COGNITIVE SCIENCE
A Multidisciplinary Journal

Check for updates

Cognitive Science 48 (2024) e13408

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13408

# Calculated Comparisons: Manufacturing Societal Causal Judgments by Implying Different Counterfactual Outcomes

Jamie Amemiya,[a] Gail D. Heyman,[b] Caren M. Walker[b]

[a]*Department of Psychology, Occidental College*
[b]*Department of Psychology, University of California, San Diego*

Received 4 October 2023; received in revised form 8 January 2024; accepted 16 January 2024

**Abstract**

How do people come to opposite causal judgments about societal problems, such as whether a public health policy reduced COVID-19 cases? The current research tests an understudied cognitive mechanism in which people may agree about what *actually* happened (e.g., that a public health policy was implemented and COVID-19 cases declined), but can be made to disagree about the counterfactual, or what *would have* happened otherwise (e.g., whether COVID-19 cases would have declined naturally without intervention) via comparison cases. Across two preregistered studies (total $N = 480$), participants reasoned about the implementation of a public policy that was followed by an immediate decline in novel virus cases. Study 1 shows that people's judgments about the causal impact of the policy could be pushed in opposite directions by emphasizing comparison cases that imply different counterfactual outcomes. Study 2 finds that people recognize they can use such information to influence others. Specifically, in service of persuading others to support or reject a public health policy, people systematically showed comparison cases implying the counterfactual outcome that aligned with their position. These findings were robust across samples of U.S. college students and politically and socioeconomically diverse U.S. adults. Together, these studies suggest that implied counterfactuals are a powerful tool that individuals can use to manufacture others' causal judgments and warrant further investigation as a mechanism contributing to belief polarization.

*Keywords:* Causal judgment; Counterfactuals; Motivated reasoning; Public health

Correspondence should be sent to Jamie Amemiya, Department of Psychology, Occidental College, 1600 Campus Drive, Los Angeles, CA 90041, USA. E-mail: amemiya@oxy.edu

Publicly available data indicate that there have been sharp fluctuations in coronavirus cases throughout the pandemic (Johns Hopkins Coronavirus Resource Center, 2022). While this trend is evident, there have been frequent debates about whether a given public health policy, such as masking or vaccination, had a causal impact on mitigating the virus (Deane, Parker, & Gramlich, 2021). In the present research, we ask: How do people come to opposite causal judgments about a societal intervention, such as whether a particular public policy impacted virus case rates during a global pandemic? Previous research has identified several important mechanisms, including exposure to different types of information and people's tendency to be more critical of information that counters their prior beliefs (i.e., confirmation bias; Xu, Coman, Yamamoto, & Najera, 2023). Here, we propose that disagreements can persist even when people agree about what *actually* happened (i.e., that a policy was implemented and virus cases subsequently declined) because they can be made to disagree about the counterfactual of *what would* have happened otherwise (i.e., the trajectory of virus cases had the policy *not* been implemented). In two preregistered studies, we examined the extent to which (a) causal judgments about a policy can be pushed in opposite directions by emphasizing comparison cases that imply different counterfactual outcomes (Study 1), and (b) people selectively highlight comparison cases to persuade *others* to reach specific causal conclusions (Study 2). The focus of this study is distinct from prior research in that it highlights how societal disagreements not only hinge on people's understanding of events that actually happened, but also on events that could have happened. Another major contribution of this research is that it examines *multiple* processes underlying belief polarization: people's *own* differential causal reasoning and their attempts to influence *others*' causal reasoning.

The current research is informed by counterfactual theories of causal judgment (Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2021; Lewis, 1973; Lucas & Kemp, 2015; Mackie, 1974; Quillien & Lucas, 2023; Woodward, 2005). According to this view, people determine causality by comparing the actual, known outcome to a counterfactual that informs what would have happened if the candidate cause had been absent. If the comparison between the actual and counterfactual outcomes reveals that the candidate cause is "difference-making," people endorse it as causal.

At least when reasoning about everyday causal phenomena (e.g., common life events, such as test performance), there is robust evidence that people rely on counterfactuals and that there is remarkable consistency in *which* counterfactuals they consider (Byrne, 2016; Gerstenberg et al., 2021; Roese & Epstude, 2017). The clearest evidence for similarities in counterfactual thinking come from studies in which people have robust prior causal knowledge and are not motivated to reach a particular causal judgment, like in the domain of intuitive physics (Gerstenberg et al., 2021). For example, when determining whether Ball A caused Ball B to go through a gate, people reliably visualize the accurate path that Ball B *would have* taken had it not been hit by Ball A (Gerstenberg, Peterson, Goodman, Lagnado, & Tenenbaum, 2017). However, this counterfactual framework has rarely been applied to understand how people make sense of complex societal problems (for important exceptions, see Kominsky, Reardon, & Bonawitz, 2021; Quillien & Barlev, 2022), which may introduce new inferential problems when making causal judgments.

Here, we consider three problems for judging the causal impact of a public health policy in a global pandemic, and also consider the practical implications for understanding societal belief polarization. First, people in these instances typically lack the relevant causal knowledge to simulate what would have happened without intervention (Amemiya, Heyman, & Walker, 2021; Caddick & Rottman, 2021; Kominsky et al., 2021). Ideally, people could observe an experiment that compares (a) virus cases when their society *implements* the policy (i.e., actual outcome) versus (b) virus cases when their society does *not* implement the policy, to generate the counterfactual outcome. Indeed, making such comparisons to generate counterfactuals is central to scientific reasoning: Scientists compare the treatment group to a control group (often referred to as the "counterfactual"; Robinson, McNulty, & Krasno, 2009) to estimate the causal effect of treatment (Leatherdale, 2019). This ability also plays a central role in intuitive science—even young children compare situations in which a variable was present to those in which it was absent to make novel causal judgments (Goddu & Gopnik, 2020; Nyhout & Ganea, 2021; Walker & Nyhout, 2020).

In the context of a novel societal problem, like a pandemic, people are unable to observe what would have happened otherwise *within* a particular society. Instead, they must *indirectly* infer the counterfactual, by comparison with another society that did not implement the policy. Indeed, during the COVID-19 pandemic, news reporters regularly compared countries that varied in their responses to judge the effectiveness of a specific country's policy (Connolly, 2020). However, comparing different countries introduces a second problem: Because no two countries are matched on all potentially relevant background variables (i.e., there is no perfect experiment), people must rely on imperfect comparisons and decide *which* background variables are most critical to match (Leatherdale, 2019). For instance, when reasoning about the effect of a policy in the United States, would the outcomes in Mexico or Australia indicate the more valid counterfactual of what would have happened otherwise? This depends on which dimension of similarity is most relevant to account for: for example, is it more important to equate the target and counterfactual on *geographical* factors (in which case, Mexico is the better counterfactual) or *socioeconomic* factors (in which case, Australia is the better counterfactual)?

Given the uncertainty about which comparisons are most relevant, a third problem arises. Specifically, people may be *susceptible* to whichever comparisons are emphasized, including those that come from partisan news channels and other biased sources (see also Kominsky & Phillips, 2019; Phillips, Luguri, & Knobe, 2015). People may also recognize that they can manipulate *others* by highlighting certain comparison cases, further exacerbating belief polarization. Indeed, there is some initial evidence that people flexibly use comparisons to imply a counterfactual outcome, at least to produce *affective* responses to societal problems (Eibach & Ehrlinger, 2006; Markman, Mizoguchi, & McMullen, 2008). For example, Markman et al. (2008) noted that to mitigate the U.S. abuse of Iraqi prisoners in Abu Ghraib, some politicians emphasized that treatment would have been worse under Saddam Hussein. Experiments indicated that people's emotions shifted in response to this comparison; participants who were prompted to think about this counterfactual outcome felt better about Abu Ghraib. Of particular relevance to our investigation, former President Trump defended his policy decisions during the pandemic by comparing U.S.'s cases to European countries with higher covid cases and ignored countries that were faring better (Smith & Edwards, 2020). Yet, no research to our knowledge has examined whether implying different counterfactuals impacts people's causal judgments about public policies, which are important for policy adherence (Moran et al., 2021).

