Polarized Imagination
Epstude, Kai; Effron, Daniel; Roese, Neal J.

Published in: Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences

DOI: 10.1098/rstb.2021.0342

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date: 2022

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
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

Author Note

Kai Epstude and Daniel A. Effron share the first authorship for this manuscript.
This research was partially supported by a grant to Daniel A. Effron from the Center for the Science of Moral Understanding at UNC Chapel Hill.
Pre-registrations, verbatim materials, data, and analysis code are posted at

https://osf.io/3m6p7/

We thank Stephanie Rodriguez and Andrei Viziteu for research assistance.

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
We propose that political partisanship not only influences the facts people believe, but also the counterfactuals (alternatives to reality) they imagine – and the conclusions they think these counterfactuals suggest. Four studies offer support ($N = 1,186$ Democrats and Republicans). Partisans were more likely to generate and endorse counterfactuals that were aligned (vs. misaligned) with their political views (Studies 1a–2). We found no evidence that Democrats and Republicans differentially preferred imagining better vs. worse alternatives to reality. Also, when partisans opposed (vs. supported) a president, they were more likely to blame that leader for a negative outcome that did not occur, but “nearly” did (Study 4). Thus, 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 grounds for blame – a pattern of motivated reasoning that may both reflect and reinforce political polarization.

(148 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. 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.

Counterfactual thoughts are mental simulations of “what might have been” – imagined alternatives to past outcomes that might have occurred if circumstances had been different. Counterfactual thoughts often support causal inferences, particularly when the counterfactual takes the form of an “if-then” conditional (Epstude & Roese 2008). Psychologically, counterfactual thoughts are consequential because 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 weigh support for or opposition to policies (e.g., Catellani & Covelli, 2013).

We propose that counterfactual thinking both *reflects* and *reinforces* partisan concerns. It reflects partisan concerns in two respects. First, partisans may judge a given counterfactual as more plausible when it aligns with their views than when it does not (Effron, 2018; Tetlock, 1998). For example, Democrats should be more likely than Republicans to accept that *If President Trump had not tightened U.S. immigration policies, then the U.S. would have a stronger labor force.* Counterfactual events, by definition, did not occur and hence cannot be
verified (Tetlock & Lebow, 2001). Thus, people are free to reach partisan conclusions about whether a counterfactual event “could have happened” without the risk of being proven wrong.

Second, partisans should be more likely to generate counterfactuals that align with their views than counterfactuals that do not. For example, when contemplating what would have happened if Trump had not tightened U.S. immigration policies, Democrats might be more likely to spontaneously imagine that *the U.S. would not currently be facing a labor shortage* whereas Republicans might imagine *crime would have risen in border cities*. Imagination is limitless in theory – but in practice, the counterfactuals that come most readily to mind are those that fit with what individuals know and expect (Roese & Epstude, 2017).

Counterfactual thinking *reinforces* partisan concerns by helping individuals reach conclusions that are aligned with their views. Not only do partisans generate and accept the plausibility of *different* counterfactuals, they also draw different conclusions from the *same* counterfactual. For example, even if Democrats and Republicans agreed that Trump “almost” provoked a war with N. Korea during his presidency, Democrats might be more likely than Republicans to see this close counterfactual as sufficient grounds for blaming Trump. More generally, people may regard a negative event that nearly happened as sufficient evidence that the relevant leader deserves blame. This hypothesis extends the idea that when people are motivated to reach a conclusion, they set lower evidentiary standards for reaching it (Dawson et al., 2002). When partisans already think poorly of a leader, they may blame that leader for a negative outcome that did not occur, but nearly did.

To summarize, we propose that (1) partisans will find a given counterfactual more plausible when it aligns with their views, (2) partisans are more likely to generate counterfactuals that align with their views, and (3) partisans are more likely to blame leaders they oppose (vs.
support) for “almost” allowing a negative event to occur. We next elaborate on these propositions, explain how they advance prior work on counterfactuals and motivated reasoning.

**Theoretical Background**

We consider two key dimensions on which counterfactuals vary. *Direction of comparison* distinguishes between whether the past could have been better (i.e., an upward counterfactual) or worse (i.e., a downward counterfactual). *Judgment impact* involves the conclusions people draw from a counterfactual. We discuss how partisanship may be related to each dimension in turn.

