RESEARCH ARTICLE

Systematic mapping of climate and environmental framing experiments and re-analysis with computational methods points to omitted interaction bias

Lukas Fesenfeld[1,2]*, Liam Beiser-McGrath[3], Yixian Sun[4], Michael Wicki[5], Thomas Bernauer[2]

1 Department of Social Sciences, University of Bern, Bern, Switzerland, **2** Department of Humanities, Social and Political Sciences, ETH Zürich, Zürich, Switzerland, **3** Department of Social Policy, London School of Economics, London, United Kingdom, **4** Department of Social and Policy Sciences, University of Bath, Bath, United Kingdom, **5** Deptartment of Civil, Environmental and Geomatic Engineering, ETH Zürich, Zürich, Switzerland

* lukas.fesenfeld@gess.ethz.ch

Abstract

Ambitious climate policy requires acceptance by millions of people whose daily lives would be affected in costly ways. In turn, this requires an understanding of how to get the mass public on board and prevent a political backlash against costly climate policies. Many scholars regard 'framing', specially tailored messages emphasizing specific subsets of political arguments to certain population subgroups, as an effective communication strategy for changing climate beliefs, attitudes, and behaviors. In contrast, other scholars argue that people hold relatively stable opinions and doubt that framing can alter public opinion on salient issues like climate change. We contribute to this debate in two ways: First, we conduct a systematic mapping of 121 experimental studies on climate and environmental policy framing, published in 46 peer-reviewed journals and present results of a survey with authors of these studies. Second, we illustrate the use of novel computational methods to check for the robustness of subgroup effects and identify omitted interaction bias. We find that most experiments report significant main and subgroup effects but rarely use advanced methods to account for potential omitted interaction bias. Moreover, only a few studies make their data publicly available to easily replicate them. Our survey of framing researchers suggests that when scholars successfully publish non-significant effects, these were typically bundled together with other, significant effects to increase publication chances. Finally, using a Bayesian computational sparse regression technique, we offer an illustrative re-analysis of 10 studies focusing on subgroup framing differences by partisanship (a key driver of climate change attitudes) and show that these effects are often not robust when accounting for omitted interaction bias.

Introduction

Climate and environmental communication is an essential lever for building public understanding of the problem's severity and increasing support for climate change (policy) solutions [1–3]. One widespread climate and environmental communication technique is framing [4–6]. Framing is an inherent part of communication and occurs when actors use messages to alter people's preferences by changing the presentation of an issue or an event [2, 7, 8]. In climate policy, politicians or other stakeholders may emphasize specific subsets of preexisting arguments–such as economic or health-benefits of climate change mitigation [5, 9, 10]–in an attempt to influence public opinion in favor of (or against) climate action (so-called emphasis framing). Communicators may also use different, but logically equivalent phrases to describe climate change mitigation (so-called equivalence framing). Yet, is framing an effective communication technique to alter public opinion about climate change–especially across different subgroups of a population? Many studies on climate and environmental communication suggest that framing can effectively influence public opinion across population subgroups as it safeguards individuals' identities by appealing to their existing values and prior beliefs [11–16]. Framing theory holds that the effectiveness of framing in altering people's attitudes varies according to whether the related information is available in individuals' memories, is accessible, and is evaluated as applicable in a given situation [7, 17]. The framing literature also builds on a bounded rationality model [18] and often assumes that citizens have limited capacity to process information systematically [7, 19–21]. From this perspective, individuals use frames as heuristics to minimize cognitive effort when forming policy attitudes [17, 22, 23].

Most framing studies on climate and environmental communication look at framing effects across population subgroups (i.e., heterogeneous framing effects). According to directional-motivated reasoning models [2], framing political messages around prior beliefs and values can reduce cognitive dissonance [24, 25] and lead to stronger framing effects on individuals' attitudes. For example, empirical studies have shown that individuals perceive frames tailored to their ideological core beliefs as less threatening. Accordingly, many studies (especially in polarized political contexts such as the United States) assume that frames aligned with citizens' ideologies and party identification are more effective at altering climate policy attitudes [2, 26–28].

Empirical evidence for the effect of framing is often generated through experiments embedded in survey-, field-, or lab studies. Such experiments, in which study participants are randomly assigned to treatments and control conditions, are seen as a gold standard for assessing the effectiveness of frames in altering public opinion [2, 3]. Typically, in such experiments, study participants are randomly confronted with differently framed messages. The aim is to assess how these different framing treatments alter respondents' climate beliefs, attitudes, and behaviors, particularly in comparison across population subgroups. For example, Bernauer and McGrath [5] as well as Bain et al. [10, 29] randomly assigned individuals to different messages that either emphasized the risks of failing to combat climate change (control frame) or highlighted different co-benefits of climate mitigation, such as economic, community building, and health benefits (treatment frames), to study if framing climate mitigation policy around co-benefits instead of risks increases public support. Others, such as Hung and Bayrak [1], have randomly varied specific terminology (e.g., climate change vs. climate crisis) used in climate communication to assess how such re-labeling affects citizens' perceptions and attitudes across different population subgroups. While many researchers presume such framing to be an effective communication technique for altering mass public opinion and behavior concerning climate change [2, 10, 12, 29–31], some scholars have expressed doubts [5, 21, 32–36]. These scholars argue that on salient and contested issues, such as climate change, people are likely to hold relatively stable, consciously formed preferences and cannot be easily

manipulated through simple framing [5, 21, 32–36]. Some also suspect a bias against reporting non-significant effects in the current framing literature [22, 36]. They point to the use of experimental designs and statistical methods that involve risks of producing noisy effects with low external validity and omitted interaction bias, especially when studying heterogeneous framing effects across population subgroups [22, 37–39].

Building on previous maps of the environmental framing and communication literature more broadly [2, 3, 6, 40–43], we provide a systematic mapping of the experimental framing literature in this field and a critical appraisal of framing effects across population subgroups reported in existing studies. The lack of publicly available replication data and the large variation of experimental designs and outcome variables present in the literature do not allow for conducting a meaningful meta-analysis that compares standardized effect sizes across different framing experiments. Thus, instead of conducting a meta-analysis, the key contributions of this paper are twofold: First, we provide a systematic mapping of existing framing experimental research on climate and environmental framing and present results of a survey with authors of these studies. Here, we show that most studies report significant subgroup effects but do not use advanced methods to check for the robustness of subgroup effects. Second, we thus illustrate the use of Bayesian sparse regression and computational methods for assessing the robustness of heterogeneous framing effects and preventing potential omitted interaction bias. Given the prominence of discussions about the robustness of partisan framing effects across ideological subgroups, we re-analyze data from a set of published studies using (different from most published work to date) LASSOplus, a machine-learning-based Bayesian sparse regression method. This more advanced computational method reduces potential overfitting of statistical models and the omitted interaction bias that can lead to non-robust and misleading subgroup framing effects in classical OLS regressions [38, 44, 45].

Results

Our systematic mapping is based on the PRISMA identification standard [46] and best-practice guidelines for systematic mappings [40, 47] (for details, see Methods). After a scoping analysis of the existing experimental framing literature in climate and environmental communication studies (for details, see Methods), we find that there exists a clear lack of experimental research on equivalence framing (i.e., logically equivalent but different descriptions of the same issue) and thus focus our review here on experimental emphasis framing studies (i.e., frames varying specific subsets of preexisting arguments). In essence, while equivalence framing could in principle be a powerful communication technique for altering people's climate and environmental attitudes and policy preferences [2, 48, 49] (e.g., by altering messages on whether a person has a 10% risk of dying or a 90% chance of surviving due to climate-induced extreme weather), in the context of real-world climate and environmental communication and respective research appears to be a less prominent and studied framing strategy. We thus focus our mapping and critical appraisal on emphasis-framing experiments.