### *The present research*

The present research was informed by the observation that different people make opposing causal judgments about societal events, like a global pandemic. Prior research has highlighted the important role of being exposed to different types of information about what *actually* happened and people's tendency to be more critical of information if it conflicts with their prior beliefs (Xu et al., 2023). Here, we focus on how people can come to opposing judgments even when they agree about the events that actually happened because they are prompted to consider different *counterfactual* outcomes via specific comparison cases.

We conducted two related studies to test this mechanism. In Study 1, we examined whether we could push participants' causal judgments about a public health policy in opposite directions by emphasizing comparison cases that imply the counterfactual that virus cases (a) would have continued to increase in the absence of intervention (supporting the causal relationship) or (b) would have declined on their own (providing evidence against the causal relationship). In Study 2, we asked people to imagine that they were politically motivated to either support or reject a particular public policy, and we examined whether they select comparison cases implying a counterfactual outcome to persuade *others*' causal judgments. For robustness, both studies included a sample of U.S. college students and a replication sample of U.S. adults who were diverse in political orientation and educational attainment.

### *Transparency and openness statement*

We report below how we determined our sample size, all data exclusions, all manipulations, and measures in the study, following the APA Journal Article Reporting Standards (Kazak, 2018). The data, code, and materials for the studies are publicly accessible at OSF: https://osf.io/ja84e/. The preregistrations for Study 1 are at https://aspredicted.org/LYT_32R (college sample) and https://aspredicted.org/HLM_WG3 (Prolific sample); the preregistrations for Study 2 are at https://aspredicted.org/QXZ_VD8 (college sample) and https://aspredicted.org/VZB_WMF (Prolific sample). Data were analyzed using R, version 4.0.4 (R Core Team, 2020), and the packages *tidyverse* (Wickham et al., 2019), *stats* (R Core Team, 2020), and *lsr* (Navarro, 2016).

## 1. Study 1

Study 1 examined the first part of our proposal, specifically, that emphasizing comparison cases implying different counterfactuals may lead to opposing causal judgments about a public health policy. This study was conducted in October 2021 (U.S. college sample) and replicated in April 2023 (Prolific sample).

## *1.1. Method*

### *1.1.1. Participants*

Participants (total $N = 240$) were 120 undergraduate students (89 women, 28 men, 2 nonbinary, 1 did not report; 59 Asian, 21 Latine/x, 19 White, 6 Middle Eastern or North African, 11 Mixed, 2 Black, 2 did not report) who attended a large, public university in the West Coast region of the United States, as well as 120 participants recruited from Prolific ($M_{\text{age}} = 45$ years, $SD = 13$; 63 women, 56 men, 1 nonbinary; 83 White, 13 Latine/x, 11 Black, 7 Asian, 5 Mixed, 1 Native American).

College participants skewed liberal in their political orientation (7 extremely liberal, 45 liberal, 16 slightly liberal, 9 moderate, 5 slightly conservative, 4 conservative, 0 extremely conservative, 21 did not identify with a political orientation, 13 did not report). Prolific participants were politically diverse (19 extremely liberal, 33 liberal, 12 slightly liberal, 9 moderate, 14 slightly conservative, 22 conservative, 10 extremely conservative, 1 did not identify with a political orientation). The Prolific sample was also diverse in educational attainment (43 high school diploma, 26 associate's or technical degree, 3 some college, 30 bachelor's degree, 14 master's degree, 4 professional degree), and their median self-reported income was $50,000.

The sample size ($n = 120$) was based on a power analysis in G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) that indicated we would need 55 participants to detect a medium effect size ($f^2 = .15$) of study condition with a power of .80 and alpha level of .05. We doubled this sample size to allow for exploratory analyses and rounded it so that there are an equal number of participants in each counterbalance (Island and U.S. states). An additional 19 participants (12 college, 7 Prolific) were dropped due to failing at least one of the comprehension checks.

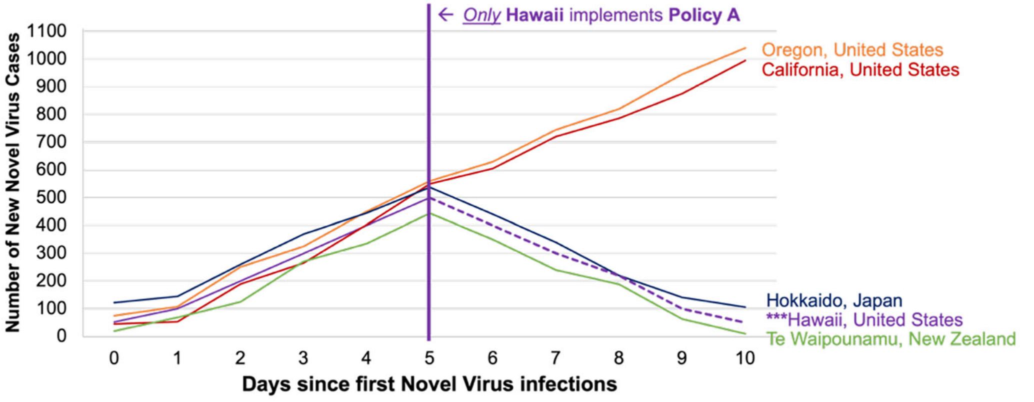
### *1.1.2. Procedure*

Participants completed the study via the Qualtrics online platform. A complete description of the procedure can be found in the online Appendix.

*1.1.2.1. Training modules:* Participants first completed two training modules designed to help them to interpret the key figure used in the experimental manipulation. These included a *locations* training module on whether the locations were U.S. states or islands (see online Appendix, Section I) and a *graph comprehension* training module covering how to interpret each element of the graph depicting the case rates in each location (see online Appendix, Section II). Participants received corrective feedback on all questions in these training sections.

*1.1.2.2. Experimental manipulation:* All participants read the following news story about a novel pandemic. In the story, only *one* location, Hawaii, implemented a public health intervention:

> In five regions of the world (Hawaii, United States; Hokkaido, Japan; Te Waipounamu, New Zealand; Oregon, United States; California, United States), there was a novel virus that began to spread. Leaders of the regions were unsure about what to do because this




*Note*. Dashed line indicates that the region **implemented** Policy A.

Fig. 1. In this version of the graph, outcomes from Oregon and California suggest that Policy A was causal for Hawaii's reduced cases, while outcomes from Hokkaido and Te Waipounamu suggest that it was not. The other version swapped the labels of Oregon and California with Hokkaido and Te Waipounamu.

> was a novel virus that had never been seen before. But, together, the leaders decided on a set of public policies, called Policy A. Hawaiian leaders chose to implement Policy A on Day 5 to see if it would reduce infections. The other four regions did not implement Policy A, waiting to see whether it worked in Hawaii.