**Partisanship and Direction of Comparison**

With regard to direction, we contrast two theoretical views that yield different predictions. Partisanship may constrain the direction of counterfactual thinking across the board. In this *directional constraint* view, liberals (e.g., contemporary American Democrats) may show a preference for upward counterfactuals whereas conservatives (e.g., contemporary American Republicans) may show a preference for downward counterfactuals. The intuition is that liberalism focuses on progress and improvement, whereas conservatism focuses on preservation of the status quo and those practices deemed traditional. A focus on improvement connects to upward counterfactuals (how preferred policy might have made things better), whereas a focus on preservation connects to downward counterfactuals (how preferred policy prevented things from getting worse). The hypothesis that reflects this constraint view is a simple relation between partisanship and counterfactual direction of comparison, such that across issues, Democrats (vs. Republicans) are more likely to affirm upward than downward counterfactuals.

An alternative view is that counterfactual thinking is flexible, bending to fit partisan beliefs on an issue-by-issue basis. In this *directional flexibility* view, individuals affirm either upward or downward counterfactuals to the extent that each aligns with their political views.
Whereas the constraint view posits that Democrats (vs. Republicans) prefer upward over downward counterfactuals, the flexibility view posits that members of both parties will prefer whichever counterfactuals in whichever direction aligns with their preferred conclusions about a particular issue. For example, given that Republicans tend to prefer lower taxes than Democrats, Republicans might be more likely than Democrats to believe that “things would have been worse if we had raised taxes” (downward) and that “things would have been better if we had cut taxes” (upward).

**Judgment Impact**

Partisanship may be related not only to the direction of counterfactual thinking, but also to the impact that such thinking has on judgments. The present research examines blame, a politically consequential moral judgment. We suggest that leaders receive blame not only for the outcomes they actually caused, but also for the outcomes one imagines they could have caused. For example, when people consider a downward counterfactual about a negative event that could have occurred on a leader’s watch, the closer people think the leader came to allowing the event to occur, the more blame they may think leader deserves. Forming negative moral judgments of people based on negative counterfactual events suggests an assimilation process (see Markman & McMullen, 2003). We propose that partisanship amplifies such assimilation. Partisans may be especially willing to blame a leader they oppose for negative events that the leader “nearly” allowed. In general, 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 would moderate the impact of counterfactuals on moral judgment.

**Prior Research on Motivated Counterfactual Thinking**
Previous research has demonstrated how people’s beliefs and motivations shape counterfactual thinking, but has not directly tested our competing hypotheses about constraint versus flexibility.

Inside and outside the political domain, people strategically deploy counterfactual thinking to reach preferred conclusions (e.g., Catellani & Covelli, 2013; 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). Here, the motivation to feel virtuous invites individuals to ponder downward counterfactuals (they feel good because they might have done worse). Further, when a lie aligned with their politics, American partisans found it easier to imagine counterfactual scenarios in which the lie was true (Effron, 2018). Here, the motivation to excuse the lie appears to have led people to accept upward counterfactuals (the lie is justified because it could have been true). Relatedly, 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). Finally, people are averse to even considering counterfactuals that seem to challenge deeply held beliefs: Fundamentalist Christians rejected counterfactuals as heretical that implied secular explanations for divine events (Tetlock et al., 2000).

Although this past work suggests that beliefs shape counterfactual thinking, the studies systematically examined neither counterfactual direction, nor impact on blame judgments. A partial exception is a study that asked liberal and conservative policy experts to judge the plausibility of historical counterfactuals that varied in direction of comparison (Tetlock & Visser, 2000). Results presented in table form showed general agreement with the flexibility view, such
there was no across-the-board association between partisanship and upward versus downward counterfactuals, but specific analytic tests of this relation were not reported. Further, it is unclear whether the results generalize beyond policy experts to partisans among the general population. We suggest that the current U.S. political climate offers an ideal opportunity in which to explore some of these unanswered questions – as well as the relation between partisanship and the judgment impact of counterfactuals – by comparing the counterfactual thinking among opposing lay partisans.

**The Present Research**

Four studies tested the relations between partisanship and the direction (Studies 1a, 1b, 2) and impact (Study 3) of counterfactual thinking among Democrats and Republicans. Studies 1a and 1b tested whether partisans would rate both upward and downward counterfactuals as more plausible when these counterfactuals aligned with their politics. Study 2 tested 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. Finally, Study 3 tested whether partisans are more likely to blame a president for a negative event when 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. Thus, in Study 1a, the directional-constraint and the directional-flexibility views make identical predictions: Democrats should favor upward counterfactuals and Republicans should favor downward counterfactuals. If we observed such a finding, it would be unclear whether it reflected a general preference among members of each party for a particular counterfactual direction, or whether it reflected a tendency to prefer whichever counterfactual aligned with their view.

In Study 1b, we reversed the linkage, such that all upward counterfactuals aligned with Republicans’ ideology and all downward counterfactuals aligned with Democrats’ ideology. Here, the constraint view makes the same prediction as before, whereas the flexibility view predicts the reverse: that Democrats will find counterfactuals more plausible if they are downward and Republicans will find counterfactuals more plausible if they are upward.

