AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2024, Vol. 153, No. 10, 2605–2625
https://doi.org/10.1037/xge0001645

Group-Based Reputational Incentives Can Blunt Sensitivity to Societal Harms and Benefits

Charles A. Dorison[1] and Nour S. Kteily[2]

[1] Management Area, McDonough School of Business, Georgetown University

[2] Department of Management and Organizations, Kellogg School of Management, Northwestern University

People's concern with maintaining their individual reputation powerfully drives judgment and decision making. But humans also identify strongly with groups. Concerns about group-based reputation may similarly shape people's psychology, perhaps especially in contexts where shifts in group reputation can have strategic consequences. Do individuals allow their concern with their group's reputation to shape their reactions to even large-scale societal suffering versus benefits? Examining both affective responses and financially incentivized behavior of partisans in the United States, five preregistered experiments (N = 7,534) demonstrate that group-based reputational incentives can weaken—and sometimes nearly eliminate—affective differentiation between present-term societal harms and benefits. This can occur even when these societal harms and benefits are substantial—including economic devastation and national security threats—and when the consequences impact *ingroup* members. Individuals' sensitivity to group-based reputation can even cause them to divert resources from more effective to less effective charities. We provide evidence that partisans care about group-based reputation in part because it holds strategic value, positioning their group to improve its standing vis-a-vis the outgroup. By allowing group-based reputational incentives to reduce their sensitivity to societal outcomes, partisans may play into the other side's cynical narratives about their disregard for human suffering, damaging bridges to cooperation.

Public Significance Statement

When groups get locked into competition, they can become highly focused on outcompeting the other side, with costly societal ramifications. In the context of Democrats and Republicans in the United States, we find that partisans are sometimes so attentive to how events make their group look vis-a-vis the outgroup (i.e., their group-based reputation) that they become less sensitive to how events impinge on present-term societal harms and benefits. This can include economic devastation, national security threats, and health care emergencies that harm the ingroup. This occurs at least partly because group-based reputational incentives hold strategic value for gaining power. Concerns with group-based reputation can even cause individuals to act in ways that bring about suboptimal societal impact if doing so avoids burnishing the outgroup's reputation. When two sides are focused on looking good rather than seeking good, prospects for cooperation diminish.

Keywords: conflict, groups, reputation, judgment and decision making, affect

Supplemental materials: https://doi.org/10.1037/xge0001645.supp

Humans are exquisitely attuned to their reputations. Some researchers even liken people to "intuitive politicians," navigating the world by managing the impressions they make on the audiences they most seek to curry favor with (Tetlock, 2002). Within psychology, most prior research on reputation focuses on individual-level reputation (Baumeister & Leary, 1995; Berman & Silver, 2022; Dorison & Heller, 2022; Jordan, 2023; Lerner & Tetlock, 1999; Schlenker & Weigold, 1992). But humans are social animals who

This article was published Online First August 29, 2024.

Rachael E. Jack served as action editor.

Charles A. Dorison https://orcid.org/0000-0002-7072-2530

Charles A. Dorison conducted data analysis. Both authors contributed to all other aspects of the research and approved the final version of the article. All data, code, materials, and preregistrations are publicly available at ResearchBox and can be accessed at https://researchbox.org/667. This work was presented at the New England Managerial Decision Making Symposium. Otherwise, it was not distributed prior to publication. The authors declare no competing interests or conflicts.

Charles A. Dorison played a lead role in data curation, formal analysis, project administration, visualization, writing–original draft, and writing–review and editing and an equal role in investigation and methodology. Nour S. Kteily played a lead role in funding acquisition and supervision, a supporting role in data curation, formal analysis, visualization, and writing–original draft, and an equal role in conceptualization, investigation, methodology, and writing–review and editing.

Correspondence concerning this article should be addressed to Charles A. Dorison, Management Area, McDonough School of Business, Georgetown University, 3700 O Street, NW Washington, DC 20057, United States. Email: Charles.dorison@georgetown.edu

belong to, identify with, and derive value from groups (Cikara et al., 2011); just as we care about our own reputations, so too do we care about the image of the groups with which we are closely identified (Cialdini et al., 1976; Snyder et al., 1986). Although understudied, a concern with the reputation of one's group can be a force for good, for example, leading people to avoid engaging in immoral behavior that reflects poorly on their group (e.g., Doosje et al., 1998) or to act in conciliatory or restorative ways to improve the ingroup's image (Kardos et al., 2019; Lickel et al., 2011; Nadler & Shnabel, 2015). But might group-based reputational incentives sometimes conflict with—and potentially offset—other valued considerations? If so, why? And just how much motivational force might they hold?

Here, we build in important ways on prior research focused on emotional responses to negative events that make the outgroup look bad (Combs et al., 2009). Across several large preregistered studies, we investigate whether group-based reputational concerns can blunt sensitivity to societal harms and benefits that both occur in the present term and that impact ingroup members (cf. Bandura, 1990; Schein & Gray, 2015, 2018).[1] To do so, we assess partisans' reactions to harms and benefits when group-based reputational incentives are (vs. are not) at stake. This allows us to conduct the first assessment of how group-based reputational incentives shape *affective differentiation*: the degree to which people feel better about events that create societal benefit over harm. Examining both affective responses—and, critically, behavior—we test when and why group-based reputational incentives can lead partisans to feel better when people suffer but it makes their group look good, or feel worse when people are spared of suffering, but it makes their group look bad. We also conduct tests of the underlying mechanism, providing novel evidence about the extent to which individuals' sensitivity to present-term societal harms versus gains may be driven by a sensitivity to the strategic consequences of group-based reputation. To the extent that partisans allow concerns with their group's reputation to shape how they react to events that impart immediate suffering or gains to society, they 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 righteousness of its conflict.

Can Group-Based Reputational Incentives Reduce Sensitivity to Societal Outcomes?

On the one hand, given the fundamental aversion to hurting others (Schein & Gray, 2015, 2018), group-based reputational incentives could lack the motivational force necessary to blunt partisans' affective differentiation between present-term events that result in societal harm versus benefit, perhaps especially when these harms and benefits are large-scale. Tolerating widespread harm—especially when those bearing the brunt of the harm include the ingroup—might be a bridge too far, no matter the implications for the ingroup's reputational standing. Individuals might consider it gauche to factor in the implications for the ingroup's reputation when they are faced with news that implies a significant setback or upside to society. For example, it would likely have felt "un-American" for a Republican to let any happiness at the increased approval rating for Republican President George W. Bush dilute their sadness after the terrorist attacks on September 11, 2001. Similarly, partisans may be loath to let reputational considerations blunt their joy about advances that generate widespread and immediate societal benefit. For example, it might have felt uncouth for Democrats to feel less positively about a successful COVID-19 vaccine because it was developed under Republican President Donald Trump. Indeed, research suggests that individuals treat certain domains as morally sacred and impervious to instrumental considerations (Tetlock et al., 2000). Consistent with this idea, in Supplemental Experiment 1, 89% of participants thought an individual who attended solely to the prevention of societal suffering was more moral than an individual who also attended to group-based reputational incentives (see Supplemental Material for details).

On the other hand, there are good reasons to posit that group-based reputational incentives *will* blunt sensitivity to present societal outcomes, either by increasing tolerance for societal suffering, decreasing appreciation for societal benefits, or both. A positive reputation can provide an important strategic advantage for a group and its members in competitive contexts, helping to deliver valued resources like power and legitimacy. Consider the conflict between Israelis and Palestinians, in which the two groups compete for third-party support over the legitimacy of their respective territorial claims (Noor et al., 2012; Shamir, 2007). Actions that substantially tarnish Israel's international reputation—for example, being seen to withhold scarce COVID-19 vaccines from Palestinians in favor of Israelis (Kingsley, 2021; Rasgon, 2021)—could have significant and rewarding strategic benefits for Palestinians (Saguy & Kteily, 2011). By bolstering the perceived legitimacy of Palestinians' grievances among the international community, Palestinians could obtain long-term value in terms of winning allies in pursuit of the effort to pressure Israel to make political concessions. Along similar lines, victories in the court of public opinion could allow Democrats or Republicans to persuade independent voters to vote for their preferred presidential candidate, thereby gaining an electoral edge over their counterpart.

Importantly, while gaining the reputational upper hand and attaining power and legitimacy can feel rewarding in its own right—it could simply feel good to "win" and to have more rather than less power—there are additional, more prosocial, reasons why groups might prioritize group-based reputation when it can impact their strategic standing. For example, gaining the reputational upper hand can also better position a group to enact policies it believes are better for the world in the long run. Viewed from this perspective, partisans who respond to reputational incentives by becoming less likely to differentiate between social harm and benefit today may be sacrificing short-term societal harm for what they believe to be long-term societal benefit.

We acknowledge that, in our studies, we cannot fully rule out the possibility that individuals may be allowing reputational incentives to blunt their affective differentiation between short-term societal harm and benefits purely because they deliberate on the long-term societal benefits that could accrue from their group's enhanced standing.[2] Nevertheless, even if some partisans allow reputational considerations to blunt their affective differentiation between present-term societal harms and benefits for potentially laudable

[1] As we return later, we separate certain societal harm and benefit in the present term from more uncertain projections about a given outcome's potential longer term societal ramifications.

[2] Note that, in supplemental analyses in Study 4, we do investigate open-ended response data to preliminarily examine whether individuals naturalistically mention deliberating on group-based reputational incentives and their impact on their group's strategic standing.

considerations of long-term societal benefits, this psychology could still have negative ramifications. For one, individuals' calculus about long-term societal benefit may be misguided, and their expectation that uncertain long-term societal benefits will outweigh guaranteed present-term societal harms may not come to pass. Indeed, individuals' ingroup biases may cause them to overestimate the extent to which their group's enhanced standing will actually be better for society, and even if they are right to believe that their group's vision for society would bring about net societal benefits, they may also run into roadblocks in implementing that vision. But beyond whether or not individuals' underlying reasons for allowing group-based reputational incentives to blunt their differentiation between present-term societal harms and gains are prosocial, this psychology might still contribute to intergroup hostility. After all, outgroups are more likely to simply see how a group responds to present-term societal harms and gains than they are to have access to the (potentially prosocial) underlying reasons. A group seen by outgroups to tolerate present societal suffering (or disapprove of societal progress) could, in the process, reinforce any sinister views held of them by those outgroups, perpetuating intergroup hostilities. We return to this issue in the General Discussion section.

The strongest support for the hypothesis that group-based reputational incentives can blunt differentiation between societal harm and gain comes from a series of clever studies conducted by Combs and colleagues (Combs et al., 2009). Using samples of U.S. undergraduates, the researchers revealed that participants reported the positive feeling of schadenfreude in response to harmful societal outcomes (e.g., troop deaths, economic harm) when those same outcomes reflected poorly on the opposition party. For example, in one study, participants belonging to both parties reported greater negative affect when reading an article detailing a presidential candidate of their party failing to stop a housing and economic crisis compared to when it was the presidential candidate of the other party, presumably because it would hurt their favored candidate in the upcoming presidential election. In other studies, too, Combs and colleagues provided valuable evidence that political gain for one's party can trigger positive affect (here, schadenfreude) even amid significant societal harm (e.g., an economic downturn).

Still, this early work leaves open consequential questions, several of which we take up in our article. For one, we generalize beyond the set of contexts that Combs and colleagues considered and test our effects among larger samples. We also consider generalizability to positive societal outcomes, that is, do individuals also allow reputational considerations with strategic implications to dampen their happiness about events that bring about positive near-term outcomes for society? Or are such effects limited only to interpretations of negative events of the type that Combs and colleagues considered?

We also shed light on at least one process underlying the affective consequences of societal events that impact the image of the ingroup. Combs and colleagues assumed that partisans' sense of schadenfreude at societal harm brought about by an outgroup was due to potential ingroup advantage, but they did not measure participants' perceptions of how societal harms impacted their group's strategic standing. We assess this psychological process by both measuring and manipulating the perceived strategic relevance of group-based reputation. Beyond *ruling in* perceived impact on their group's strategic standing as a process linking reputation to the blunting of differentiation between present-term societal harms and gain, we are also able to *rule out* several alternative explanations that were not considered in prior work. For example, participants may simply have felt better when the outgroup fumbled (at the expense of harm to society) versus succeeded (with the benefit of gains to society) not because of strategic gain for their ingroup but because outgroup failures confirmed their preexisting worldview about ingroup superiority/outgroup inferiority. We also address related alternative explanations, showing that our effects cannot be explained by mere generalized dislike of the outgroup or by conclusions drawn about group leaders' competence. Finally, we consider not only whether group-based reputational considerations impact individuals' affective reactions to societal events but further whether they have the motivational force necessary to impact behavior. In particular, we test whether partisans allow group-based reputational considerations to cause them to act in ways that have a suboptimal impact on society in the present term (i.e., by diverting resources from a more effective to a less effective charity).