Given that the four other locations did *not* intervene, these cases could serve as the implied counterfactual for what Hawaii's virus cases would have looked like had it not intervened (i.e., the "control groups"). The other locations were similar to Hawaii on one of two dimensions: Either they were also U.S. states (Oregon and California) or they were also islands (Hokkaido and Te Waipounamu). Participants read information that implied one set of locations (either the U.S. states or the islands) was the more relevant comparison case, given their shared background characteristics:

> Scientists started to examine the virus and reported that the virus could have similar trajectories in [island regions/the United States], but completely different trajectories in regions that are not [islands/part of the United States]. The scientists weren't sure about the direction of the trajectories in these different places—the virus may decline on its own or continue to increase. But the scientists strongly believed that, whatever the direction, the virus trajectories would follow similar paths within [island/U.S.] regions. This is because [island regions/states in the U.S.] have similar health care systems, economies, and populations that are unique from [non-island/non-U.S.] regions.

Participants then observed the virus trajectories as presented in Fig. 1. In the *Causal* condition (here, the U.S. states, Oregon and California), the emphasized comparison cases *supported* the causal impact of Policy A because these data implied the counterfactual that Hawaii's virus trajectory would have kept increasing without intervention. In the *Not Causal* condition (here the islands, Hokkaido and Te Waipounamu), the emphasized comparison cases provided evidence *against* the causal impact of Policy A because they implied the counterfactual that Hawaii's trajectory would have declined naturally. We counterbalanced whether the U.S. states or islands were causal by creating another figure that swapped the labels of Oregon and California with Hokkaido and Te Waipounamu.

Importantly, we designed the pattern of the trajectories to assess the extent to which people care about counterfactuals above and beyond other factors that are important for making causal judgments. Specifically, we presented mixed evidence regarding patterns of *covariation* (see Cheng, 1997). On the one hand, there was strong *temporal* covariation between events that occurred in Hawaii across both the *Causal* and *Not Causal* conditions: Hawaii's implementation of Policy A was *immediately* followed by a stark decline in virus cases. If people *only* consider temporal covariation information about target events when making causal judgments, and do not counterfactual information, there should be no condition differences. Rather, participants in both conditions should make strong judgments that Policy A caused Hawaii's decline in cases because of the strong temporal covariation between the target events that occurred in Hawaii.

However, if people consider the most relevant counterfactual outcome for Hawaii and consequently focus on specific regions (i.e., only U.S. states or only islands), there are condition differences in the patterns of covariation *within each region*. For example, if focusing only on the U.S. states in Fig. 1 (i.e., the *Causal* condition), there is strong covariation between policy implementation and virus cases. In contrast, if focusing only on the islands (i.e., the *Not Causal* condition), there is no covariation between these events. Thus, if people make stronger causal judgments in the *Causal* versus *Not Causal* condition, this suggests that people determine relevant regions using counterfactual reasoning and the covariation *within* each region further shapes their causal judgments.

### *1.1.3. Dependent measures*

Participants were asked to make a *causal judgment* about the policy intervention: "On Day 5, Hawaii's virus cases began to decrease. How likely is it that this decrease was caused by Policy A?" (0 = extremely unlikely to 100 = extremely likely). Participants were then asked to explain their causal judgment with the question, "Why do you think that?" Causal judgment questions always came *before* the counterfactual judgment (see below), to avoid explicitly prompting participants to use counterfactual information (see the *causal* condition in Gerstenberg et al., 2017 for a similar approach).

Participants next made a *counterfactual judgment*, which informed our interpretation of the causal judgment data: "Imagine that Hawaii had *not* implemented Policy A and instead had done nothing. Would the number of new cases after Day 5 have (A) decreased, (B) stayed about the same, or (C) kept increasing?" Each option was accompanied by a graph illustrating the trend (see online Appendix, Section V). Participants who selected Option A—that virus cases would have decreased anyway—indicates they believed the policy did *not* make a difference. On the other hand, participants who selected Options B or C inferred that the policy was difference-making, as cases would have been worse without intervention.

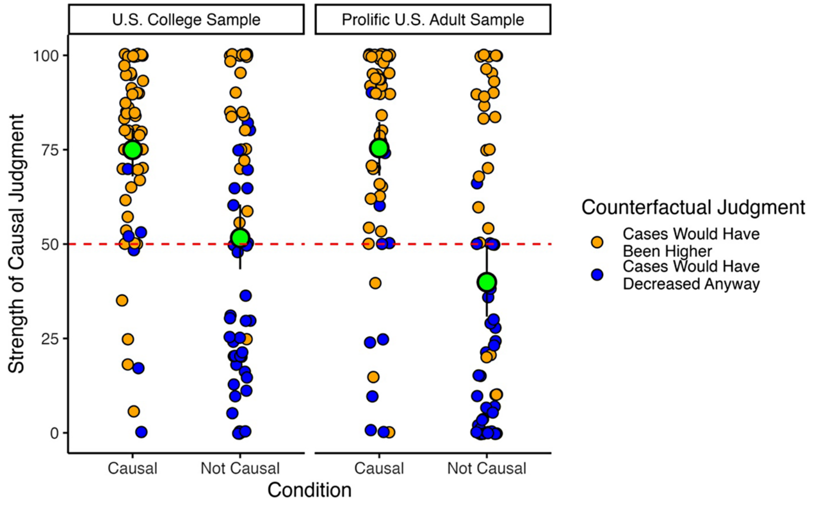



Fig. 2. Participants' causal and counterfactual judgments about the public policy based on whether the emphasized comparison cases supported the causal relationship via the implied counterfactual, split by participant sample (college vs. Prolific). Raw data (in orange and blue), means (in green), and 95% CIs around the condition means are graphed.

Finally, all participants completed a final set of manipulation check questions and demographic information (see online Appendix, Sections VI and VII). To assess participants' general attitudes about public mandates, we assessed both political orientation and beliefs that public mandates are generally effective, asking, "Public mandates (e.g., requiring people to wear masks and social distance) are effective in reducing the number of new virus cases in a society" (0 = definitely disagree to 100 = definitely agree). U.S. college participants were significantly more likely to agree with this item ($M = 89$) than Prolific adults ($M = 73$), $t(238) = 5.09$, $p < .001$. To be included in the analyses, participants needed to pass all of the final manipulation check questions (recall that an additional 12 college students and 7 Prolific participants were dropped to failing at least one check).

### *1.2. Results*

#### *1.2.1. Causal judgments*

We ran a linear regression that predicted participants' causal judgments (on a scale of 0−100) by condition (1 = *Causal*, 0 = *Not Causal*), and included the emphasized dimension as a covariate (1 = islands, −1 = U.S. states). The emphasized dimension was not related to judgments, nor did it interact with condition in any analyses for Study 1 or 2, and thus we will focus only on condition differences. As shown in Fig. 2 (college sample on left, Prolific sample on right), participants' causal judgments were significantly higher in the *Causal* condition relative to the *Not Causal* condition, college sample: $B = 23.33$, $SE = 5.48$, $p < .001$, 95% CI [12.48, 34.19], $\beta = .37$; Prolific sample: $B = 35.53$, $SE = 6.07$, $p < .001$, 95% CI [23.51, 47.56], $\beta = .48$. Participants in the *Causal* condition endorsed Policy A as causal (college: $M = 74.97$, 95% CI [68.58, 81.36]; Prolific: $M = 75.43$, 95% CI [68.08, 82.79]), whereas college participants were ambivalent and Prolific participants rejected causality in the *Not Causal* condition (college: $M = 51.63$, 95% CI [43.02, 60.25], Prolific: $M = 39.90$, 95% CI [30.59, 49.21]).

### 1.2.2. Counterfactual judgments

We color-coded participants' counterfactual judgments in Fig. 2, in which orange datapoints indicate that participants judged that Policy A was difference-making (i.e., they selected Options B or C which both show that Hawaii's cases would have been worse without the policy; 60% of the college sample and 58% of the Prolific sample picked Option C, the more severe alternative that aligned with one set of counterfactuals, only 5% of the college sample and 3% of the Prolific sample chose Option B), while blue datapoints indicate that participants reasoned that it was *not* difference-making (i.e., they selected Option A which show that Hawaii's cases would have declined anyway; 35% of the college sample and 39% of the Prolific sample picked Option A).

