Polarized Imagination:
Partisanship Influences the Direction and Consequences of Counterfactual Thinking

Kai Epstude
University of Groningen

Daniel A. Effron
London Business School

Neal J. Roese
Kellogg School of Management, Northwestern University

Correspondence should be addressed to:
Kai Epstude
University of Groningen
Department of Psychology
Grote Kruisstraat 2/1
9717TS Groningen
The Netherlands
Email: k.epstude@rug.nl

Abstract

Four studies examine how political partisanship qualifies previously-documented regularities in people's counterfactual thinking ($N = 1,186$ Democrats and Republicans). First, whereas prior work finds that people generally prefer to think about how things could have been better instead of worse (i.e., entertain counterfactuals in an upward vs. downward direction), Studies 1a–2 find that partisans are more likely to generate and endorse counterfactuals in whichever direction best aligns with their political views. Second, previous research finds that the closer someone comes to causing a negative event, the more blame that person receives; Study 3 finds that this effect is more pronounced among partisans who oppose (vs. support) a leader who "almost" caused a negative event. Thus, partisan reasoning may influence which alternatives to reality people will find most plausible, will be most likely to imagine spontaneously, and will view as sufficient grounds for blame.

(149 words)

KEYWORDS: Counterfactual thinking, mental simulation, political partisanship, motivated reasoning, moral judgment

Polarized Imagination:

Partisanship Influences the Direction and Consequences of Counterfactual Thinking

Today's politically polarized climate provides ample examples of a social-psychological truism: Partisans can witness the same situation but disagree on what actually happened (Ditto et al., 2019; Leeper & Slothus, 2014). The present research explores a subtler manifestation of political polarization: Partisans considering the same situation may disagree on what *could have* happened. In other words, partisanship may not only influence the *facts* people believe, but also the *counterfactuals* they imagine – and the conclusions they draw from these counterfactuals.

Counterfactuals are mental simulations of "what might have been" – imagined alternatives to past outcomes that might have occurred if circumstances had been different Byrne, 2016; Roese, 1997). Counterfactuals often support causal inferences, particularly when the counterfactual takes the form of an "if-then" conditional (Epstude & Roese 2008). Psychologically, counterfactual thoughts are consequential in that they help individuals to learn from mistakes and to plan for the future (Epstude & Roese, 2011; Roese & Epstude, 2017). Politically, they allow societies to judge the effectiveness of past actions, assign praise or blame to leaders, and determine how much to support various policies (e.g., Catellani & Covelli, 2013).

Yet counterfactuals also provide fertile ground for partisan reasoning. Counterfactual events, by definition, did not occur and hence cannot be verified (Tetlock & Lebow, 2001). People may imagine counterfactual events that fit with their beliefs and motivations, asserting it "almost happened" without the risk of being proven wrong. We propose that, when people reflect counterfactually about political events, political partisanship predicts *what they imagine* as well as *what they infer* – in other words, both the content and the conclusions of their counterfactual thinking.

With respect to *content*, we propose that partisanship predicts preferences for imagining a better or worse alternative to reality. For example, when considering what the U.S. economy would be like if President Trump had not cut taxes in 2017, Democrats might be more likely to imagine that the economy would have been better, whereas Republicans might be more likely to imagine it would have been worse. With respect to *conclusions*, we propose that people not only endorse *different* counterfactuals, but also draw different inferences from the *same* counterfactual. In general, people may regard a negative event that nearly happened as sufficient evidence that the relevant leader deserves blame. But more specifically, partisanship may moderate this relation. For example, even if Democrats and Republicans agreed that Trump "almost" provoked a war with North Korea in 2017, Democrats might be more likely than Republicans to see this close counterfactual as sufficient grounds for blaming Trump. By examining the content and conclusions of counterfactual thinking, we reveal an underappreciated source of political polarization. The next sections develop our theorizing about counterfactual content and conclusions.

Partisanship and Counterfactual Content

Prior research identifies several empirical regularities in the counterfactual thoughts that people generate and endorse. One regularity concerns a counterfactual's *direction of comparison* (Epstude & Roese, 2008; Roese, 1997). Most counterfactual thoughts focus on how the past could have been better instead of worse – i.e., counterfactuals tend to be *upward* rather than *downward* (Roese & Epstude, 2017). This preference for upward counterfactuals connects to what people want and strive for (Gamlin et al., 2020). Upward counterfactuals suggest ways of achieving a better state of affairs and hence reflect a functional orientation toward personal improvement (Roese & Epstude, 2017). Counterfactual thinking is typically activated by

unexpected negative events (e.g., failing a test), and upward counterfactual thinking helps people determine how to avoid such events in the future (e.g., "I would have passed if only I had studied"). Downward counterfactuals can help people feel better about negative events (e.g., White & Lehman, 2005), but in practice are rarely generated.

A second empirical regularity is that personal knowledge and preferences shape counterfactual thinking (e.g., Byrne, 2016). An important source of such knowledge and preference is political partisanship. Accordingly, partisanship predicts which counterfactuals people find most plausible (Tetlock & Visser, 2000). For example, partisans are more likely to believe that a lie "could have been true" if it aligns with their political views (Effron, 2018).

However, previous work neither predicts nor tests how these two empirical regularities may relate to each other. When partisans consider how a political event or policy could have been different, do they gravitate towards upward counterfactuals (Roese & Epstude, 2017) or towards counterfactuals in whichever direction happens to fit with their political views (Tetlock & Visser, 2000)? We predict the latter. In other words, we propose that political partisanship creates an important boundary condition on people's general preference for upward counterfactual thinking.

There are at least two reasons to predict partisan effects on the direction of counterfactual thinking. First, different partisans are motivated to reach different conclusions. Whereas laypeople's counterfactual thinking may be motivated by a desire to improve following a negative event, resulting in upward counterfactual thoughts that point the way to betterment, partisans' counterfactual thinking may instead be more motivated by a desire to justify and defend their political views, resulting in more flexibility about the direction of counterfactual thinking. Much like people selectively search their memories for evidence consistent with

preferred beliefs (e.g., Santioso et al, 1990) or test the impact of selected economic choices in

line with their political preference (Caddick & Rottman, 2021), partisans may thus selectively

imagine counterfactuals in whichever direction is consistent with preferred political conclusions.

Second, different partisans have different knowledge, beliefs, and assumptions. As a result, they

may accept and generate different counterfactual thoughts – not through motivated reasoning,

but through rational (Bayesian) thinking. In practice, these processes are difficult to disentangle

and may operate in concert (Tappin et al., 2020; Tetlock & Levi, 1982). Based on these theorized

processes, we hypothesized:

> ***Hypothesis 1a (H1a):*** *Partisans flexibly accept and generate either upward or downward*
>
> *counterfactuals that are consistent with the preferred ideological stance or inconsistent*
>
> *with the opposed ideological stance.*
>
> The alternative hypothesis we tested was:
>
> ***Hypothesis 1b (H1b):*** *Regardless of their preferred ideological stance, partisans will*
>
> *accept and generate upward counterfactuals more than downward counterfactuals.*

Partisanship and Counterfactual Conclusions

Partisanship may be related not only to the direction of counterfactual thinking, but also

to the conclusions people draw from counterfactuals in a particular direction. Our focus is on

blame, a politically consequential moral judgment linked to counterfactual thinking (see Byrne,

2016). Another empirical regularity about counterfactual thinking is that the closer someone

comes to causing a negative event, the more blame that person receives (Johnson, 1986; Miller &

McFarland, 1986). Citizens might blame their leader more for bringing their nation within

minutes of an avoidable nuclear war than for bringing them within months of such a war, even

though in both cases no war occurred. At the same time, it is ambiguous how much

counterfactual closeness should figure into blame judgments. *How* much more blame does a leader deserve for bringing a nation within minutes versus months of a nuclear war? We propose that this ambiguity offers a degree of flexibility that partisan reasoning can exploit. Specifically, the closer people think a negative event came to occurring, the more they may blame the relevant leader – but especially if they oppose (vs. support) that leader. In this way, partisanship may moderate the impact of counterfactual closeness on moral judgment.

This conceptualization extends the idea that when people are motivated to reach a conclusion, they set lower evidentiary standards for reaching it (Dawson et al., 2002). For example, participants examined the qualifications of a disliked person less thoroughly than those of a person they liked (Ditto & Lopez, 1992). Facts provide stronger evidence than counterfactuals. But when a particular conclusion aligns with a partisan's views, he or she may be less likely to require factual evidence of the conclusion; counterfactual evidence may suffice. For example, counterfactual thinking played a bigger role in partisans' judgments of media hypocrisy when the partisans were motivated to dismiss the media as hypocritical than when they were not (Helgason & Effron, 2022).