Present Research

We investigate whether group-based reputational incentives can blunt partisans' affective differentiation between present societal harm and benefit. In doing so, we provide a proof-of-principle test of the impact of group-based reputation on feelings about societal harm/gain rather than seeking to make general claims about the precise relative importance of group-based reputation and societal harm/gain broadly writ. Still, across experiments, we investigate whether group-based reputational incentives *have no effect on* individuals' affective differentiation between societal harm and gain, *weaken* differentiation, or even *eliminate* differentiation, stress testing our hypothesizing across a broad range of scenarios, including scenarios that vary in the starkness of the societal stakes. Building on past research that began to investigate how attention to strategic consequences influences affective reactions to societal harm (Combs et al., 2009), the present research allows for a direct test of how much attention to reputation "closes the gap" in reactions to negative *and positive* societal outcomes and whether such effects extend to behavior.

Our experiments proceed as follows. In the context of an economic stimulus package at a local peak in the COVID-19 pandemic, Experiments 1A and 1B provide an initial test of whether group-based reputational incentives reduce affective differentiation between societal harm and benefit for Republicans and Democrats, respectively. Experiments 2A, 2B, and 3 replicate and extend Experiment 1 in a new context (national security) while also addressing alternative explanations. Experiment 4 investigates the extent to which attention to strategic benefits underpins the effect of reputational incentives on affective differentiation between societal harm and benefit. Finally, Experiment 5 turns from affective responses to financially incentivized behavior, assessing whether partisans act in ways that have a suboptimal present-term societal impact on account of group-based reputational incentives.

We also conducted five supplemental experiments. Supplemental Experiment 1 (mentioned above) reveals that participants evaluate an individual who attended solely to the prevention of societal suffering as more moral than an individual who also attended to group-based reputational incentives. Supplemental Experiments 2–3 provide manipulation checks for Experiments 1–3, confirming their internal validity. Supplemental Experiment 4 replicated Experiment 4 with a revised measure of the mediator. Finally, Supplemental Experiment 5

provides a proof-of-concept demonstration that, in cases where relative reputational (vs. societal) stakes are sufficiently large, reputation might not merely *impact* individuals' affective experience of societal harms and gains but even cause them to feel overall *happier* in scenarios involving worse societal outcomes (but better reputational ones) versus those involving better societal outcomes (but worse reputational ones). Full details for all supplemental experiments are available in the Supplemental Material. In addition, Supplemental Experiment 4 is described in detail in the Experiment 4 Results section, and Supplemental Experiment 5 is mentioned in the General Discussion section.

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in all experiments (Simmons et al., 2012). For all experiments, we first conducted pilot studies to generate estimated effect sizes. We then conducted a priori power analyses for each experiment and chose sample sizes that ensured at least 90% statistical power to detect effect sizes obtained in the relevant pilot. In doing so, we recruited a diverse sample of adults in the United States and included at least 200 participants per experimental condition in all experiments. All experiments were conducted on Amazon Mechanical Turk and were preregistered on https://aspredicted.org (Logg & Dorison, 2021). All data, materials, preregistrations, and analysis code can be found on Researchbox.com and can be accessed at https://researchbox.org/667. This research was approved by the host university's institutional review board (Number: STU00213032). We included a detailed debriefing procedure at the end of studies in which participants read fictional news stories.

Experiments 1A and 1B

Experiments 1A and 1B begin to test whether group-based reputational incentives can reduce affective differentiation between present societal harms and benefits for Republicans and Democrats, respectively. To do so, we investigate whether partisans feel better about societal harms when it makes their group look good and vice versa (i.e., feel worse about societal benefits that make one's group look bad).

Method

The preregistrations for Experiments 1A and 1B are available at https://aspredicted.org/YXM_MPC and at https://aspredicted.org/RF5_XFS.

Participants

We conducted an experiment with a sample of 991 Republicans (Experiment 1A: $M_{age} = 42.07$, $SD = 12.96$) and 1,023 Democrats (Experiment 1B: $M_{age} = 38.10$, $SD = 12.00$). In Experiment 1A, 506 participants reported their gender as female, 482 as male, two as other/prefer not to disclose, and one did not report their gender. In Experiment 1B, 540 participants reported their gender as female, 472 as male, 10 as other/prefer not to disclose, and one did not report their gender. As preregistered, we included only the 915 Republicans (92.3%) and 973 Democrats (95.1%) who passed an attention check before random assignment and did not ask to have their data removed. Experiment 1 was conducted in August 2020 at the time of the first spending bill aimed at addressing widespread unemployment at the onset of the COVID-19 pandemic (Cochrane, 2020).

Procedure

In all experiments, participants always first gave informed consent, answered a simple attention check before random assignment (for full text, see Supplemental Material), and completed a set of demographics (age, gender, and political party identification). To measure party identification, we followed the procedure used by 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 et al., 2006; for full text, see online materials).

After completing this opening set of questions, participants learned that they would read an ostensibly real news story and answer questions about it on the following pages. Full text for all news stories is available in the Supplemental Material. Participants were randomly assigned to one of four between-subjects experimental conditions in a 2 (Societal Benefit vs. Societal Harm) × 2 (Low-Reputation Salience vs. High-Reputation Salience) design.

All participants read a news story in which a public figure made a pessimistic prediction, suggesting that there would be major economic harm 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 whose party affiliation was not mentioned. 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 1 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 salience 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 societal harm 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 1 million per week. The article thus concluded that the public figure's forecasts were spot on. In the societal benefit 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 read that homelessness rates had remained stable and that jobless claims had fallen below 1 million per week. We took a conservative approach, not directly mentioning reputational consequences for the predictor or their impact on their group's strategic standing. Based on our theorizing about individuals' sensitivity to group-based reputational standing, we reasoned that perceivers would nevertheless pick up on reputational consequences and their potential strategic impact.

One additional point regarding the experimental manipulation merits note. In this and all experiments, participants responded to societal gains (vs. harms) that we expected people across the political spectrum to agree were worthwhile. While in many cases competing groups want different outcomes (e.g., legalized status of abortion), here, we purposefully selected outcomes about which there was bipartisan consensus about the end (minimizing economic suffering) even when there might be differences in opinion about the optimal means to achieve it. Supplemental Experiments 2–3 provided manipulation checks confirming the internal validity of the experimental manipulations. In these supplemental experiments, participants read the vignettes from Experiments 1–3 and reported how much harm (vs. gain) was experienced by society and how the vignette reflected on group reputation. Results provided robust and consistent support that participants perceived large differences in both societal and reputational outcomes (see Supplemental Material for further details).

Participants then answered questions regarding their affective reactions to the news story. These answers constituted the primary dependent variable in this experiment. We measured affect using a composite of 10 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, 0 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 (αs = .90 for both experiments after reverse-scoring negative items). While we report results collapsing across this index in the main text, details broken down by positive versus negative affect are presented in the Supplemental Material.[3]

Results

Across Experiments 1–3, we maintained the same 2 (Reputation Salience: High vs. Low) × 2 (Societal Outcome: Benefit vs. Harm) design. Experiments 1–3 tested two overarching hypotheses. First, we expected people to show a robust pattern of responding more positively to societal benefits than harms. We expected this to manifest in a clear simple effect of societal outcome (benefit vs. harm) when the predictor was an economist; thus, no reputation was at stake (i.e., the low-reputation salience condition).

Second, and of more central interest, was whether individuals would be sensitive to reputation in a context with large-scale and immediate societal consequences (including for ingroup members). Put another way, we were interested in whether group-based reputational incentives could dampen affective differentiation between societal harms and benefits. We predicted that they would and that this would manifest in an interaction in which for partisan (vs. nonpartisan) predictors, the affective differentiation between societal benefits versus harms would be blunted because societal outcomes were in opposition to reputational outcomes for an ingroup leader. As such, in addition to the interaction (the primary hypothesis test), we also report the simple effect of societal outcome when reputation is at stake (i.e., the high-reputation salience condition). Given the identical designs and hypotheses for Experiments 1A and 1B, we report them together below. Here and throughout the article, we also include details about individual subexperiments where relevant or where results diverge. Full regression tables are presented in the Supplemental Material.[4]

We found support for both hypotheses in Experiment 1. First, when the predictor was an economist (i.e., the low-reputation salience condition), we found a simple effect of societal outcome: Unsurprisingly, participants reported a far more positive overall response for societal benefit (vs. harm), and this effect was quite substantial in size, both when collapsing across both the Republican and Democrat subexperiments, $M_{\text{harm}} = -3.34$ versus $M_{\text{benefit}} = 0.32$, $t(893) = 26.45$, $p < .001$, Cohen's $d = 1.72$, and when considering each of the Republican, Experiment 1A: $M_{\text{harm}} = -2.87$ versus $M_{\text{benefit}} = 0.84$, $t(435) = 18.62$, $p < .001$, Cohen's $d = 1.75$, and Democrat, Experiment 1B: $M_{\text{harm}} = -3.37$ versus $M_{\text{benefit}} = -0.15$, $t(451) = 19.75$, $p < .001$, Cohen's $d = 1.76$, subexperiments individually. Participants showed clear affective differentiation between present harms and benefits to society when reputation was not at stake.

Second and supporting our key theoretical prediction, this affective differentiation was 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 subexperiments yielded robust evidence for an interaction between societal outcome and reputational salience (interaction: $b = 0.81$, $SE = 0.19$, $t = 4.24$, $p < .001$). In Experiment 1A (i.e., among Republicans), we found significant evidence to support this hypothesis (interaction: $b = 1.18$, $SE = 0.26$, $t = 4.49$, $p < .001$), and in Experiment 1B (i.e., among Democrats), we found marginally significant evidence (interaction: $b = 0.44$, $SE = 0.26$, $t = 1.71$, $p = .088$).

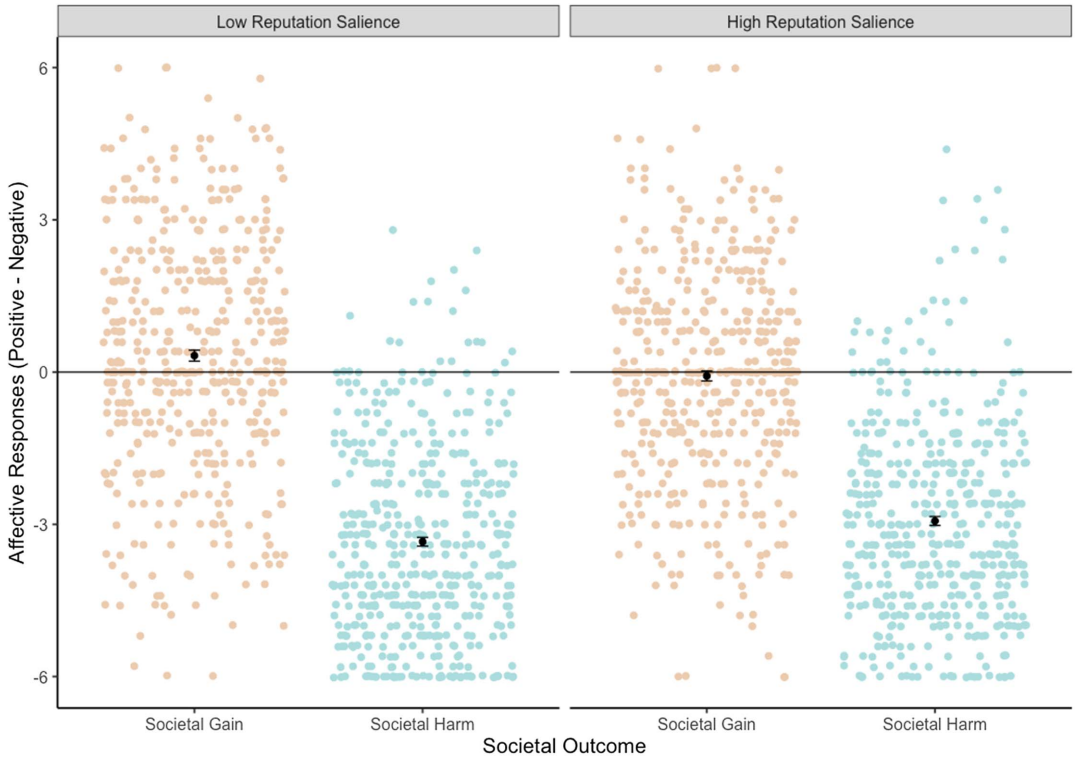
Thus, reputational considerations blunted individuals' preferences for good over bad societal outcomes, although this pattern was marginally significant when considering Democrats alone. That is, partisans appeared to let reputational considerations influence their (degree of) preference for present-term beneficial over harmful

[3] While results tended to be slightly stronger for positive (vs. negative) affect, both showed similar patterns of results when considering the evidence across experiments. More specifically, the hypothesized interaction between societal outcome and reputation salience was statistically significant for Experiments 1–3 when analyzing positive affect in isolation; it was statistically significant in Experiments 2–3, but not Experiment 1, when analyzing negative affect in isolation. Experiments 4–5 did not include the composite measure, so these analyses cannot be conducted for those samples. See Supplemental Information for full regression results and details.

