text{bad material}} = 38.91$ vs. $M_{\text{good material}} = 57.52$, $t(292) = 8.07$, $p < .001$, Cohen's $d = 0.94$) and on the affect composite ($M_{\text{bad material}} = -1.79$ vs. $M_{\text{good material}} = 0.87$, $t(291) = 11.65$, $p < .001$, Cohen's $d = 1.36$).

We next tested our focal interaction prediction. While we hypothesized that participants would report a more positive overall response for positive material outcomes in the low reputation salience condition, we theorized that this effect would be attenuated (and potentially even eliminated) when the positive material outcome hurt an ingroup leader's (in this case,

President Biden's) reputation. This was in fact the case: analyses yielded strong evidence for an interaction between reputation salience and material outcomes for both the single general affect item ($b = 16.75$, $se = 3.09$, $t = 5.43$, $p < .001$) and the collapsed 10-item scale ($b = 1.52$, $se = 0.31$, $t = 4.85$, $p < .001$).

As depicted in Figure 3, the reputation effect was quite substantial in magnitude. That is, relative to the low reputation conditions, in the high reputation salience conditions Democrats reported more similar affective responses when the material outcome was negative (but the reputational outcome was positive) vs. when the material outcome was positive (but the reputational outcome was negative). Indeed, this represented an approximately 55% (composite measure) to 90% (single item slider) reduction in the gap in affective responses between low vs. high threat to the United States from Iran (Single item: $d_{\text{low reputation salience}} = 0.94$ vs. $d_{\text{high reputation salience}} = 0.11$; Composite: $d_{\text{low reputation salience}} = 1.36$ vs. $d_{\text{high reputation salience}} = 0.62$). Of note, for the single item measure, responses were *nearly identical*—and were not significantly different—when the material outcome was negative (but the reputational effect for Biden was positive) and when the material outcome was positive (but the reputational effect for Biden was negative).

In addition to the four conditions described above, we also included two additional conditions aimed at addressing a potential alternative explanation. We theorized that—despite material upsides—Democrats would be less happy when Biden (vs. a CIA director with unknown partisan membership) made a pessimistic prediction that turned out to be inaccurate because the inaccurate prediction would have negative effects on Democrats' reputation. Alternatively, however, it could be the case that this interaction is driven less by group-based reputational considerations per se than by reducing confidence in Biden's general competence with respect to foreign policy and ability to handle related threats (e.g., from China and Russia).

That is, participants might view Biden's incorrect prediction as reflecting negatively on the U.S.'s ability to handle foreign policy challenges effectively. If so, then their negative affect might be due not to reputation but to a calculation that, as a result, the U.S. would now face greater material challenges (despite the good material news about Iran's weaker-than-predicted capabilities).⁴

To address this potential alternative explanation, we needed to separate the reputational effects of Biden's prediction from the broader material consequences of his (in)accuracy. To do this, we included in our randomized design two additional conditions that were identical to the high reputation salience conditions described above, with one exception. Specifically, the article participants read included additional text indicating that while citizens often attend to and care about the president's views on foreign policy (i.e., that reputation is at stake), the reality behind the scenes is that he is not in fact closely involved in the day-to-day operations, which are largely handled by the CIA and military officials (i.e., that the president's inaccuracy has less impact on the U.S.'s material ability to manage foreign policy challenges than might be expected). With this design, the text including Biden's prediction would still reveal materially impactful information about Iran's capabilities, but would no longer provide a basis for drawing broader material conclusions about U.S. foreign policy competence. We theorized that if we again obtained a comparable interaction even when we replaced the two high reputation salience conditions described above with these two supplemental conditions, then it would yield greater confidence that the interaction is driven by reputational considerations.

⁴ Of note, in Study 3a we aimed to amplify Biden's involvement in national security in order to provide a more stringent test of our theorizing (because any reputational gain for Republicans was coupled with additional material harm in the form of his incompetence at handling related threats). Here, we took the inverse approach—downplaying Biden's involvement in national security—in order to de-couple reputational benefits for Biden from material benefits in the form of competence of handling related threats.

As depicted in Figure 3, we find evidence to support this theorizing. Specifically, when replacing the two primary high reputation salience conditions with the two supplemental Biden conditions, we again find significant evidence that the reputation effect attenuates the effect of material outcomes for both the single general affect item ($b = 15.55$, $se = 3.27$, $t = 4.76$, $p < .001$) and for the combined index ($b = 1.21$, $se = 0.33$, $t = 3.71$, $p < .001$). Of note, the interaction with the primary conditions was almost equal in magnitude to the interaction with the supplemental conditions (Single item: $b_{\text{primary}} = 16.75$ vs. $b_{\text{supplemental}} = 15.55$; Composite: $b_{\text{primary}} = 1.52$ vs. $b_{\text{supplemental}} = 1.21$), suggesting that implications for general competency in handling other foreign threats were not the primary driver of the interaction.