Extending these ideas, we propose that the closeness of a downward counterfactual seems like a more compelling reason to blame someone when the target of blame is someone a partisan opposes (vs. supports). Specifically, when partisans think poorly of a leader, they may be more inclined to blame that leader for a negative outcome that did not occur, but nearly did. In this sense, partisanship sets the evidentiary standards people use for drawing conclusions about blame. Based on these theorized processes, we hypothesized:

Hypothesis 2 (H2): *The positive relationship between the closeness of an undesirable counterfactual event and blame is stronger when partisans oppose (vs. support) the target of blame.*

Prior Research on How Beliefs and Motivations Shape Counterfactual Thinking

Inside and outside the political domain, people's beliefs and motivations predict their counterfactual thoughts (e.g., Catellani & Covelli, 2013; Effron, 2018; Milesi & Catellani, 2012; Spellman & Mandel, 1999). For example, when motivated to prove their moral character, people will invent "counterfactual transgressions" – bad deeds they imagine they could have done, but did not actually do (Effron et al., 2012, 2013). Policy experts found historical counterfactuals (e.g., how the Cold War could have turned out differently) to be more plausible if those counterfactuals aligned with their views (Tetlock, 1998, Tetlock & Visser, 2000). And fundamentalist Christians were averse to even considering counterfactuals that challenged their religious beliefs (Tetlock et al., 2000).

Unlike the present research, however, these prior studies systematically examined neither counterfactual direction, nor the relation between counterfactual closeness and blame judgments. We further advance this prior research by demonstrating how partisanship qualifies previously-documented regularities in counterfactual thinking. First, we find that the effect of partisanship on counterfactual thinking is so strong that it overrides people's general preference to for upward (vs. downward) counterfactual thoughts (Roese & Epstude, 2017). Second, we demonstrate that partisanship moderates people's general tendency to use counterfactual closeness as a cue for whom to blame. Finally, by testing our hypotheses with a range of contemporary political issues, we reveal that partisan divisions extend beyond what people believe has actually happened; they also encompass what people think could have happened, and who is to blame for it.

The Present Research

Four studies tested our hypotheses among Democrats and Republicans. Testing H1a, Studies 1a and 1b examined whether partisans would rate both upward and downward counterfactuals as more plausible when these counterfactuals aligned with their politics, and Study 2 examined whether partisans would be more likely to generate counterfactuals in whichever direction (upward vs. downward) was more aligned with their views on a given political issue. These studies also tested alternative hypothesis H1b, that partisans would show a general preference for upward counterfactual thinking. Finally, testing H2, Study 3 examined whether partisans are more likely to blame a president for a negative event when they think it "almost" occurred – especially when they oppose that president.

Open Practices

We pre-registered all studies, determined stopping rules for data collection before running each study, and report all measures, conditions, and data exclusions. Verbatim study materials, data, analysis code, and links to pre-registration documents are posted at https://osf.io/3m6p7/?view_only=3c12b1bb164344dcb4cb1f4791d624bf

Studies 1a and 1b

Studies 1a and 1b examined how partisans judge the plausibility of political counterfactuals. Democrats and Republicans rated the plausibility of six counterfactuals: half upward and half downward, half aligned and half misaligned with participants' political views. In Study 1a, the upward counterfactuals aligned with Democrats' views and all downward counterfactuals aligned with Republicans' views. In Study 1b, we reversed the linkage, such that all upward counterfactuals aligned with Republicans' ideology and all downward counterfactuals aligned with Democrats' ideology. Flipping the connection between partisanship and

counterfactual direction tests H1a (that the partisan thinker will accept either an upward or downward counterfactual depending on its alignment with their politics) and H1b (that partisans will instead prefer upward over downward counterfactuals).

Method

Participants

We recruited American partisans from Prolific Academic (see Online Supplement for screening criteria and exclusions). Study 1a's final sample size was $N = 201$ (101 Democrats, 100 Republicans; 97 men, 103 women, and 1 non-binary person; M age = 40 years, $SD = 13$; 83% White, 7% Black, 5% American Indian or Alaska Native, 6% Latino/Latina) and Study 1b's was $N = 192$ (100 Democrats and 92 Republicans; 93 men, 93 women, 5 nonbinary people, and 1 person who declined to indicate gender; M age = 35 years, $SD = 14$; 66% White, 12% Asian, 11% Black, 6% Latino/a, and the remainder other races and ethnicities).

Procedure

Both studies presented six political topics with corresponding counterfactuals (see Online Supplement, Table S1). Three counterfactuals were upward and three downward, with counterbalancing of direction with topic. In Study 1a, all upward counterfactuals aligned with Democrats' views (e.g., "If Trump had not passed the tax cuts, then the economy would currently be much better") and all downward counterfactuals aligned with Republicans' views (e.g., "If Trump had not passed the tax cuts, then the economy would currently be much worse"). In Study 1b, flipped this pairing, such that all downward counterfactuals aligned with Democrats' views (e.g., "If Trump had been able to pass even bigger tax cuts, then the economy would currently be much worse") and all upward counterfactuals aligned with Republicans'

views (e.g., "If Trump had been able to pass even bigger tax cuts, then the economy would currently be much better").

Measures

Participants rated the plausibility of each counterfactual via 3 items: agreement, appropriateness, and plausibility ($\alpha = .91$ in both Studies 1a and 1b). As an ancillary measure, participants also rated how angry the counterfactual made them (see Online Supplement for results). As a manipulation check, participants rated the counterfactuals' compatibility with their beliefs. Response options ranged from *extremely appropriate, plausible* etc. (5) to *extremely inappropriate, implausible* etc. (1). For exploratory purposes, participants rated the importance of each topic and reported strength of party identification (using a 4-item scale adapted from Leach et al., 2008); relevant findings appear in the Online Supplement.

Results

Analytic Strategy

We submitted each measure to a mixed regression model with fixed effects for condition (1 = aligned, 0 = misaligned), fixed effects for the six political issues, and random intercepts for participants to account for the repeated-measures design.[1] The mixed models in this and all subsequent studies were computed in Stata 16 using the *mixed* command, which assumes an independent variance-covariance structure, employs maximum likelihood estimation (ML), and tests coefficients against the z distribution. We pre-registered one-tailed significance tests of directional predictions (conclusions were identical with two-tailed tests).

Manipulation Check

---

[1] In this and all subsequent studies, we pre-registered item as a fixed effect due to the small number of items (i.e., $k <= 8$), but the conclusions were always identical when we instead treated item as a random effect.

Confirming the success of our manipulation, participants rated the counterfactuals that were meant to be aligned with their politics as more compatible with their beliefs than the counterfactuals that were meant to be misaligned with their politics. This result emerged in both Study 1a (aligned: ($M = 3.78$, $SD = .86$; misaligned: $M = 2.40$, $SD = .99$), $b = 1.38$, $SE = .07$, $z = 19.23$, $p < .001$, $d_z = .97$, and Study 1b (aligned: $M = 3.33$, $SD = .91$; misaligned: $M = 2.26$, $SD = .86$), $b = 1.08$, $SE = .07$, $z = 14.54$, $p < .001$, $d_z = .81$.

Counterfactuals Seemed More Plausible When Aligned With One's Politics

In Study 1a, participants thought a counterfactual was more plausible when it was aligned with their politics ($M = 3.86$, $SD = .77$) than when it was misaligned with their politics ($M = 2.70$, $SD = .92$), $b = 1.16$, $SE = .06$, $z = 18.78$, $p < .001$, $d_z = .98$. This finding confirms our main prediction for this study. However, it is unclear whether the results reflect a preference for upward counterfactuals among Democrats or a more general tendency for partisans to prefer whichever direction of counterfactual happens to appeal to their partisan view.

Study 1b disentangles these interpretations by reversing which counterfactual direction aligned with which party's views. The results showed a flexible acceptance of either upward or downward counterfactuals by partisans. That is, participants rated a counterfactual as more plausible when it was aligned with their politics ($M = 3.46$, $SD = .83$) than when it was misaligned with their politics ($M = 2.47$, $SD = .77$), $b = .99$, $SE = .07$, $z = 14.87$, $p < .001$, $d_z = .80$. Note that this relation is the same as in Study 1a, despite the fact that in Study 1b, upward counterfactuals were aligned with Republicans' views and downward counterfactuals were aligned with Democrats' views. Taken together, Studies 1a and 1b suggest that partisans prefer whichever counterfactual direction aligns with their views.

Preference for Upward vs. Downward Counterfactuals Depended on Partisanship

Recall the previous research documents a general preference for upward over downward counterfactual thinking (e.g., Roese & Epstude, 2017). Were the partisan effects in the present studies strong enough to swamp this general preference? Figure 1's results suggest that the answer is yes. We analyzed these results by submitting plausibility ratings to a mixed model with fixed effects for the counterfactual's direction (1 = up, 0 = down), participants' political party (1 = Republican, 0 = Democrat), and their interaction, fixed effects for item, and random intercepts for participant. We then computed the simple slope of counterfactual direction for each political party (these analyses were not pre-registered).