[4] We do not report a "main effect of reputation" in Experiments 1–3 because it is uninterpretable in our context. The reason is because we expected the reputation factor to have opposite effects depending on the level of societal outcome. When the societal outcome is positive, in our studies, this typically means the reputational outcome is negative; the reverse is also true (when the societal outcome is negative, in our studies, this typically means that the reputational outcome is positive). As such, there is not really an interpretable "effect" of reputation across conditions, as reputation has a different direction depending on the level of societal outcome.

Figure 1

Partisans' Preferences for Good Versus Bad Societal Outcomes (Experiment 1)



Note. In the context of domestic economic policy, preferences were significantly blunted when the good versus bad societal outcomes were yoked, respectively, to bad versus good reputational outcomes for an ingroup leader. This pattern was statistically significant for Republicans (Experiment 1A) and marginally significant for Democrats (Experiment 1B). Error bars represent 1 *SE*, and colored dots represent raw data. *SE* = standard error. See the online article for the color version of this figure.

societal outcomes. Rather than simply focusing on whether the policy had wrought economic destruction on society as a whole, their affective responses factored in how these respective outcomes made their group look.

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 societal outcomes here was grave, representing the presence or absence of economic devastation amid a global pandemic. Although we thought it would be noteworthy if, even within this context, individuals' affective responses were shaped by reputation, we expected that we would still observe a more substantial effect of societal outcome, with individuals overall preferring good over bad societal outcomes in both the partisan and economist conditions. As depicted in Figure 1, this was in fact the case: Although affective differentiation between societal harm and benefit was smaller when reputation was on the line (vs. when it was not), individuals felt more positive about societal benefit versus harm in both conditions, simple effect in the high-reputation salience condition: $M_{bad} = -2.93$ versus $M_{good} = -0.08$, $t(932) = 21.69$, $p < .001$, Cohen's $d = 1.41$. Collapsing across Experiments 1A and 1B, the sensitivity to societal outcomes was reduced by approximately 20% in high-reputation salience condition relative to the low-reputation salience condition (i.e., from $d = 1.72$ to $d = 1.41$), suggesting that despite a substantial effect of reputational considerations, the relative size of the effect of societal outcomes was much larger in this context.

Discussion

Experiment 1 provided initial evidence that reputational incentives can blunt partisans' affective differentiation between societal harms and benefits. This pattern was statistically significant when collapsing across parties and for Republicans alone and was marginally significant for Democrats alone. At least in Experiment 1's specific context, partisans' affective reactions were driven more strongly by societal outcomes than by reputational considerations; in Experiments 2–3, we diversify contexts and continue examining the relative effect sizes of societal outcomes and group-based reputational incentives.

Experiments 2A and 2B

Experiment 2 tests whether reputational incentives can blunt partisans' sensitivity to current societal harms and benefits in a new domain: national security. Experiment 2A explores this possibility among Republicans. Experiment 2B is a parallel experiment with

Democrats, reported separately because of a few noteworthy design differences (described below).

Experiment 2A Method

The preregistration for Experiment 2A is available at https://aspredicted.org/NZ9_587.

Participants

We conducted an experiment with a sample of 600 Republicans (M_{age} = 43.95, SD = 13.12). In Experiment 2A, 296 participants reported their gender as female, 301 as male, two as other/prefer not to disclose, and one did not report their gender. As preregistered, we included only the 581 Republicans (96.8%) who passed an attention check and did not ask to have their data removed. Experiment 2A was conducted in April 2021.

Procedure

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 societal 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 Experiment 1, we manipulated both the salience of reputational outcomes (high vs. low) and the valence of the societal outcome (harm vs. benefit in the form of news about Iran's actual cyber capabilities). To manipulate reputation salience, we mimicked the design of Experiment 1 by manipulating the partisan identity of the predictor—in this case, either Democratic President Biden (high-reputation salience) or Central Intelligence Agency (CIA) Director Burns (low-reputation salience; Burns' partisan group membership was not mentioned, and he was confirmed unanimously by the U.S. Senate). To manipulate societal outcome, participants learned that a new independent report either confirmed (societal benefit) or contradicted (societal harm) the optimistic 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 sophisticated (societal benefit) or more sophisticated (societal harm) than previously believed. Thus, relative reputational gain for Republicans (i.e., Biden looking bad) co-occurred with 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. First, we measured it using a single item of overall positivity versus negativity on a 100-point sliding scale anchored at *extremely negative* and *extremely positive*. Second, we used the same combined index of 10 specific emotion items from Experiment 1, including five positive items and five negative items. The index again achieved a high level of reliability (α = .91 after reverse-scoring negative items).

While now in the context of national security rather than domestic economic policy, the design parallels Experiment 1: Participants read a news story in which a public figure makes a public prediction with reputational consequences. Here, we also sought to provide a stringent test of our theorizing about the role of reputation by further raising the countervailing societal 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 societal downsides: Not only were Iran's cyber capabilities more sophisticated, but the United States faced the additional harm 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). These heightened societal stakes make our empirical test of the ability of reputation to blunt affective differentiation between societal outcomes more conservative, a point we return to in Experiment 2B when discussing perceived competence of a leader.

Experiment 2A Results

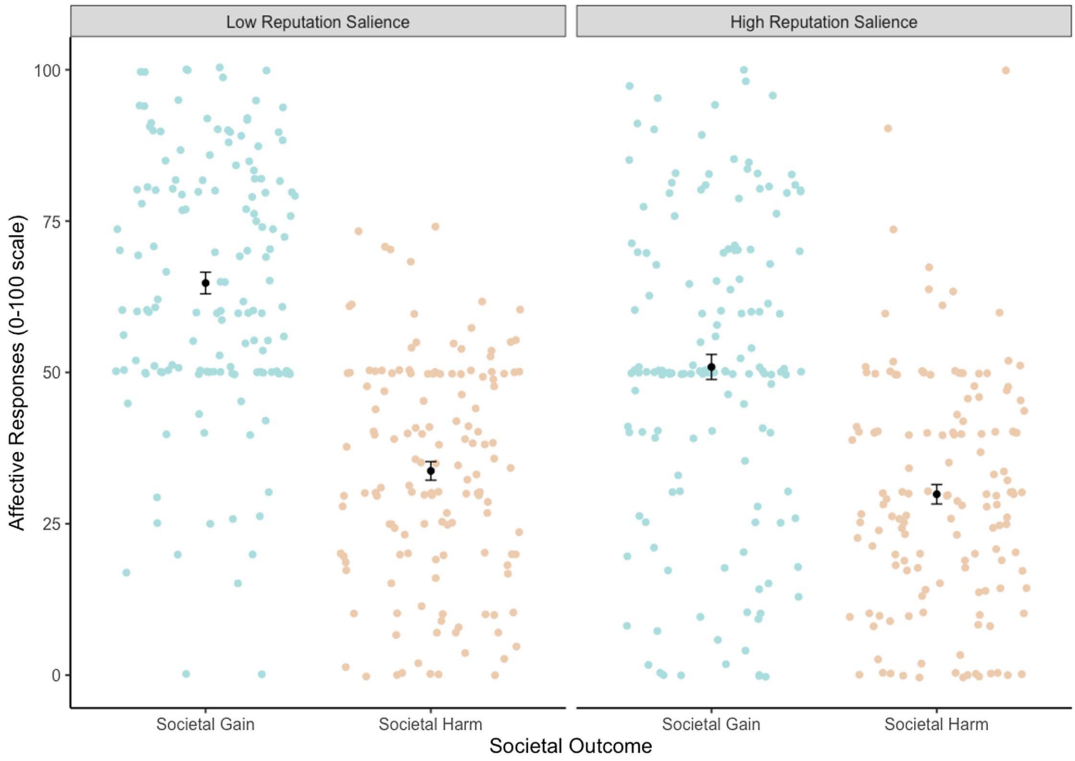
As in Experiment 1, we first examined the simple effect of societal outcomes in the low-reputation conditions. Unsurprisingly, when the predictor was CIA Director Burns (i.e., the low-reputation salience condition), we found a clear and sizeable effect of societal outcome: Participants responded much more positively to societal benefit versus harm, for both the single general affect item, M_{harm} = 33.74 versus $M_{benefit}$ = 64.77, $t(280) = 13.17$, $p < .001$, Cohen's $d = 1.55$, and the 10-item composite, $M_{harm} = -2.15$ versus $M_{benefit} = 1.79$, $t(284) = 16.55$, $p < .001$, Cohen's $d = 1.95$.

Building on Experiment 1, we theorized that this sensitivity to societal outcomes would be significantly attenuated when societal gain helped an outparty partisan leader's reputation (in this case, President Biden's). As depicted in Figure 2, this interaction prediction was confirmed: Reputational gains blunted affective responses to societal benefits for both the single general affect item (interaction: $b = 10.01$, $SE = 3.55$, $t = 2.82$, $p = .0049$) and for the combined index (interaction: $b = 0.97$, $SE = 0.34$, $t = 2.86$, $p = .0044$). While individuals continued to feel more positively about societal gains (vs. harms) in the high-reputation salience condition, single-item slider: $M_{bad\ societal}$ = 29.88 versus $M_{good\ societal}$ = 50.91, $t(277) = 8.03$, $p < .001$, Cohen's $d = 0.93$; 10-item composite: $M_{bad\ societal} = -2.33$ versus $M_{good\ societal} = 0.63$, $t(272) = 12.35$, $p < .001$, Cohen's $d = 1.43$, this represented an approximately 30%–40% reduction in the gap in affective responses between low versus high threat to the U.S. from Iran (Single item: $d_{low\text{-}reputation\ salience} = 1.55$ vs. $d_{high\text{-}reputation\ salience} = 0.93$; Composite: $d_{low\text{-}reputation\ salience} = 1.95$ vs. $d_{high\text{-}reputation\ salience} = 1.43$). Conceptually replicating Experiment 1 in a new context, the results provided converging evidence that reputational considerations can reduce sensitivity to societal outcomes.

Unlike in Experiment 1, here, the interaction appeared primarily driven by individuals feeling less positively about societal benefits when there were reputational downsides than it was by

Figure 2

Republicans' Preferences for Good Versus Bad Societal Outcomes (Experiment 2A)



Note. In the context of national security, preferences were blunted when the good versus bad societal outcomes were yoked, respectively, to bad versus good reputational outcomes for an outgroup leader (i.e., President Biden). Error bars represent 1 *SE*, and colored dots represent raw data. *SE* = standard error. See the online article for the color version of this figure.

individuals feeling less negatively about societal harms when there were reputational upsides. While these simple slope analyses are presented in the Supplemental Materials for all studies, we make note of this distinction where relevant across experiments and return to discuss it further in the General Discussion section. For now, it is worth noting that in both Experiments 1 and 2A, reputational incentives reduced overall differentiation between societal harm and benefit.

Experiment 2B Method

The preregistration for Experiment 2B is available at https://aspredicted.org/4NW_67L.

Participants

We conducted an experiment with a sample of 903 Democrats (M_{age} = 40.05, *SD* = 12.60). In Experiment 2B, 433 participants reported their gender as female, 462 as male, and eight as other/prefer not to disclose. As preregistered, we included only the 869 Democrats (96.2%) who passed an attention check and did not ask to have their data removed. Experiment 2B was also conducted in April 2021.

Procedure

Experiment 2B was a parallel experiment 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 Experiment 2A and two supplemental conditions (described later).