In total, we identified 121 emphasis-framing experimental studies published in 46 peer-reviewed journals between 2007 and 6/2020. All studies use an experimental design to assess the effects of different types of emphasis framing treatments on an individual's climate and environmental beliefs, attitudes, and behaviors (see Methods Table B in S1 Text for the complete list of studies). While most studies considered in our mapping relate specifically to climate change, some studies also include treatment groups and dependent variables related to other environmental issues, such as air pollution. We decided to include all these studies to increase the scope of our findings. According to the experimental stimuli used in these 121 studies, we classified them into six climate and environmental emphasis framing research

categories. These are issue and solution frames, value- and norm-based frames, re-labeling frames, psychological distance frames, consensus and uncertainty frames, and source cue frames (for further details, see below and in Table A in S1 Text).

Potential risk of over-reporting significant framing effects

Our first goal is to systematically map existing emphasis framing experiments on climate and environmental issues. Fig 1 provides an overview of our mapping (for further details, see Methods and Table B in S1 Text). Approximately 92% (n = 111) of the framing studies we reviewed report significant main framing effects. Only 7% (n = 9) report non-significant main effects, and 1% (n = 1) do not report any main effects. Around 20% (n = 24) of all studies do not report and discuss any heterogeneous treatment effects (e.g., interactions between participants' characteristics, such as party ideology, and framing treatments). In contrast, 70% (n = 85) of all reviewed studies identify at least one significant subgroup effect, while 10% (n = 12) report no significant subgroup effects at all. In other words, 88% (n = 85) of all studies that report on heterogeneous treatment effects (n = 97) find some significant subgroup effects.

Besides the overview provided in Fig 1, we also observe some temporal trends in the data (for further details, see Table B in S1 Text and Fig B-F in S1 Text). First, the total number of published emphasis-framing experimental studies per year has substantially increased since 2007. While in 2007 there was only one published framing experimental study, in 2019 we identified 18 framing experimental studies in our mapping. Over the review period, the share of studies that employ large-n (n>1000), non-convenience and population-representative samples has increased. However, while the size and representativeness of study samples has grown over time, studies have rarely reported statistical power calculations, especially for estimation of sub-group effects. Thus, while an increasing number of studies report significant subgroup effects (see Fig E in S1 Text), there is a risk that the estimation of these effects is underpowered. Moreover, the vast majority (120 out of 121) reviewed studies use classical linear or logistic regression models to estimate main and sub-group effects. Only one of the reviewed articles has used more advanced statistical methods, such as Bayesian sparse regression methods, to control for potential omitted interaction biases and double-check the robustness of estimated heterogenous effects. Below, we discuss the purpose and application of these methods in more depth (see "Critical appraisal and re-analysis of framing studies").

Survey-based experiments have experienced the largest growth rate over the review period–from 0 studies using this study design type to 16 studies using survey experiments in 2019. In contrast, field- and lab-based experiments or the combination of field and lab-based experiments with survey experiments stays at a very low level (around 0–1 studies per year). While US-focused studies have the largest overall share of all experiments, they also experienced the strongest average growth rate per year between 2007 and 2020. Yet, over time also the number of (internationally) comparative framing experiments increased–but at a lower average growth rate per year. Most of these comparative framing experiments compared the US with another country, often from Europe. Even though in recent years some framing experiments have been conducted in developing and emerging economies, such as Brazil, China, or India, we still see a substantial lack of framing experiments in the developing country context.

Moreover, we can identify some trends in terms of the framing types being studied (see Table A in S1 Text and Fig B in S1 Text). In the first period from 2008 to 2013, psychological distance frames belong to the most widely studied framing types. For example, some of these earlier framing studies [50, 51] varied the spatial, social, and temporal distance of climate and environmental impacts to assess whether people support ambitious mitigation more when they perceive climate and environmental change as a proximate problem. From 2013 onwards,

Fig 1. Overview of 121 emphasis framing experimental studies in the field of climate and environmental politics, economics and psychology published between 2007 and 6/2020. Note: Panel designs (i.e., repeated measurements for the same study participants at two or more points in time) are used to study whether framing effects vary over time (e.g. how long the effect of a one-time exposure lasts). Competing frames are used to emphasize competing arguments in a debate (e.g., pro and contra climate mitigation messages).

https://doi.org/10.1371/journal.pclm.0000297.g001

issue and solution frames were the most widely studied frames, with a peak of up to 10 of these experimental studies published in 2016. Issue and solution frames often emphasize environmental risks and co-benefits of environmental protection or climate mitigation. For example, some studies [10, 29] in this category highlight that emphasizing co-benefits of climate mitigation (such as technological innovation, green jobs, community building, or health improvements) could foster public support for ambitious mitigation policies. The second most widely studied framing type in our sample of reviewed studies, especially since 2013, are value- and normative-based frames that emphasize values and social norms, and attribute responsibility for environmental problems and solutions. In recent years, also source-cue frames, varying the sender of a message, and re-labeling frames, changing specific words (e.g., global warming vs. climate change) in a message, have been studied more widely. In contrast, research interest in frames varying the degree of consensus or uncertainty about climate change existence and impacts has decreased over time.

Finally, the number of articles including published replication data has increased since 2007 but none of these studies were preregistered. In the last three years, the majority of

published framing experiments makes replication data publicly available. The number of articles that reported non-significant main or heterogeneous treatment effects has also increased, especially since 2013. For example, before 2013 none of the reviewed studies reported any non-significant heterogeneous framing effects, while in 2019 at least four studies did so. However, the number of reported non-significant effects has increased only very slowly, and most published studies still report only significant framing effects.

Author survey suggests that bundling non-significant and significant effects eases publication

One concern that arises in view of the large proportion of studies finding statistically significant framing effects is that there may be a file-drawer problem, where mainly significant effects are published, and non-significant ones not [22]. We implemented an online survey to assess how the authors of published framing experiments experienced the publishing process and dealt with non-significant framing effects they encountered (for further details, see Section VI in S1 Text). We contacted all 173 authors of the 121 publications via email and received a total of 63 responses (a response rate of 36% of all authors and around 52% of all reviewed studies; most often the lead author of each study responded to our survey). We find that around 80% (n = 50) of all respondents also identified non-significant effects in their framing experiments. Around 76% (n = 38) of this subset of authors tried to publish their results, including non-significant effects in peer-reviewed journals. And again, only 63% (n = 24) of this sub-subset of authors who tried to publish non-significant effects (or 48% of survey respondents that identified them) were able to successfully publish studies with non-significant effects. However, according to these authors, in most cases, publishing their findings was only possible when non-significant results were bundled together with other significant effects (for further details, see Methods). Therefore, the observed gap between the small number of published non-significant framing effects (see Fig 1 above) and the substantially larger number of identified non-significant framing effects reported by the surveyed authors strongly suggests a potential publication bias towards significant treatment results.

Lack of publicly available data makes it difficult to formally assess file-drawer problem

Formally assessing the potential existence and magnitude of a 'file-drawer problem' would require public access to the data and re-analyses of the original study results as part of a meta-analysis. However, only 23% (n = 28) of the 121 articles we reviewed made their data publicly available. In addition, out of those 93 reviewed articles whose data was not published, we obtained data for 29 studies by contacting authors via email (i.e., overall, we could not get access to the data of more than 53% (n = 64) of all reviewed studies). The large number of experiments that report significant framing effects without publishing data or making replication data available on request thus raises significant barriers for researchers attempting to assess the robustness of published results. For example, extra and often unsuccessful efforts to obtain access to data increase the costs of systematically re-analyzing existing studies, assessing their results' robustness, and estimating the size of the potential file-drawer problem.

Finally, the lack of publicly available data and large variance in experimental designs prevents meaningful meta-analyses on the distribution and magnitude of average framing effects. For example, while some experimental designs use a (placebo) control group, others only compare effects for different framing groups. Also, the manifold larger and smaller variations in treatment wording and design make a proper comparison of effect sizes in a meta-analysis very difficult. Besides systematic mapping of the work in this area, we thus focus our

illustrative critical appraisal on one key area of interest to many researchers: heterogeneous framing effects.

Making inferences about framing effects by sub-group

As mentioned above, climate and environmental communication researchers are often interested in how framing effects vary across population subgroups. Druckman and McGrath [2], for instance, note that "rather than continually testing the impact of one frame after another, the literature would benefit from [...] investigating which types of messages resonate in light of motivations and particular prior beliefs, values and identities." For example, in view of the possibility of directional-motivated reasoning, one prominent argument in the climate and environmental communication literature is that frames aligning with peoples' prior beliefs reduce cognitive dissonance [24, 25] and are thus more effective at shifting public opinion about climate change.