The results showed that in Study 1a, when *upward* counterfactuals were aligned with Democrats' views, upward counterfactuals were rated as more plausible than downward counterfactuals among Democrats (see Figure 1's top panel; M_{up} = 4.03 vs. M_{down} = 2.38, SDs = 1.05 and 1.12, respectively), b = 1.65, SE = .08, z = 19.42, $p < .001$, consistent with the general pattern found in previous research. However, Republicans showed the reverse pattern, rating downward counterfactuals as more plausible than upward counterfactuals (M_{up} = 3.02, M_{down} = 3.70, SDs = 1.26 and 1.15, respectively, b = –.67, SE = .09, z = 7.90, $p < .001$. The interaction between party and direction was significant, b = –2.32, SE = .12, z = 19.33, $p < .001$.

The results were the mirror image in Study 1b, in which *downward* counterfactuals were aligned with Democrats' views (see Figure 1's bottom panel). Now Republicans showed the usual pattern, rating upward counterfactuals as more plausible than downward counterfactuals (M_{up} = 3.30, M_{down} = 2.77, SDs = 1.18 and 1.24, respectively), b = .26, SE = .09, z = 2.84, p = .005, whereas Democrats showed the reverse pattern, rating down counterfactuals as more plausible than upward counterfactuals (M_{up} = 2.20, M_{down} = 3.86, SDs = .98 and 1.17,

respectively), $b = -1.67$, $SE = .09$, $z = 19.21$, $p < .001$. The interaction was significant, $b = 1.93$,

$SE = .13$, $z = 15.38$, $p < .001$.

In short, the results of Studies 1a and 1b indicate that the plausibility of an upward or

downward counterfactual depends on how well it fits with partisan political views, supporting

H1a. This effect was apparently strong enough to override people's general tendency to prefer

upward over downward counterfactuals, thus offering no support for H1b.

Figure 1

Mean Plausibility Judgments, ± 95% CI, by Political Party in Studies 1a and 1b

(figure continues on next page)

Study 1b: upward counterfactuals aligned with Republicans
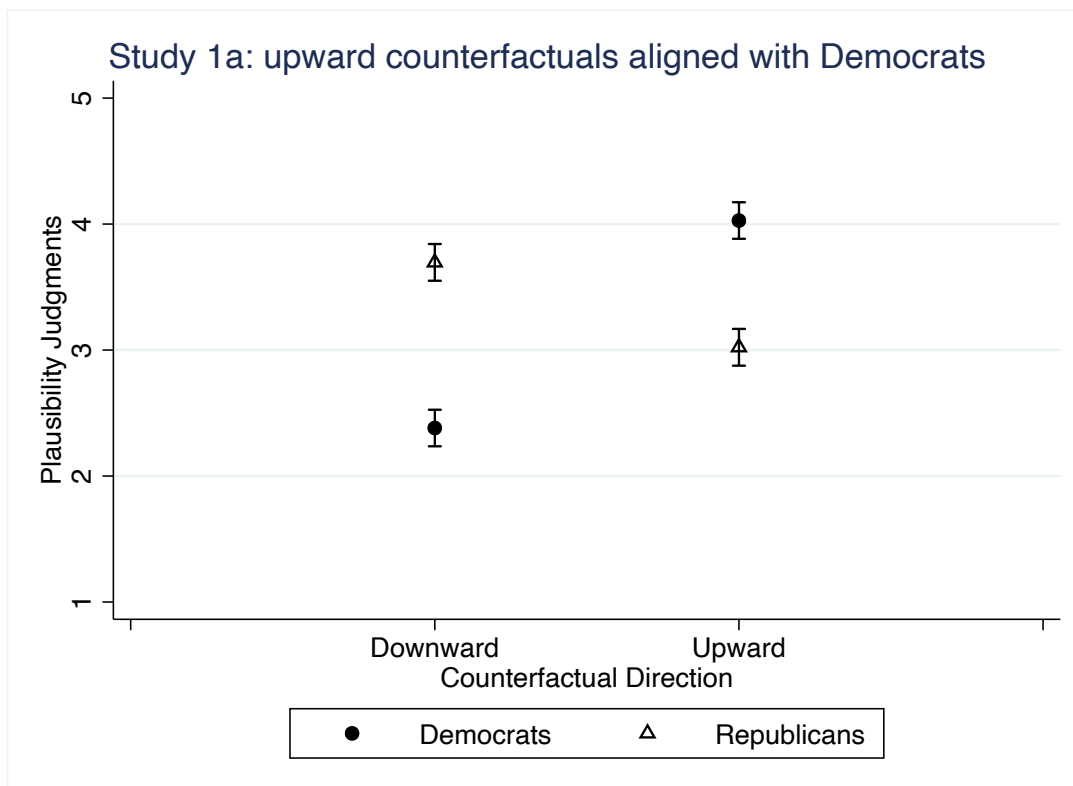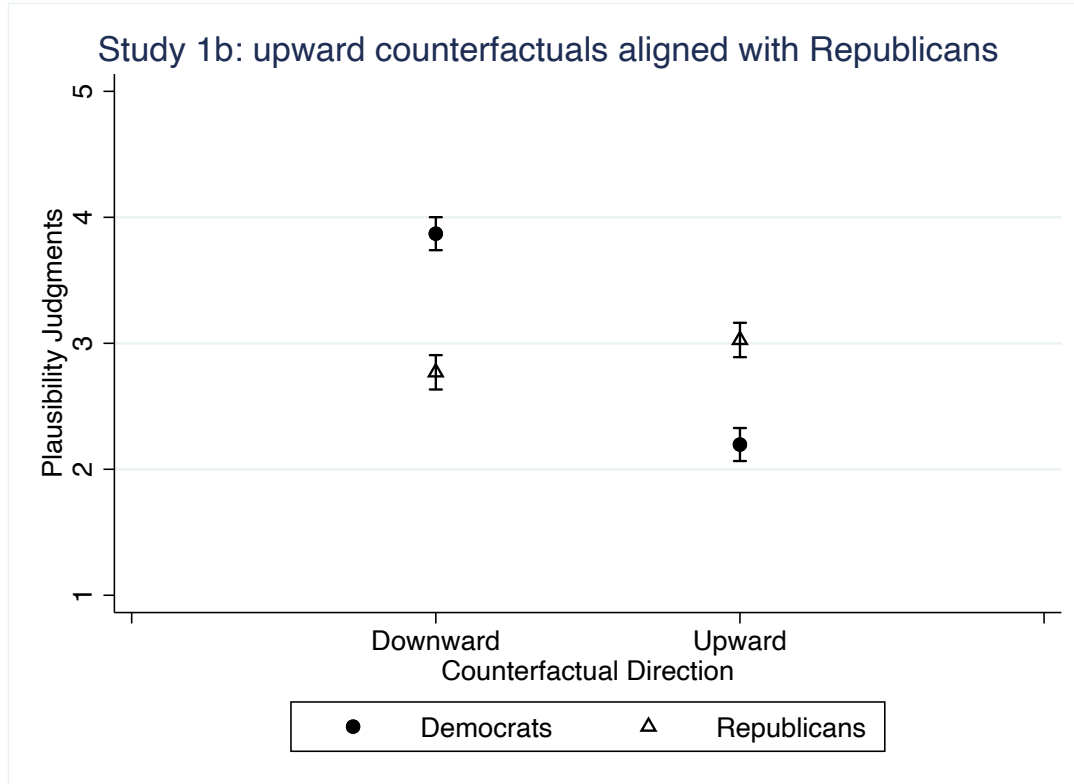
Note. *M*s are predictive margins from the mixed regression models described in the main text.

Study 2

Study 2 provided a further test of H1a and H1b. Participants wrote counterfactuals about eight political events, and we counterbalanced whether the most salient upward or downward counterfactuals aligned with Democrats' or Republicans' views. Whereas Studies 1a and 1b showed that partisans flexibly *accept* pre-written counterfactuals in whichever direction aligns with their politics, Study 2 tests whether partisans will also *generate their own* counterfactuals in the direction aligned with their politics (H1a) – or whether they will consistently generate upward instead of downward counterfactuals (H1b).

Method

Participants

We recruited American partisans from Prolific Academic (see Online Supplement for screening criteria and exclusions). The final sample was $N = 190$ (100 Democrats, 90 Republicans; 152 women, 33 men, and 5 nonbinary people; M age = 26 years, $SD = 8$; 76% White, 9% Black, 6% Asian, 6% Latina/o, remainder other races).

Procedure

Participants read 8 descriptions of contentious political issues. After each issue, we asked participants to complete a partially written counterfactual statement (free response). Specifically, each statement provided participants with a counterfactual's antecedent (i.e., "if [something had been different] …") and participants needed to write the consequent (i.e. "then …"; see Online Supplement, Table S2). The statements varied such that in half, upward (vs. downward) counterfactuals aligned with Democratic views and in the other half they aligned with Republican views. For example, one item was, "If Senate Republican's hadn't blocked Obama's appointee for the Supreme Court …" Here, an upward consequent aligns more with Democrats' views (e.g., "then things would have been better") whereas a downward consequent aligns more with Republicans' views (e.g., "then things would have been worse"). Another item was, "If Republicans had been able to pass the tax cut earlier than 2017…". Here, an upward consequent aligns more with Republicans' views, whereas a downward consequent aligns more with Democrats' views. In addition to this counterbalancing of ideological-direction alignment, we also counterbalanced whether the antecedent was an action or an inaction.