**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; \( M \text{ age} = 35 \text{ years}, SD = 14; 66\% \text{ White, 12}\% \text{ Asian, 11}\% \text{ Black, 6}\% \text{ Latino/a, and the remainder other races and ethnicities}.\)

**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 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. 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

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: \(M = 2.26, SD = .86\)), \(b = 1.08, SE = .07, z = 14.54, p < .001, d_z = .81\).

Plausibility

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 (constraint view), or a more general tendency for

---

\footnote{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.}
partisans to prefer whichever direction of counterfactual happens to appeal to their politics (flexibility view). The reason is that in Study 1a, upward counterfactuals always aligned with Democrats’ views, whereas downward counterfactuals always aligned with Republicans’ views.

Study 1b disentangles these two interpretations by reversing which counterfactual direction aligned with which party’s views. The results were consistent with the flexibility view. That is, participants rated a counterfactual as more plausible when it aligned with their politics ($M = 3.46, SD = .83$) than when it 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, rather than having a consistent preference for counterfactuals in any one direction.

Examining how plausible Democrats and Republicans found upward and downward counterfactuals emphasizes this point (see Table 1). We submitted 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 ($M_{up} = 4.03$ vs. $M_{down} = 2.38, SDs = 1.05$ and $1.12$, respectively), $b = 1.65$, $SE = .08$, $z = 19.42$, $p < .001$ – and vice versa among Republicans ($M_{up} = 3.02, M_{down} = 3.70, SDs = 1.26$ and $1.15$, respectively), $b = -.67$, $SE = .09$, $z = 7.90$, $p < .001$. The interaction between party and direction was significant, $b$
By contrast, in Study 1b, when downward counterfactuals were aligned with Democrats’ views, Democrats rated downward counterfactuals as more plausible than upward counterfactuals ($M_{up} = 2.20$, $M_{down} = 3.86$, $SDs = .98$ and $1.17$, respectively), $b = -1.67$, $SE = .09$, $z = 19.21$, $p < .001$ – and vice versa among Republicans ($M_{up} = 3.30$, $M_{down} = 2.77$, $SDs = 1.18$ and $1.24$, respectively), $b = .26$, $SE = .09$, $z = 2.84$, $p = .005$. The interaction was significant, $b = 1.93$, $SE = .13$, $z = 15.38$, $p < .001$.

**Table 1**

<table>
<thead>
<tr>
<th>Counterfactual Direction</th>
<th>Studies 1a and 1b: Plausibility Judgments by Political Party</th>
<th>Democrat</th>
<th>Republican</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Study 1A – upward aligned with Democrats</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Upward</td>
<td>4.03</td>
<td>(1.05)</td>
<td>3.02</td>
</tr>
<tr>
<td>Downward</td>
<td>2.38</td>
<td>(1.12)</td>
<td>3.70</td>
</tr>
<tr>
<td></td>
<td>Study 1B – upward aligned with Republicans</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Upward</td>
<td>2.20</td>
<td>(1.17)</td>
<td>3.03</td>
</tr>
<tr>
<td>Downward</td>
<td>3.86</td>
<td>(0.98)</td>
<td>2.77</td>
</tr>
</tbody>
</table>

**Discussion**

Together, the results of Studies 1a and 1b demonstrate that the plausibility of an upward or downward counterfactual depends on how well it fits with their politics. These findings support the flexibility (vs. constraint) view.

**Study 2**

Study 2 asked participants to generate counterfactuals about eight political events, counterbalancing whether upward or downward counterfactuals aligned with Democrats’ or Republicans’ views. Given the previous studies’ support for the flexibility (vs. constraint) view,
we expected participants to generate more upward (vs. downward) counterfactuals for issues where upward (vs. downward) counterfactuals would support their political views.

**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 statements about contentious political issues. For each, we presented the antecedent of a counterfactual statement (i.e., the “if” part) and asked them to complete the consequence (i.e. the “then” part) in the form of an open response (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 counterfactual 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**

Consistent with the flexibility view, participants were more likely to generate counterfactuals in whichever direction was more aligned with their partisan views. As Table 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 or upward was similar across conditions (11.97% vs. 11.71%).

**Table 2**

*Study 2: Percentage of Upward and Downward Counterfactuals Generated in Each Condition*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Direction of Self-Generated Counterfactual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upward</td>
</tr>
<tr>
<td>Downward aligned</td>
<td>20.53%</td>
</tr>
<tr>
<td>Upward aligned</td>
<td>66.45%</td>
</tr>
<tr>
<td>Total</td>
<td>43.49%</td>
</tr>
</tbody>
</table>

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 flexibility hypothesis predicts 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, which is exactly what the results revealed, $OR = 19.17, SE = 3.45, z = 16.40, p < .001$.