We ran a logistic regression predicting participants' counterfactual judgments (1 = difference-making, 0 = not difference-making) by condition, and similarly included the emphasized dimension as a covariate. Participants were significantly more likely to infer that Policy A was difference-making in the *Causal* condition relative to the *Not Causal* condition, college sample: $B = 2.61$, $SE = 0.51$, $p < .001$, 95% CI [1.68, 3.68], $OR = 13.57$, Prolific sample: $B = 1.74$, $SE = 0.42$, $p < .001$, 95% CI [0.94, 2.59], $OR = 5.69$. This effect is readily apparent in Fig. 2, in which there are more orange relative to blue datapoints in the *Causal* condition than in the *Not Causal* condition.

### 1.2.3. Relation between causal and counterfactual judgments

A linear regression predicting causal judgments from counterfactual judgments indicated the two were strongly correlated, college sample: $B = 47.65$, $SE = 4.33$, $p < .001$, 95% CI [39.07, 56.23], $\beta = .71$, Prolific sample: $B = 55.06$, $SE = 4.95$, $p < .001$, 95% CI [45.26, 64.85], $\beta = .72$, such that participants who inferred that cases would have been worse without intervention made strong causal judgments (college: $M = 80.01$, 95% CI [75.27, 84.76], Prolific: $M = 79.18$, 95% CI [73.23, 85.13]), while participants who inferred that cases would have declined anyway made weak causal judgments ($M = 32.26$, 95% CI [24.82, 39.71], Prolific: $M = 24.26$, 95% CI [16.62, 31.89]).

### 1.2.4. Secondary and exploratory analyses

Two independent coders coded participants' open-ended explanations for their causal judgments. The most common theme to emerge was mentioning the comparison cases that we had emphasized (e.g., in the *island* version, "Because island health care systems are fairly similar, so if a policy wasn't implemented in Hawaii, we would expect to see results like what we see in Japan and New Zealand. And because we don't see similar reactions in Hawaii, we can strongly infer it was due to the policy"), Cohen's K = 0.88 for college sample; Cohen's K = 0.85 for Prolific sample (56% of college sample's explanations; 49% of Prolific sample's explanations). A smaller percentage of participants discussed the comparison cases that were *not* emphasized (e.g., in the *island* version, "Every state [that is] part of the United States saw a similar decrease in cases even though they didn't implement Policy A"), Cohen's K = 0.74 for college sample; Cohen's K = 0.88 for Prolific sample (10% of college sample's explanations; 18% of Prolific sample's explanations). The fact that people most commonly cited comparisons in their explanations provides further evidence that people consider counterfactual outcomes when making causal judgments.

A minority of participants focused solely on what happened in Hawaii (e.g., "Because the line started to decrease right after day 5"), Cohen's K = 0.77 for college sample; Cohen's K = 0.72 for Prolific sample (18% of college sample's explanations; 16% of Prolific sample's explanations) or proposed or alluded to a causal mechanism for the policy (e.g., "Policy probably affected the amount of exposure from the virus"), Cohen's K = 0.64 for college sample; Cohen's K = 0.89 for Prolific sample (9% of college sample's explanations; 16% of Prolific sample's explanations). This also provides support for the counterfactual account in that few people relied solely on the strong temporal covariation between the actual events that happened in Hawaii.

Finally, we explored the relationship between participants' existing beliefs (i.e., liberal political orientation and beliefs that public mandates are generally effective) and causal judgments. We standardized and averaged these variables given their correlation (college: $r = .46$, $p < .001$; Prolific: $r = .55$, $p < .001$), with higher scores indicating a greater orientation toward endorsing public mandates. We found that these pre-existing beliefs were correlated with stronger causal judgments among the college students in the *Not Causal* condition, $r = .36$, $p = .02$ (but not with causal judgments in the *Causal* condition, $r = .11$, $p = .50$), while these beliefs were correlated with stronger causal judgments among the Prolific participants in the *Causal* condition, $r = .45$, $p < .001$ (but not with causal judgments in the *Not Causal* condition, $r = .04$, $p = .78$). Given that there are many differences between these samples, including the population and time the study was run, we hesitate to make strong inferences about why the correlations between pre-existing beliefs and judgments emerged in different conditions. Nonetheless, the results converge on the finding that people are less likely to make the predicted causal judgments when the implied counterfactual outcome goes against their prior beliefs.

### *1.3. Supplemental study*

We report the results of a supplemental study ($N = 43$ U.S. college students) that addressed several alternative explanations. Similar to Study 1, this supplemental study presented a graph of five trajectories (one target town, four comparison cases that implied different counterfactual outcomes) and participants (a) rated whether a public health intervention implemented in the target town caused their trajectory and (b) made counterfactual judgments about what would have happened without intervention.

However, there were several important differences. First, we presented a different societal problem (a novel insect infestation and population rates of insect bites) and a different public health intervention (evacuation from the town), which allowed us to test for the generalizability of the effect. Second, while we presented a target trajectory (Hillsbrook,

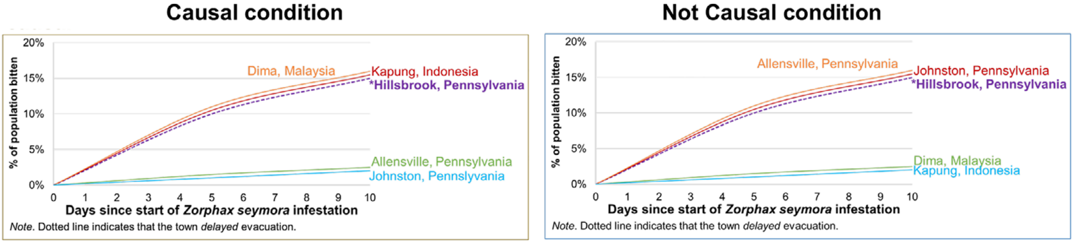



Fig. 3. Causal and Not Causal condition in the supplemental study.

Pennsylvania, whose population evacuated) and four trajectories that could be used to generate the counterfactual (towns which *delayed* evacuation; two other towns in Pennsylvania: Allensville, Pennsylvania; Johnston, Pennsylvania; and two towns in Southeast Asia: Kapung, Indonesia; Dima, Malaysia), the study did *not* emphasize any of the cases to participants. Instead, we examined whether participants would *spontaneously* privilege the data that offered a better-controlled comparison (i.e., other Pennsylvanian towns over Southeast Asian towns). Here, the *Causal* condition was the version in which the other Pennsylvanian towns indicated that Hillsbrook's insect bite trajectory would have been worse without evacuation (while the Southeast Asian towns suggested it would have stayed low regardless), while the *Not Causal* condition was the version in which the other Pennsylvanian towns indicated that Hillsbrook's insect bite trajectory would have stayed low regardless (while the Southeast Asian towns suggested it would have been worse). Third, the shapes of the trajectories were different, which allowed us to generalize the findings to unique patterns of data. Finally, we did not provide any graph training prior to presenting the graph, which allowed us to better understand how people may use this information in the real world, in the absence of scaffolding. See Fig. 3 for the data that were presented in the *Causal* and *Not Causal* conditions.