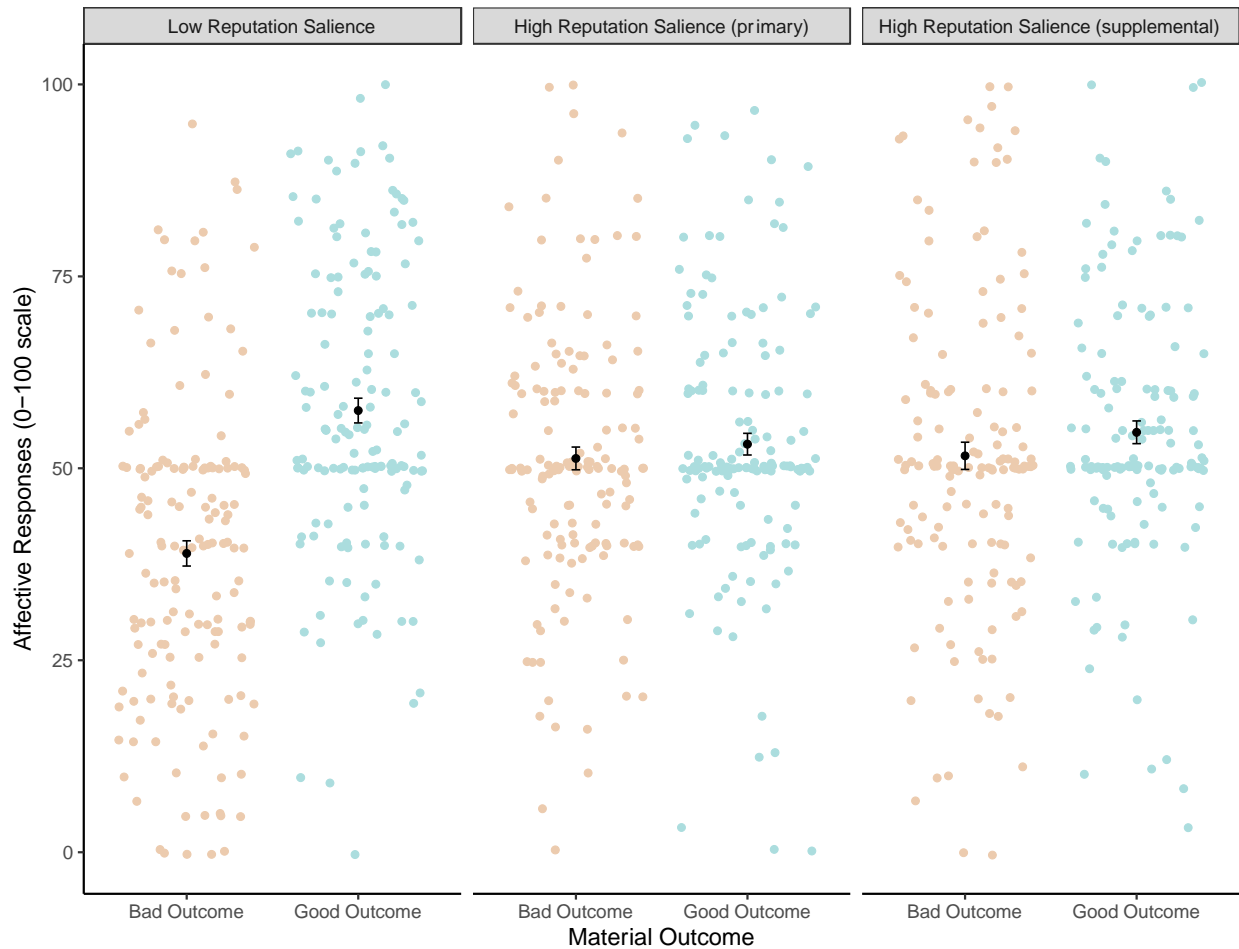


Figure 3. In the context of national security, Democrats’ preferences for good vs. bad material outcomes were effectively eliminated when the good vs. bad material outcomes were yoked, respectively, to bad vs. good reputational outcomes for an ingroup leader (i.e., President Biden). This pattern of results held in two supplemental conditions which addressed an alternative explanation based on non-reputational considerations (Study 3b; see main text for details). Error bars represent 1 SE and colored dots represent raw data.

Taken together, the results from Study 3 not only replicated the effects of reputation identified in Study 2, but also extended Study 2 by demonstrating that, in certain cases, reputational stakes can matter almost as much as material outcomes. Of note, the fact that the relative effect of material versus reputational considerations was larger among Republicans (where material concerns regarding Iran’s threat co-occurred with the additional material risk of presidential incompetence) than among Democrats (where material concerns about Iran occurred

absent concerns about presidential incompetence) suggests that the relative size of material versus reputational considerations in shaping affect may depend on the relative balance of material versus reputational stakes. Thus, in assessing the practical consequentiality of the fact that reputation shapes responses to materially good versus bad outcomes, it is worth considering not only the effect size of reputation, but also the gravity of the material considerations in question. Although Study 3 starkly highlights that reputation can matter just as much material outcomes in some cases, even a relatively small effect of reputation in shaping responses to material outcomes is noteworthy when the difference between good and bad material outcomes in question is stark.

Study 4

Our studies thus far suggest the power of reputational considerations in shaping responses to material outcomes. In Study 4, however, we sought to address an alternative possibility that could account for some of the prior results (albeit not our results from Study 1): that partisans are deriving pleasure from having their anticipated view of the world confirmed, rather than reputational gains relative to the outgroup (e.g., Kaiser, Vick, & Major, 2004; Major, Kaiser, O'Brien, & McCoy, 2007). That is, perhaps Democrats, for example, feel better when Pelosi or Biden (vs. an economist or CIA Director) is right about a pessimistic prediction not because of their sensitivity to Democrats' reputation viz. Republicans but simply because Pelosi or Biden being right about things is more consistent with their ideological worldview. To disentangle the relative contributions of reputation and worldview confirmation, we thus needed to examine a context in which reputational gain for the ingroup occurs as a consequence of events that disconfirm one's worldview. We designed Study 4 (July 2021: Study 4 = 926 Republicans) to address worldview confirmation as a possible alternative explanation.

In Study 4, Republicans were randomly assigned to one of four between-subjects experimental conditions. Across conditions, participants read an ostensibly real news story (adapted from Study 1) in which they learned that a leading public figure had made a private, pessimistic prediction about a new domestic economic policy. Specifically, participants read that the leading public figure had privately noted reservations about using targeted efforts to reduce unemployment (i.e., an approach advocated in the new policy). As in Studies 2-3, we manipulated both the salience of the reputational outcome (high vs. low) and the valence of the material outcome (negative vs. positive effects of the policy) in a fully between-subjects design. To manipulate reputation salience, the predictor was either Democratic President Biden (high reputation salience) or Council of Economic Advisors Chair Rouse⁵ (low reputation salience). To manipulate the material outcome, participants learned that a new report found either that the policy had reduced unemployment and efficiently allocated resources (good material outcome) or had no effect on unemployment and inefficiently allocated resources (bad material outcome). Specifically, participants read either that over one million families had received nutritional assistance that they badly needed (good material outcome) or that there had been almost no reduction in the number of families receiving nutritional assistance that they badly needed (bad material outcome).

As in Study 1, the story made clear in all conditions that while the policy was put into effect while the public figure was in office, the policy was in fact developed and implemented independently of the public figures (and of their private opinions about the policy's likely effectiveness). Still, the text noted that the public figure was credited or blamed with the policy's

⁵ While Chair Cecilia Rouse is in fact female, we gave her a male-gendered name (Charlie) to avoid introducing a confound for gender in the study. Of the 500 participants who were in the low reputation salience condition (and thus read about Rouse), none commented on the name change in an open-ended text box at the end of the study.

success or failure by the public (i.e., reputation was always at stake for the public figure; note though that this impinged on group-based reputational considerations for the participant only when Biden was the public figure in question).

The key conceptual change in Study 4 compared to Studies 2-3 was the dissociation between worldview confirmation and reputational incentives. Unlike Studies 2-3, Study 4 created a contrast between the correctness of the private prediction and the public perception. Specifically, in the good material outcome conditions, the leader made an incorrect private prediction (i.e., incorrectly assuming the policy would be a failure), but was nevertheless publicly credited for the success of the policy; in the bad material outcome conditions, the leader made a correct private prediction (i.e., correctly assuming the policy would be a failure), but was nevertheless publicly blamed for the policy. Republicans' worldview should be confirmed to a greater extent in the good material/bad reputational outcome condition (in which Biden made an incorrect prediction, the policy was effective, and Biden was publicly credited for it) compared to the bad material/good reputational outcome condition (in which Biden made a correct prediction, the policy was ineffective, and Biden was publicly blamed for it).⁶ Thus, Republican participants experience reputational *gain* when their worldview is *disconfirmed* and vice versa, disentangling worldview confirmation from reputational gain.