In the four primary conditions, Democrats were randomly assigned in a 2 (Reputation Salience: High vs. Low) × 2 (Societal Outcome: Benefit vs. Harm) 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 Experiment 2A, in Experiment 2B, 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 societal 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 Experiment 2A: with a single global evaluation and with a combined index of 10 specific

emotion items (five positive and five negative; $\alpha = .84$ after reverse-scoring negative items).

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 societal 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 United States's ability to handle foreign policy challenges effectively. If so, then their negative affect might be due not to reputational damage but to a calculation that, despite the positive news about Iran's weaker-than-predicted capabilities, the United States would be facing greater threats overall (because Biden's incorrect prediction revealed incompetence on the foreign policy change).

To address this potential alternative explanation, we needed to separate the reputational effects of Biden's prediction from the broader consequences of his (in)accuracy. To do this, we included two additional conditions in our randomized design that were identical to the high-reputation salience conditions described above, with one exception. Specifically, the article that the 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 underappreciated 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 United States's 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 in fact specifically driven by reputational considerations rather than by concerns about leader competence.[5]

Experiment 2B Results

We again began by examining the simple effect of societal outcomes in the low-reputation conditions. As predicted, participants in the low-reputation salience conditions demonstrated strong sensitivity to societal outcomes: They reported a more positive affective response when the pessimistic prediction turned out to be incorrect (i.e., resulting in societal gain: Iran's cyber capabilities are, in fact, *not* growing more sophisticated) compared to when the prediction was confirmed (i.e., resulting in societal harm: Iran's cyber capabilities are *indeed* growing more sophisticated). This was true both for responses on the single item measuring general global affect, $M_{harm} = 38.91$ versus $M_{gain} = 57.52$, $t(292) = 8.07$, $p < .001$, Cohen's $d = 0.94$, and on the 10-item affect composite, $M_{harm} = -1.79$ versus $M_{gain} = 0.87$, $t(291) = 11.65$, $p < .001$, 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 societal gain in the low-reputation salience condition, we theorized that this effect would be attenuated (and potentially even eliminated) when societal gain hurts 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 societal outcomes for both the single general affect item (interaction: $b = 16.75$, $SE = 3.09$, $t = 5.43$, $p < .001$) and the collapsed 10-item scale (interaction: $b = 1.52$, $SE = 0.31$, $t = 4.85$, $p < .001$). Strikingly, individuals no longer demonstrated a statistically significant sensitivity to societal outcomes when affect was measured with the single-item slider, $M_{harm} = 51.29$ versus $M_{gain} = 53.14$, $t(289) = 0.91$, $p = .37$, Cohen's $d = 0.11$. Sensitivity to societal outcomes persisted—but was meaningfully reduced—as measured by the 10-item composite, $M_{harm} = -0.97$ versus $M_{gain} = 0.16$, $t(290) = 5.28$, $p < .001$, Cohen's $d = 0.62$. To put these results in perspective, the blunting through reputational incentives represented an approximately 55%–90% reduction in differentiation between societal harm and gain in the form of 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. $d_{\text{high-reputation salience}} = 0.62$).

Finally, we did not find evidence to support the notion that the interaction was driven by inferences regarding Biden's competencies (rather than by reputational concerns, as theorized). Specifically, when replacing the two primary high-reputation salience conditions with the two supplemental Biden conditions (in which we specified that his incorrect prediction had little bearing on U.S. foreign policy competencies), we again find significant evidence that the reputation effect attenuates the effect of societal outcomes for both the single general affect item (interaction: $b = 15.55$, $SE = 3.27$, $t = 4.76$, $p < .001$) and for the combined index (interaction: $b = 1.21$, $SE = 0.33$, $t = 3.71$, $p < .001$). The interaction with the primary conditions was almost equal in magnitude to the interaction with the supplemental conditions (single item: $b_{primary} = 16.75$ vs. $b_{supplemental} = 15.55$; composite: $b_{primary} = 1.52$ vs. $b_{supplemental} = 1.21$), suggesting that implications for general competency in handling other threats were not the primary driver of the interaction. Results are depicted in Figure 3.

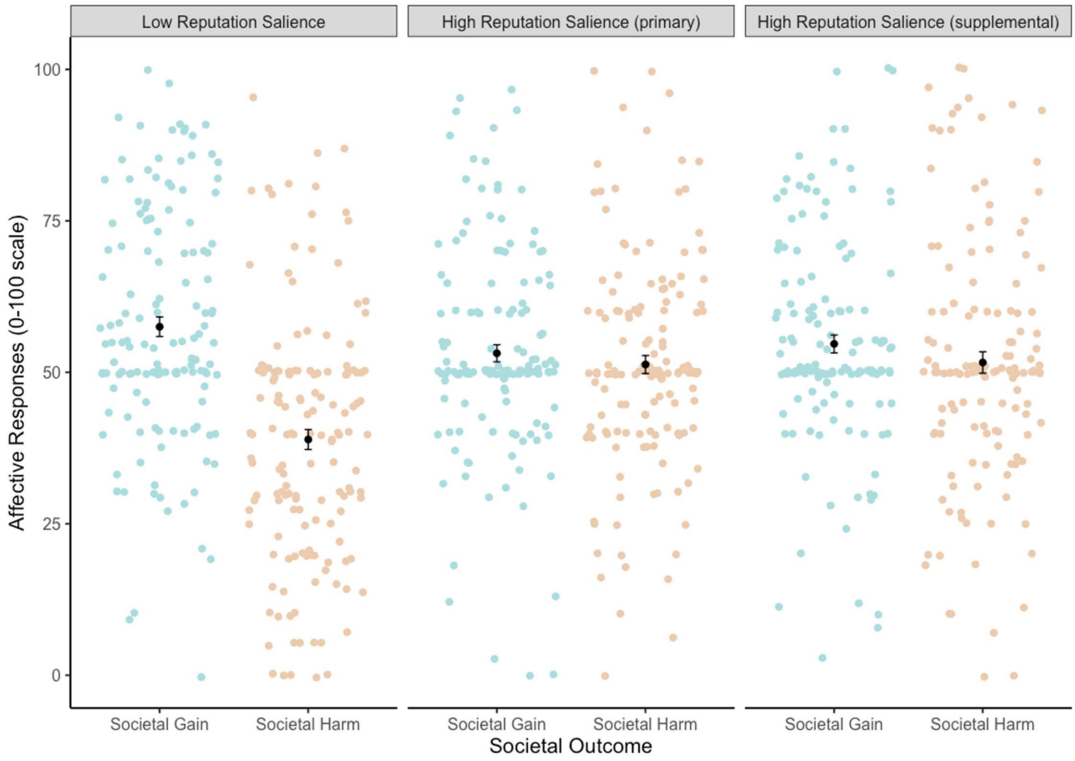
Discussion

Experiment 2 demonstrated that, in certain cases, reputational stakes can nearly eliminate affective differentiation between societal harm and gain. The fact that the relative effect of societal outcomes versus group-based reputational incentives was larger among Republicans (where societal harm co-occurred with the additional material risk of presidential incompetence) than among Democrats (where societal harm occurred without the additional material risk of

[5] Theoretically, it could be the case that affective responses are rooted in general concerns about the leader's competence in *other* domains in which he matters, but we did not think that such inferences were likely to account for our results.

Figure 3

Democrats' Preferences for Good Versus Bad Societal Outcomes



Note. In the context of national security, preferences were nearly eliminated when the good versus bad societal outcomes were yoked, respectively, to bad versus good reputational outcomes for an ingroup leader (i.e., President Biden). This pattern of results held in two supplemental conditions that addressed an alternative explanation based on nonreputational considerations (Experiment 2B; see main text for details). Error bars represent 1 *SE* and colored dots represent raw data. *SE* = standard error. See the online article for the color version of this figure.

presidential incompetence) suggests—albeit tentatively—that the relative size of societal versus reputational considerations in shaping affect may depend on the relative balance of societal versus reputational stakes.

Experiment 3

Experiment 3 seeks to address an alternative possibility that could account for some of the prior results: Perhaps partisans are not feeling happier (unhappier) because of reputation per se, but because they are having their expectations confirmed (disconfirmed; e.g., Kaiser et al., 2004; Major et al., 2007). While there are several ways that a worldview could be confirmed (or disconfirmed), Experiment 3 pushes against a worldview confirmation hypothesis by ruling out one major potential basis for worldview confirmation: an ingroup leader making a correct prediction. That is, perhaps Democrats, for example, feel better (worse) when Pelosi or Biden (vs. an economist or CIA Director) is right (wrong) about a pessimistic prediction not because of their sensitivity to Democrats' reputation relative to Republicans but simply because Pelosi or Biden being right (wrong) about things is more consistent with their political expectations, with expectancy violations yielding affective costs. Therefore, to disentangle the relative contributions of reputation and having one's expectations about an ingroup leader's accuracy confirmed, we needed to examine a context in which reputational gain for the ingroup occurs because of events that disconfirm one's expectations on this dimension.[6]

[6] We note that it is possible that an ingroup leader's reputation being bolstered might itself plausibly confirm one's worldview, insofar as individuals might expect their leader's actions to burnish their reputation by virtue of their general favorability toward their leader. We are unable to entirely rule this out, given that reputational benefit would inherently be connected to worldview reinforcement for those who expect their leader's reputation to always be burnished, and it would therefore not be possible to easily disentangle these two aspects experimentally. At the same time, it is not in fact obvious that individuals *will* have a strong worldview/expectation that their leader's reputation will be burnished. After all, individuals are likely aware that counterpartisans will regularly attempt to impugn their leader's reputation, and it is common knowledge that leaders, especially incumbents like Biden, often face reputation-damaging criticism from various quarters. Indeed, it has been widely reported that, despite confidence in his abilities to be president, Democrats are currently concerned about how allowing Biden to be more active on the interview circuit might damage his reputation during the reelection campaign (Klein, 2024). In any case, even though we cannot definitively rule out *all* aspects of worldview confirmation in our current design, we nevertheless deemed that disentangling reputation from the expectation of one's ingroup leader being right was a worthwhile test.

Method

The preregistration for Experiment 3 is available at https://aspredicted.org/YL6_422.

Participants

We conducted an experiment with a sample of 973 Republicans (M_{age} = 44.31, SD = 12.84). In Experiment 3, 510 participants reported their gender as female, 461 as male, and two as other/prefer not to disclose. As preregistered, we included only the 926 Republicans (95.2%) who passed an attention check and did not ask to have their data removed. Experiment 3 was conducted in July 2021.

Procedure

Republicans were randomly assigned to one of four between-subjects experimental conditions. Across conditions, participants read an ostensibly real news story 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 Experiments 1–2, we manipulated both the salience of the reputational outcome (high vs. low) and the valence of the societal outcome (harmful vs. beneficial 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 Rouse[7] (low-reputation salience). To manipulate the societal outcome, participants learned that a new report found either that the policy had reduced unemployment and efficiently allocated resources (societal gain) or had no effect on unemployment and inefficiently allocated resources (societal harm). Specifically, participants read either that over 1 million families had received nutritional assistance that they badly needed (societal gain) or that there had been almost no reduction in the number of families receiving nutritional assistance that they badly needed (societal harm). The story made clear in all conditions that, while the policy was put into effect while the public figure was in office (thus impacting their reputation), the policy was in fact developed and implemented independently of the public figures (and of their private opinions about the policy's likely effectiveness). The text noted that the public figure was credited or blamed with the policy's 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 Experiment 3 compared to Experiments 1–2 was the dissociation between confirming expectations regarding the "wrongness" of an outgroup's leader and reputational incentives. Unlike Experiments 1–2, Experiment 3 created a contrast between the correctness of the private prediction and the public perception. Specifically, in the societal gain conditions, the leader made an incorrect private prediction (i.e., incorrectly assuming the policy would be a failure) but happened to be publicly credited for the success of the policy; in the societal harm conditions, the leader made a correct private prediction (i.e., correctly assuming the policy would be a failure) but happened to be publicly blamed for the policy. Republicans' expectations should be confirmed to a greater extent in the societal gain/bad reputational outcome condition (in which Biden made an incorrect prediction) compared to the societal harm/good reputational outcome condition (in which Biden made a correct prediction). Thus, Republican participants experience reputational *gain* when their expectations are *dis*confirmed and vice versa, disentangling confirmation of one's expectations from reputational gain.