Despite these differences with Study 1, the results of the supplemental study were essentially the same. As shown in Fig. 4, we found that participants' causal judgments were significantly higher in the *Causal* condition relative to the *Not Causal* condition, $B = 24.16$, $SE = 9.60$, $p = .02$, 95% CI [4.76, 43.55], $\beta = .37$, indicating that participants *spontaneously* privilege more relevant comparison cases (i.e., other Pennsylvanian towns) when making causal judgments. We also found that participants in the *Causal* condition made stronger counterfactual judgments that rates of insect bites would have been worse had Hillsbrook not evacuated (measured with a continuous scale), $B = 22.82$, $SE = 8.30$, $p = .01$, 95% CI [6.05, 39.58], $\beta = .39$. Finally, participants' causal and counterfactual judgments were highly correlated, $r = .76$, $p < .001$, again providing support for the counterfactual account of causal judgment. In combination with Study 1, these data indicate that, regardless of the domain, emphasis on certain comparison cases, trajectory shape, and graph training, people's societal causal judgments are influenced by (relevant) comparison cases that indicate a certain counterfactual outcome.

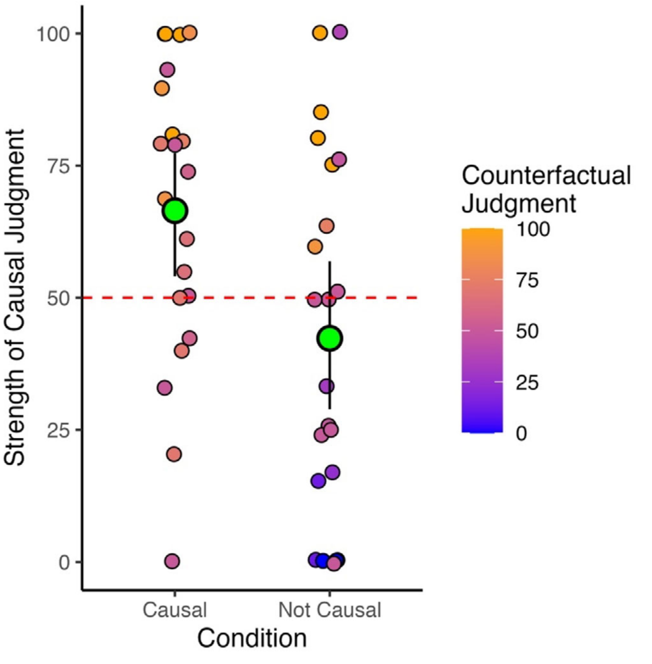



Fig. 4. Participants' causal and counterfactual judgments about the public policy in the supplemental study. Raw data (from orange to blue), means (in green), and 95% CIs around the condition means are graphed.

### 1.4. *Study 1 discussion*

Results from Study 1 support our proposal that emphasizing comparisons to imply different counterfactual outcomes produces opposing societal causal judgments. Participants in the *Causal* condition overwhelmingly endorsed the effectiveness of Policy A, while many in the *Not Causal* condition rejected Policy A's causal impact.

Our results also provide support for the counterfactual account of causal judgment. Despite the strong temporal covariation between Policy A's implementation and Hawaii's subsequent change in trajectory, we found that participants' causal judgments tracked whether the comparison cases implied the *counterfactual* that this change would or would not have happened without intervention. It was also possible that all participants would *reject* Policy A's causal impact, given that all comparison cases were visually present, including those that provided counterevidence against the policy's effectiveness. Yet, the robust condition effect suggests that people are not only sensitive to implied counterfactual outcomes, but that they privilege the comparison cases that introduce fewer *confounds*. Although there was some evidence that people's prior beliefs relate to their acceptance of this information, overall, we find that people's societal causal judgments are impacted by which counterfactual outcomes are implied.

Importantly, we are not suggesting that counterfactual accounts *replace* other accounts of causal judgment, such as those emphasizing people's attention to covariation (Cheng, 1997; see also Gong & Bramley, 2023; Griffiths & Tenenbaum, 2005, 2009; Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008). Indeed, it is likely that counterfactual *and* covariation information within the relevant region jointly influenced people's causal judgments. Moreover, it is possible that people's causal judgments would have been even more polarized (i.e., weaker in the *Not Causal* condition) if the temporal covariation information in our study did not so strongly indicate causality across both conditions.

## 2. Study 2

While Study 1 established people's sensitivity to comparison cases, Study 2 tested whether people are aware that they can use this approach to manipulate *others*' causal judgments to support a particular agenda. If so, this would provide further evidence that manipulations of counterfactual reasoning may be an important mechanism underlying belief polarization. To test this hypothesis, a new sample of participants was given the same information presented in Study 1 (i.e., target Hawaii data and data on counterfactuals that either supported or provided evidence against a causal relationship) and told to imagine that they were politicians who are either in favor of or in opposition to the policy. We were interested in whether people would selectively show comparison cases that imply the counterfactual outcome supporting their assigned policy preference. This study was conducted in January 2022 (college sample) and replicated in April 2023 (Prolific sample).

### *2.1. Method*

#### *2.1.1. Participants*

Participants were a sample of 120 undergraduate students (101 women, 18 men, 1 did not report; 69 Asian, 16 White, 14 Latine/x, 14 Mixed, 2 Middle Eastern or North African, 2 Black) who attended a large, public university in the West Coast region of the United States, as well 120 participants recruited from Prolific ($M_{\text{age}} = 43$ years, $SD = 13$; 61 women, 57 men, 1 nonbinary, 1 reported "other"; 73 White, 24 Black, 16 Latine/x, 3 Mixed, 1 Asian, 3 did not report). This sample size ($n = 120$) was based on a power analysis in G*Power (Faul et al., 2009) that indicated we would need 61 participants to detect a medium to large effect size ($w = .40$) of study condition with a power of .80 and alpha level of .05. We roughly doubled this sample size to allow for exploratory analyses. An additional 39 participants (19 college, 20 Prolific) were dropped due to failing at least one of the comprehension checks.

Similar to Study 1, college participants were more liberal in their political orientation (12 extremely liberal, 36 liberal, 24 slightly liberal, 17 moderate, 3 slightly conservative, 4 conservative, 19 did not identify with a political orientation, 5 did not report). Prolific participants were more politically diverse (25 extremely liberal, 31 liberal, 13 slightly liberal, 8 moderate, 8 slightly conservative, 26 conservative, 9 extremely conservative) and were also diverse with respect to educational attainment (40 high school diploma, 16 associate's or technical degree, 3 some college, 46 bachelor's degree, 12 master's degree, 3 professional degree). The median self-reported income among Prolific participants was $50,000.

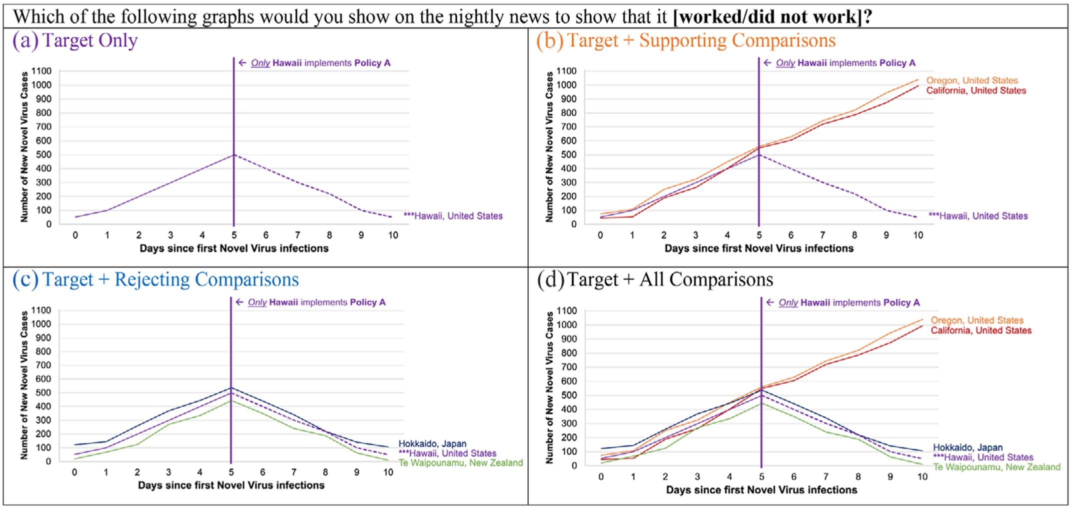



Fig. 5. Dependent measure for Study 2. Participants selected which type of data to use to support their causal claim.