After participants responded to the 8 statements, they responded to the dependent measure. Specifically, participants viewed their prior responses to each statement, and indicated whether their responses focused on how the situation could have been better (indicating an

upward counterfactual), worse (indicating a downward counterfactual), or neither. Finally,

participants rated the importance of each topic as well as their party-identification strength using

the same scales as in the previous studies (see Online Supplement for relevant results).

Results and Discussion

Supporting H1a, participants were more likely to generate counterfactuals in whichever

direction was more aligned with their partisan views. As Figure 2 shows, participants generated

more upward counterfactuals (66.45%) than downward counterfactuals (21.84%) when it was

upward counterfactuals that aligned with their views. When downward counterfactuals aligned

with their views, they instead generated more downward (67.50%) than upward (21.84%)

counterfactuals. The proportion of counterfactuals classified as neither downward nor upward

was similar across conditions (11.97% vs. 11.71%). Note that these results suggest that partisan

thinking in this context overrides the general preference for upward over downward

counterfactual thinking suggested by prior research, thus offering no support for H1b.

Figure 2

Study 2: Percentage of Upward and Downward Counterfactuals Generated in Each Condition, ± 95% CI
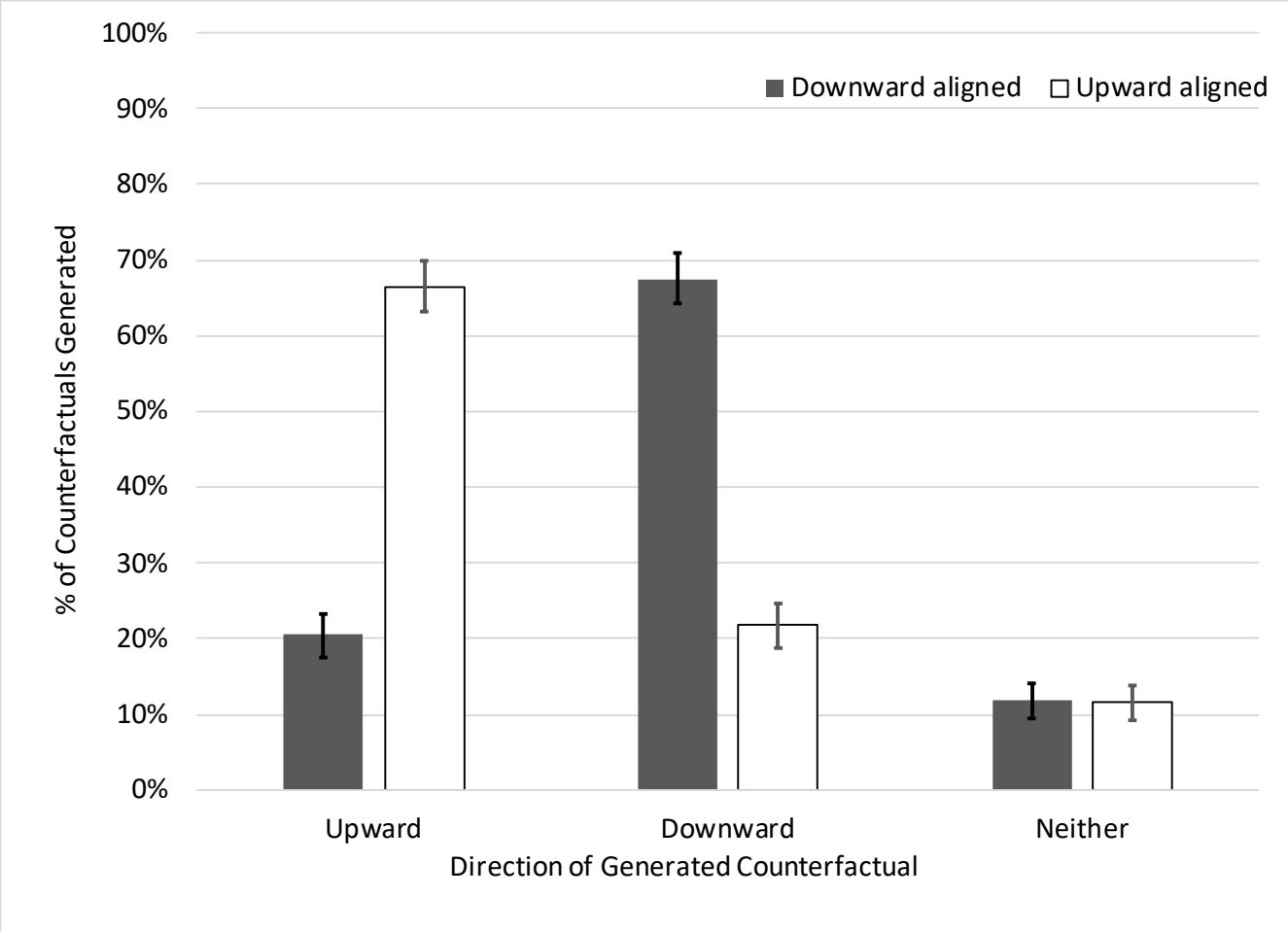
To test the significance of this pattern, we submitted the dependent measure (counterfactual direction; upward = 1, downward = 0; "neither up nor down" responses omitted) to a mixed logistic regression model with fixed effects for condition (1 = upward aligned, 0 = downward aligned), fixed effects for the eight political topics, and random intercepts for participants. The results showed a significant effect of condition with an $OR > 1$, indicating that people are more likely to generate upward (vs. downward) counterfactuals when upward (vs. downward) counterfactuals are aligned with their politics, $OR = 19.17$, $SE = 3.45$, $z = 16.40$, $p < .001$.

Disaggregating the results by political party in an exploratory analysis showed no evidence of a general preference for upward counterfactuals among members of either party (see Figure 3). When upward counterfactuals aligned with Democrats' views, they were far more likely to generate upward counterfactuals (82.75%) than downward counterfactuals (7.00%), but when downward counterfactuals aligned with Democrats' views, they were far more likely to generate downward counterfactuals (82.25%) than upward counterfactuals (9.75%), $OR = 119.52$, $SE = 81.11$, $z = 7.05$, $p < .001$ for the effect of condition when analyzing Democrats' data with the mixed model described above. Republicans showed the same result: When upward counterfactuals aligned with their views, they were more likely to generate upward (48.33%) than downward (38.33%) counterfactuals – but when downward counterfactuals aligned with their views, it was downward counterfactuals that they generated more frequently than upward (51.11% vs. 32.50%), $OR = 12.80$, $SE = 5.22$, $z = 6.25$, $p < .001$ for the effect of condition when analyzing Republicans' data. (Interestingly, similar to Studies 1a and 1b, the effect of condition was stronger among Democrats than Republicans. That is, when we analyzed all the data with the mixed model described above, adding a fixed effect for party [1 = Republican, 0 = Democrat]

and its interaction with condition, the interaction was significant, $OR = .02$, $SE = .01$, $z = 12.55$, $p < .001$). In summary, Study 2 showed that participants were more likely to generate counterfactuals in whichever direction that was consistent with their partisan views on a particular issue, thus supporting H1a over H1b.