Our analysis plan proceeded in lockstep with Studies 2-3, with the overarching hypothesis that a similar pattern of results would emerge despite the fact that any effect of worldview confirmation would oppose any effect of reputation. Despite this potential oppositional force, a near-identical pattern of results emerged. First, participants in the

⁶ Note that we refer to the perspective of the Republican participant when we say “good” and “bad” reputational outcome (rather than the perspective of the public figure). That is, a good group-based reputational outcome for a Republican is when Biden suffers reputational damage, and vice-versa.

conditions involving low reputation salience responded much more positively to good vs. bad material outcomes, for both the single general affect item ($M_{\text{bad material}} = 38.69$ vs. $M_{\text{good material}} = 51.86$, $t(459) = 6.66$, $p < .001$, Cohen's $d = 0.62$) and the 10-item composite ($M_{\text{bad material}} = -1.61$ vs. $M_{\text{good material}} = 0.50$, $t(459) = 10.86$, $p < .001$, Cohen's $d = 1.01$). However, and more critically, this preference for the positive over negative material outcomes was significantly attenuated when the positive outcome helped an outparty partisan leader's (in this case, again President Biden's) reputation, for both the single general affect item ($b = 11.54$, $se = 2.86$, $t = 4.03$, $p < .001$) and the 10-item composite ($b = 1.25$, $se = 0.28$, $t = 4.53$, $p < .001$).

Of note, and as depicted in Figure 4, the reputation effect was again quite substantial in magnitude, despite the potential oppositional force of worldview confirmation: it represented an approximately 60-90% reduction in the gap in affective responses between low vs. high threat to the United States from Iran (Single item: $d_{\text{low reputation salience}} = 0.62$ vs. $d_{\text{high reputation salience}} = 0.07$; Composite: $d_{\text{low reputation salience}} = 1.01$ vs. $d_{\text{high reputation salience}} = 0.40$). Indeed, in the high reputation salience condition, Republicans reported nearly identical affective responses for the general affect item when the policy was a failure (but Biden was blamed: Mean = 39.16) and when the policy was a success (but Biden was credited: Mean = 40.79).

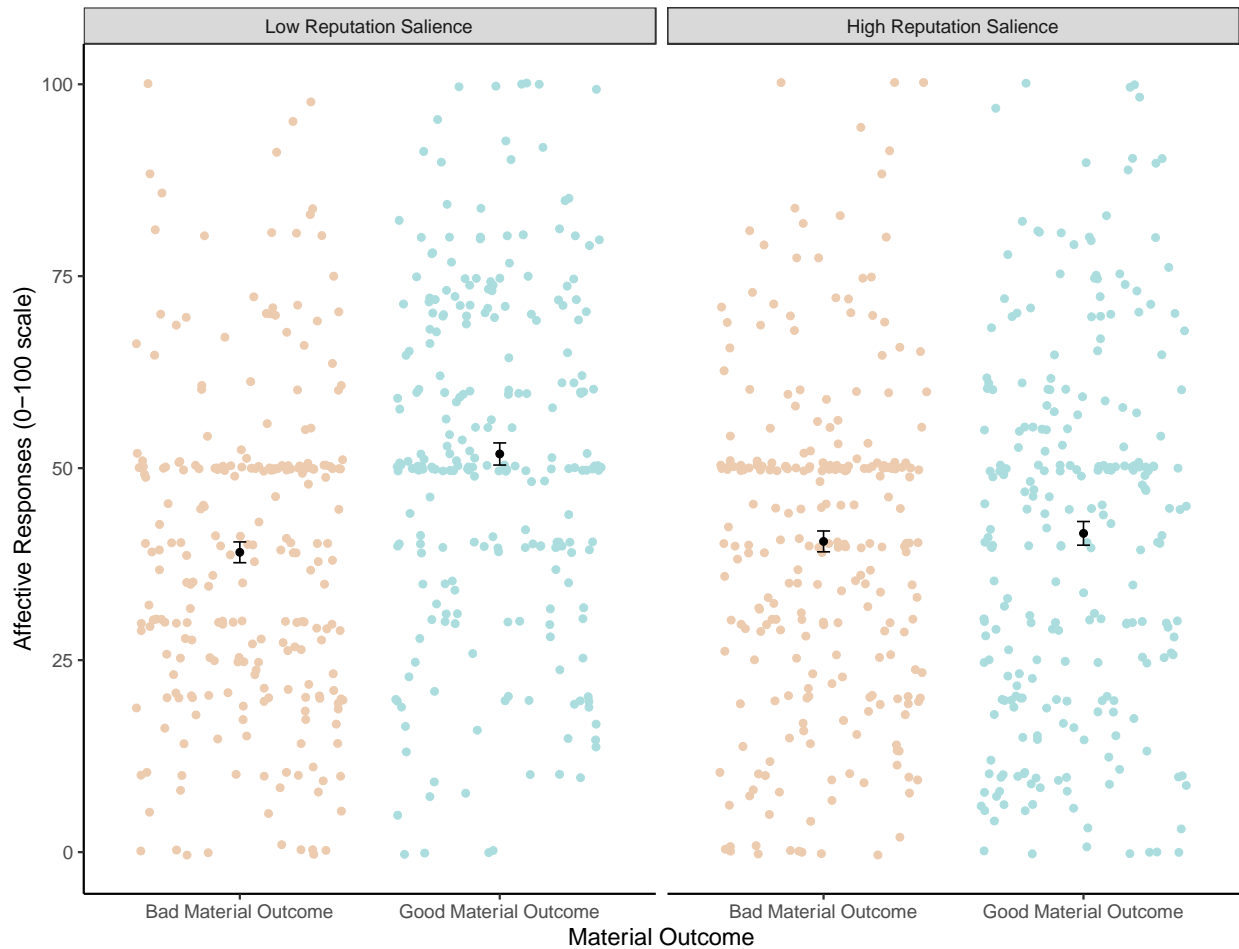


Figure 4. In the context of national security, Republicans’ preferences for good vs. bad material outcomes were eliminated when the good vs. bad material outcomes were yoked, respectively, to bad vs. good reputational outcomes for an outgroup leader (i.e., President Biden). This pattern held despite the fact that any positive affective consequences of reputational gain could, in principle, have been counteracted by negative affective consequences of worldview disconfirmation (Study 4). Error bars represent 1 SE and colored dots represent raw data.