Disaggregating the results by political party in an exploratory analysis showed no evidence that Democrats had a stronger preference for upward versus downward counterfactuals, or that Republicans had the reverse preference (see Table 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 as 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, $OR = .02, SE = .01, z = 12.55, p < .001$).
Table 3

Percentage of Upward Counterfactuals Generated

<table>
<thead>
<tr>
<th>Participants’ Party</th>
<th>Condition</th>
<th>Direction</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Up</td>
<td>Down</td>
<td>Neither</td>
<td></td>
</tr>
<tr>
<td>Democrat</td>
<td>Upward aligned</td>
<td>82.75%</td>
<td>7.00%</td>
<td>10.25%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Downward aligned</td>
<td>9.75%</td>
<td>82.25%</td>
<td>8.00%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>46.25%</td>
<td>44.62%</td>
<td>9.12%</td>
<td></td>
</tr>
<tr>
<td>Republican</td>
<td>Upward aligned</td>
<td>48.33%</td>
<td>38.33%</td>
<td>13.33%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Downward aligned</td>
<td>32.50%</td>
<td>51.11%</td>
<td>16.39%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>46.25%</td>
<td>44.62%</td>
<td>9.12%</td>
<td></td>
</tr>
</tbody>
</table>

These results are consistent with the flexibility (vs. constraint) view. Participants were more likely to generate counterfactuals in whichever direction that was consistent with their partisan views a particular issue, with no overall preference across issues for Democrats to prefer one type of counterfactual direction versus the other.

Study 3

Whereas Studies 1a–2 showed that partisanship predicts the direction of counterfactual thinking, Study 3 examines the relation between partisanship and judgment impact. 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 N. 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” people thought 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 reversed when people considered a president they supported. In other words, we predicted a stronger positive relationship between counterfactual closeness and blame when participants had opposed (vs. supported) the relevant president.

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 1). 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 blame, \(b = .42 SE = .02, z = 22.43, p < .001\), however (as shown by the interaction term reported above), it was significantly attenuated.

\(^2\) 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.
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 2 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. 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 relation between closeness and praise judgments regardless of whether participants supported or opposed the president, but this relation was significantly larger when participants opposed the president, $b = −.12, SE = .02, z = 7.29, p <$
.001, than when they supported the president, $b = -0.05$, $SE = .02$, $z = 3.01$, $p = .003$. 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 2**

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

![Graph showing relationship between counterfactual closeness and praise](image)

*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 3). 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 4), 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 3

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 4

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
Polarized Imagination

and 1b), selectively generate counterfactuals that align with their views (Study 2), and deploy counterfactuals that support preferred moral judgments about leaders (Study 3).

Together, these findings are consistent with our proposal that counterfactual thinking both reflects and reinforces partisanship. It reflects partisanship in that the alternatives to reality that people find plausible and spontaneously imagine will tend to align with their political views (Studies 1a–2). What would have happened if Trump had not passed his tax cut, or Biden had not withdrawn troops from Afghanistan? Nobody knows for sure, and what people imagine will depend on their politics.

Counterfactual thinking also reinforces partisanship by facilitating moral judgments that align with one’s politics. Is “almost” starting a war (but not actually starting it) sufficient grounds for blaming a president? Our results suggest that partisans are more likely to think the answer is “yes” if they oppose that president (Study 3). In this way, partisans may not only generate different counterfactuals (Study 2), they may reach different conclusions based on the same counterfactual – conclusions that probably serve to bolster their partisan perspective.

Our research makes several theoretical contributions. First, Studies 1a–2 distinguish between two theoretical views on how partisanship relates to the direction of counterfactual thinking. The directional-constraint view suggests an across-the-board preference among Democrats for upward counterfactuals among Republicans for downward counterfactuals. The directional-flexibility view suggests that alignment with partisan beliefs matters more and creates shifting preferences for upward or downward counterfactuals depending on this issue. The results supported the flexibility view over the constraint view. Partisans endorsed and generated counterfactuals in whichever direction best aligned with their views, with no overall preference.
across issues for Democrats or Republicans to prefer one type of counterfactual direction versus the other.

Second, 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 partisan motives shape counterfactual thinking among people 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 clearly do not support the possibility that, when it comes to counterfactual thinking, conservatives show more partisan bias than do liberals.

**Conclusion**

As Tetlock and Visser (2000) observed, “counterfactual thinking is often heavily theory-driven” (p. 174). Our results suggest that these theories are partisan. 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


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


https://doi.org/10.1017/S0003055400400043

https://doi.org/10.1348/014466600164417