Results

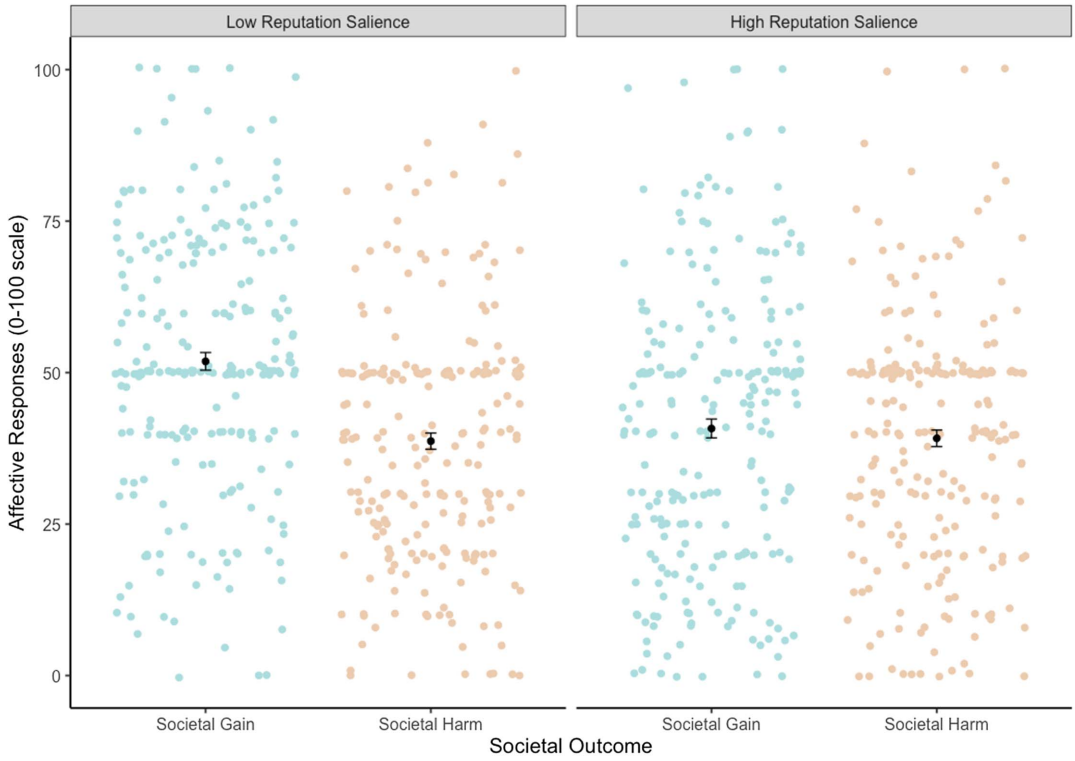
Our analysis plan proceeded as in Experiments 1–2, with the overarching hypothesis that a similar pattern of results would emerge even though any effect of confirming one's expectations regarding the wrongness of an outparty leader could, if operative, oppose any effect of reputation. Despite this potential oppositional force, a near-identical pattern of results emerged. First, we again found a simple effect of societal outcome in the low-reputation salience conditions: Participants in the conditions involving Council of Economic Affairs Chair Rouse responded much more positively to societal gain versus harm, for both the single general affect item, M_{harm} = 38.69 versus M_{gain} = 51.86, $t(459) = 6.66$, $p < .001$, Cohen's $d = 0.62$, and the 10-item composite, $M_{harm} = -1.61$ versus $M_{gain} = 0.50$, $t(459) = 10.86$, $p < .001$, Cohen's $d = 1.01$. However, and more critically, we again found evidence for an interaction: This sensitivity to societal outcomes was significantly attenuated when it conflicted with an outparty partisan leader's (in this case, again President Biden's) reputation, for both the single general affect item (interaction: $b = 11.54$, $SE = 2.86$, $t = 4.03$, $p < .001$) and the 10-item composite (interaction: $b = 1.25$, $SE = 0.28$, $t = 4.53$, $p < .001$). Mirroring the results of Experiment 2B, participants no longer demonstrated a statistically significant differentiation between societal harm and gain when affect was measured with the single-item slider, M_{harm} = 39.16 versus M_{gain} = 40.78, $t(456)$ = 0.79, p = .43, Cohen's d = 0.07, although this differentiation persisted—but was meaningfully reduced—as measured by the 10-item composite, $M_{harm} = -1.45$ versus $M_{gain} = -0.60$, $t(459) = 4.34$, $p < .001$, Cohen's d = 0.40.

As depicted in Figure 4, the reputation effect was again quite substantial. It represented an approximately 60%–90% reduction in the gap in affective responses between low versus high threat to the United States from Iran (single item: $d_{low\text{-}reputation\ salience} = 0.62$ vs. $d_{high\text{-}reputation\ salience} = 0.07$; composite: $d_{low\text{-}reputation\ salience} = 1.01$ vs. $d_{high\text{-}reputation\ salience} = 0.40$). Indeed, in the high-reputation salience condition, Republicans reported nearly identical affective responses for the general affect item when the policy was a failure and hurt society as a whole, including the ingroup (but Biden was blamed: $M = 39.16$), as compared to when the policy was a success and benefited society as a whole, including the ingroup (but Biden was credited: $M = 40.79$).

[7] While Chair Cecilia Rouse is female, we gave her a male-gendered name (Charlie) to avoid introducing a confound for gender in the experiment. Of the 500 participants in the low-reputation salience condition (and thus read about Rouse), none indicated familiarity with Rouse or commented on the name change in an open-ended text box at the end of the experiment.

Figure 4

Republicans' Preferences for Good Versus Bad Societal Outcomes



Note. In the context of domestic economic policy, preferences were nearly eliminated when the good versus bad societal outcomes were yoked, respectively, to bad versus good reputational outcomes for an outgroup leader (i.e., President Biden). This pattern held even though any positive affective consequences of reputational gain could, in principle, have been counteracted by negative affective consequences of disconfirming prior expectations about an outgroup/ingroup leader (Experiment 3). Error bars represent 1 *SE*, and colored dots represent raw data. *SE* = standard error. See the online article for the color version of this figure.

Discussion

Experiments 1–3 provided consistent evidence that group-based reputational incentives can reduce (or even largely eliminate) affective differentiation between societal harms and gains, even when these outcomes influence the ingroup. Moreover, this pattern did not seem to be driven by affective costs (or benefits) of disconfirming (or confirming) prior expectations about an outgroup/ingroup leader being wrong/right about the world.

Experiment 4

Experiment 4 begins to investigate why group-based reputational incentives temper sensitivity to present-term societal outcomes. While Experiments 2B and 3 attempted to support the role of reputation by ruling out alternative explanations, Experiment 4 attempts to rule in an underlying mechanism for why individuals care about reputational incentives: perceived strategic benefits for one's group.

Adapting the methodology from Experiment 2B, Democrats read a news story in which a public figure correctly predicted an increasing national security threat from Iran (thereby bolstering the predictor's reputation) and reported their affective reaction. Thus, all participants were assigned to a condition that included societal harm. They were again randomly assigned such that there were low- versus high-reputational stakes (i.e., the predictor was either a nonpartisan public official or a party leader). In Experiment 4 (and unlike prior studies), we built on this design by measuring self-reported attention to strategic benefits for one's ingroup and testing whether this attention mediated the effect of reputational incentives on affective reactions.

Prior studies were relatively subtle in identifying the strategic benefits for one's group: Strategic benefits were often implied as a consequence of reputational gain but were not directly noted. This was because we expected that participants would be impacted by the strategic implications associated with reputational stakes even when these were not spelled out. Nevertheless, if our theoretical account is correct—that is, if individuals' affect is influenced by reputational considerations because of its strategic consequences for the ingroup—then making the strategic implications of reputation more salient should further amplify our effect.

Thus, in Experiment 4, in addition to measuring stated attention to strategic benefits, we also manipulated them. That is, we included a new experimental condition in which we make the strategic benefits for one's group more salient. We expected that Democrats would report a more positive affective reaction when their group's leader

(vs. a nonpartisan public official) accurately predicted an increased security threat from Iran, but we additionally expected that such effects would be even stronger when the strategic benefits were further spotlighted.

Method

The preregistration for Experiment 4 is available at https://aspredicted.org/LGY_TCZ.

Participants

We conducted an experiment with a sample of 748 Democrats (M_{age} = 41.35, SD = 12.85). In Experiment 4, 438 participants reported their gender as female, 307 as male, and three as other/prefer not to disclose. As preregistered, we included only the 704 Democrats (94.1%) who passed an attention check and did not ask to have their data removed. Experiment 4 was conducted in December 2022 (i.e., after the 2022 midterm election in which President Biden and Democrats maintained power in the Senate but lost power in the House of Representatives and just weeks before the incoming Congress took power).

Procedure

Experiment 4 adapted the procedure from Experiment 2B. After answering the same individual difference and demographic questions as in prior experiments, Democrats in all conditions read a news story in which a public figure correctly predicted an increasing national security threat from Iran (thereby burnishing the predictor's reputation). Democrats then reported their affective reactions after reading the news story (here, using only the slider scale from 0 to 100).

Democrats were randomly assigned to one of three between-subjects experimental conditions. The first two conditions were identical to the societal harm conditions from Experiment 2B, thereby manipulating reputation salience (low vs. high). In the low-reputation salience condition (i.e., the CIA Condition), the public figure was CIA Director Burns. In the high-reputation salience condition (i.e., the baseline Biden condition), the public figure was President Biden. As in Experiment 2B, we predicted that the Democrats would report more positive affective reactions when President Biden (vs. CIA Director Burns) accurately predicted the increased threat from Iran.

Beyond this main effect of experimental condition on affect, Experiment 4 also examines the underlying mechanism of perceived strategic benefits for one's group. To do so, on the page following the affect slider, we included two Likert items explicitly asking participants whether they considered strategic consequences for the Democratic Party when reading the news story: "When reading, how much (if at all) did you think about how the events described in the news article would influence the strategic standing of Democrats as compared to Republicans?" and "When reading, how much (if at all) did you think about how the events in the news article would give Democrats the upper hand on Republicans?"; interitem correlation = .87. Both items were answered on 7-point scales ranging from 1 (*not at all*) to 7 (*very much*). We predicted that Democrats would report higher attention to strategic standing in the Biden (vs. CIA) condition and that this difference would mediate the effect of condition on affect.

In addition to the two conditions described above, we included a third experimental condition: the high-reputation salience + strategic salience condition (i.e., the strategic salient Biden condition). The goal of this condition was to directly make top-of-mind the potential strategic implications of a boost to the group's reputation. In this way, this third condition was designed to "manipulate the mediator," experimentally manipulating the proposed mediator (sensitivity to strategic consequences associated with reputation) by making the consideration of strategic consequences for the ingroup salient and testing for effects on the outcome variable of interest (i.e., affective responses).

In this third condition, the public figure was again President Biden—but the news story also made salient the potential consequences of Biden's improved standing in the eyes of the public for the strategic goals of the Democratic party. More concretely, the headline of the news story included the following subheading: "Analysts suggest Biden's astute perspective could impact Democrats' political position." At the end of the news story, participants in this condition also read that the boost to Biden's reputation could boost Democrats' political position and afford an advantage with the incoming Congress.

We had two specific predictions relating to this new, third experimental condition. First, we predicted that participants would report *even more* positive affective reactions in the strategic salient Biden condition compared to the baseline Biden condition (which we expected itself would generate more positive affective reactions to the CIA condition, replicating Study 2B). Second, we expected a parallel pattern for self-reported attention to strategic benefits: that participants would report even more attention to strategic benefits in the strategic salient Biden condition to the baseline Biden condition and that this difference in attention to strategic benefits would mediate the difference between the baseline Biden condition and the strategic salient Biden condition on affect.

Results

Our analysis proceeds in lockstep with the procedure described above: We first replicate and extend our analyses of the two experimental conditions from Experiment 2B, then analyze the novel third experimental condition. All analyses and statistical estimates described below were fit in a single overarching model including contrasts between all three conditions to ensure that we appropriately account for all of the variance; however, we report pairwise contrasts separately for clarity.