Researchers, therefore, typically split their sample into groups based upon respondent characteristics and then re-estimate their statistical models to assess, for example, whether the framing effect is more (or less) significant for Democrats or Republicans in the United States. While this is a valid approach for generating *descriptive* insights regarding variation in treatment effects, researchers occasionally slip into the use of causal language. For instance, some of the reviewed studies state that "issue frames can lead Republicans and those on the political right to view climate change policy as less important" [52] or highlight that "Republicans [...] increase their support if Republican politicians take leadership roles in supporting proposed bills, and if air quality benefits are emphasized" [53].

Crucially, however, random assignment of treatment does not guarantee the identification of heterogeneous causal effects [54]. In this section, we therefore outline descriptive and causal inference in the exploration of sub-group effects and the fundamental challenges to inference through the example of omitted interaction bias. We also examine how incorporating additional interaction effects in a manner suggested by previous research [38, 44, 45, 55] affects the estimates of previously published subgroup effects. Finally, we reflect on these challenges in the pursuit of causal inference for sub-group effects, while also highlighting recent work that allows for more principled descriptive examination of sub-group effects.

One example of a threat to causal inference from considering sub-group effects in isolation is *omitted interaction bias* [38, 44, 45, 55]. Omitted interaction bias occurs where differences between the sub-groups on other characteristics, such as age, education, and income, also result in heterogeneous treatment effects that are left unmodelled, which are absorbed by the included interaction effect. While such sub-group analysis is valid for generating descriptive inferences about how treatment effects vary across these individuals, the typical approach of researchers is to interact the treatment with the sub-group of interest, which is vulnerable to omitted interaction bias. The lack of randomized experimental manipulation and/or adjustment for other potential heterogeneous treatment effects diminish the ability of researchers to draw causal inferences about how treatment effects vary across subgroups.

In the existing literature, however, researchers often draw causal inferences from the results of sub-group analyses, where the sub-group membership is not randomized and/or other sub-group effects are left unmodelled. For example, some studies conduct sub-group analyses where they estimate treatment effects separately for Democrats and Republicans. However, party identification is not a trait that can be randomly assigned. Thus, not accounting for other characteristics that may also moderate the effect of treatments associated with party identification (e.g., age, education, gender) could lead to biased estimates. However, instead of making a descriptive statement about significant differences in framing effects between Democrats and

Republicans, many studies make causal inferences and policy recommendations about which frames most successfully shift support across different partisan groups [53].

As many sub-group characteristics of interest, such as partisan identification, are difficult if not impossible to manipulate experimentally, researchers likely need to adjust for other heterogeneous framing effects in order to make causal claims. Yet, several studies [37–39, 56–58] have shown that standard specification choices and statistical methods (e.g., ordinary least squares [OLS] regressions) can run the risk of producing non-robust and noisy heterogeneous framing results because of overfitted models, even in perfectly randomized experiments [37, 38, 54]. If researchers wish to exclude this potential risk to their causal inferences, then they need to assess how sensitive published heterogeneous framing effects are to model misspecification. Note however, that these may not be all potential threats to causal inference, see for instance Bansak [54] for a generalized framework for estimating causal moderation effects and possible methods for conducting sensitivity analysis in this context.

Recent research points towards the use of machine learning in estimating such heterogeneous framing effects across population sub-group effects and prevent *omitted interaction bias* [38, 44, 45, 55] (for further details, see Methods). With uncertainty about the true data generation process, the number of relevant heterogeneous treatment effects and conditional effects amongst relevant covariates results in large numbers of parameters to be estimated. In such circumstances, OLS estimation faces problems of statistical efficiency. Machine learning methods for variable selection, such as Lasso, overcome this problem by setting parameters with little predictive power to zero. This improves statistical efficiency by excluding non-meaningful heterogeneous effects, thereby reducing the set of heterogeneous effects estimated. As our systematic mapping (see above) shows, the vast majority (120 out of 121) reviewed studies used classical linear or logistic regression models to estimate subgroup effects and does not make use of more advanced methods to assess or mitigate omitted interaction bias, such as LASSOplus [38].
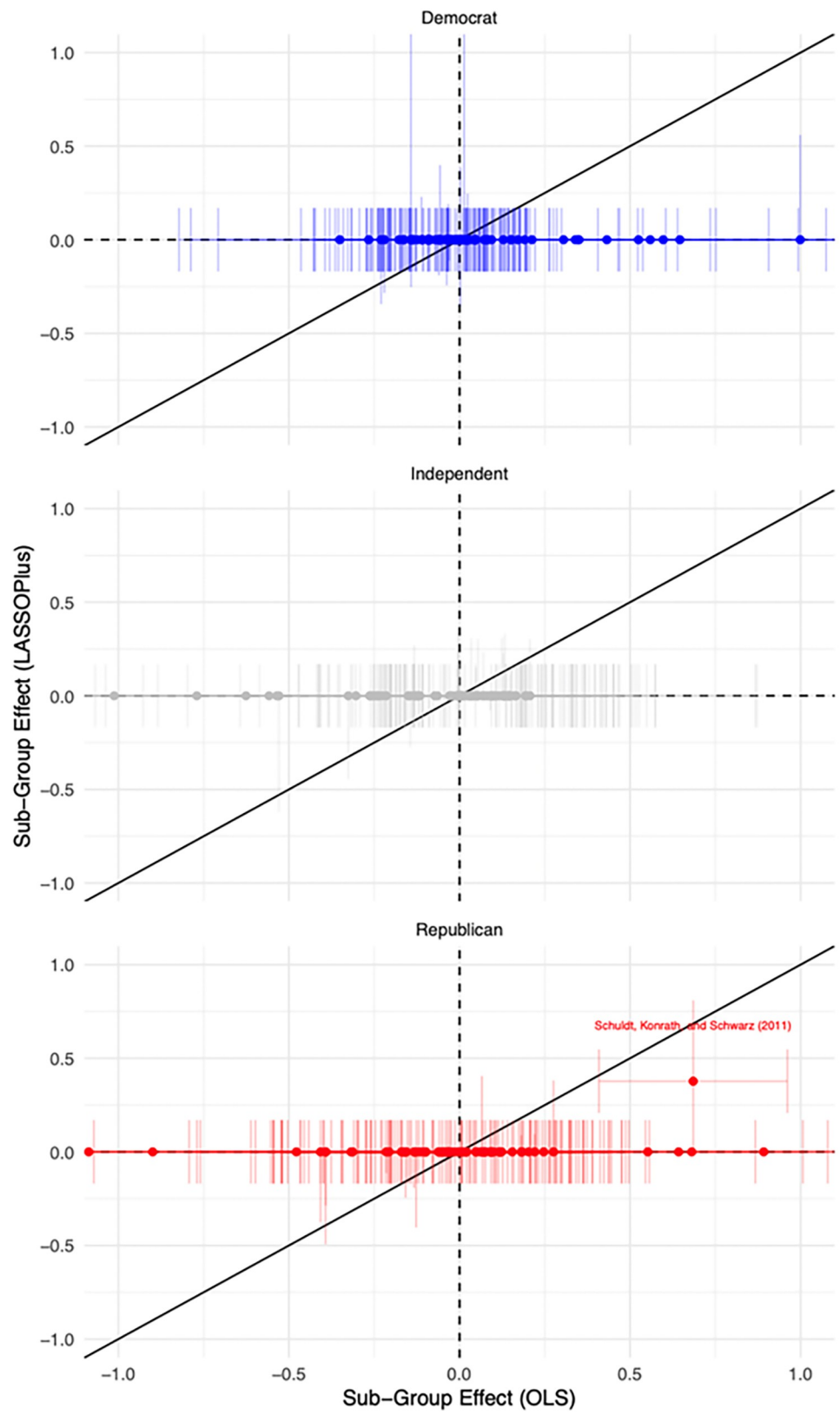
To re-assess published framing effects along these lines, we employed one such method, LASSOplus [38]. This estimator was chosen as it is explicitly tailored to estimating heterogeneous treatment effects, as is often the goal of framing studies, with substantial simulation evidence showing its improvement upon OLS for this task. LASSOplus has also been shown to be a superior approach compared to other sparse regression techniques that seek to overcome overfitting and omitted interaction bias [38]. LASSOplus allows for simultaneous estimation of sub-group effects for all included pretreatment covariates (e.g., age, education, income, and gender) and regularizing insignificant effects to avoid overfitting (for further details, see Methods). Many studies covered by our mapping and review focus on the politically polarized country case of the United States (see Fig 1) and examine how framing effects vary by respondents' partisanship. Hence, we decided to illustrate the sensitivity of heterogeneous framing effects by assessing the robustness of partisan subgroup effects for studies with publicly available data. From the 28 studies with publicly available data, 10 studies focus on partisan subgroup effects in the US context. We thus concentrated our re-analysis of partisan subgroup effects on these 10 available studies.