Figure 3

Study 2: Percentage of Upward and Downward Counterfactuals Generated in Each Condition, by Democrats and Republicans, ± 95% CI
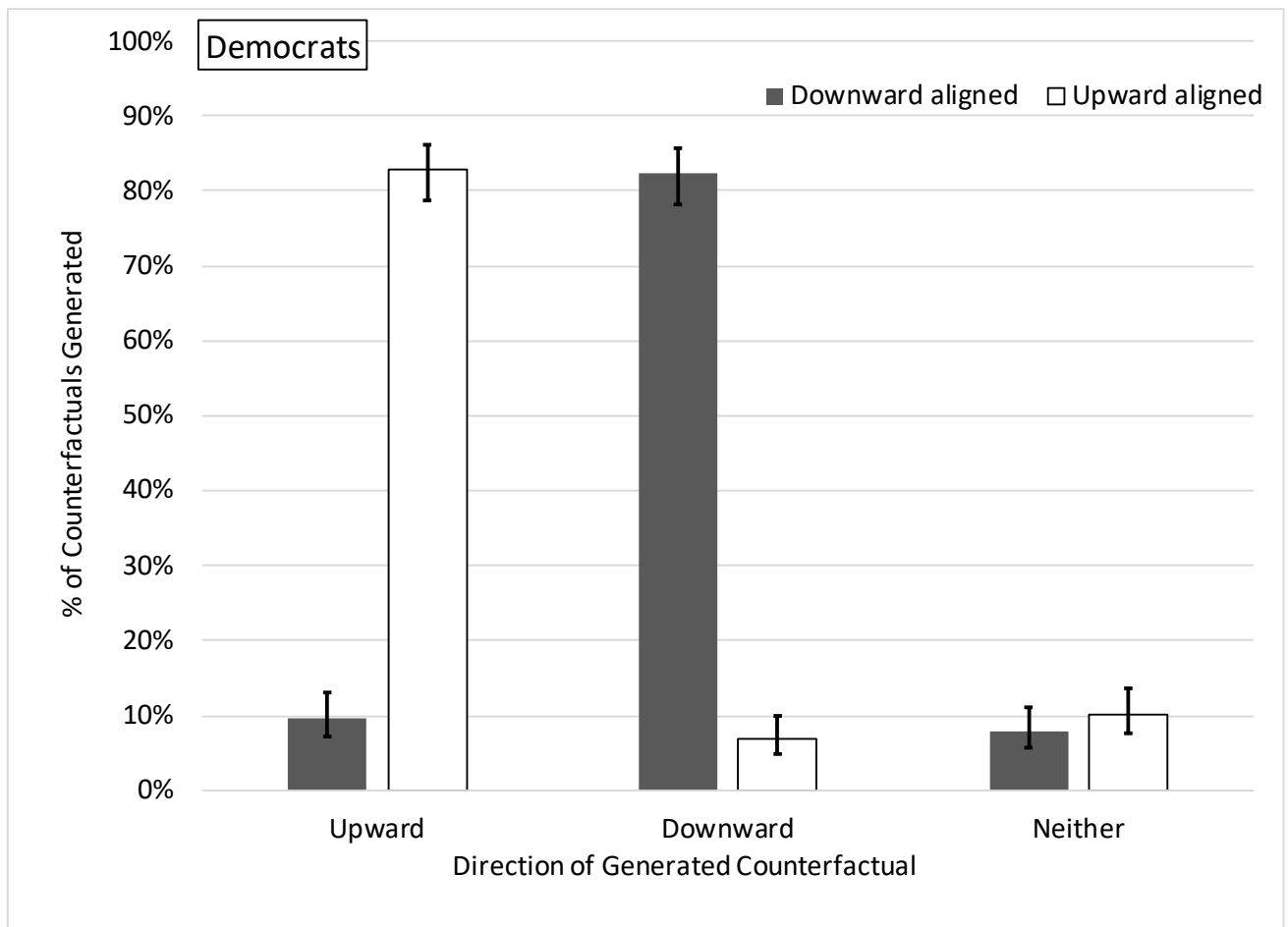
(figure continues on next page)
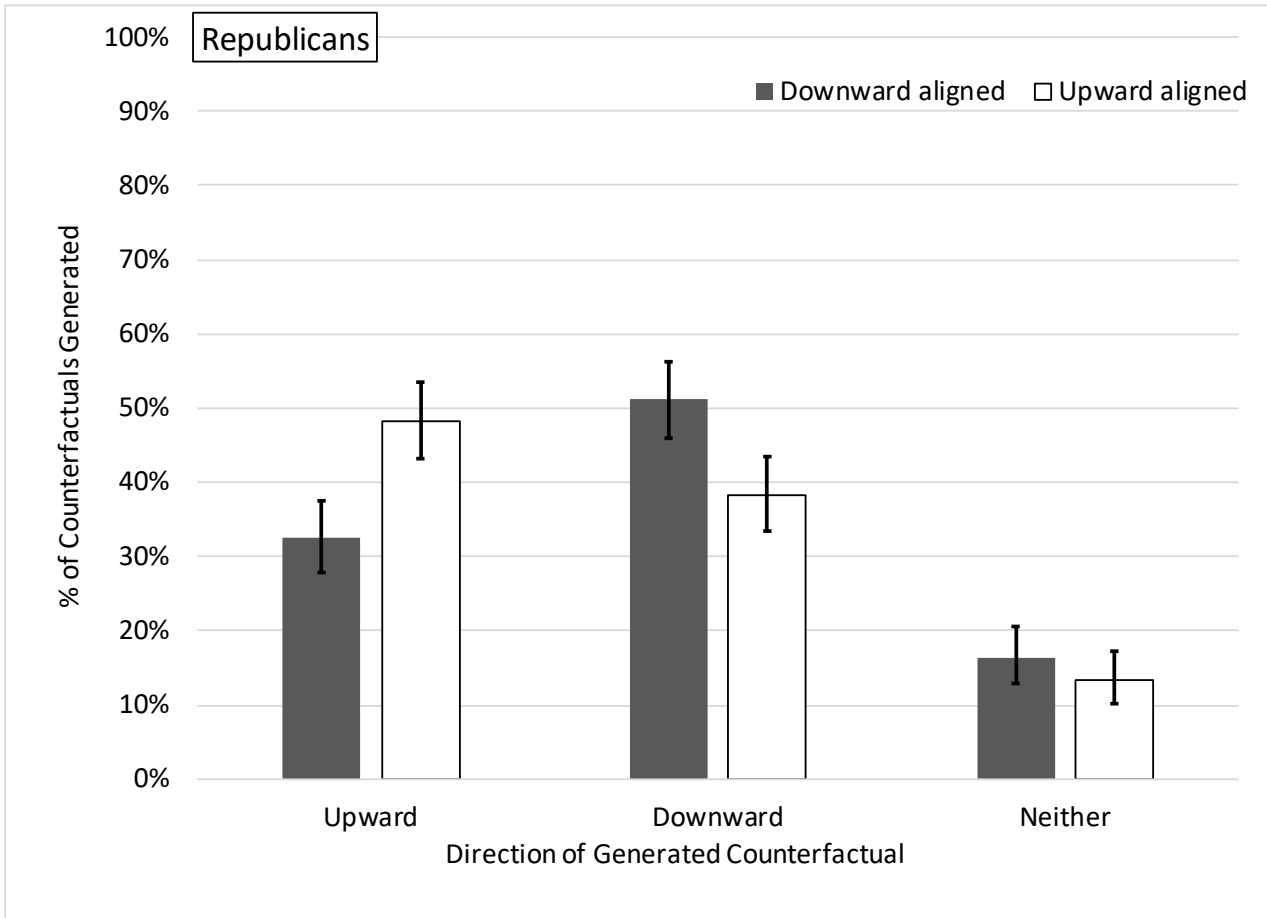
Study 3

Whereas the previous studies showed that partisanship predicted the direction of comparison of counterfactual thinking (H1a and H1b), Study 3 tested whether partisanship would moderate the relationship between judgments of counterfactual closeness and blame (H2). How much will partisans blame a leader for a negative event that "almost occurred" on that leader's watch? Study 3 asked American partisans to consider downward counterfactual events, some of which could have occurred during the Trump presidency (e.g., war with North Korea), and some of which could have occurred during the Biden presidency (e.g., renewed war with the Taliban). We then assessed the relationship between how close partisans thought these events

came to occurring and how much they blamed the leader who was president when the events

could have occurred. We expected that the closer people thought the event came to occurring, the

more they would blame the president – but especially if they opposed (vs. supported) that

president.

Method

Participants

After applying our pre-registered exclusion criteria (see Online Supplement), the final

sample was $N = 603$ American participants recruited from Prolific Academic (595 of whom

provided demographics; 354 women, 234 men, and 7 nonbinary; M age = 30 years, $SD = 11$; 305

Trump voters and 304 Biden voters; 74% White, 12% Black, 6% Latino/a, 4% Asian, remainder

other races and ethnicities).

Procedure

Participants read 8 brief descriptions of negative political events that did not happen, such

as the U.S. and North Korea going to war in the summer of 2017 (see Online Supplement, Table

S3). We chose these counterfactual events because we expected variance in how "close"

participants would think the events came to occurring. Half of the events plausibly could have

occurred during the term of a president participants supported (i.e., Biden or Trump, depending

on participants' politics), whereas the other half could have occurred during the term of a

president participants opposed (i.e., Trump or Biden, depending on participants' politics).

Participants evaluated how close the event came to occurring (1 = *not close at all* to 7 =

extremely close) and how much the president at the time (i.e., Trump or Biden) should be blamed

or praised for "nearly" allowing or causing the negative event. Then we administered some

exploratory analyses (emotional reactions, issue importance, how good or bad the counterfactual

outcome would have been, and political party identification), which we discuss in the Online

Supplement. Participants also reported demographics.

Results

Analytic Approach

We submitted each dependent measure to a mixed regression model with fixed effects for

the president being judged (1 = supported; 0 = opposed), counterfactual closeness (1-7 scale),

their interaction, fixed effects for the 8 items, and random intercepts for participants.[2]

Blame

Our main hypothesis was that when people considered a president they opposed, the

closer they believed a negative event came to occurring under his watch, the more they would

blame him, and that this effect would be attenuated (or even reversed) when people considered a

president they supported. In other words, we predicted a stronger positive relation between

counterfactual closeness and blame when participants had opposed (vs. supported) the relevant

president (H2).