### 2.1.2. *Procedure*

The full procedure for Study 2 appears in the online Appendix. Study 2 used a similar design as Study 1: Participants first completed *location* and *graph comprehension* training modules and received corrective feedback. Participants then read the novel virus news story, which again included attention checks with corrective feedback. Rather than making causal judgments, however, participants were randomly assigned to a motivated reasoning condition, in which they were prompted to imagine themselves as a politician who either (a) strongly believed that Policy A would be effective (*Motivated to Support* condition), or (b) strongly believed that Policy A would be *ineffective* (*Motivated to Reject* condition). Participants were then given the graph with the data for all five locations. After being prompted to describe the data in their own words, they selected *which* data to show on the nightly news to support their assigned position.

### 2.1.3. *Dependent measure*

Participants were allowed to show one of four figures on the nightly news: (a) Target only, (b) Target + Supporting Comparisons, (c) Target + Rejecting Comparisons, and (d) Target + All Comparisons (note that the graphs were not labeled with these titles; see Fig. 5). We were interested in whether participants would use comparison cases to their advantage, such that they would choose to show supporting comparison cases implying the counterfactual that virus cases would have been worse in the *Motivated to Support* condition (i.e., Option B), but would show the rejecting comparison cases implying that the counterfactual that virus cases would have declined anyway in the *Motivated to Reject* condition (i.e., Option C). We also included an option that only showed the declining trend of Hawaii (i.e., Option A) to assess whether participants prefer to include comparison cases when given the choice not to

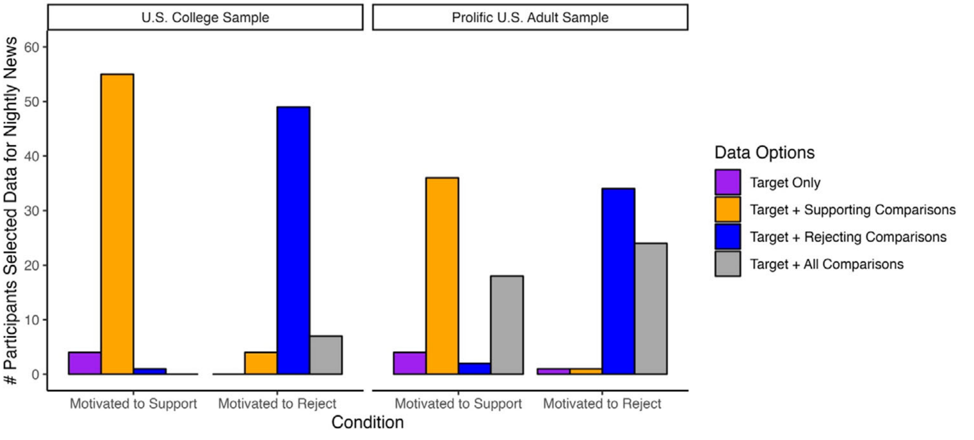



Fig. 6. Participants' data selection based on their motivation to support or reject a policy, split by sample (college vs. Prolific).

do so. This was particularly of interest for the *Causal* condition, in which showing Option A would *also* suggest a causal relationship due to the strong temporal covariation between policy implementation and virus cases.

We asked participants to justify their data selections, with the question: "Imagine that someone saw all of the data for all 5 cities and asked you to explain why you chose to show this particular graph. What would you say to justify your decision?" We were interested in whether participants would spontaneously generate post hoc justifications that had characteristics of Study 1's manipulation, such as claiming that Hawaii is more similar to other U.S. states than to other islands, and thus the U.S. states provide the most relevant comparison case.

As in Study 1, all participants completed a final set of manipulation check questions and demographic information. Participants needed to pass all of the final manipulation check questions to be included in the analyses (19 college and 20 Prolific participants were dropped).

### 2.2. *Results*

Fig. 6 shows people's data selections by the motivated reasoning condition. Data selections for both samples varied significantly by condition, college sample: $\chi^2(3) = 101.16$, $p < .001$, Cramer's V $= .92$; Prolific sample: $\chi^2(3) = 64.21$, $p < .001$, Cramer's V $= .73$. Participants most commonly chose one of the two data options that *selectively* showed comparison cases (Options B or C; 91% of college students' selections; 61% of Prolific adults' selections), whereas 6% of college students and 35% of Prolific adults chose to show all comparison cases (mostly in the *Motivated to Reject* condition), and 3% of college students and 4% of Prolific adults chose to show only what happened in Hawaii (mostly in the *Motivated to Support* condition). Before describing the condition effects—which were robust across samples—we note that more Prolific adults showed *all* of the data compared to the college students (e.g., 0 college students showed all of the data in the *Motivated to Support* condition, while 18 Prolific adults did), with some of these participants expressing concerns about hiding information. For example, one Prolific participant stated, "This is the information we have and need to judge on. To show only some of the datapoints would [be] using partial data in order to lie." Nonetheless, the majority of participants from both samples selectively highlighted comparison cases in ways that aligned with the condition manipulation.

We next ran two sets of logistic regressions: one that predicted selecting the *supporting* comparison cases (Option B), and one that predicted selecting the *rejecting* comparison cases (Option C). As hypothesized, participants in the *Motivated to Support* condition were more likely to show the supporting comparison cases than participants in the *Motivated to Reject* condition, college sample: $B = 5.15$, $SE = 0.74$, $p < .001$, 95% CI [3.85, 6.80], $OR = 173.16$, Prolific sample: $B = 4.50$, $SE = 1.04$, $p < .001$, 95% CI [2.88, 7.41], $OR = 90.38$. Also as expected, participants in the *Motivated to Reject* condition were more likely to show the rejecting comparison cases than participants in the *Motivated to Support* condition, college sample: $B = 5.57$, $SE = 1.06$, $p < .001$, 95% CI [3.90, 8.50], $OR = 262.82$, Prolific sample: $B = 3.64$, $SE = 0.77$, $p < .001$, 95% CI [2.35, 5.51], $OR = 38.17$.

### 2.2.1. Secondary analyses

We next examined participants' justifications for the data they chose to show. We were specifically interested in the participants (91% of college students, 61% of Prolific adults) who selectively showed comparison cases and whether they would generate post hoc explanations that their selected locations were more similar to Hawaii. Recall that, in this study, we did *not* say that the U.S. states or islands were more similar to Hawaii. Nonetheless, 41% of college participants ($n = 45$; Cohen's K = 0.73) and 53% of Prolific adults ($n = 39$; Cohen's K = 0.86) who showed data selectively generated explanations for why the locations that *they* highlighted were more relevant. In line with the experimental manipulation used in Study 1, participants in Study 2 generated reasons why the U.S. states were more relevant (e.g., "…different countries or regions of the world have different mindsets [e.g., the U.S. is more individualistic than many other places]…The other locations in different countries could be taking other measures or have different styles of living that may be affecting their virus cases, instead of Policy A"). However, they also came up with explanations for why the islands were more relevant if those data supported their claims (e.g., "Because there's a distinct difference in pattern[s] between island regions and non-island regions, so the decrease is possibly cause[d] by factors that are related to their geography. Therefore, the other two [U.S.] states that have different geographic conditions do not have the value to be used as a reference to be compared with the effect of Policy A in Hawaii").