Study 5

Studies 1-4 provide robust evidence that reputational stakes shape partisans’ affective responses to material suffering. Strikingly, reputational stakes sometimes even eliminated affective preferences for positive over negative material outcomes. That being said, in our prior studies, participants were responding after the fact to events that had already taken place.

Whereas participants might feel better about bad material outcomes for the world when it helps

their group's reputation, will they also *act* in ways that help entrench material suffering on account of group-based reputational considerations? In Study 5 (N = 1,192 Democrats), we use a pre-registered, incentive-compatible design to test whether partisans are less willing to actively intervene to stop material suffering when doing so would yield reputational gains for the outgroup.

Democrats were randomly assigned to one of two between-subjects experimental conditions (low vs. high reputation salience). Democrats randomly assigned to the low reputation salience condition read an ostensibly real news story in which they learned that Americares was a charity focused on helping with ground logistics for distributing a COVID-19 vaccine. Democrats randomly assigned to the high reputation salience condition read the same story, but also learned that President Trump (a Republican) would be credited with the logistical success of vaccine distribution. Specifically, the article's headline mentioned that President Trump was eager to take credit for logistics and that President Trump had publicly stated that people would look back at vaccine logistics and praise him for the job he did.

Participants then had the opportunity to donate their study bonus to Americares. We hypothesized that participants would be less likely to donate their bonus in the high reputation salience condition because they would be hesitant to contribute to a cause that could eventually shed a positive light on the Republican effort to stem coronavirus. This hypothesis was supported: Democrats were significantly (albeit slightly) less likely to donate their study bonus in the high reputation salience condition (M = 51%) compared to the low reputation salience condition (M = 57%), $b = 0.24$, $se = .12$, $t(1116) = 2.03$, $p = .043$; this difference represents an approximately 12% relative reduction in probability of donating to a cause that could dramatically relieve material suffering on the basis of averting reputational gain to the outgroup.

General Discussion

In contexts marked by intense group competition, attending to the reputational standing of the ingroup relative to the outgroup can take on critical importance. But can reputational considerations blunt even partisans' preference for avoiding widespread material suffering?

Five pre-registered experiments reveal that partisans in the United States are more willing to tolerate widespread material harm—even that also borne by ingroup members—when it serves the ingroup's reputational standing relative to the outgroup. We provide evidence (1) among both Democrats and Republicans; (2) across both affective responses and (financially-incentivized) behavior; (3) both when reacting to past harm and evaluating opportunities to stop future harm; and (4) after ruling out explanations based on worldview confirmation. Strikingly, incentives for reputational standing can weaken—and sometimes even eliminate—preferences for positive vs. negative material outcomes. Moreover, our work suggests that individuals might sometimes forgo opportunities to avert material suffering if avoiding harm comes at the expense of burnishing the outgroup's reputation.

Our work makes multiple theoretical contributions. First, our work sheds new light on a psychological mechanism reinforcing group-based conflict. Prior work in intergroup relations focuses on processes such as intergroup animus, dehumanization, and negative meta-perceptions as important contributors to cycles of conflict and associated material suffering (e.g., Kteily, Hodson, & Bruneau, 2016; Lees & Cikara, 2020). Here, we highlight that individuals may sometimes let the world 'burn' in order to gain a reputational upper-hand over the outgroup. By focusing more on making the other side look bad than on averting material suffering, partisans may play into the other side's cynical narratives about their disregard for the sanctity of human life, damaging potential bridges to cooperation and increasing the outgroup's conviction in the

rightness of its conflict. When two sides are more focused on being right than seeking good, prospects for peace diminish.

Relatedly, our work contributes to research considering individuals' enjoyment of harm to others and their willingness to engage in taboo tradeoffs. Indeed, while both our work and prior work on schadenfreude investigate contexts in which partisans in conflict experience positive affect in response to material harm (e.g., Leach, Spears, Branscombe, & Doosje, 2003), our work makes at least three key advances over this prior body of research. First, on a conceptual level, prior research on schadenfreude examines harm that is concentrated to outgroup members. In the present research, we examine such dynamics when the material harm is widespread (*including* the ingroup). Indeed, in a further supplemental study (described in the *SI*), we provide evidence that such effects persist even when the material harm is concentrated further still, targeting only populations the participant's ingroup traditionally prioritizes (i.e., low income and minority populations, for Democrats). Second, on a methodological level, prior research on schadenfreude has traditionally focused on situations in which the material suffering is relatively minor, such as in an outgroup political leader falling off a bicycle or an outgroup losing a competitive sports game (for a notable exception, see Combs et al., 2009). In the present work, we examine contexts with dramatic material suffering in the form of economic devastation and national security threats to the United States. Third, prior work has defined schadenfreude as "the pleasure of passively seeing others suffer" (Leach, Spears, Branscombe, & Doosje, 2003; see also Nietzsche 1887/1967). In contrast, we examined contexts in which participants had the option to actively intervene to avert material harm from occurring (Study 5). Finally, while not a direct test of taboo trade-offs, our work suggests that partisans are at least somewhat willing to

tolerate widespread material harm to others when doing so carries concomitant reputational gains.

Finally, our work also extends research on the power of reputation to drive judgment and decision making in group contexts. Prior work highlights that individual-level reputational considerations can substantially shape behavior (e.g., Dorison, Umphres, & Lerner, 2021; Jordan, Hoffman, Nowak, & Rand, 2016; Schlenker & Weigold, 1992; Lerner & Tetlock, 1999; Tetlock, 2002). Previous research also suggests that individuals are sensitive to their group's reputation, for example expressing concern over maintaining the ingroup's moral standing in the eyes of others (Noor, Shnabel, Halabi, & Nadler, 2012; Rothschild, Landau, Molina, Branscombe, & Sullivan, 2013). However, whereas prior work has often focused on the ways that groups act prosocially in order to maintain or restore their reputation (e.g., by offering intergroup apologies; Wohl, Hornsey, & Bennett, 2012), our work illuminates new contours of the darker side of sensitivity to group-based reputation (see also Saguy & Kteily, 2011).

Limitations and future directions

Despite the robustness of our findings across different scenarios and dependent variables, a few key limitations and future directions of the present work merit note. First, in most cases, participants attended more heavily to material outcomes than to reputational outcomes. While our work suggests that reputation can matter substantially even when the material stakes are quite high, we do not claim that reputational outcomes are the only (or even the most important) factor in driving partisans' reactions to outcomes in the world. Still, even a relatively small effect of reputation in shaping responses to material outcomes is noteworthy when the difference between good and bad material outcomes in question is stark.