As predicted, we observed a significant interaction between judgments of counterfactual

closeness and whether participants supported or opposed the relevant president, $b = -.12$, $SE =$

$.03$, $z = 4.70$, $p < .001$. Decomposing this interaction with simple slopes revealed the predicted

pattern (shown in Figure 4; for violin plot, see Figure S5 in the Online Supplement). When

people considered the president they opposed, there was a strong positive relationship between

counterfactual closeness and blame, $b = .54$, $SE = .02$, $z = 28.02$, $p < .001$. When people

considered the president they supported, there was also a positive relation between closeness and

---

[2] We coded whether participants supported or opposed each president based on prescreen data indicating whom they had voted for (see Online Supplement). However, the conclusions were the same when we instead coded based on participants' responses to a question about presidential support administered after the dependent measures.

blame, *b* = .42 *SE* = .02, *z* = 22.43, *p* < .001, however (as shown by the interaction term reported

above), it was significantly attenuated.

Figure 4

Study 3: Stronger Relationship Between Counterfactual Closeness and Blame When Participants

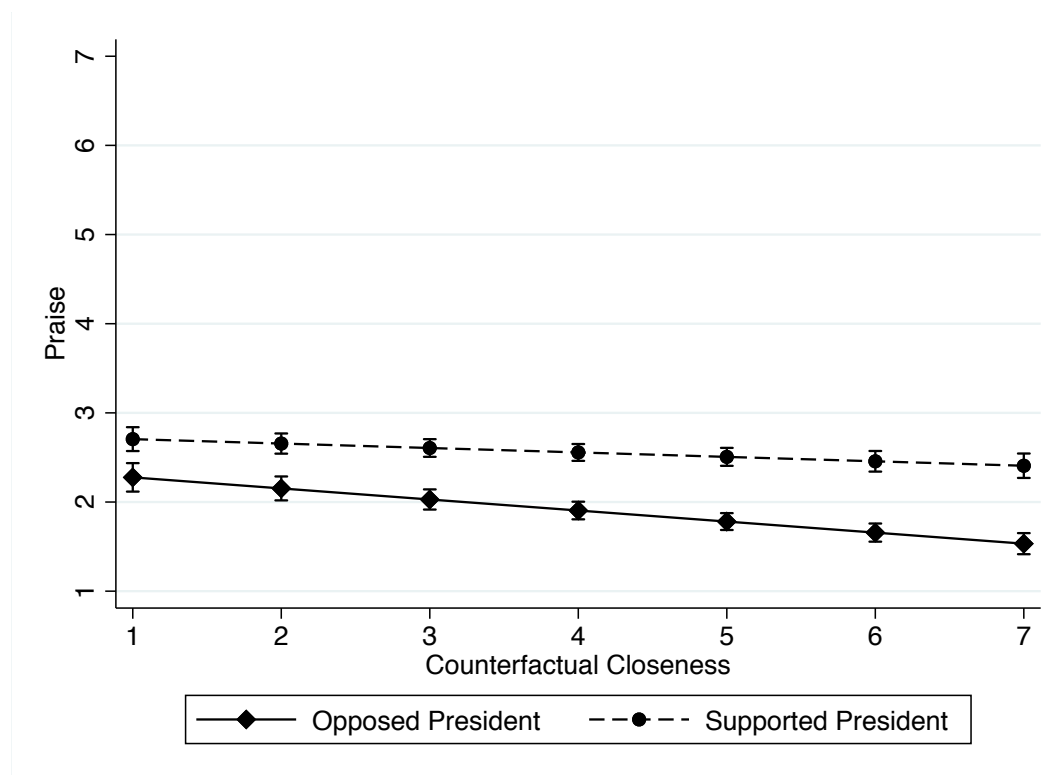*Judged a President They Opposed (vs. Supported)*
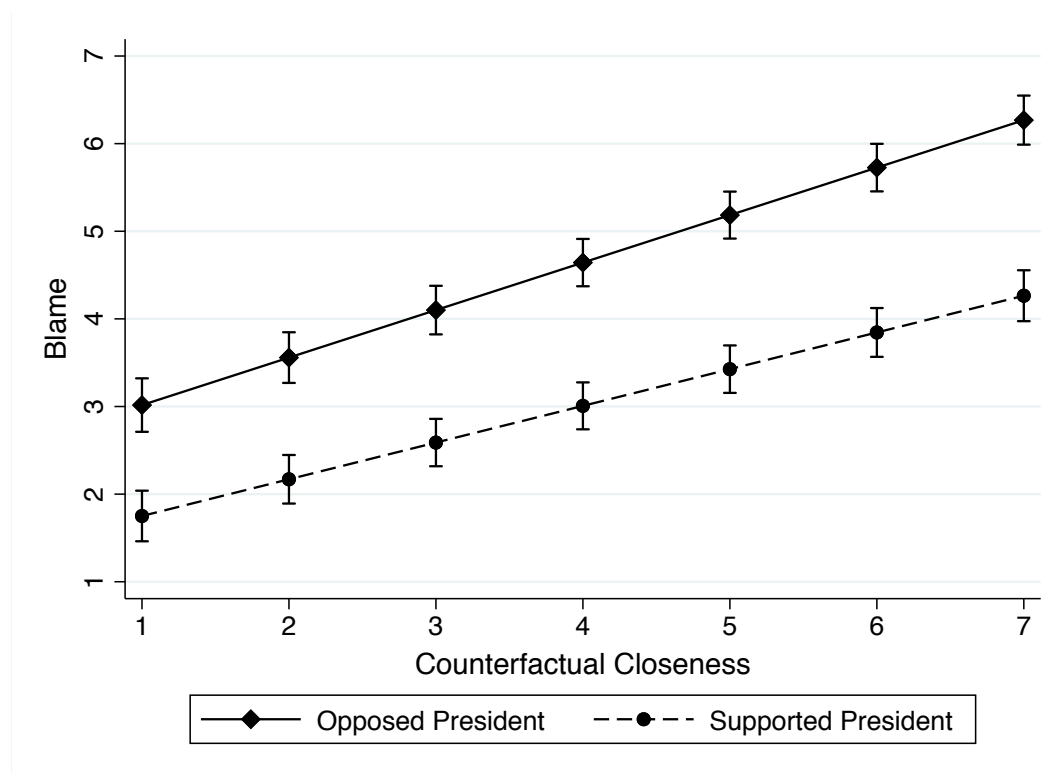
Note. The values are predictive margins, with 95% CIs, from the mixed regression model.

Praise

The results for praise judgments complemented the results for blame judgments. As Figure 5 shows, the closer participants thought a negative event had come to occurring, the less they praised the respective president, especially if they opposed that president (see Figure S6 in the Online Supplement for violin plot). This pattern was significant, as shown by the interaction term in the mixed model described above, $b = .07$, $SE = .02$, $z = 3.29$, $p < .001$. There was a negative relationship between closeness and praise judgments regardless of whether participants supported or opposed the president, but this relationship was significantly stronger when participants opposed the president, $b = -.12$, $SE = .02$, $z = 7.29$, $p < .001$, than when they supported the president, $b = -.05$, $SE = .02$, $z = 3.01$, $p = .003$. We pre-registered a tentative prediction that greater closeness could mean *more* praise (for avoiding the negative event) from partisans who supported the relevant president, but the results did not support this prediction.

Figure 5

Study 3: Stronger Relationship Between Counterfactual Closeness and Praise When Participants Judged a President They Opposed (vs. Supported)

Note. The values are predictive margins, with 95% CIs, from the mixed regression model.

Moderation by Political Orientation

Exploratory analyses found that political orientation moderated the results in different ways for our two dependent measures.

Recall that the closer participants thought a negative event had come to occurring, the more they *blamed* the relevant president, but only if they opposed that president. This effect was entirely driven by *Trump* voters (see Figure 6). When we added a dummy code for the specific president participants supported to the mixed model described above (0 = Trump supporters, 1 = Biden supporters), we found a significant three-way interaction between counterfactual

closeness, whether participants were judging a president they supported (coded 1) versus

opposed (coded 0), and whether participants were Trump voters (coded 0) or Biden voters

(coded 1), $b = .23$, $SE = .05$, $z = 4.38$, $p < .001$.

Recall also that that the closer participants thought a negative event had come to

occurring, the less they *praised* the relevant president – but only if they opposed that president.

This effect was entirely driven by *Biden* voters (see Figure 7), as shown by a significant three-

way interaction when submitting the praise measure to the mixed model just described, $b = .19$,

$SE = .05$, $z = 4.15$, $p < .001$.