Sub-group effect estimates are sensitive to the exclusion of other sub-group effects

Fig 2 displays re-estimated subgroup effects for Democrats, Independents, and Republicans for the ten studies with publicly available data. We compare effects estimated using both classical OLS and LASSOplus [38].

While all of the original studies report significant partisan subgroup effects when using OLS (see x-axis of Fig 2), we find that for the vast majority of re-analyzed studies (9 out of 10)

partisan sub-group effects are not statistically distinguishable from zero when using LASSO-plus (see y-axis of Fig 2 and for further details, see Methods section on the exact treatment design of each study). In addition to assessing the robustness of published subgroup effects by partisanship, we also explored other potential subgroup effects (e.g., by age, education, income, gender) that were not the focus of the original studies. However, also in this explorative analysis of heterogeneous framing effects, we do not find support for robust variation in framing effects across different subgroups (see Table C-L in S1 Text).

Overall, our illustrative re-analysis of heterogeneous framing effects highlights the sensitivity of sub-group effects to the inclusion of other relevant sub-group effects. What implications do the findings of this critical appraisal of published framing effects across partisan subgroups have in the real world?

First, in general, our re-analysis suggests that other variables that correlate with partisanship also interact with framing treatments and that the significant partisan subgroup effects detected in the original studies may be absorbing other key interactions, such as between frames and age. For instance, the re-analysed studies by Saunders (id73) and Hardisty et al. (id74) both indicate significant interaction effects between their respective framing treatments and partisanship. In our critical appraisal with LASSOplus, we could, however, not replicate this result but instead found that age significantly interacted with the frames. While LASSO-plus penalizes very small effect sizes to prevent overfitting and reduce omitted interaction bias, it does so less often for larger effect sizes. This implies that, for instance, Hardisty et al.'s finding that cost framing changed preferences among Republicans but not Democrats might still be existent but substantially smaller than originally expected. In this case, the strong subgroup effect for age in the reanalysis suggests that observed partisan effect could in fact be attributed to underlying differences in age among partisan groups.

Second, our analysis and discussion encourage researchers to think exactly about what type of inference they wish to make when estimating sub-group effects. Do they wish to simply *describe* how a treatment varies across different groups or do they seek to claim that respondents' characteristics *cause* them to respond differently to the treatment. If the latter, then researchers need to reflect upon and discuss the inferential challenges for doing so and use statistical approaches such as those discussed here to assess the sensitivity of our results. If the former, then researchers should be upfront and clear with this decision, as well as use other explorative and descriptive approaches toward treatment effect heterogeneity [59, 60] to understand their findings.

Third, our analysis warrants us to think more carefully about the robustness of hypothesized subgroup effects, but also about the substantial meaning of effect sizes. To illustrate this point, we discuss the results of our re-analysis for two re-labeling framing experimental studies published by Schuldt et al.–for one [61] partisan sub-group effects were still significant and for the other [62] not. In Schuldt, Konrath and Schwarz (id86) [61], where reframing has a significant effect amongst Republicans, after adjusting for additional heterogenous treatment effects. The original study and our re-analysis indicate that Republicans were more likely to agree that the phenomenon is real when it was referred to as climate change rather than global warming.

However, Democrats were not affected by the specific question-wording. In contrast, we did not observe such significant differences with the other re-labeling framing experiment results by Schuldt, Enns and Cavaliere [62]. The original study [62] makes a causal inference about subgroup effects by stating that 'the US public is more likely to doubt the existence of global warming than climate change—and that Republicans are driving the effect". However, in this case the LASSOplus re-analysis does not support such causal inference. Why would only one of the two re-labeling experiments that varied the terms "climate change" and "global warming" lead to differences in partisan sub-group effects after adjusting for other heterogenous treatment effects? Here, it is key to consider the differences in outcome variables across the two studies, namely belief in the existence of climate change/global warming in Schuldt, Konrath and Schwarz [61] and concern about climate change/global warming in Schuldt, Enns and Cavaliere [62]. In the case of the first Schuldt study [61], effect sizes for Republicans (at least in the representative sample in 2011) were substantial. However, in the second study by Schuldt et al. [62] effect sizes of re-labeling on the concern outcome were not as large. This example thus illustrates that researchers should be careful when interpreting subgroup effects for different outcome variables in framing studies. The choice of outcome variables can affect the size of identified effects and their robustness. Additionally, given the demands of statistical power in reliably estimating sub-group effects, researchers should take care to engage in ex-ante power analysis to ensure sufficient power in the design stage of their studies or use ex-post power analysis to highlight potential challenges when exploring sub-group effects after data has been collected.

Fourth, our re-analysis can also help us to more carefully interpret the robustness of sub-group results from different study designs and timings. For example, Soutter and Mottus [63] as well as Schuldt et al. [64] point out that differences in results between their replicated re-labeling experiments were likely to stem from differences in sample composition and the large time difference between implementation of the original and replicated experiments. Given that LASSOplus penalizes very small effect sizes to prevent overfitting, it also cautions researchers to over-interpret smaller effect sizes across subgroups (e.g., partisanship) and more carefully reflect how the specific research design (e.g., self-reported or behavioral outcomes, question wording, survey or field experiment etc.) and timing (e.g., framing salience in public discourse at the time of the study) affect the experimental results.

This discussion highlights the importance of paying more attention to the size of estimated framing effects. As we show, larger effect sizes reduce the risk of drawing false causal inferences due to omitted interaction bias (e.g., falsely attributing a detected subgroup completely to differences across partisan groups rather than also age, education, or gender differences). Ultimately, policymakers and climate communicators care about the substantive effect sizes of different framing options when planning their communication strategies.

Discussion

Our study suggests that researchers on climate communication need to reconsider dominant approaches and methods used in this field to identify robust and meaningful framing effects. Our systematic mapping, author survey, and re-analysis point to a lack of publicly available replication data and a potential file-drawer problem in the existing publications of climate and environmental communication research. Hence, journals should encourage researchers to publish studies with null results and make their data publicly available. While our systematic mapping indicates that data availability was a larger issue in the past, the increasing trend of authors that now publish replication data is encouraging. It shows that researchers and journals have taken first steps to respond to the 'replication crises'.

We also showed that most reviewed studies do not employ advanced statistical methods to check for the robustness of their subgroup effects estimated with classical methods (e.g., linear or logistic regressions). Building on this finding, we illustrate the potential limitations in the identification of sub-group framing effects, due to methodological approaches that fail to account for potential omitted interaction bias. Our re-analysis of partisan framing effects demonstrates that accounting for other plausible heterogenous treatment effects, through more fully specified models, results in few robust sub-group effects. This suggests that statistically significant sub-group effects may be driven by omitted interaction bias, an important challenge that must be addressed if researchers wish to make *causal inferences* about treatment effects for specific sub-groups. Ideally, researchers cannot only randomly vary the framing treatment but also the moderator, for which they aim to investigate interaction effects. However, for many important moderators, such as partisanship, it is difficult or even impossible to induce random variation. Accordingly, researchers interested in *causal inferences* about sub-group effects ensure sufficient statistical power to identify these effects, in the presence of unmodelled heterogenous treatment effects, and measure those potential confounders that might cause omitted interaction bias. Combining this with statistical methods that perform well in multidimensional settings [38, 44, 45], typically based on machine learning algorithms, allows increasing confidence in causal interpretations of sub-group effects in the presence of potential model misspecification and omitted interaction bias. While not presently common in the field, this approach would increase the robustness and credibility and generalizability of findings based on framing experiments. It would also make researchers reflect more carefully about the size of detected framing effects.