That being said, our work did reveal that reputational considerations can sometimes have dramatic effects, factoring in almost as much as material outcomes. For example, in Study 3b, Democrats' affective responses were almost equally driven by whether the outcome improved vs. undermined Biden's reputational standing as by whether Iran's cyber capabilities were accelerating vs. stagnating. Additionally, in Study 4, Republicans reported nearly identical affective responses when an economic policy was a failure (but President Biden was blamed) compared to when an economic policy was a success (but President Biden was credited).

Indeed, our theorizing suggests that, in certain contexts, reputational considerations might matter even more than material outcomes. Our scenarios tended to feature highly consequential material outcomes. But, when material stakes are sufficiently small and/or reputational stakes sufficiently large, partisans might in fact be happier in scenarios involving worse material outcomes (but better reputational ones) versus those involving better material outcomes (but worse reputational ones). We conducted a second supplemental study as a proof-of-concept test of this theorizing. In a study conducted in the weeks following President Biden's inauguration, we found support for this corollary: Republicans reported greater positive affect when a Biden supporter injured an innocent store owner more severely (but the incident was covered in the media) compared to when the same Biden supporter injured an innocent store owner less severely (but the incident was not covered) (see *SI* for details). Future research could consider more formally assessing the magnitude of reputational and material effects on individuals' affect and behavior at varying levels of reputational and material stakes.

Second, one alternative possibility that could account for some of the results is that partisans are deriving pleasure from having their anticipated view of the world confirmed (e.g., Kaisor, Vick, & Major, 2004; Major, Kaiser, O'Brien, & McCoy, 2007). That is, perhaps

Democrats feel better when Pelosi (vs. an economist) is right about a pessimistic prediction not because of their sensitivity to Democrats' reputation but simply because Pelosi being right about things is more consistent with their ideological worldview. Whereas this alternative cannot be entirely ruled out for some of our studies, it is difficult to apply to our results from several others. For example, explanations based on worldview confirmation may struggle to explain the pattern of results in Study 4, where worldview confirmation was designed to be in direct opposition to the reputational incentives. Further, worldview confirmation is traditionally studied as responses to events that have already occurred; however, in Study 5, we examined a context in which partisans made active, future-oriented choices not to intervene to stop harm from occurring. Still, future work assessing tradeoffs between reputational and material outcomes would benefit from considering the role of worldview confirmation even more systematically.

Third, while our findings were robust across a variety of methodological features (e.g., Democrats and Republicans, affective responses and financially-incentivized behavior), it is worth considering how other methodological factors may amplify vs. attenuate sensitivity to reputational considerations. One factor worth considering more systematically is the extent to which individuals factor reputational considerations similarly when faced with contexts involving material harm versus material benefit: If a good material outcome is paired with a reputational loss and a bad material outcome is paired with an equivalently-sized reputational gain, are participants equally moved by reputation under both scenarios?

Although we did not set up our studies to focus specifically on this potential symmetry, our results suggest that reputation can matter in both contexts, but might matter even more in the context of good versus bad material outcomes. In Studies 1, 2, and 3b reputational effects were approximately equal in the context of good and bad material outcomes. In Study 2, for example,

the degree to which reputational benefit counteracted participants' sadness at widespread economic suffering was comparable to the degree to which reputational benefit diminished participants' *happiness* at widespread economic benefit. But reputational considerations appeared stronger in the context of good (vs. bad) material outcomes in Studies 3a and 4; in Study 3a, for example, reputational loss (i.e., Biden vs. a non-partisan CIA director being wrong) did more to stifle Republicans' happiness about Iran's weaker-than-predicted cyber capabilities than reputational gain did to buffer their sadness about Iran's stronger-than-predicted cyber capabilities. Future research should explore this potential asymmetry further, as well as consider whether participants differentiate between reputational gain to the ingroup and reputational loss to the outgroup (or between reputational loss to the ingroup and gain to the outgroup).

Along similar lines, prior research on moral judgment has identified key differences in harms of omission vs. commission (e.g., Ritov & Baron, 1999; Spranca, Minsk, & Baron, 1991). In the present studies, while we did test whether reputational incentives would decrease willingness to intervene to stop material harm (i.e., omission; Study 5), we did not test whether reputational incentives would increase willingness to intervene to actively *create* material harm (i.e., commission). Future research could explore whether partisans are even willing to proactively bring about widespread material suffering when group-based reputational incentives are on the line.

Fourth, with the exception of Study 5, our studies directly asked participants to respond to purportedly real articles. While the present studies prioritized internal validity, it would be useful to complement this approach with observational data collected outside of a controlled laboratory setting. For example, one such potential study could examine emotional/linguistic cues on social media during conflicts to examine whether reputationally-damaging actions by the

outgroup elicit positivity, even when this comes at the expense of ingroup suffering. Future work could also go beyond the context of partisanship in the United States to examine other contexts where similar dynamics might be at play, such as the Israeli-Palestinian conflict or sectarian conflict in Iraq.

Finally, while the present studies tested sensitivity to reputational stakes in the context of widespread material harm, future research could examine the interpersonal and intergroup consequences of this sensitivity. For example, how do participants evaluate co-partisans who are vs. are not willing to prioritize reputational gain over material suffering? Further, how much do individuals think the other side attends to reputational over material outcomes, and are these (meta-) perceptions accurate? If they are in fact overly pessimistic—such that people falsely assume the other side is more willing to see the world suffer in order to gain reputation than they actually are—might we be able to correct their meta-perceptions and thereby reduce intergroup animosity? Related research on false polarization and group-based meta-perceptions suggest that this could be a fruitful avenue for intervention (e.g., Dorison, Minson, & Rogers, 2019; Lees & Cikara, 2020; Moore-Berg, Ankori-Karlinsky, Hameriri, & Bruneau, 2020; Ruggeri et al., 2021).

Conclusion

Affectively-charged conflicts can impede cooperation, creating situations in which group members attend heavily to their group's relative standing. We highlight the critical role of group-based reputational incentives in reinforcing conflict by decreasing attention to material suffering, thus contributing to a vicious cycle of conflict escalation.

Method

All studies were pre-registered, conducted on Amazon Mechanical Turk, and were approved by Northwestern University's Institutional Review Board (IRB number

STU00213032). Due to space constraints, full methodological details for all studies—including full text for all vignettes—are available in the Supplementary Information (*SI*).

Sample sizes were set in advance and determined by power analyses derived from pilot studies. Participants always first gave informed consent, answered an attention check (for full text, see *SI*), and completed a set of demographics (age, gender, and political party identification). To measure party identification, we followed the elicitation used by Pew Research Center (Pew Research Center, 2015). Specifically, we first asked participants whether they considered themselves more of a Democrat or a Republican. If the participant indicated Independent, we then asked the following question: “As of today, do you lean more Democrat or Republican?” After the party identification question, participants completed a variety of exploratory individual difference measures, including ingroup glorification and ingroup attachment (Roccas, Klar, & Liviatan, 2006).