Finally, we examined the role of prior beliefs (i.e., the average of participants' political orientation and beliefs about public mandates). As in Study 1, these two beliefs were correlated in both samples (college: $r = .44$, $p < .001$; Prolific: $r = .59$, $p < .001$). Prior beliefs did not predict participants' data selections in either sample or condition, suggesting that participants were using their experimentally assigned political view to guide their data selections.

## 2.3. Study 2 discussion

Study 2 indicates that people will selectively use comparison cases when prompted to influence *others*' causal judgments. Furthermore, participants recognized *which* cases strengthen or weaken others' causal inferences—that is, cases that do or do not imply the counterfactual outcome would have been different, respectively. Notably, when asked to justify their decision to selectively highlight certain cases, people readily came up with reasons why the comparison case they showed was more relevant than the case they withheld (e.g., stating how Hawaii is more similar to other U.S. states than to other islands). Thus, people were able to spontaneously generate the justifications that we used in Study 1 that led to opposing causal judgments.

## 3. General discussion

While prior research indicates that people can reach opposing causal judgments because of disagreements about what *actually* happened, the current research finds that this can be due to the fact that people are made to disagree about what *would have* happened otherwise. Moreover, politically motivated individuals may exacerbate this problem by selectively highlighting comparison cases for others and generating post hoc justifications about the ostensible relevance of those cases. Notably, these effects were robust across samples of college students and politically and socioeconomically diverse U.S. adults. These results point to a previously understudied mechanism of belief polarization—differential *counterfactual* reasoning—that is worthy of future investigation, as claims about what would have happened are more difficult to falsify than claims about what actually happened or what will happen (Teigen, Kanten, & Terum, 2011).

Theoretically, this study sought to test the extent to which a counterfactual theory of causal judgment describes people's reasoning about complex societal events. Study 1 provided robust evidence that counterfactual reasoning matters in these contexts: Despite the fact that many participants endorsed the general effectiveness of public policies, and received information that virus cases declined immediately following policy implementation, they *still* tempered their causal judgments if the comparison cases suggested that the decrease would have occurred without intervention. Thus, our research provides further evidence that counterfactuals underscore people's causal judgments across domains (Gerstenberg et al., 2017; Kominsky & Phillips, 2019; Phillips et al., 2015), including complex societal events. Moreover, we found that people privileged the counterfactuals that were more similar to the target case, suggesting that, like scientists, reasoners seek to reduce confounds (Leatherdale, 2019). This tendency aligns with developmental research finding that even very young learners already prefer unconfounded evidence when making causal inferences (Köksal, Sodian, & Legare, 2021; Schulz & Bonawitz, 2007).

Importantly, we also found evidence that people recognize that comparison cases can be used to persuade others: When reasoning from the perspective of a politician who is motivated to make a specific causal claim, the majority of participants selectively highlighted cases that implied the supporting counterfactual. One participant explicitly acknowledged their ability to deceive others, saying, "*If I wanted to be sneaky, I would talk about how I'm referring to the U.S. and [say that] the decline in [the island] regions may be due to a different, regional factor*." These findings support functional accounts of counterfactual thinking, which suggest that people consider alternatives that facilitate their goals (Epstude & Roese, 2011; Roese &

Epstude, 2017). Here, we showed that flexibility in counterfactual reasoning extends to social goals such as persuasion.

Given that people can selectively highlight comparison cases to persuade others, it will be important to assess the downstream consequences for the individuals receiving this evidence. One possibility is that receivers' pre-existing beliefs play a role in how likely they are to show skepticism toward such selective evidence presentation. In line with this possibility, we found that people were less likely to make the predicted causal judgments when the implied counterfactual contrasted with their prior beliefs. Relatedly, an important future direction would be to assess people's counterfactual judgments *prior* to showing any comparison cases, which would more directly capture how they simulate the counterfactual on their own. It would be interesting to assess how these initial judgments interact with the data that is presented to them, for example, if they become even more resistant to comparison cases that indicate a different counterfactual outcome. Another possibility is that receivers' own experience in selectively presenting comparisons to others might make them more epistemically vigilant when hearing these kinds of arguments (see Ding, Lim, & Heyman, 2022).

Another important next step in this line of research is to examine how people reason about more complex, real-world patterns of data. In the current studies, we presented participants with data on a novel global pandemic that were highly schematic (e.g., there was an immediate decline in virus cases following intervention). Although we demonstrated in the supplemental study that the results replicate with a completely different pattern of data (i.e., a curve that flattens after intervention rather than declines), a more realistic scenario would have a greater delay in the impact of an intervention, and the curve may flatten and then show a slow decline. In addition, it would be interesting to ask people directly what patterns of data would indicate a causal impact of a policy intervention.

While our study focused on a global pandemic, the present account may help to explain belief polarization about a wide variety of societal problems. Consider, for example, disagreements about whether climate change is caused by human activity versus natural processes. Climate change skeptics may point to evidence suggesting the counterfactual that the climate would have changed *regardless* of human activity (e.g., data from other points in history when the climate also changed). In fact, environmental activist websites list such arguments as one of the most common points from climate change skeptics (Rainforest Alliance, 2021). In a similar vein, people may point to different counterfactual outcomes when debating whether social inequalities are caused by structural constraints versus natural traits of groups (Amemiya, Mortenson, Heyman, & Walker, 2023). For example, people may argue that removing structural constraints makes a difference for inequality, while others may point to instances in which inequality persisted regardless of structural changes. We contend that our framework is relevant for understanding belief polarization about any societal problem in which there are competing causal explanations and the counterfactual outcome is uncertain.

### *3.1. Conclusion*

This research sought to explain disagreements in people's causal judgments about societal events, such as a public health intervention during a global pandemic. We propose one mechanism may be that people can be made to have different counterfactuals in mind via comparison cases, leading them to make opposing causal judgments. Moreover, motivated individuals may exacerbate this problem by emphasizing different comparisons. We found evidence for this hypothesis both in our ability to manipulate participants' causal judgments by selectively highlighting comparison cases (Study 1) and in participants' use of this same strategy to manipulate others' judgments (Study 2). This framework offers new insight into why belief polarization can occur even when people agree on the events that actually happened and informs our understanding of belief polarization more broadly (Amemiya et al., 2023).

## Acknowledgments

The writing of this article was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number F32HD098777 and the National Science Foundation Award Numbers 2203810 and 2317714 awarded to Jamie Amemiya. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation.

## Open Research Badges

This article has earned Open Data, Open Materials and Preregistered Research Design badges. Data, materials and the preregistered design are available at https://osf.io/ja84e/ and https://aspredicted.org/LYT_32R, https://aspredicted.org/HLM_WG3, https://aspredicted.org/QXZ_VD8, https://aspredicted.org/VZB_WMF.

## References

Amemiya, J., Heyman, G. D., & Walker, C. M. (2021). How people make causal judgments about unprecedented societal events. In *Proceedings of the 43rd Annual Conference of the Cognitive Science Society*. CogSci.

Amemiya, J., Mortenson, E., Heyman, G. D., & Walker, C. M. (2023). Thinking structurally: A cognitive framework for understanding how people attribute inequality to structural causes. *Perspectives on Psychological Science*, *18*, 259–274. https://doi.org/10.1177/17456916221093593

Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*(1), 135–157. https://doi.org/10.1146/annurev-psych-122414-033249

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

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*(2), 367–405. https://doi.org/10.1037/0033-295X.104.2.367

Connolly, K. (2020). *Coronavirus: How to tell which countries are coping best with Covid*. BBC News. Retrieved from https://www.bbc.com/news/world-europe-54391482

Deane, C., Parker, K., & Gramlich, J. (2021). *A year of U.S. public opinion on the coronavirus pandemic*. Pew Research Center.