On this basis, researchers should move beyond null hypothesis testing to examine whether the treatment effect estimated is substantively meaningful [65–67]. Equivalence tests are a prominent approach for doing so. Originating in biostatistics, but increasingly used also in the social sciences, "two one-sided tests" (TOSTs) allow researchers to formally assess whether the estimated treatment effect is statistically significantly different from a non-meaningful effect specified by the researcher (see Section II in S1 Text for an illustration of this approach). While placing a greater burden on the researcher, by having to explicitly specify a meaningful effect and conducting additional analyses, this approach would increase readers' confidence that the framing effects identified are substantial and worthy of attention of policymakers.

Finally, exploring effective climate and environmental communication strategies requires new transdisciplinary research collaborations that embrace the real-world complexities of communication.

First, while different types of frames have been subject to empirical evaluation, our review shows that most of these framing experiments were embedded in surveys at one point in time and in one specific country, mostly the United States. The lack of comparative and panel designs strongly limits the external validity of results. That is because framing is likely to unfold its effects over time [7, 68] and vary by context [22, 69, 70]. Moreover, messages emphasizing only one side of a political argument may lead to artificially large framing effects and reduce the external validity of experiments [22, 69–71]. In reality, different political elites employ multiple combined and competing rational and emotional cues, building on voice, imagery, and written text [72]. Due to filter bubbles, motivated reasoning, and confirmation biases [2], in reality people from different societal or partisan groups might not be confronted with particular types of frames or only in combination with strong competing messages and adverse source cue interpretations. In this sense, the communication context and individual-level heterogeneity interact in many ways with the effectiveness of the specific framing approach (see Table A in S1 Text for different types of frames) [22]. Future research could account for this by reconsidering established methods of studying heterogenous framing

effects in single survey-experiments. While there is certainly room for using one-time survey-experiments, we believe that climate and environmental communication research would greatly benefit from embracing more comparative and panel approaches that assess the combination of more realistic treatments (e.g., randomizing competing messages that combine different types of frames and use both rational and emotional cues) across different contexts, subgroups and periods of time.

Second, we believe that field-experiments are very useful for studying how framing interventions affect both attitudes and behaviors in real-world settings and across subgroups [73]. However, currently only a few studies directly compare stated attitudes and revealed behaviors in both survey- and field-experimental environments. For example, Levine and Kline [73] show that two different climate risk frames increased people's stated concern about climate change in a survey-embedded setting but that these two frames also decreased participants' revealed engagement in political action in field-experiments. These results point to a puzzling divergence between stated attitudes and revealed behaviors and caution against an overly optimistic view about the behavioral effects of framing. Future research, combining field- and survey-experiments, should particularly check for the differences in framing effect sizes between stated attitudes and revealed behaviors for different population subgroups.

Third, our mapping indicates a lack of mixed-method approaches that combine different types of randomized experiments (e.g., survey-, field-, lab-based) with quasi-experimental (e.g., synthetic control), computational (e.g., natural language processing), sensory (e.g., eye tracking), and/or qualitative methods (e.g., cognitive interviews). These mixed-methods approaches would allow analysts to better elucidate the moderating and mediating factors that influence how different individuals process information and react (differently) to framing treatments. Applying a variety of methods can also help to reduce omitted interaction bias and advance our theoretical understanding of when, how, and why different frames effectively change public opinion. For example, field-experiments and natural language processing techniques (e.g., automated text analysis of open-ended survey responses [74]) could be combined with qualitative interviews and sensory approaches to reassess the role of emotion in climate and environmental communication and the effects of message tailoring across real political campaigns [75].

In doing so, researchers should follow established experimental best-practice standards. The most important of these are preregistration of study designs, publication of replication materials, and advanced post-design solutions to prevent over-reporting of weak effects (for further details, see Section II in S1 Text). Overall, future climate and environmental communication research can benefit from critically reflecting on the limits of framing, especially for certain subgroups, to provide more useful recommendations to climate change communicators and policymakers.

Methods

A systematic mapping of framing studies

In line with the "Preferred Reporting Items for Systematic reviews and Meta-Analyses" (PRISMA) [46] and the best-practice guidelines for systematic mapping outlined by James et al. [40] and the Collaboration for Environmental Evidence [47], we developed a review protocol (see Section III in S1 Text) and systematically reviewed framing studies in the field of environmental politics, economics and psychology according to the following three steps. We also provide a critical appraisal and reanalysis of a subset of published framing experiments that focus on heterogeneous framing effects across partisan subgroups (see below).

First, we conducted a scoping analysis of environment-related framing experiments published in a peer-reviewed scientific journal in Google Scholar, Web of Science, and personal

databases (from Dr. Fesenfeld) using the following search string: (("emphasis fram*" OR "issue fram*" OR "policy fram*" OR "refram*" OR "fram* experiment" OR "information treatment" OR "communication" OR "message" OR "priming" OR "persuasive information" OR "argument") AND (("survey" AND "experiment") OR ("field" AND "experiment") OR ("lab*" AND "experiment")) AND ("climate change" OR "environment"))

In addition, we used a forward and backward snowball technique to identify relevant framing experiments using citations and the reference lists of the reviewed articles. We limited the scope to studies that were published before or in 6/2020. We only identified relevant studies published between 2007 and 6/2020.

Second, during our scoping analysis, and in line with the PRISMA standard [46], we identified 121 peer-reviewed articles in 46 social science journals that we classified as framing experimental studies in the field of environmental politics, economics and psychology (see PRISMA scheme in Fig 3 below). The PRISMA standard aims to report systematic mappings transparently and comprises an evidence-based minimum set of reporting items. We defined the so-called Population, Intervention, Outcome and Study Design (PIOS) criteria for the inclusion and exclusion of articles in our systematic mapping (see Table 1 below) [3]. We only included studies that randomly varied (emphasis) framing treatments and assessed their effects on an individual's climate change and environmental beliefs, attitudes, or behaviors. Our main systematic review analysis included studies that varied the information's connotation but excluded so-called equivalence framing experiments. We however included studies that used both emphasis framing and equivalence framing experiments but then only reviewed the results from the emphasis framing experiment of these studies.