CIA Condition Versus Baseline Biden Condition

In a first set of analyses, we examined our first two experimental conditions: the CIA condition and the baseline Biden condition. Directly replicating Experiment 2B and conceptually replicating the other prior experiments, Democrats reported less negative affect in the baseline Biden condition compared to the CIA condition ($M_{CIA} = 45.10$ vs. $M_{baseline\ Biden} = 52.67$, $t = 4.35$, $p < .001$, Cohen's d = 0.39), that is, participants' negative affect when a leader's prediction of societal harm indeed came to pass was buffered when that leader belong (vs. did not belong) to their ingroup, and the accuracy of his prediction could therefore have positive reputational

ramifications for the group. New to Experiment 4 was the Likert scale measurement of attention to strategic consequences associated with reputation. Participants also reported considering strategic incentives for Democrats less often in the CIA condition ($M = 2.32$) than in the baseline Biden condition ($M = 3.38$: $t = 7.19$, $p < .001$, Cohen's $d = 0.66$). We tested for mediation using the Lavaan package in R (Rosseel, 2012). The independent variable was the experimental condition, the mediator was a consideration of strategic incentives (using the Likert items), and the dependent variable was self-reported affect. Notwithstanding limitations of correlational designs testing mediation (Bullock et al., 2010), the model provided evidence for a significant and positive indirect effect (Indirect effect = 3.93, $SE = 0.71$, $z = 5.57$, $p < .001$). The indirect effect was moderate in magnitude: Differences in consideration of strategic incentives explained approximately 52% of the total effect of condition on affective responses. Results are depicted in Figure 5 (top panel).

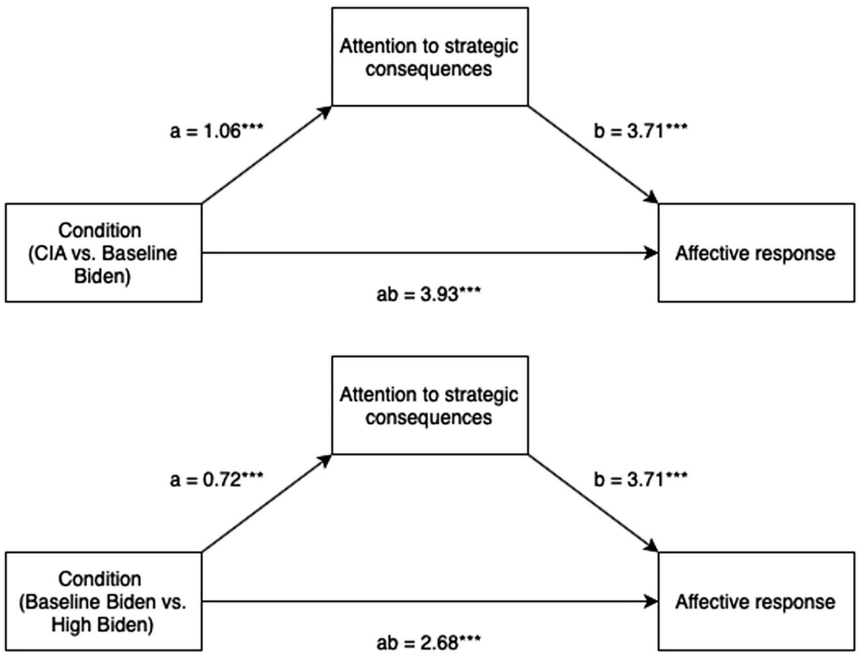
Baseline Biden Condition Versus Strategic Biden Condition

In a second set of analyses, we conducted an analogous set of tests with the new third condition: the strategic salient Biden condition. Participants reported even more positive affective responses in the strategic salient Biden condition compared to the baseline Biden condition ($M_{\text{baseline Biden}} = 52.67$ vs. $M_{\text{strategic Biden}} = 61.06$, $t = 4.79$, $p < .001$, Cohen's $d = 0.45$). Given that the only difference between these conditions was that the strategic consequences of a boost to Biden's reputation was made explicit, this result supports the notion that partisans are influenced by strategic boosts associated with reputational gains, even when this comes at the cost of current societal harm. In addition and confirming the internal validity of the experimental manipulation, participants reported higher levels of attention to strategic incentives in the strategic salient Biden condition compared to the baseline Biden condition ($M_{\text{baseline Biden}} = 3.38$ vs. $M_{\text{strategic Biden}} = 4.10$, $t = 4.86$, $p < .001$, Cohen's $d = 0.44$). As predicted, this difference in consideration of strategic incentives mediated the effects of condition on self-reported affect (indirect effect = 2.68, $SE = 0.63$, $z = 4.26$, $p < .001$). The effect was again moderate in magnitude, explaining approximately 32% of the total effect of condition on affective responses. Results are depicted in Figure 5 (bottom panel).

Replication (Supplemental Experiment 4)

Experiment 4 provided evidence that stated attention to strategic benefits for one's ingroup mediated the effect of reputational incentives on affective reactions. Yet, the measure of attention to strategic consequences mentioned not only the standing of the individuals' ingroup but also its relative position to the outgroup, for example, "When reading, how much (if at all) did you think about how the events described in the news article would influence the strategic standing of Democrats as compared to Republicans?". Contrary to our theorizing, it could be the case that, by mentioning

Figure 5
Democrats' Aversion to Negative Societal Outcomes



Note. In the context of national security, aversion to negative societal outcomes was reduced when negative societal outcomes were yoked to good reputational outcomes for an ingroup leader (i.e., President Biden). This pattern was underpinned by attention to the strategic benefits of such reputational gain (Experiment 4). CIA = Central Intelligence Agency.
*** $p < .001$.

Democrats *in comparison to* Republicans, the items are making the intergroup *conflict* more salient. Therefore, the salience of intergroup conflict—rather than, as we propose, concern with the ingroup's strategic standing—could, in principle, be driving our results. We conducted Supplemental Experiment 4 to address this possibility.

Supplemental Experiment 4 was a replication and extension of Experiment 4 with a modification to rule out an alternative possibility. Participants were again self-reported Democrats (N = 589) who read a news story in which a leading public figure predicted that Iran's cyber capabilities were more sophisticated than previously believed (the news story was minimally adapted to reflect the fact that the replication took place later in Biden's presidency). Again, they then learned that a new independent report confirmed the leading public figure's prediction. Simplifying Experiment 4, participants were randomly assigned to one of two between-subjects experimental conditions (CIA vs. Biden). We did not include the strategic salient Biden condition in this replication study because it did not bear on the central hypothesis of testing robustness across wording of the proposed mediator.

They key difference between Experiment 4 and Supplemental Experiment 4 was the measure of attention to strategic benefits (i.e., the proposed mediator). In Experiment 4, this measure also included a reference to the outgroup, running the risk that we might have primed intergroup conflict between Democrats and Republicans (rather than capturing sensitivity to ingroup strategic standing, as we sought to). In Supplemental Experiment 4, it did not; therefore, we more cleanly isolated sensitivity to ingroup strategic standing. For example, participants were asked "When reading, how much (if at all) did you think about how the events described in the news article would influence the strategic standing of Democrats?"

Despite the revised wording of the mediating items, results closely replicated those from the comparable contrasts in Experiment 4. Participants reported more positive affective reactions in the Biden (vs. Burns) condition ($M_{CIA} = 43.28$ vs. $M_{Biden} = 49.00$, $t = 3.42$, $p < .001$, Cohen's $d = 0.28$). Further, they reported greater attention to strategic consequences in the Biden (vs. Burns) condition ($M_{CIA} = 2.83$ vs. $M_{Biden} = 3.91$, $t = 7.64$, $p < .001$, Cohen's $d = 0.61$), and attention to strategic consequences again mediated the effect of condition on affect (indirect effect = 3.28, $SE = 0.67$, $z = 4.88$, $p < .001$). The effect size was similar in magnitude, such that self-reported differences in consideration of strategic incentives explained approximately 57% of the total effect of condition on affective responses (compared to approximately 52% in Experiment 4).

Discussion

Experiment 4 and its replication provide converging evidence for the underlying psychology driving the effect of group-based reputational incentives on sensitivity to societal outcomes. Via tests both measuring and manipulating our proposed mediator, we obtained evidence consistent with the claim that partisans' affective responses to reputational stakes (even in the face of societal harms) are driven by sensitivity to the strategic value of reputation for the ingroup's standing relative to the outgroup.

Experiments 5A and 5B

Experiments 1–4 provide converging evidence that reputational stakes shape partisans' affective responses to societal events, that they do so in part because of attention to strategic consequences, and that they can do so quite substantially—sometimes even nearly eliminating affective differentiation between immediate societal harms and benefits.

Experiment 5 has two primary goals. First, Experiment 5 turns from affective responses to behavior. Whereas participants might feel better about societal harm when it helps their group's reputation, will they also act in ways that have a suboptimal societal impact on account of group-based reputational considerations? To test this possibility, we created a paradigm in which partisans made financially incentivized donation choices between educational charities. We then experimentally manipulated whether participants did or did not have group-based reputational incentives for the success of the less effective charity. We investigated whether the presence of group-based reputational incentives could shift partisans' donation choices away from the more effective charity and toward the less effective charity whose success would have strategic upsides for the ingroup. Second, Experiment 5 further probes mechanism. While Experiment 4 revealed that *amplifying* strategic stakes can strengthen the tendency for group-based reputation to blunt sensitivity to present-term societal harm, Experiment 5 additionally tests whether *downplaying* strategic stakes can have an analogous effect. That is, we again take a manipulate-the-mediator approach to probe mechanism.

We recruited partisans and assessed their preference for donating to a more (vs. less) effective charity. Depending on the experimental condition, we varied whether we made salient or downplayed the extent to which the success of the more effective charity would engender a strategic disadvantage to the ingroup (by bolstering the outgroup's candidate ahead of the presidential election). We expected that participants would be more likely to donate to the less effective charity when the salience of strategic incentives was amplified versus downplayed. Looked at from the other perspective, downplaying the strategic upside to the ingroup of donating to the less effective charity would increase the percentage of participants donating to the charity that would have a better societal impact. For exploratory analyses, we also included a condition in which such strategic incentives were present but were neither made explicitly more salient nor downplayed. We tested these confirmatory and exploratory hypotheses for both Republicans (Experiment 5A) and Democrats (Experiment 5B).

Method

The preregistrations for Experiments 5A and 5B are available at https://aspredicted.org/6VW_4XS and at https://aspredicted.org/6VW_4XS.

Participants

We conducted an experiment with a sample of 964 Republicans (Experiment 5A: $M_{age} = 45.07$, $SD = 12.75$) and 1,332 Democrats (Experiment 5B: $M_{age} = 43.24$, $SD = 13.30$). In Experiment 5A, 488 participants reported their gender as female, 464 as male, and two as other/prefer not to disclose. In Experiment 5B, 796 participants reported their gender as female, 501 as male, 10 as other/prefer not to disclose, and 25 did not answer. As preregistered, we included only the 907 Republicans (94.1%) and 1,201 Democrats (90.2%) who passed three attention checks before random assignment and

who self-identified as a member of the Republican or Democratic party for Experiments 5A and 5B, respectively. Experiment 5 was conducted in February 2024.

Procedure

As in prior experiments, participants first gave informed consent before filling out the same demographic measures as in prior studies.

After demographics and three simple attention checks (which occurred prior to random assignment), participants learned that they would see a description of two charities (Child Success and Open Book, referred to as Charity 1 and Charity 2 hereinafter for simplicity). We informed participants that both charities focus on improving educational opportunities for kids and were currently undergoing fundraising campaigns. We selected educational charities for children as the study context because child education is something that is broadly valued across the partisan spectrum. During the study, participants were led to believe that both the charities and their donations were real; although they learned after the study was over that the charities were not real, participants were actually given an additional study bonus and a link to donate to a real charity similar to the one they learned about during the study.

All participants then read more detail about the two charities. Charity 1 (which focused on efforts to improve science, technology, engineering, and mathematics education among K–12 students) was always described as meeting expectations. More concretely, participants learned that an independent organization had awarded them two out of four stars on "impact and results" and three out of four stars on "accountability and finance." In contrast, Charity 2 (focused on efforts to improve literacy among elementary school students) was always described as underperforming expectations. More concretely, participants learned that the same independent organization had awarded them just one out of four stars on "impact and results" and only two out of four stars on "accountability and finance." Stars and descriptions were adapted from Charity Navigator, an independent organization that assesses charities worldwide.