Ding, X. P., Lim, H. Y., & Heyman, G. D. (2022). Training young children in strategic deception promotes epistemic vigilance. *Developmental Psychology*, *58*(6), 1128–1138. https://doi.org/10.1037/dev0001350

Eibach, R. P., & Ehrlinger, J. (2006). “Keep your eyes on the prize”: Reference points and racial differences in assessing progress toward equality. *Personality and Social Psychology Bulletin*, *32*(1), 66–77. https://doi.org/10.1177/0146167205279585

Epstude, K., & Roese, N. J. (2011). When goal pursuit fails: The functions of counterfactual thought in intention formation. *Social Psychology*, *42*(1), 19–27. https://doi.org/10.1027/1864-9335/a000039

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Gerstenberg, T., Goodman, N., Lagnado, D., & Tenenbaum, J. B. (2021). A counterfactual simulation model of causal judgments for physical events. *Psychological Review*, *128*, 936–975. https://doi.org/10.31234/osf.io/7zj94

Gerstenberg, T., Peterson, M. F., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2017). Eye-tracking causality. *Psychological Science*, *28*(12), 1731–1744. https://doi.org/10.1177/0956797617713053

Goddu, M. K., & Gopnik, A. (2020). Learning what to change: Young children use “difference-making” to identify causally relevant variables. *Developmental Psychology*, *56*(2), 275–284. https://doi.org/10.1037/dev0000872

Gong, T., & Bramley, N. R. (2023). Continuous time causal structure induction with prevention and generation. *Cognition*, *240*, 105530. https://doi.org/10.1016/j.cognition.2023.105530

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction☆. *Cognitive Psychology*, *51*(4), 334–384. https://doi.org/10.1016/j.cogpsych.2005.05.004

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review*, *116*(4), 661–716. https://doi.org/10.1037/a0017201

Johns Hopkins Coronavirus Resource Center. (2022). *United States—Covid-19 overview*. Retrieved from https://coronavirus.jhu.edu/region/united-states

Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist*, *73*(1), 1–2. https://doi.org/10.1037/amp0000263

Köksal, Ö., Sodian, B., & Legare, C. H. (2021). Young children’s metacognitive awareness of confounded evidence. *Journal of Experimental Child Psychology*, *205*, 105080. https://doi.org/10.1016/j.jecp.2020.105080

Kominsky, J. F., & Phillips, J. (2019). Immoral professors and malfunctioning tools: Counterfactual relevance accounts explain the effect of norm violations on causal selection. *Cognitive Science*, *43*(11), e12792. https://doi.org/10.1111/cogs.12792

Kominsky, J. F., Reardon, D., & Bonawitz, E. (2021). Intuitive judgments of “overreaction” and their relationship to compliance with public health measures. *Journal of Applied Research in Memory and Cognition*, *10*(4), 542–553. https://doi.org/10.1016/j.jarmac.2021.11.001

Leatherdale, S. T. (2019). Natural experiment methodology for research: A review of how different methods can support real-world research. *International Journal of Social Research Methodology*, *22*(1), 19–35. https://doi.org/10.1080/13645579.2018.1488449

Lewis, D. (1973). Causation. *Journal of Philosophy*, *70*(17), 556. https://doi.org/10.2307/2025310

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review*, *115*(4), 955–984. https://doi.org/10.1037/a0013256

Lucas, C. G., & Kemp, C. (2015). An improved probabilistic account of counterfactual reasoning. *Psychological Review*, *122*(4), 700–734. https://doi.org/10.1037/a0039655

Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Oxford University Press.

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*(3), 650–654. https://doi.org/10.1016/j.jesp.2007.03.005

Moran, C., Campbell, D. J. T., Campbell, T. S., Roach, P., Bourassa, L., Collins, Z., Stasiewicz, M., & McLane, P. (2021). Predictors of attitudes and adherence to COVID-19 public health guidelines in Western countries: A rapid review of the emerging literature. *Journal of Public Health*, *43*(4), 739–753. https://doi.org/10.1093/pubmed/fdab070

Navarro, D. J. (2016). *Learning statistics with R: A tutorial for psychology students and other beginners* (0.5.2). University of Adelaide.

Nyhout, A., & Ganea, P. A. (2021). Scientific reasoning and counterfactual reasoning in development. In Jeffrey J. Lockman (Ed.), *Advances in child development and behavior* (Vol. 61, pp. 223–253). Elsevier. https://doi.org/10.1016/bs.acdb.2021.04.005

Phillips, J., Luguri, J. B., & Knobe, J. (2015). Unifying morality's influence on non-moral judgments: The relevance of alternative possibilities. *Cognition*, *145*, 30–42. https://doi.org/10.1016/j.cognition.2015.08.001

Quillien, T., & Barlev, M. (2022). Causal judgment in the wild: Evidence from the 2020 U.S. Presidential Election. *Cognitive Science*, *46*(2), e13101. https://doi.org/10.1111/cogs.13101

Quillien, T., & Lucas, C. G. (2023). Counterfactuals and the logic of causal selection. *Psychological Review*. https://doi.org/10.1037/rev0000428

Rainforest Alliance. (2021). *6 Claims Made by Climate Change Skeptics—And How to Respond*. Rainforest Alliance. https://www.rainforest-alliance.org/everyday-actions/6-claims-made-by-climate-change-skeptics-and-how-to-respond/

R Core Team. (2020). R: A language and environment for statistical computing. (4.0.4) [Computer software]. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Robinson, G., McNulty, J. E., & Krasno, J. S. (2009). Observing the counterfactual? The search for political experiments in nature. *Political Analysis*, *17*(4), 341–357. https://doi.org/10.1093/pan/mpp011

Roese, N. J., & Epstude, K. (2017). The functional theory of counterfactual thinking: New evidence, new challenges, new insights. In James M. Olson (Ed.), *Advances in experimental social psychology* (Vol. *56*, pp, 1–79). Elsevier. https://doi.org/10.1016/bs.aesp.2017.02.001

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, *43*(4), 1045–1050. https://doi.org/10.1037/0012-1649.43.4.1045

Smith, A., & Edwards, E. (2020). *Trump tried to justify rising Covid cases by pointing to Europe. Experts say he's wrong*. Retrieved from https://www.nbcnews.com/health/health-news/trump-tried-justify-rising-covid-cases-pointing-europe-experts-say-n1246786

Teigen, K. H., Kanten, A. B., & Terum, J. A. (2011). Going to the other extreme: Counterfactual thinking leads to polarised judgements. *Thinking & Reasoning*, *17*(1), 1–29. https://doi.org/10.1080/13546783.2010.537491

Walker, C. M., & Nyhout, A. (2020). Asking "why?" and "what if?": The influence of questions on children's inferences. In Lucas Payne Butler, Samuel Ronfard, Kathleen H. Corriveau (Eds.), *The questioning child: Insights from psychology & education* (pp. 252–280). Cambridge University Press.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Woodward, J. (2005). *Making things happen: A theory of causal explanation*. Oxford University Press.

Xu, S., Coman, I. A., Yamamoto, M., & Najera, C. J. (2023). Exposure effects or confirmation bias? Examining reciprocal dynamics of misinformation, misperceptions, and attitudes toward COVID-19 vaccines. *Health Communication*, *38*(10), 2210–2220. https://doi.org/10.1080/10410236.2022.2059802

**Supporting Information**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Online Appendix