In contrast to emphasis framing, equivalence framing uses different, but logically equivalent phrases to label and describe an issue. An example for an equivalence frame would be to state that a person has a 10% risk of dying or a 90% chance of surviving due to climate-induced extreme weather. The rationale for focusing our review on emphasis frames is that equivalence frames are a less prominent strategy in climate and environmental communication (research) and policymaking [34, 48, 76–78]. Policymakers typically vary the emphasis on a specific subset of relevant arguments in a policy debate, rather than using logically equivalent phrases to alter public opinion. As mentioned, equivalence framing could be a powerful communication technique to alter people's climate change and environmental attitudes and policy preferences [2, 48]. Yet, in the context of real-world climate and environmental communication and respective research is appears to be a less prominent framing strategy. Our search results support this assumption. For example, in the Web of Science and Google Scholar we only identified four experimental studies related to equivalence framing in climate and environmental communication [50, 79–81] (using the following search string: (("equivalence fram*") AND ("climate change" OR "environment") AND ("experiment")). Overall, there seems to exist a clear lack of experimental research on equivalence framing in climate and environmental communication and we thus focus our review here on experimental emphasis framing studies. Moreover, studies that did not use a survey-, lab- or field-experimental design or did not focus on the environmental domain were not included in the review. For example, some studies use discourse narrative analysis or observational survey analysis to detect framing effects. These studies were not included because they do not use an experimental design. Finally, while our search string was successful at identifying most relevant framing experimental studies in the field of climate and environmental communication that fulfilled our Population, Intervention, Outcome and Study Design (PIOS) criteria, with our search string we also identified a large number of irrelevant studies from other fields of research, such as engineering, computer science, or education research, which we excluded.
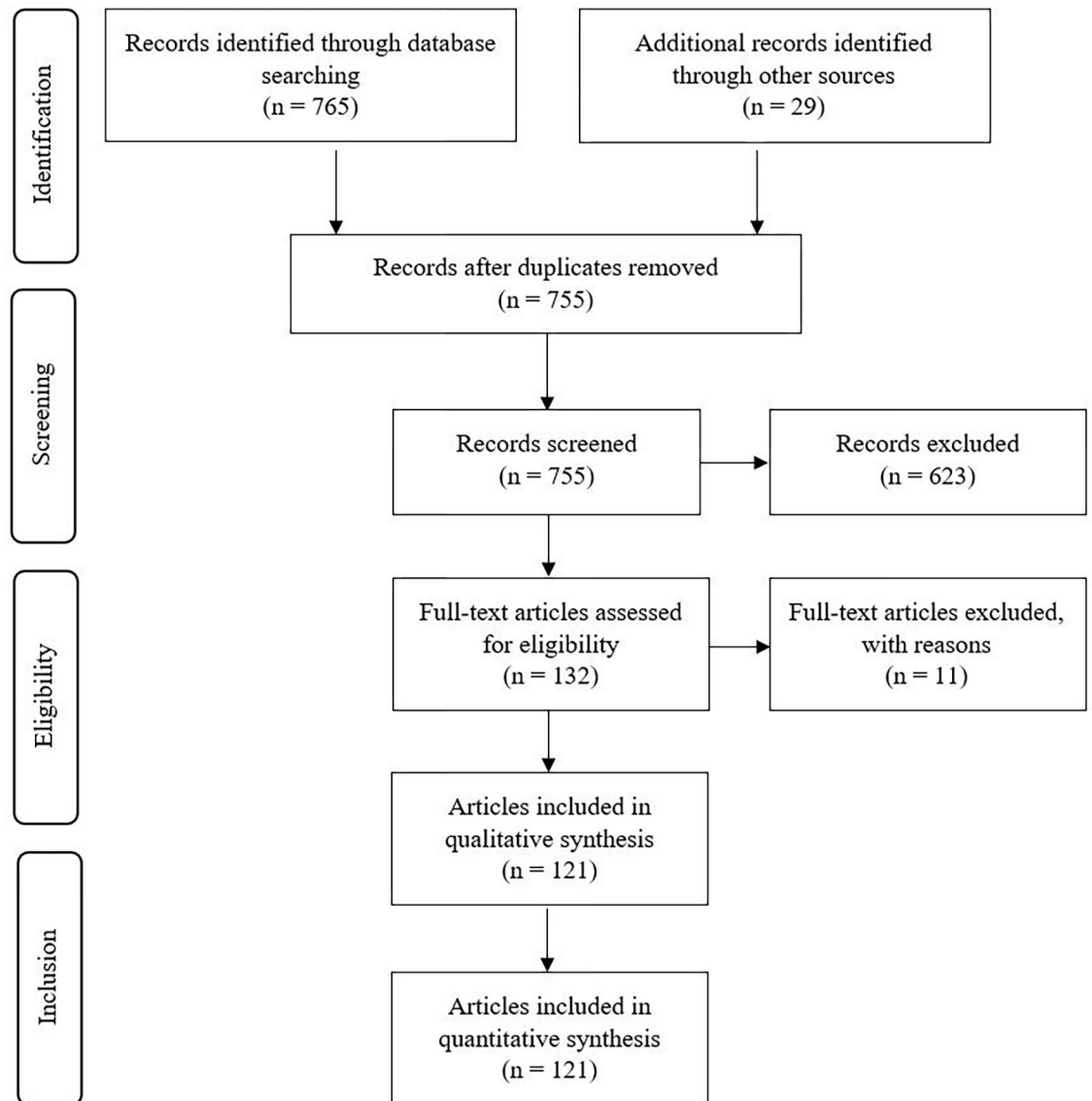
Fig 3. Prisma flow diagram. Third, we systematically analyzed those 121 articles by coding each of the articles according to the following criteria: a) Significant main treatment effect: To what extent did the experiment report any type of significant main treatment effect? If the study included multiple treatment groups, and at least one of these had a significant effect on the outcome variable, we coded the study as reporting significant main effects. If no main effect was reported, we marked this category as not applicable. b) Significant heterogeneous treatment effect: To what extent did the experiment report any type of significant heterogeneous treatment effect? If the study included multiple treatment groups, and at least one had a significant heterogeneous effect for population subgroups, we coded the study as reporting significant heterogeneous effects. If no heterogeneous effect was reported, we marked this category as not applicable. c) Comparative research design: To what extent did the experiment use a comparative research design? If the study focused on more than one country case, we coded the study as a comparative research design. d) Case: In which countries were the experiment(s) conducted?. e) Panel research design: To what extent did the experiment use a panel research design? We coded the study as panel research design if the study was conducted at multiple points in time (at least two data collection waves). f) Experimental design setting: What type of experimental design did the experiment use? We coded whether the study used a field-, survey- or lab-experimental design or a combination of those experimental design types. g) Competing frames: To what extent did the study use different, competing frames? We coded studies as using competing frames if they used one-sided messages and employed frames that emphasize competing arguments and subsets of information. h) Method used: To what extent did the study use an advanced statistical method to check for the results' robustness? We coded studies using an advanced computational method if they employed LASSOplus, LASSO, Ridge Regression, or Kernel regularized least squares. i) Sample type: To what extent did the study use a convenience/population non-representative sample not or a non-convenience/ population representative sample? We coded studies as convenience/non-representative sample if they study did not use a probability-based, stratified or controlled quota sampling methods to aim at representing the target population. Most of the time convenience samples in our review had a sample size of below 500 and were based on student samples. Population representative/

non-convenience samples were mostly larger (n >1000). j) Published data: To what extent did the study make the data publicly available? We only coded studies as publicly available material if the authors had deposited the data in a public repository, such as Harvard Dataverse.

https://doi.org/10.1371/journal.pclm.0000297.g003

We trained three research assistants as coders. In addition, three of the authors also coded articles and double-checked the coding results. In the case of coding-related uncertainty, we asked coders to make comments. The authors then independently looked at these comments and came to an individual decision. Subsequently, the authors discussed these pending cases to make a final decision. Overall, we estimated an inter-coder reliability of 0.7933, implying a substantial agreement between coders.

We also qualitatively analyzed the sample articles and inductively created six framing-type groups, as presented in Fig 1 of the paper. Namely, these are "Issue/Solution Frames", "Value/Norm/Attribution Frames", "Re-Labelling Frames", "Psychological Distance Frames", "Consensus/Uncertainty Frames", and "Source Cue Frames". The definition of each category and relevant examples are listed in Table A in S1 Text. To clarify, the objective of making this typology was to identify the central focus of the treatment conditions in each framing experiment. For those studies that contained two types of manipulations, we coded 0.5 for each category.

Critical appraisal and re-analysis of framing studies

As described in the main body, most framing experiments study heterogeneous treatment effects across population subgroups. We thus conducted a critical appraisal to test the robustness and substantial relevance of reported heterogeneous framing effects by re-analyzing ten typical and widely cited studies that reported heterogeneous framing effects across partisanship subgroups and made replication material publicly available. These re-analyses are not representative of all published framing experiments. Unfortunately, however, the quantity of publicly available data material is very limited, so we could not fully assess existing results' robustness. In this sense, the re-analysis's primary goal was to investigate empirically and potentially verify our suspicion of potential bias against the reporting of non-significant effects. We also intended that the re-analysis process could familiarize applied researchers using advanced statistical methods that they may use in future communication research to check the robustness and relevance of effects.

In line with the original studies, we first used classical ordinary least squares (OLS) regressions in our re-analysis to replicate the original study results. To check for the robustness and substantive relevance of framing effects, we went beyond using these standard linear regression methods and employed a recently developed Bayesian method for variable selection in high-dimensional settings, LASSOplus [38]. Here, our premise is that robust framing effects should

Table 1. Population, intervention, outcome and study design (PIOS) criteria for the inclusion and exclusion of articles.