Study 1. We conducted a study with a sample of 1016 Democrats ($M_{age} = 40.55$, $SD = 12.75$, 47% female). As pre-registered, we included only the 981 Democrats (96.6%) who passed an attention check and did not ask to have their data removed.

Participants learned that they would read an ostensibly real news story and answer questions about it on the following pages. Participants were randomly assigned to one of four between-subjects experimental conditions in a 2 (good material outcome vs. bad material outcome) x 2 (low reputation salience vs. high reputation salience) design.

All participants read a news story in which President Biden made a correct private prediction about a new economic policy that had passed all necessary bureaucratic obstacles before he came into office and did not require presidential approval. To manipulate the material outcome, participants learned that a new report found either that the policy had reduced

unemployment and efficiently allocated resources (good material outcome) or had no effect on unemployment and inefficiently allocated resources (bad material outcome). In both cases, Biden's private prediction was correct (i.e., he made an optimistic prediction in the good material outcome condition and made a pessimistic prediction in the bad material outcome condition). To manipulate the reputational outcome, participants learned either that President Biden had received public credit in the news media for his correct private prediction (high reputation) or had not received public credit for his correct private prediction (low reputation). Specifically, participants read either that media reporting had emphasized his private correct prediction regarding the policy and that public perception of Biden (as reflected in public polling) had improved as a result (high reputation) or that media reporting had emphasized his lack of involvement and that public perception of Biden had remained unchanged as a result (low reputation).

The primary dependent variable was self-reported affect, which was measured in two ways in a counterbalanced order. First, we measured it using a single item of overall positivity vs. negativity on a 100-point sliding scale anchored at "Extremely negative" and "Extremely positive." Second, we measured it using a composite of ten specific emotion items. All items were measured on a 7-point Likert scale anchored at "Not at all" and "Very much." We measured positive affect as the average of five items (happy, relieved, enthusiastic, glad, excited). We also measured negative affect as the average of five items (upset, sad, distressed, concerned, uneasy). We created a combined index by subtracting the average of the negative affect scores from the average of the positive affect scores. Thus, zero indicated equal levels of positive and negative affect, positive scores indicated more positive experiences, and negative

scores indicated more negative experiences. The index achieved a high level of reliability ($\alpha = .90$ after reverse-scoring negative items).

Study 2. We conducted a study with a sample of 991 Republicans (Study 2a: $Mage = 42.07$, $SD = 12.96$, 51% female) and 1023 Democrats (Study 2b: $Mage = 38.10$, $SD = 12.00$, 53% female). As pre-registered, we included only the 915 Republicans (92.3%) and 973 Democrats (95.1%) who passed an attention check and did not ask to have their data removed.

As in Study 1, participants learned that they would read an ostensibly real news story and answer questions about it on the following pages. Participants were randomly assigned to one of four between-subjects experimental conditions in a 2 (low reputation salience vs. high reputation salience) x 2 (good material outcome vs. bad material outcome) design.

All participants read a news story in which a public figure made a pessimistic prediction, suggesting that there would be major economic damage if a policy they were advocating for was not implemented (i.e., speedy reopening by states of coronavirus-shuttered businesses for Republicans; new government stimulus package for Democrats). In the low reputation salience condition, the public figure was a group of leading economists. In the high reputation salience condition, the public figure was either (then) Senate Majority Leader Mitch McConnell (for Republicans) or House Speaker Nancy Pelosi (for Democrats). Participants read that the public figure warned that, if the policy they were advocating for was not implemented, homelessness rates would rise by 30% in the next month and jobless claims would remain above one million per week. In the low reputation salience condition, there were no reputational consequences for the ingroup if the (economist's) prediction was right or wrong; in contrast, in the high reputational salience condition, the partisan leader being right or wrong about the policy would reflect well or poorly on the ingroup. To amplify the reputational stakes in the high reputation

saliency condition, we included language from the outgroup implying that the ingroup figure's prediction would be wrong. In particular (depending on condition), McConnell and Pelosi's concerns were dismissed by an outgroup Senator (Democrat Kristin Gillibrand or Republican Lindsey Graham, respectively) as merely "alarmist politics."

In the negative material outcome condition, the negative forecast came to pass. Specifically, participants read that a new report on the U.S. economy commissioned by the Federal Reserve confirmed that homelessness rates had increased by 30% and that the report cited data from the Labor Department indicating that jobless claims had persisted above one million per week. The article thus concluded that the public figure's forecasts were spot on. In the positive material outcome condition, participants learned that the forecasts were off base. Specifically, participants learned about the same report from the Federal Reserve (and data from the Labor Department), but that homelessness rates had remained stable and that jobless claims had fallen below one million per week. Of note, we did not directly mention reputational consequences for the predictor; we assumed that these would be implied.

Participants then answered questions regarding their affective reactions to the news story. These answers constituted the primary dependent variable in this study. The affect index was composed of answers to the same ten questions used in Study 1, although this study did not include the single-item general affect item. The index achieved a high level of reliability (alphas = .90 for both studies after reverse-scoring negative items).

Study 3. We conducted a study with a sample of 903 Democrats (Study 3a: $M_{age} = 40.05$, $SD = 12.60$, 50% female) and 600 Republicans (Study 3b: $M_{age} = 43.95$, $SD = 13.12$, 49% female). As pre-registered, we included only the 869 Democrats (96.2%) and 581 Republicans (96.8%) who passed an attention check and did not ask to have their data removed.

As in Studies 1-2, participants learned that they would read an ostensibly real news story and answer questions about it on the following pages. Republicans in Study 3a were randomly assigned to one of four between-subjects experimental conditions in a 2 (material outcome: good vs. bad) x 2 (reputation: high vs. low) design. Democrats in Study 3b were randomly assigned to one of six between-subjects experimental conditions. The randomized design was identical to Study 3a, with the addition of two supplementary high reputation salience conditions (described further in the main text). Greater detail regarding the experimental manipulations and procedure are provided in the main text above and complete text for all news stories are available in the *SI*.

Republicans and Democrats answered the same questions regarding their affective reactions to the news story from Study 1, including both the single item of overall positivity vs. negativity and the ten-item composite, in a counterbalanced order. The composite again achieved a high level of reliability (alphas = .91 and .84 for Studies 3a and 3b, respectively, after reverse-scoring negative items).

Study 4. We conducted a study with a sample of 973 Republicans ($M_{age} = 44.31$, $SD = 12.84$, 52% female). As pre-registered, we included only the 926 Republicans (95.2%) who passed an attention check and did not ask to have their data removed.