Beyond reading a description of the two charities, participants also learned that there was a public official who had been a vocal supporter of Charity 1 (and whose reputation was therefore assumed to be impacted by the charity's success in meeting the stated fundraising goal). Unlike prior experiments, we did not manipulate the identity (and associated political relevance) of the public official: This public official was always the outgroup's leading presidential candidate at the time of the experiment (Joe Biden and Donald Trump for Experiments 5A and 5B, respectively). In all conditions, participants learned that the outcome of the fundraising campaign for Charity 1 (i.e., Child Success, the more effective charity) would reflect on the outgroup leader's reputation. As such, this created a tension between prioritizing societal good (by donating to the more effective charity) versus prioritizing group reputation (by avoiding donating to the charity associated with the outgroup leader). For example, Republicans in Experiment 5A read the following:

> Perhaps of note, Joe Biden is on the board of Child Success. Recently, he has made several statements and appearances on their behalf. He's also been a visible part of their recent fundraising campaign, and recently stated that he was committed to making sure that they met their fundraising goal. Child Success is currently behind track on this fundraising goal, and a failure to meet the target would represent a major embarrassment to Biden's image.

The key experimental manipulation was the manipulation of the salience of strategic consequences as a consequence of reputation. In the salient condition, participants read the following text, which amplified strategic incentives for the ingroup by noting the electoral consequences for the opposing presidential candidate:

> With an election less than 12 months away, being publicly associated with a failure in the fundraising campaign would look bad for Biden, feeding into negative perceptions of his ability to achieve results. It could also reinforce perceptions that Biden's policy decisions are poorly thought out. Taken together, a public failure of a fundraising campaign that Biden has linked his image to in the runup to the election is likely to hinder his re-election chances.

In the downplayed condition, they instead read the following text, which had the opposite purpose of diminishing the salience of strategic consequences as a function of reputation:

> Being publicly associated with a failure in the fundraising campaign would look bad for Biden, feeding into negative perceptions of his ability to achieve results. It could also reinforce perceptions that Biden's policy decisions are poorly thought out. However, with almost a full year until the election, it is highly unlikely that the success or failure of Child Success' fundraising campaign would have any meaningful impact on Biden's re-election chances. News cycles change quickly, and work by political scientists shows that voters are affected much more by events that occur just before voting than they are by events that take place many months before. Taken together, a public failure of a fundraising campaign Biden has linked his image to almost a full year before the election is unlikely to hinder his re-election chances.

Finally, in the neutral condition, they read just the following text, which did not mention strategic incentives as a consequence of reputation and image:

> Being publicly associated with a failure in the fundraising campaign would look bad for Biden, feeding into negative perceptions of his ability to achieve results. It could also reinforce perceptions that Biden's policy decisions are poorly thought out.

The key dependent variable in Experiment 5 was the participant's binary decision donation. They could either donate to Charity 1 (therefore prioritizing immediate societal gain, at the expense of potentially bolstering an outgroup leader's reputation) or to Charity 2 (therefore prioritizing avoiding potentially bolstering an outgroup leader's reputation, at the expense of forgoing societal gains by donating to the less effective charity). Critically, because we hold the candidate and the impact of the donation on their reputation constant—while varying the reputation-associated strategic consequences for the ingroup—we can rule out the possibility that any effects are due to participants' mere dislike of the candidate. That is, any differences across condition in donation to Charity 1 cannot be accounted for by any differences across condition in the evaluation of or the desire to be associated with the outgroup candidate. Rather, the only difference between conditions was the extent to which the strategic consequences for the ingroup's odds of winning the upcoming presidential election were made salient.

Finally, as an exploratory measure, we measured attitudes toward the outgroup's presidential candidate. Participants indicated their attitudes on a 100-point slider anchored at 0 (*very cold*) and 100

(*very warm*).[8] We also collected measure of ingroup attachment and ingroup glorification, as in prior experiments.

Results

In the downplayed condition, participants were willing to prioritize immediate societal gain: Partisans donated to Charity 1 (i.e., the more effective charity, which was associated with an outgroup leader) 60.3% of the time, a likelihood greater than chance, $t(695) = 5.58$, $p < .001$. This result is consistent with those from Experiments 1–4 demonstrating that partisans affectively differentiate between societal harms and gains.

Our key question, however, was whether such willingness would be blunted when the strategic consequences as function of reputation were made salient instead of downplayed. We conducted a logistic regression, regressing donation choice on the experimental condition. The focal contrast was between the salient and downplayed conditions. We expected donation rates to Charity 1 (i.e., the more effective charity but also the one associated with the outgroup's presidential candidate) to be lower in the salient condition compared to the downplayed condition. This was what we found: Partisans donated to the more effective charity only 51.4% of the time when the strategic consequences for their group's electoral outcomes were made salient, a significantly lower rate than 60.3% in the downplayed condition, $z = 3.33$, $p < .001$. The rates of donation in the salient condition suggest that when participants were made to reflect on the strategic downsides for their group, they were more or less at chance (i.e., 50%) between donating to a more versus less effective charity focused on child education. The difference between the salient and downplayed conditions was marginally significant for Republicans alone ($z = 1.77$, $p = .078$) and statistically significant for Democrats alone ($z = 2.91$, $p = .0036$); there was no significant interaction with party when we added party as a moderator variable ($z = .54$, $p = .59$).

In an exploratory analysis, we also assessed donation rates within the neutral condition. We expected donation rates to fall somewhere between the two primary conditions, but we did not have strong predictions for its relative placement. Donation rates to Charity 1 in this neutral condition did in fact fall between the other two conditions: Participants donated to the more effective (but politically unpalatable) charity 53.1% of the time. This likelihood was significantly different from the downplayed condition ($z = 2.74$ $p = .0062$), providing additional evidence that strategic consequences was one of the original drivers for why individuals attend to group-based reputational incentives. It was not significantly different from the salient condition ($z = 0.62$, $p = .54$), perhaps suggesting that—at least in this context—partisans may have assumed, at baseline, that group-based reputational incentives would have strategic consequences. We return to this possibility in the General Discussion section, especially as it relates to the results from Experiment 4 (where amplifying strategic incentives significantly shifted affective responses).

Discussion

Experiment 5 provided evidence for two central hypotheses. First, group-based reputational incentives can shift behavior (beyond affective responses), such that individuals become more likely to act in ways that generate suboptimal present-term societal outcomes when reputational incentives are in play (we note that these results were significant overall and among Democrats alone but marginally significant among Republicans alone). Second, one mechanism through which reputational incentives may operate to blunt differentiation between societal harms and gains is by increasing attention to downstream strategic consequences for the ingroup.

General Discussion

When groups get locked into competition, they can become highly focused on outcompeting the other side, with costly societal ramifications (Sidanius et al., 2007). Here, we consider one novel manifestation of such psychology. We consider how reputational stakes, with the potential to impact the relative strategic standing of the ingroup vis-a-vis the outgroup, impact affective differentiation between present societal harms and benefits. In the context of the competitive relationship between Democrats and Republicans, five preregistered experiments revealed that group-based reputational incentives shape partisans' responses to current societal outcomes, including economic devastation, national security threats, and health care emergencies. Strikingly, incentives for reputational standing can weaken—and sometimes even nearly eliminate—affective differentiation between present-term societal harms and benefits. Moreover, our work suggests that concerns with group reputation can cause individuals to *act* in ways that bring about suboptimal societal impact if doing so avoids burnishing the outgroup's reputation. Our work also provided evidence for the psychological process underlying these effects. A set of analyses that involved both measuring and manipulating our proposed mediator provided compelling evidence that one reason group-based reputational incentives hold motivational force is because they yield strategic gains for one's group.

Theoretical Contributions

Our work expands the scope of research on individuals' sensitivity to reputation. Prior research makes clear that individuals are strongly attuned to their individual reputations (Berman & Silver, 2022; Dorison, 2024; Lerner & Tetlock, 1999; Tetlock, 2002). For example, individuals become more likely to pay the costs to punish transgressors when their decision to punish or not punish are being observed by others (Jordan & Kteily, 2023) and are more likely to avoid politically incongruous information when observed by ingroup members (Moore et al., 2023). Along similar lines, people are more likely to engage in uncalculated cooperation—that is, to "cooperate without looking" at the costs of cooperation—when their decision to look (or not) is observable to interaction partners (Jordan et al., 2016). Research on moral self-licensing even shows that individuals actively seek to strategically create evidence of their virtue when they need a license to act less than virtuously (Effron & Conway, 2015). But whereas research on sensitivity to individual reputation and its manifestations is well-developed, there is substantially less research on individuals' sensitivity to group-based reputation and, especially, its downstream consequences. This is surprising, given the well-known importance that individuals place on their identification with social groups (Turner et al., 1979). Moreover, most of the work that has been done on group-based reputation has tended to focus on

[8] This measure did not moderate the primary results, nor did controlling for it impact the statistical significance or interpretation of them.

relatively prosocial implications. For example, research shows that individuals care about their group being seen as moral and can become more motivated (including via moral emotions like guilt or shame) to act in conciliatory or restorative ways to improve the ingroup's image (e.g., Nadler & Shnabel, 2015; Kardos et al., 2019; Lickel et al., 2011).

But humans are also highly attuned to their group's status relative to competitive outgroups (Cikara et al., 2011). A group's reputation can often have an outsized impact on its strategic standing relative to competitive outgroups (Saguy & Kteily, 2011), helping to deliver valued resources like power, legitimacy, and allyship. For example, important events that bolster the ingroup's reputation or harm the outgroup's can significantly shape electoral outcomes or the willingness of third parties to provide material or diplomatic support (consider, e.g., how events in Israel/Palestine that shape the parties' respective reputations can impact other states' willingness to stand by each of the two sides by providing economic aid or political cover in the United Nations).

Given this, we reasoned that individuals may come to place a high premium on maintaining their groups' reputational standing, especially when strategic implications are salient. Indeed, our findings highlight an important—and darker—consequence of this attunement to group reputation: the blurring of individuals' preference for outcomes that provide widespread societal benefits over harms. Strikingly, this blunted differentiation between societal gains and harms, when the societal gains are yoked to reputational damage and the societal harms to reputational gains, even extends to cases where the present-term societal impact directly affects the ingroup (and ingroup-associated causes) and not just the rest of society more broadly. Our work therefore broadens the scope of research on reputation and group-based reputation specifically, highlighting new avenues worthy of greater study and theoretical development.[9]

By showing that group-based reputation incentives can cause people to be more tolerant of present societal harm that directly impacts their *own* group as well as the outgroup, our work (along with Combs et al., 2009) also extends important existing research investigating individuals' enjoyment at outgroup pain (i.e., schadenfreude; e.g., Cikara et al., 2011; Leach et al., 2003). Furthermore, our work also contributes to the literature on intergroup conflict by suggesting a new psychological mechanism with the potential to reinforce conflict. When a group comes across as focusing more on making the other side look bad than on averting immediate societal suffering (or securing immediate societal gain), this 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 righteousness of its conflict. For example, if Democrats and Republicans come to believe that their outparty is prioritizing ingroup interests over societal harms/benefits—say, wishing for an economic calamity under the outgroup's leadership; this may come to harden their own hostility toward the outgroup and increase their own willingness to prioritize their ingroup over other considerations.

This could be true even if partisans are (more laudably) prioritizing group-based reputation over short-term societal harm because they calculate that they might use the strategic gains reputation affords to put policies in place that ensure greater long-term benefit for society. After all, individuals are unlikely to have better access to the outgroup's emotional expressions and actions than they are to their underlying reasoning, and even if they did, they might question its sincerity. When two sides believe one another to be more focused on looking good than seeking good, prospects for cooperation likely diminish. Whereas our research offers much reason to imagine that the patterns we observe could feed intergroup conflict, future work would do well to directly investigate how partisans perceive outgroup members who allow reputational incentives to shape how they respond to societal harms and gains and how they respond in kind.

Limitations and Future Directions

Our work has several limitations, which in turn serve as foundations for further study. While we found across multiple contexts (e.g., Republicans and Democrats; affective responses and financially incentivized behavior) that group-based reputational incentives can shape individuals' reactions to societal outcomes, we also found notable variation. Reputational incentives merely *weakened* affective differentiation between societal harm and benefit in Experiments 1 and 2A but nearly *eliminated* it in Experiments 2B and 4. Indeed, Supplemental Experiment 5 provides a proof-of-concept test that reputational incentives can *reverse* such affective differentiation when societal stakes are sufficiently small and reputational stakes are sufficiently big, that is, under certain conditions, individuals might not merely fail to prefer the outcome that is materially better for society on account of group-based reputational concerns, they might even favor the outcome that's materially worse for society. Future research should explore how the relative balance between the size of societal and reputational stakes determines how strongly reputational incentives shape affective responses to societal harms and gains.