PIOS element	Description	Inclusion criteria
Population	The body of climate and environmental framing studies	Is it a defined academic article in a peer-reviewed journal on the issue of climate and/or environmental communication?
Intervention	An experimentally-varied (emphasis) framing intervention	Is the study using one or more experimentally varied (emphasis) frames based on textual and/or visual messages
Outcome	Individuals climate and environmental beliefs, attitudes, or behaviors	Is the study about how individuals climate and environmental beliefs, attitudes, or behaviors.
Study design	An experimental study design	Is the study using a field-, lab- or survey-based experimental research design?

https://doi.org/10.1371/journal.pclm.0000297.t001

be detectable through different statistical methods. In essence, robust framing effects should be detectable using classical linear regressions and more advanced computational sparse regression techniques that regularize weak and noisy effects [37, 38, 82]. LASSOplus belongs to a family of advanced computational sparse regression methods developed to test the robustness and substantial relevance of (heterogeneous) experimental effects [37, 38, 82]. Such methods constitute variable selection, through the use of regularization (i.e. setting parameter estimates to zero that do not significantly contribute to predicting the outcome of interest). It thereby reduces the risk of over-fitting the estimation model. This is necessary given the large potential number of sub-group effects that could be explored when assessing the heterogeneity of treatment effects. In other words, LASSOplus penalizes weak and noisy effects to increase efficiency and thereby lessens the risk of false positives. We use LASSOplus as it is designed explicitly for the estimation of heterogeneous treatment effects. LASSOplus allows for the estimation and selection of multiple effects simultaneously, without engaging in potentially arbitrary sub-setting of data. This approach allows the researcher to include many interaction effects to avoid potential omitted interaction bias while simultaneously preventing overfitting the model [38, 44, 45]. Such approaches are supported by many Monte Carlo experiments demonstrating their performance compared to naive regression approaches [37, 38, 82].

In sum, compared to classical linear regressions, this method provides more conservative and robust estimates with credible intervals. It also permits the estimation of interaction effects that can be interpreted independently of their lower-order terms (for further details about LASSOplus–e.g., its Bayesian prior structure and regularization parameters–refer to the original methodological paper [38]). It is important to note that these more advanced computational methods are no 'manna from heaven' to draw causal inference. However, they can complement existing statistical methods (e.g., OLS regressions) to identify the most robust framing effects by preventing omitted interaction bias.

In our implementation of LASSOplus for each model we collected 1000 samples of the posterior distribution, using a burn-in period of 1000 samples and thinning with 10 samples. The posterior chains were inspected by the researchers to assess and ensure convergence.

In the following, we summarize the ten studies we re-analyzed in Fig 2. The re-analyzed studies include analyzed data by Stokes and Warshaw (id17); Christenson, Goldfarb and Kriner (id35); Singh and Swanson (id41); id57: Schuldt, Enns and Cavaliere (id57); Bolsen, Leeper and Shapiro (id71); Saunders (id73); Hardisty, Johnson and Weber (id74); Bolsen and Druckmann (id83); Schuldt, Konrath and Schwarz (id86); and DeGolia, Hiroyasu and Anderson (id115). Please refer to the original studies for further details about the theoretical expectations and experimental design and the supplementary information for full regression tables (see Tables C-L in S1 Text).

Stokes and Warshaw (id17) use a survey experiment to study effects on public support for different renewable portfolio standards bills by varying information about the bill's residential electricity costs, jobs and pollution effects, as well as climate change framing and source cues. Their results indicate that all of these factors are important drivers of public support. Focusing on partisan differences in source cue framing effects, the authors indicate that Democratic (Republican) respondents are more likely to support Democratic (Republican) state legislators' bills. Using LASSOplus, we, however, cannot find any significant subgroup effects among partisan differences.

Christenson, Goldfarb and Kriners' (id35) study use a US nationally representative sample to conduct a survey experiment testing how information about economic and environmental costs and benefits affects fracking support. Their results provide limited evidence of motivated partisan reasoning as framing effects are most considerable for respondents with conflicting partisanship and climate change beliefs. Using LASSOplus, we can confirm that the effects only show limited evidence of motivated partisan reasoning as we could not fully reproduce these significant effects.

Singh and Swanson (id41) study the effect of the different issue- and source-cue frames on US citizens' perceived importance of climate change policy. While the original study shows that the framing conditions did not affect climate policy's perceived importance, it reports several significant subgroup effects for different ideological groups of the sample. Using LASSO-plus, we can confirm the main treatment effects' null findings but cannot reproduce any of the original heterogonous treatment effects.

The study by Schuldt, Enns and Cavaliere (id57) uses data from a probability-based survey experiment conducted among 1461 US adults in 2016 to test the prediction that Demoracts and Republicans respond differently whether global warming vs climate change exists. Their results show that, in the United States, Republicans are more concerned about the term "climate change" than the term "global warming". However, our re-analysis using LASSOplus did not confirm these significant heterogeneous treatment effects reported by the original study.

The study by Bolsen, Leeper and Shapiro (id71) uses a framing experiment to test whether messages highlighting social norms or mentioning science in communication affect respondents' willingness to take action against and beliefs about global warming. Their results show that the norm- and science-based interventions strongly affect attitudes about global warming, support for policies that would reduce carbon emissions, and behavioral intentions to take voluntary action. These effects partly differ depending on party preference. However, running LASSOplus regressions does not confirm these marginally significant differences in effects between ideological groups for the confidence outcome reported by the original study.

Saunders (id73) studies Anthropogenic Global Warming conspiracy beliefs by testing both phrasing (global warming vs climate change) and motivated partisan reasoning. Results indicate that, in line with the theoretical expectations, for the case of climate change, trust moderates hoax beliefs among Republicans. However, this is not the case for global warming, where trust does not moderate conspiracy endorsement among Republicans. Re-analyzing data using LASSOplus does not confirm any of the significant main or heterogeneous treatment effects reported by the original study.

Hardisty, Johnson and Weber (id74) conducted a framing experiment among 889 Americans to assess how labelling charges for environmental costs as either an earmarked tax or an offset included as a surcharge for emitted carbon dioxide affects consumer choices. Their results indicate that cost framing changed preferences for both respondents self-identifying as Republicans and Independents. Democrats' preferences were not significantly affected by these frames. Conversely, the LASSOplus results did not confirm any significant main or heterogeneous treatment effects reported by the original study.

The study by Bolson and Druckmann (id83) investigates the role of partisan group identity and the politicization of science in weakening the impact of a scientific-consensus-based message about human-induced climate change in the United States. Based on OLS regressions, the original study found that partisan identity and politicized messages can alter the effects of messages about the scientific consensus regarding the negative impacts of climate change. However, the LASSOplus results did not confirm any significant main or heterogeneous treatment effects reported by the original study, which employed standard linear regression techniques.

The study by Schuldt, Konrath and Schwarz (id86) uses a survey experiment with 2267 US respondents to test how different wording of global climate change (global warming vs climate change) affects whether individuals perceive the phenomenon to be real or not. Their results indicate that, as expected, Republicans were more likely to endorse that the phenomenon is real when it was referred to as climate change rather than global warming. In contrast, Democrats were not affected by the specific question-wording. This study deems an exception amongst all the re-analyzed studies as it is the only study where reframing has a significant effect amongst Republicans when using LASSOplus.

DeGolia, Hiroyasu and Anderson (id115) use a survey experiment with a two (economic, ecological) by two (gain, loss) factorial design to evaluate how different types of benefit and loss attribute frames in environmental communication affect public support. Among other sub-group analyses, their results indicate that ecological and economic frames differed based on individuals' political ideologies and environmentalism, as conservatives were most responsive to economic messaging. At the same time, liberals were most responsive to ecological framing. For this study, we also did not find any significant subgroup effects when using LASSOplus.

Author survey

To assess how the authors of published framing experiments experienced the publishing process and dealt with non-significant framing effects they encountered, we implemented an online survey using the survey platform Qualtrics. We contacted all 173 authors of the 121 publications via email and received a total of 63 responses (a response rate of 36% of all authors and around 52% of all reviewed studies; please note that most of the time the lead authors of each study responded to our survey).