As in prior studies, participants learned that they would read an ostensibly real news story and answer questions about it on the following pages. Participants were randomly assigned to one of four between-subjects experimental conditions in a 2 (low reputation salience vs. high reputation salience) x 2 (good material outcome vs. bad material outcome) design.

Study 4 was an adapted version of Study 1 for Democrats and to address alternative explanations based on worldview confirmation. As in Study 1, participants read a news story regarding a new economic policy that had passed all necessary bureaucratic obstacles before the

public figure came into office, and thus did not require approval. To manipulate reputation salience, the predictor was either Democratic President Biden (high reputation salience) or Council of Economic Advisors Chair Rouse (low reputation salience). To manipulate the material outcome, participants learned that a new report found either that the policy had reduced unemployment and efficiently allocated resources (good material outcome) or had no effect on unemployment and inefficiently allocated resources (bad material outcome).

In Study 4, the public figure always made a pessimistic prediction about the policy. Thus, when the material outcome was positive, this meant that he was privately incorrect; in contrast, when the material outcome was negative, this meant that he was privately correct.

Whereas in Study 1 Biden was either publicly credited or received a neutral reputational outcome, here the reputational outcome was either positive (reputational credit) or negative (reputational blame). Specifically, in the negative material outcome conditions (in which Biden/Rouse privately, and correctly, predicted that the policy would be a failure), participants read that (1) the blame for the policy has been overwhelmingly attributed to Biden/Rouse; (2) media reporting has emphasized that the policy came to fruition during his tenure; and (3) as a consequence, public opinion polling suggests that public perception of him has worsened as a result. In contrast, in the positive material outcome conditions (in which Biden/Rouse privately, and incorrectly, predicted that the policy would be a failure), participants read that (1) the credit for the policy has been overwhelmingly attributed to Biden/Rouse; (2) media reporting has emphasized that the policy came to fruition during his tenure; and (3) as a consequence, public opinion polling suggests that public perception of him has improved as a result.

As described in the main text, the set-up of the vignette thus created a tension between worldview confirmation (i.e., Republicans' beliefs that Biden would make an incorrect

prediction) and reputational stakes (i.e., Republicans' incentives for Biden to look bad in the eyes of the general public).

Affect was measured in two ways, identically to Studies 1 and 3. The affect index again achieved a high level of reliability ($\alpha = .85$ after reverse-scoring negative items).

Study 5. We conducted a study with a sample of 1192 Democrats ($M_{age} = 40.03$, $SD = 12.79$, 54% female). As pre-registered, we included only the 1117 participants (93.7%) who passed an attention check and did not opt (post-debriefing) to have their data removed.

As in prior studies, participants learned that they would read an ostensibly real news story and answer questions about it on the following pages. Participants were randomly assigned to one of two between-subjects experimental conditions (low reputation salience vs. high reputation salience). In the low reputation salience condition, participants then read an article with the following headline: "Multiple priorities considered for COVID-19 vaccine strategy, including marketing and logistics." The article (1) described how a vaccine could provide hope to the United States after widespread harm from COVID-19; (2) noted that there were two competing priorities: marketing and ground logistics; and (3) described Americares as an organization focused on helping with ground logistics. Participants in the high reputation salience condition read the identical news story, except the headline included the additional sentence "President Trump eager to take credit for logistics" and the article included a final paragraph that indicated that in a recent interview, President Trump had said that, as President, his job was to oversee the logistics operation and ensure its success and that "people are going to look back and say Trump did something that had never ever been done before." We wrote this additional paragraph to raise the reputational stakes (here, an outgroup President receiving acclaim) for Democrats.

On the following page, participants then learned that for participating in the survey, they were eligible to receive a \$.50 bonus. They also learned that they had the opportunity to donate the bonus (or any part of it) to Americares in order to help with COVID-19 vaccine logistics. The choice of whether to donate the bonus (and how much) served as the primary dependent variable in this study. Of note, participants kept any money they did not donate.

References

- Abramowitz, A. I., & Webster, S. (2016). The rise of negative partisanship and the nationalization of US elections in the 21st century. *Electoral Studies*, 41, 12-22.
- Amodio, D. M., & Cikara, M. (2021). The social neuroscience of prejudice. *Annual Review of Psychology*, 72, 439-469.
- Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, 2(3), 412-414.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497.
- Cassese, E. C. (2021). Partisan Dehumanization in American Politics. *Political Behavior*, 43(1).
- Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and them: Intergroup failures of empathy. *Current Directions in Psychological Science*, 20(3), 149-153.
- Cikara, M., & Fiske, S. T. (2012). Stereotypes and schadenfreude: Affective and physiological markers of pleasure at outgroup misfortunes. *Social Psychological and Personality Science*, 3(1), 63-71.
- Combs, D. J., Powell, C. A., Schurtz, D. R., & Smith, R. H. (2009). Politics, schadenfreude, and ingroup identification: The sometimes happy thing about a poor economy and death. *Journal of Experimental Social Psychology*, 45(4), 635-646.
- Corasaniti, N. & Rutenberg, J. (2021). Republicans pushed to restrict voting. Millions of Americans pushed back. *New York Times*.
<https://www.nytimes.com/2020/12/05/us/politics/2020-election-turnout.html>. Accessed 30 August 2021.
- Davidai, S., & Ongis, M. (2019). The politics of zero-sum thinking: The relationship between political ideology and the belief that life is a zero-sum game. *Science Advances*, 5(12), eaay3761.
- De Waal, F. B. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annual Review of Psychology*, 59, 279-300.
- Dorison, C. A., Minson, J. A., & Rogers, T. (2019). Selective exposure partly relies on faulty affective forecasts. *Cognition*, 188, 98-107.
- Dorison, C.A., Umphres, C., & Lerner, J.S. (2021). Staying the course: decision makers who escalate commitment are trusted and trustworthy. *Journal of Experimental Psychology: General*.
- Druckman, J. N., & Levendusky, M. S. (2019). What do we measure when we measure affective polarization?. *Public Opinion Quarterly*, 83(1), 114-122.
- Esses, V. M. (2021). Prejudice and discrimination toward immigrants. *Annual Review of Psychology*, 72, 503-531.
- Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., ... & Druckman, J. N. (2020). Political sectarianism in America. *Science*, 370(6516), 533-536.
- Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science*, 11, 563-588.

- Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: an evolutionary analysis and empirical review. *Psychological Bulletin*, 136(3), 351.
- Greene, J. D. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. Penguin.
- Hare, B. (2017). Survival of the friendliest: Homo sapiens evolved via selection for prosociality. *Annual Review of Psychology*, 68, 155-186.
- Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, 72, 207-240.
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual review of psychology*, 53(1), 575-604.
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22, 129-146.
- Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology a social identity perspective on polarization. *Public Opinion Quarterly*, 76(3), 405-431.
- Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, 59(3), 690-707.
- Jordan, J. J., Hoffman, M., Bloom, P., & Rand, D. G. (2016). Third-party punishment as a costly signal of trustworthiness. *Nature*, 530(7591), 473-476.
- Jordan, J. J., Hoffman, M., Nowak, M. A., & Rand, D. G. (2016). Uncalculating cooperation is used to signal trustworthiness. *Proceedings of the National Academy of Sciences*, 113(31), 8658-8663.
- Jordan, J. J., & Rand, D. G. (2020). Signaling when no one is watching: A reputation heuristics account of outrage and punishment in one-shot anonymous interactions. *Journal of Personality and Social Psychology*, 118(1), 57.
- Kaiser, C. R., Vick, S. B., & Major, B. (2004). A prospective investigation of the relationship between just-world beliefs and the desire for revenge after September 11, 2001. *Psychological Science*, 15(7), 503-506.
- Kalin, M., & Sambanis, N. (2018). How to think about social identity. *Annual Review of Political Science*, 21, 239-257.
- Kingsley, P. (2021). Israel gives vaccine to far-off allies, as Palestinians wait. *New York Times*. <https://www.nytimes.com/2021/02/23/world/middleeast/israel-palestinians-vaccine-diplomacy.html>. Accessed 30 August 2021.
- Kteily, N., Bruneau, E., Waytz, A., & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology*, 109(5), 901.
- Kteily, N., Hodson, G., & Bruneau, E. (2016). They see us as less than human: Metadehumanization predicts intergroup conflict via reciprocal dehumanization. *Journal of Personality and Social Psychology*, 110(3), 343.
- Kudesia, R. S. (2021). Emergent Strategy from Spontaneous Anger: Crowd Dynamics in the First 48 Hours of the Ferguson Shooting. *Organization Science*.

- Kurzban, R., Burton-Chellew, M. N., & West, S. A. (2015). The evolution of altruism in humans. *Annual Review of Psychology*, *66*, 575-599.
- Leach, C. W., Spears, R., Branscombe, N. R., & Doosje, B. (2003). Malicious pleasure: Schadenfreude at the suffering of another group. *Journal of Personality and Social Psychology*, *84*(5), 932.
- Leach, C. W., & Spears, R. (2008). "A vengefulness of the impotent": The pain of in-group inferiority and schadenfreude toward successful out-groups. *Journal of Personality and Social Psychology*, *95*(6), 1383.
- Lees, J., & Cikara, M. (2020). Inaccurate group meta-perceptions drive negative out-group attributions in competitive contexts. *Nature Human Behaviour*, *4*(3), 279-286.
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*(2), 255.
- Major, B., Kaiser, C. R., O'Brien, L. T., & McCoy, S. K. (2007). Perceived discrimination as worldview threat or worldview confirmation: Implications for self-esteem. *Journal of Personality and Social Psychology*, *92*(6), 1068.
- Martherus, J. L., Martinez, A. G., Piff, P. K., & Theodoridis, A. G. (2021). Party animals? Extreme partisan polarization and dehumanization. *Political Behavior*, *43*(2), 517-540.
- McCarty, N., Poole, K. T., & Rosenthal, H. (2016). *Polarized America: The dance of ideology and unequal riches*. MIT Press.
- Moore-Berg, S. L., Ankori-Karlinisky, L. O., Hameiri, B., & Bruneau, E. (2020). Exaggerated meta-perceptions predict intergroup hostility between American political partisans. *Proceedings of the National Academy of Sciences*, *117*(26), 14864-14872.
- Nietzsche, F. (1887/1967). *On the genealogy of morals*. New York: Random House.
- Noor, M., Shnabel, N., Halabi, S., & Nadler, A. (2012). When suffering begets suffering: The psychology of competitive victimhood between adversarial groups in violent conflicts. *Personality and Social Psychology Review*, *16*(4), 351-374.
- Pew Research Center (2015). Appendix A: Why partisans include 'leaners.'
<https://www.pewresearch.org/politics/2015/11/23/appendix-a-why-partisans-include-leaners/>. Accessed 30 August 2021.
- Rasgon, A. (2021). Israel's vaccine success unleashes debate on Palestinian inequities. *New York Times*. <https://www.nytimes.com/2021/02/04/world/middleeast/israel-palestinians-vaccine.html>. Accessed 30 August 2021.
- Richeson, J. A., & Sommers, S. R. (2016). Toward a social psychology of race and race relations for the twenty-first century. *Annual Review of Psychology*, *67*, 439-463.
- Ritov, I., & Baron, J. (1999). Protected values and omission bias. *Organizational Behavior and Human Decision Processes*, *79*(2), 79-94.
- Roccas, S., Klar, Y., & Liviatan, I. (2006). The paradox of group-based guilt: modes of national identification, conflict vehemence, and reactions to the in-group's moral violations. *Journal of Personality and Social Psychology*, *91*(4), 698.
- Rothschild, Z. K., Landau, M. J., Molina, L. E., Branscombe, N. R., & Sullivan, D. (2013).

- Displacing blame over the ingroup's harming of a disadvantaged group can fuel moral outrage at a third-party scapegoat. *Journal of Experimental Social Psychology*, 49(5), 898-906.
- Ruggeri, K., Većkalov, B., Bojanić, L., Andersen, T. L., Ashcroft-Jones, S., Ayacaxli, N., ... & Folke, T. (2021). The general fault in our fault lines. *Nature Human Behaviour*, 1-11.
- Saguy, T., & Kteily, N. (2011). Inside the opponent's head: Perceived losses in group position predict accuracy in metaperceptions between groups. *Psychological Science*, 22(7), 951-958.
- Schlenker, B. R., & Weigold, M. F. (1992). Interpersonal processes involving impression regulation and management. *Annual Review of Psychology*, 43(1), 133-168.
- Shamir, J. (2007). Public opinion in the Israeli-Palestinian Conflict: From Geneva to Disengagement to Kadima to Hamas. *United States Institute of Peace*.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. Available at SSRN 2160588.
- Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, 27(1), 76-105.
- Svolik, M. W. (2019). Polarization versus democracy. *Journal of Democracy*, 30(3), 20-32.
- Tetlock, P. E. (2000). Cognitive biases and organizational correctives: Do both disease and cure depend on the politics of the beholder?. *Administrative Science Quarterly*, 45(2), 293-326.
- Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, 78(5), 853.
- Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: intuitive politicians, theologians, and prosecutors. *Psychological Review*, 109(3), 451.
- Wohl, M. J., Hornsey, M. J., & Bennett, S. H. (2012). Why group apologies succeed and fail: Intergroup forgiveness and the role of primary and secondary emotions. *Journal of Personality and Social Psychology*, 102(2), 306.