Relatedly, one additional factor worth considering more systematically is the extent to which reputation has similar or different force in blunting sensitivity to societal harms versus benefits. In Experiments 1 and 2B, reputational effects were approximately equal in the context of societal benefit and harm, but reputational considerations appeared stronger in the context of societal gains (vs. harms) in Experiments 2A and 3. We report greater detail of these analyses in the Supplemental Materials, but future work should consider whether there is robust evidence for an asymmetry in the effect of reputation in the context of societal gains versus harms. Future research should also 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) at varying levels of reputational and societal stakes.

Another interesting open question is whether people explicitly walk through the logic of strategic costs and benefits as they react to reputation-relevant societal harms and gains. On the one hand, it is

[9] While there is a relative lack of research in psychology on how individuals are influenced by group-based reputational concerns, there is a body of research outside of psychology (in management) studying how firms manage their reputation. That research is focused on how to burnish or protect the reputation of a given organization by individuals who are squarely attending to the interests of their organization. In our work, individuals are presumably seeking to manage *tradeoffs* between what is good for society and what is good for their group, in a context where they might reasonably care about both aspects. Moreover, as we discuss further in the main text, our research—in contrast to most research on corporate reputation—has direct implications for intergroup hostility and conflict in society.

possible that, even when the strategic effects on their group are not overtly mentioned in the experimental materials, a large proportion of individuals explicitly reflect on the strategic calculus for their group of events that have reputational upsides (or downsides) despite present-term societal harms (or benefits). On the other hand, it is possible that individuals are not explicitly thinking through the strategic consequences for their group as they react to societal outcomes but have nevertheless, over time, internalized a heuristic sensitivity to events' group-based strategic impacts, influencing their affect (cf. Jordan & Rand, 2020, for a similar example of how individuals can internalize a sensitivity to individual reputation, sometimes heuristically acting as though they are being observed even when they are not).

To gain initial (albeit limited) insight into this question, we enlisted three research assistants to code the open-ended text responses given by participants in Experiment 4 directly after giving their affect ratings (but before being asked about attention to strategic incentives). Coding of open-ended responses preliminarily revealed that such considerations may not always be explicitly top-of-mind: Although participants' affective reactions were impacted by between-condition differences in strategic incentives rooted in group-based reputation—and recognized the strategic impact when they were later directly asked about it—they did not naturalistically report considering strategic consequences when they were first asked in our open-ended responses to broadly comment on why they had responded as they did to the affective items. They did so just 4% of the time. Even still, we are cautious about inferring evidence of absence from absence of evidence, especially given that people may not be reliable narrators of their own internal experiences. It may be that participants *did* think about the reputational incentives' impact on the strategic standing of their group, but they failed to report it in the open-ended response section either out of mere "laziness" (preferring not to expend the effort to complete the open-ended response) or because of social desirability considerations (not wanting to appear "calculated"). See full analysis details and further discussion in the Supplemental Material, including more discussion of the limits on the inferences we can draw from these analyses.

Another valuable direction for future research would be to more deeply consider what individuals seek to do with the strategic advantages that positive reputational events afford. In certain cases (e.g., when a reputational boost comes with a bump in electoral chances), events that improve the group's relative reputational standing may come with the increased power, including the power to shape social policies and material reality. Power may be affectively rewarding for several reasons, including because it is associated with status and freedom from constraints. But for at least some individuals, group power may be affectively rewarding precisely because it allows the ingroup to structure society in ways that they believe are materially better in the long run. For example, Democrats likely believe that Joe Biden winning an election over Donald Trump and gaining power would have significant positive ramifications to society, including things like a fairer economy and an enhanced social safety net; it is conceivable that Democrats with this mindset could calculate that an event that caused great suffering today but dented Trump's reputation might be worth it in the long run if it put Democrats in a position to reduce suffering going forward. Most of our experiments did not involve upcoming elections that would straightforwardly lend themselves to this logic, and the coding data described above did not clearly suggest that most individuals are engaging in such deliberative calculations most of the time. Still, it may be that some individuals—perhaps especially those who are more consequentialist in orientation—allow long-term calculations like this to influence how reputational considerations shape their affective differentiation between immediate-term societal harm and benefit.

Finally, it would be valuable to get a more precise sense of the extent to which such reasoning informs individuals' differentiation between present-term societal harms and gains. Nevertheless, even if a substantial proportion of partisans *do* engage in such long-term calculus, it bears mentioning that the psychology we uncover in our work might still have negative ramifications for society. For one, individuals may be mistaken in their calculus, overestimating the societal value that their group's policies would bring about relative to the present-term suffering. Second, even if they are accurately calibrated on the potential long-term gain of their group's policy plans, they may encounter obstacles in actually implementing these policies. Third, as mentioned above, even if partisans' underlying intentions are pure, outgroup members might fail to appreciate that this internal calculus is occurring and draw more cynical and conflict-promoting conclusions when they see partisans prioritizing strategy over present-term societal impact.

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 differentiation to present societal harms and benefits, thus potentially contributing to cycles of conflict escalation.

Constraints on Generality

All experiments were conducted with online samples of U.S. participants who self-identified as Republicans or Democrats. While we believe that the dynamics of group-based reputational incentives would apply to other highly charged partisan contests, this is not tested in the present article. In addition, while we investigated reactions to several societal events, others may present boundary conditions. Our experiments used artificially designed stimuli to maximize experimental control. It would be useful to complement the present methodological 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. Finally, 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, including conflicts like those between Israel/Palestine, India/Pakistan, and Russia/Ukraine.

References

Bandura, A. (1990). Selective activation and disengagement of moral control. *Journal of Social Issues*, *46*(1), 27–46. https://doi.org/10.1111/j.1540-4560.1990.tb00270.x

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Berman, J. Z., & Silver, I. (2022). Prosocial behavior and reputation: When does doing good lead to looking good? *Current Opinion in Psychology*, *43*, 102–107. https://doi.org/10.1016/j.copsyc.2021.06.021

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*(4), 550–558. https://doi.org/10.1037/a0018933

Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology*, *34*(3), 366–375. https://doi.org/10.1037/0022-3514.34.3.366

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. https://doi.org/10.1177/0963721411408713

Cochrane, E. (2020). Administration reaches out to Democrats on stimulus bill. *New York Times*. Retrieved August 30, 2021, from https://www.nytimes.com/2020/08/12/us/politics/coronavirus-relief-bill.html

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. https://doi.org/10.1016/j.jesp.2009.02.009

Doosje, B., Branscombe, N. R., Spears, R., & Manstead, A. S. (1998). Guilty by association: When one's group has a negative history. *Journal of Personality and Social Psychology*, *75*(4), 872–886. https://doi.org/10.1037/0022-3514.75.4.872

Dorison, C. A. (2024). *Reputational rationality theory* [Working Paper]. https://osf.io/preprints/psyarxiv/q27sj

Dorison, C. A., & Heller, B. H. (2022). Observers penalize decision makers whose risk preferences are unaffected by loss-gain framing. *Journal of Experimental Psychology: General*, *151*(9), 2043–2059. https://doi.org/10.1037/xge0001187

Effron, D. A., & Conway, P. (2015). When virtue leads to villainy: Advances in research on moral self-licensing. *Current Opinion in Psychology*, *6*, 32–35. https://doi.org/10.1016/j.copsyc.2015.03.017

Jordan, J. J. (2023). A pull versus push framework for reputation. *Trends in Cognitive Sciences*, *27*(9), 852–866. https://doi.org/10.1016/j.tics.2023.06.005

Jordan, J. J., Hoffman, M., Nowak, M. A., & Rand, D. G. (2016). Uncalculating cooperation is used to signal trustworthiness. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *113*(31), 8658–8663. https://doi.org/10.1073/pnas.1601280113

Jordan, J. J., & Kteily, N. S. (2023). How reputation does (and does not) drive people to punish without looking. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *120*(28), 1–11. https://doi.org/10.1073/pnas.2302475120

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), 57–88. https://doi.org/10.1037/pspi0000186

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. https://doi.org/10.1111/j.0956-7976.2004.00709.x

Kardos, P., Leidner, B., Castano, E., & Lickel, B. (2019). The benefits of collective responsibility: How ingroup reputation concern motivates prosociality in intergroup contexts. *European Journal of Social Psychology*, *49*(1), 93–109. https://doi.org/10.1002/ejsp.2506

Kingsley, P. (2021). Israel gives vaccines to far-off allies, as Palestinians wait. *New York Times*. Retrieved August 30, 2021, from https://www.nytimes.com/2021/02/23/world/middleeast/israel-palestinians-vaccine-diplomacy.html

Klein, E. (2024). Democrats have a better option than Biden. *New York Times*. Retrieved June 12, 2024, from https://www.nytimes.com/2024/02/16/opinion/ezra-klein-biden-audio-essay.html

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), 932–943. https://doi.org/10.1037/0022-3514.84.5.932

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*(2), 255–275. https://doi.org/10.1037/0033-2909.125.2.255

Lickel, B., Steele, R. R., & Schmader, T. (2011). Group-based shame and guilt: Emerging directions in research. *Social and Personality Psychology Compass*, *5*(3), 153–163. https://doi.org/10.1111/j.1751-9004.2010.00340.x

Logg, J. M., & Dorison, C. A. (2021). Pre-registration: Weighing costs and benefits for researchers. *Organizational Behavior and Human Decision Processes*, *167*, 18–27. https://doi.org/10.1016/j.obhdp.2021.05.006

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), 1068–1086. https://doi.org/10.1037/0022-3514.92.6.1068

Moore, M., Dorison, C. A., & Minson, J. A. (2023). The contingent reputational benefits of selective exposure to partisan information. *Journal of Experimental Psychology: General*, *152*(12), 3490–3525. https://doi.org/10.1037/xge0001463

Nadler, A., & Shnabel, N. (2015). Intergroup reconciliation: Instrumental and socio-emotional processes and the needs-based model. *European Review of Social Psychology*, *26*(1), 93–125. https://doi.org/10.1080/10463283.2015.1106712

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. https://doi.org/10.1177/1088868312440048

Pew Research Center. (2015). *Appendix A: Why partisans include 'leaners.'* Retrieved August 30, 2021, from https://www.pewresearch.org/politics/2015/11/23/appendix-a-why-partisans-include-leaners/

Rasgon, A. (2021). Israel's vaccine success unleashes debate on Palestinian inequities. *New York Times*. Retrieved August 30, 2021, from https://www.nytimes.com/2021/02/04/world/middleeast/israel-palestinians-vaccine.html

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), 698–711. https://doi.org/10.1037/0022-3514.91.4.698

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

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. https://doi.org/10.1177/0956797611412388

Schein, C., & Gray, K. (2015). The unifying moral dyad: Liberals and conservatives share the same harm-based moral template. *Personality and Social Psychology Bulletin*, *41*(8), 1147–1163. https://doi.org/10.1177/0146167215591501

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, *22*(1), 32–70. https://doi.org/10.1177/1088868317698288

Schlenker, B. R., & Weigold, M. F. (1992). Interpersonal processes involving impression regulation and management. *Annual Review of Psychology*, *43*(1), 133–168. https://doi.org/10.1146/annurev.ps.43.020192.001025

Shamir, J. (2007). *Public opinion in the Israeli-Palestinian conflict: From Geneva to disengagement to Kadima and Hamas*. United States Institute of Peace. Retrieved August 30, 2021, from https://www.usip.org/publications/2007/06/public-opinion-israeli-palestinian-conflict-geneva-disengagement-kadima-and

Sidanius, J., Haley, H., Molina, L., & Pratto, F. (2007). Vladimir's choice and the distribution of social resources: A group dominance perspective. *Group Processes & Intergroup Relations*, *10*(2), 257–265. https://doi.org/10.1177/1368430207074732

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). *A 21 word solution*. SSRN.

Snyder, C. R., Lassegard, M., & Ford, C. E. (1986). Distancing after group success and failure: Basking in reflected glory and cutting off reflected failure. *Journal of Personality and Social Psychology*, *51*(2), 382–388. https://doi.org/10.1037/0022-3514.51.2.382

Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review*, *109*(3), 451–471. https://doi.org/10.1037/0033-295X.109.3.451

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), 853–870. https://doi.org/10.1037/0022-3514.78.5.853

Turner, J. C., Brown, R. J., & Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology*, *9*(2), 187–204. https://doi.org/10.1002/ejsp.2420090207

Received June 29, 2023
Revision received June 12, 2024
Accepted June 25, 2024 ■