It is unclear whether this pattern of results reflects something about Biden versus Trump

supporters, or about the specific political issues or negative counterfactual events to which the
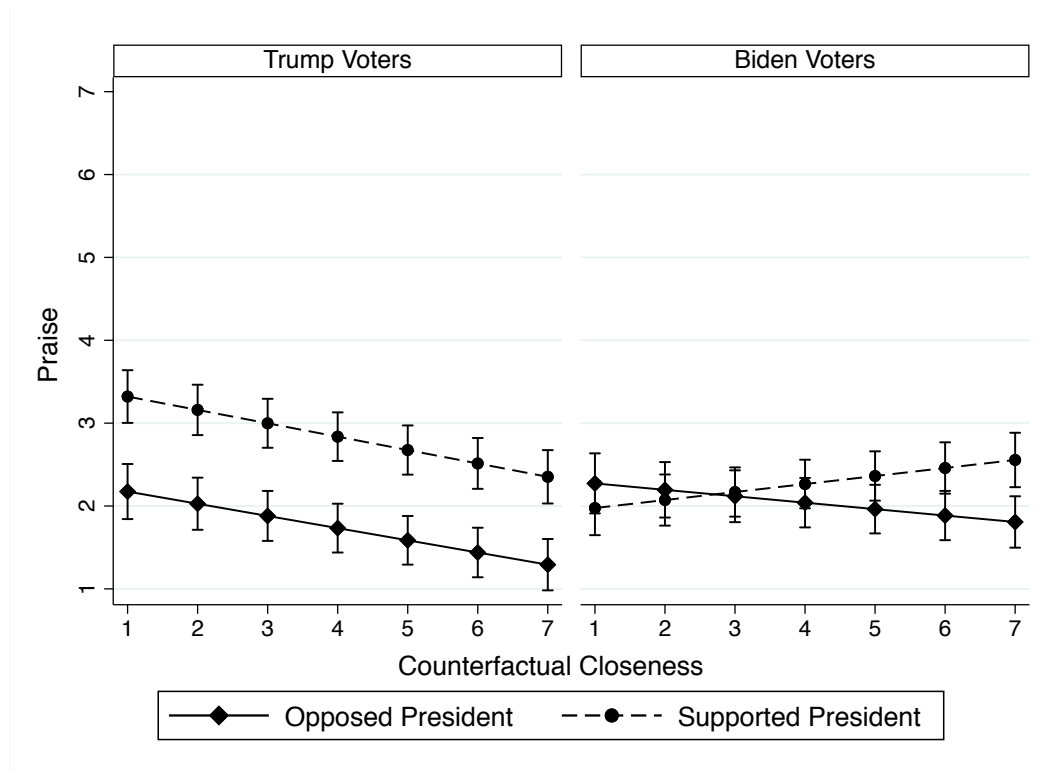
stimuli referred.

Figure 6

Study 3: Trump Voters Drive the Predicted Pattern for the Blame Measure
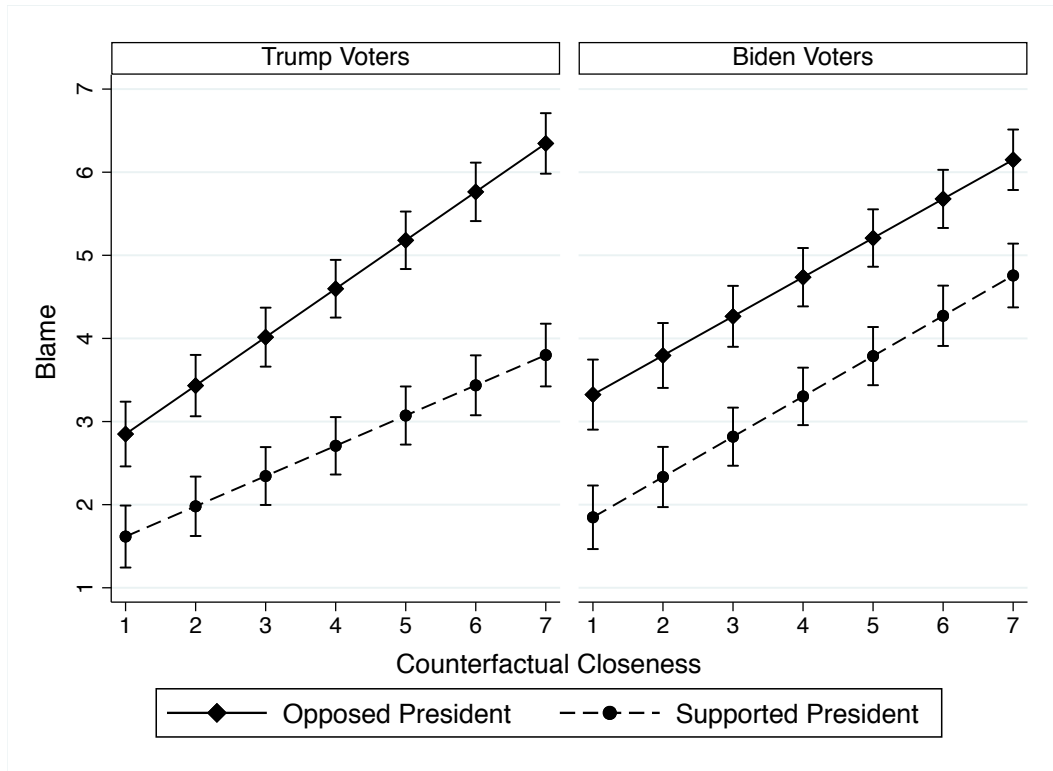
Note. The values are predictive margins, with 95% CIs, from the mixed regression model.

Figure 7

Study 3: Biden Voters Drive the Predicted Pattern for the Praise Measure

Note. The values are predictive margins, with 95% CIs, from the mixed regression model.

Discussion

Study 3 suggests that partisanship moderates the relationship between counterfactual thinking and moral judgments of blame and praise. The closer a negative event came to occurring on a president's watch, the more harshly partisans blamed that president, particularly when they had opposed him. When partisans already think poorly of a leader, they are more likely to blame that leader for a negative outcome that did not occur, but nearly did.

General Discussion

Our four studies shed new light on how partisan beliefs relate to counterfactual thinking. Partisans find a given counterfactual more plausible when it aligns with their views (Studies 1a and 1b), selectively generate counterfactuals that align with their views (Study 2), and deploy counterfactuals that support preferred moral judgments about leaders (Study 3). In sum, partisanship predicts both the content and the conclusions of counterfactual thoughts.

Our research makes several theoretical contributions. Our main contribution is to demonstrate how partisanship qualifies two empirical regularities in the counterfactual-thinking literature. First, whereas previous research demonstrated an overwhelming preference for upward over downward counterfactuals (e.g., Roese & Olson, 1993), Studies 1a–2 found a complete reversal of this preference when downward counterfactuals aligned with participants' views. That is, partisans in our studies flexibly generated and endorsed counterfactuals in whichever direction best aligned with their political views on a particular issue (supporting H1a over H1b). One explanation is that prior research tended to examine situations in which people were motivated by a desire to discover "how things could be better," whereas partisans tend to be more motivated by a desire to justify and defend their political views.

The second empirical regularity we qualify is that the closer someone comes to causing a negative event, the more blame that person receives (e.g., Miller & McFarland, 1986). Study 3 replicated this effect, but also showed that it is more pronounced among partisans who oppose (vs. support) a leader who "almost" caused a negative event (supporting H2). One explanation is that when people dislike a leader, they lower their standards for what constitutes evidence of that leader's blameworthiness, giving more weight to imagined events – what *could have* happened under the leader's watch.

In short, partisan reasoning may influence which alternatives to reality people will find most plausible, will be most likely to spontaneously imagine, and will view as sufficient grounds for blame – thus creating important boundary conditions on previously-documented effects.

Another theoretical contribution is that Study 3 advances understanding of counterfactual thinking's role in moral judgment (see Byrne, 2017). In some cases, downward counterfactual thinking connects to more-lenient moral judgments (Markman et al., 2008) – a *contrast* effect. For example, participants felt licensed to act in a less-than-virtuous manner after they reflected on the sinful actions they could have (but did not) performed (Effron et al., 2012; Effron et al., 2013). In other cases, downward counterfactual thinking results in harsher moral judgments (Miller et al., 2005) – an *assimilation effect*. Study 3 suggests that the extent to which downward counterfactual thinking produces harsher moral judgments depends on partisanship. When partisans disliked a president, downward counterfactual thinking was more tightly associated with blaming that president. That is, the closer people thought a negative event came to occurring, the more likely they were to blame the president, especially if the president was opposed by the partisan. Our findings thus raise the possibility that motivation influences how much of an assimilation effect result from downward counterfactual thinking. Future research should further examine this possibility.

Third, our results contribute to a debate about whether conservatives are more prone to cognitive biases than are liberals (cf. Baron & Jost, 2019; Brandt & Crawford, 2020; Ditto et al., 2019). Our results suggest that partisanship connects to counterfactual thinking among people at both ends of the political spectrum (i.e., Democrats and Republicans). That said, our results contain nuance. In Studies 1a–2, Democrats and Republicans alike were more inclined to endorse and generate counterfactuals that were aligned (vs. misaligned) with their views – but

this effect was larger among Democrats. In Study 3, people's tendency to *blame* a president they opposed for negative events that nearly happened was larger among Trump supporters than Biden supporters – but participants' tendency to *praise* a president they supported for having averted negative events was larger among Biden than Trump supporters. Future research should assess the generality of these patterns and pinpoint why they emerge. However, our results do not support the possibility that, when it comes to counterfactual thinking, conservatives show more partisan bias than do liberals.

As noted, our results are consistent with the idea that partisans engage in motivated counterfactual thinking. That is, the content and conclusions of their counterfactual thinking may reflect their desire to justify their political beliefs and to blame leaders they oppose. However, like most partisan effects in political psychology (Tappin et al., 2020), ours could also be explained by non-motivated processes. For example, Republicans could be more likely than Democrats to think that "things would have been better" without a particular Democratic policy because Republicans have been exposed to more information about that policy's shortcomings. Meanwhile, Democrats could be more likely than Republicans to blame Trump for "almost" causing war with North Korea because only Democrats are more likely to have the prior that Trump makes bad decision. Of course, the priors and indeed the information to which partisans have been exposed may themselves have motivated origins, which illustrates the challenge of distinguishing motivated from purely cognitive processes (see Tetlock & Levi, 1982). In practice, both types of processes may work together (Kunda, 1990).

Conclusion

As Tetlock and Visser (2000) observed, "counterfactual thinking is often heavily theory-driven" (p. 174). Our results add new nuance to the recognition that partisanship constitutes an

important aspect of theory-driven counterfactual thinking. People's political views predict which

alternatives to reality they will find most plausible, will be most likely to spontaneously imagine,

and will view as sufficient evidence of a conclusion. Partisans do not only disagree about facts –

they disagree about counterfactuals and their implications for moral judgment. In today's

political climate, it is not just our attitudes that are polarized – it is also our imaginations.

References

Baron, J., & Jost, J. T. (2019). False equivalence: Are liberals and conservatives in the United States equally biased? *Perspectives on Psychological Science,* 14, 292-303. https://doi.org/10.1177/1745691618788876

Brandt, M. J., & Crawford, J. T. (2020). Worldview conflict and prejudice. *Advances in Experimental Social Psychology, 61*, 1-66. https://doi.org/10.1016/bs.aesp.2019.09.002

Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology, 67*, 7.1-7.23. doi: 0.1146/annurev-psych-122414-033249

Byrne, R. M. J. (2017). Counterfactual thinking: From logic to morality. *Current Directions in Psychological Science, 26*, 314-322. https://doi.org/10.1177/0963721417695617

Caddick, Z. A., & Rottman, B. M. (2021). Motivated Reasoning in an Explore-Exploit Task. *Cognitive Science*, *45*(8), e13018. https://doi.org/10.1111/cogs.13018

Catellani, P., & Covelli, V. (2013). The strategic use of counterfactual communication in politics. *Journal of Language and Social Psychology, 32(4),* 480-489. https://doi.org/10.1177/0261927X13495548

Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated reasoning and performance on the Wason selection task. *Personality and Social Psychology Bulletin*, *28*(10), 1379-1387. https://doi.org/10.1177/014616702236869

Ditto, P. H., Liu, B. S., Clark, C. J., Wojcik, S. P., Chen, E. E., Grady, R. H., Celniker, J. B., & Zinger, J. F. (2019). At least bias is bipartisan: A meta-analytic comparison of partisan bias in liberals and conservatives. *Perspectives on Psychological Science, 14*, 273-291. doi: 10.1177/1745691617746796

Effron, D. A. (2018). It could have been true: How counterfactual thoughts reduce condemnation

    of falsehoods and increase political polarization. *Personality and Social Psychology*

    Bulletin, 44, 729-745. https://doi.org/10.1177/0146167217746152

Effron, D. A., Miller, D. T., & Monin, B. (2012). Inventing racist roads not taken: The licensing

    effect of immoral counterfactual behaviors. *Journal of Personality and Social*

    Psychology, 103, 916-932. https://doi.org/10.1037/a0030008

Effron, D. A., Monin, B., & Miller, D. (2013). The unhealthy road not taken: Licensing

    indulgence by exaggerating counterfactual sins. *Journal of Experimental Social*

    Psychology, 49, 573-578. https://doi.org/10.1016/j.jesp.2012.08.012

Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality*

    and Social Psychology Review, 12, 168-192. https://doi.org/10.1177/1088868308316091

Epstude, K., & Roese, N. J. (2011). When goal pursuit fails: The functions of counterfactual

    thought in intention formation. *Social Psychology, 42*, 19-27.

    https://doi.org/10.1027/1864-9335/a000039

Gamlin, J., Smallman, R., Epstude, K., & Roese, N. J. (2020). Dispositional optimism weakly

    predicts upward, rather than downward, counterfactual thinking: A prospective

    correlational study using episodic recall. *PLoS ONE,* 15(8): e0237644.

    https://doi.org/10.1371/journal.pone.0237644

Helgason, B., & Effron, D. A. (2022). From critical to hypocritical: Counterfactual thinking

    increases partisan disagreement about media hypocrisy. *Journal of Experimental Social*

    Psychology, 101, 104308. https://doi.org/10.1016/j.jesp.2022.104308

Johnson, J. T. (1986). The knowledge of what might have been: Affective and attributional

    consequences of near outcomes. *Personality and Social Psychology Bulletin, 12*, 51-62.

    https://doi.org/10.1177/0146167286121006

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480–

    498. https://doi.org/10.1037/0033-2909.108.3.480

Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B.,

    Ouwerkerk, J. W., & Spears, R. (2008). Group-level self-definition and self-investment:

    A hierarchical (multicomponent) model of in-group identification. *Journal of Personality*

    and Social Psychology, 95(1), 144-165. https://doi.org/10.1037/0022-3514.95.1.144

Leeper, T. J., & Slothuus, R. (2014). Political parties, motivated reasoning, and public opinion

    formation. *Advances in Political Psychology*, 35, 129-156. doi: 10.1111/pops.12164

Markman, K. D., Mizoguchi, N., & McMullen, M. N. (2008). "It would have been worse under

    Saddam:" Implications of counterfactual thinking for beliefs regarding the ethical

    treatment of prisoners of war. *Journal of Experimental Social Psychology, 44,* 650-654.

    https://doi.org/10.1016/j.jesp.2007.03.005

Milesi, P., & Catellani, P. (2011). The day after an electoral defeat: Counterfactuals and

    collective action. *British Journal of Social Psychology, 50*, 690-706.

    https://doi.org/10.1111/j.2044-8309.2011.02068.x

Miller, D. T., & McFarland, C. (1986). Counterfactual thinking and victim compensation: A test

    of norm theory. *Personality and Social Psychology Bulletin, 12,* 513-519.

    https://doi.org/10.1177/0146167286124014

Miller, D. T., Visser, P. S., & Staub, B. D. (2005). How surveillance begets perceptions of dishonesty: The case of the counterfactual sinner. *Journal of Personality and Social Psychology, 89*, 117-128. https://doi.org/10.1037/0022-3514.89.2.117

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin, 121*, 133-148. https://doi.org/10.1037/0033-2909.121.1.133

Roese, N. J., & Epstude, K. (2017). The functional theory of counter- factual thinking: New evidence, new challenges, new insights. *Advances in Experimental Social Psychology, 56*, 1-79. https://doi.org/10.1016/bs.aesp.2017.02.001

Roese, N. J., & Olson, J. M. (1993). The Structure of Counterfactual Thought. Personality and Social Psychology Bulletin, 19(3), 312–319. https://doi.org/10.1177/0146167293193008

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology, 59*(2), 229–241. https://doi.org/10.1037/0022-3514.59.2.229

Spellman, B. A., & Mandel, D. R. (1999). When possibility informs reality: Counterfactual thinking as a cue to causality. *Current Directions in Psychological Science, 8*, 120-123. https://doi.org/10.1111/1467-8721.00028

Tappin, B. M., Pennycook, G., & Rand, D. G. (2020). Thinking clearly about causal inferences of politically motivated reasoning: Why paradigmatic study designs often undermine causal inference. *Current Opinion in Behavioral Sciences, 34,* 81-87. https://doi.org/10.1016/j.cobeha.2020.01.003

Tetlock, P. E. (1998). Close-call counterfactuals and belief-system defenses: I was not almost wrong but I was almost right. *Journal of Personality and Social Psychology, 75*, 639-652. https://doi.org/10.1037/0022-3514.75.3.639

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,* 853-970. https://doi.org/10.1037/0022-3514.78.5.853

Tetlock, P. E., & Lebow, R. N. (2001). Poking counterfactual holes in covering laws: Cognitive styles and historical reasoning. *American Political Science Review, 95*, 829-843. https://doi.org/10.1017/S0003055400400043

Tetlock, P. E., & Levi, A. (1982). Attribution bias: On the inconclusiveness of the cognition-motivation debate. *Journal of Experimental Social Psychology, 18*(1), 68-88. https://doi.org/10.1016/0022-1031(82)90082-8

Tetlock, P. E., & Visser, P. S. (2000). Thinking about Russia: Plausible pasts and probable futures. *British Journal of Social Psychology, 39*, 173-196. https://doi.org/10.1348/014466600164417

White, K., & Lehman, D. R. (2005). Looking on the bright side: Downward counterfactual thinking in response to negative life events. *Personality and Social Psychology Bulletin, 31,* 1413-1424. https://doi.org/10.1177/0146167205276064