In addition to a number of closed-ended questions about the publishing experience of experimental framing studies in the field of climate and environmental communication (see Section VI in S1 Text), we also included an open-ended question in the survey to understand under which circumstances authors were able to publish non-significant results. In total, we have received 22 open-ended answers and several of these answers explicitly highlighted the importance of bundling non-significant and significant results together to increase the publication chance. For example, one authors stated "It helped that although our positive frames had no effect, our negative frames did". Other authors stated the following: "Non-significant effects are publishable if the experiment is not the only object of interest in the study. Put another way, of course it's hard to publish null effects if that is the only hypothesis". Also, one author mentions: "As we had included a couple of other interesting analysis and additional minor experiments into the study design, it was easy to report the null effect together with other very interesting findings from the data.". Also the following author states: "There were multiple analyses in the paper—some null, some significant.". Another author stated: "While writing up and reporting both results, I think it helped in the publishing process that there was a significant effect found in addition to the null effect.". Another author mentioned: "We had other effects that were significant, though." and yet one other author highlighted: "It helped that although our positive frames had no effect, our negative frames did."

We interpreted these qualitative statements as clear evidence that bundling non-significant and significant results together could yield higher publication chances. Moreover, some authors also pointed out that publishing non-significant results would be easier if these results are counter-intuitive. For example, one author stated "So that's how we "sold" the non-significant effect: as going against the conventional wisdom."

Supporting information

S1 Text. Fig A-H, Table A-L.
(DOCX)

Author Contributions

Conceptualization: Lukas Fesenfeld, Liam Beiser-McGrath, Yixian Sun, Michael Wicki, Thomas Bernauer.

Data curation: Lukas Fesenfeld, Liam Beiser-McGrath, Michael Wicki.

References

1. Hung LS, Bayrak MM. Comparing the effects of climate change labelling on reactions of the Taiwanese public. Nat Commun. 2020; 11: 1–6. https://doi.org/10.1038/s41467-020-19979-0 PMID: 33247144

2. Druckman J, McGrath M. The evidence for motivated reasoning in climate change preference formation. Nat Clim Chang. 2019; 9: 111–119. https://doi.org/10.1038/s41558-018-0360-1

3. Badullovich N, Grant WJ, Colvin RM. Framing climate change for effective communication: A systematic map. Environ Re. 2020; 15: 123002.

4. Nisbet MC. Communicating Climate Change: Why Frames Matter for Public Engagement. Environ Sci Policy Sustain Dev. 2009; 51: 12–23.

5. Bernauer T, McGrath LF. Simple reframing unlikely to boost public support for climate policy. Nat Clim Chang. 2016; 6: 680–683.

6. Schäfer MS, O'Neill S. Frame analysis in climate change communication. Oxford research encyclopedia of climate science. 2017.

7. Chong D, Druckman J. Framing Theory. Annu Rev Polit Sci. 2007; 10: 103–126. https://doi.org/10.1146/annurev.polisci.10.072805.103054

8. Scheufele D, Iyengar S. The State of Framing Research: A Call for New Directions. The Oxford Handbook of Political Communication. New York: Oxford University Press; 2014. pp. 1–26. https://doi.org/10.1093/oxfordhb/9780199793471.013.47

9. Aklin M, Urpelainen J. Debating clean energy: Frames, counter frames, and audiences. Glob Environ Chang. 2013; 23: 1225–1232. https://doi.org/10.1016/j.gloenvcha.2013.03.007

10. Bain PG, Hornsey MJ, Bongiorno R, Jeffries C. Promoting pro-environmental action in climate change deniers. Nat Clim Chang. 2012; 2: 603–603. Available: http://dx.doi.org/10.1038/nclimate1532

11. Bolderdijk J, Gorsira M, Keizer K, Steg L. Values Determine the (In) Effectiveness of Informational Interventions in Promoting Pro-Environmental Behavior. PLoS One. 2013; 8: 1–7. https://doi.org/10.1371/journal.pone.0083911 PMID: 24367619

12. Graham T, Abrahamse W. Communicating the climate impacts of meat consumption: The effect of values and message framing. Glob Environ Chang. 2017; 44: 98–108. https://doi.org/10.1016/j.gloenvcha.2017.03.004

13. Borgstede C Von, Andersson M, Hansla A. Value-Congruent Information Processing: The Role of Issue Involvement and Argument Strength Value-Congruent Information Processing: The Role of Issue Involvement and Argument Strength. Basic Appl Soc Psych. 2014; 36: 461–477. https://doi.org/10.1080/01973533.2014.958226

14. Nilsson A, Hansla A, Heiling JM, Bergstad CJ, Martinsson J. Public acceptability towards environmental policy measures: Value-matching appeals. Environ Sci Policy. 2016; 61: 176–184. https://doi.org/10.1016/j.envsci.2016.04.013

15. Schultz PW, Zelezny L. Reframing Environmental Messages to be Congruent with American Values. Hum Ecol Rev. 2003; 10: 126–136.

16. Boomsma C, Steg L. The effect of information and values on acceptability of reduced street lighting. J Environ Psychol. 2014; 39: 22–31.

17. Nelson T, Oxley ZM, Clawson RA. Toward a psychology of framing effects. Polit Behav. 1997; 19: 221–246.

18. Jones BD. Bounded Rationality. Annu Rev Polit Sci. 1999; 2: 297–321. https://doi.org/10.1016/B978-0-08-097086-8.93012–5

19. Zaller J, Feldman S. A simple theory of the survey response: Answering questions versus revealing preferences. Am J Pol Sci. 1992; 36: 579–616.

20. Eagly AH, Chaiken S. The psychology of attitudes. Harcourt Brace Jovanovich College Publishers; 1993.

21. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. Communication and persuasion. New York: Springer; 1986. pp. 1–24.

22. Druckman J. Political Preference Formation: Competition, Deliberation, and the (Ir)relevance of Framing Effects. Am Polit Sci Rev. 2004; 98: 671–686.

23. Iyengar S. Is anyone responsible? How television frames political issues. Chicago: University of Chicago Press; 1991.

24. Festinger L. A theory of cognitive dissonance. Stanford university press; 1962.

25. Kunda Z. The case for motivated reasoning. Psychol Bull. 1990; 108: 480. https://doi.org/10.1037/0033-2909.108.3.480 PMID: 2270237

26. Wolsko C, Ariceaga H, Seiden J. Red, white, and blue enough to be green: Effects of moral framing on climate change attitudes and conservation behaviors. J Exp Soc Psychol. 2016; 65: 7–19.

27. Hart PS, Nisbet EC. Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. Communic Res. 2012; 39: 701–723.

28. Baumer EPS, Polletta F, Pierski N, Gay GK. A simple intervention to reduce framing effects in perceptions of global climate change. Environ Commun. 2017; 11: 289–310.

29. Bain PG, Milfont TL, Kashima Y, Bilewicz M, Doron G, Gardarsdottir RB, et al. Co-benefits of addressing climate change can motivate action around the world. Nat Clim Chang. 2016; 6: 154–157.

30. Zhou J. Boomerangs versus Javelins: How Polarization Constrains Communication on Climate Change. Env Polit. 2016; 25: 788–811. https://doi.org/10.1080/09644016.2016.1166602

31. Bolderdijk J, Steg L, Geller ES, Lehman PK, Postmes T. Comparing the effectiveness of monetary versus moral motives in environmental campaigning. Nat Clim Chang. 2013; 3: 413–416. https://doi.org/10.1038/nclimate1767

32. Druckman J, Leeper TJ. Learning more from political communication experiments: Pretreatment and its effects. Am J Pol Sci. 2012; 56: 875–896.

33. Slothuus R. When can political parties lead public opinion? Evidence from a natural experiment. Polit Commun. 2010; 27: 158–177. https://doi.org/10.1080/10584601003709381

34. Fesenfeld L. The Political Feasibility of Transformative Climate Policy–Public Opinion about Transforming Food and Transport Systems. ETH Zurich. 2020. https://doi.org/10.3929/ethz-b-000425564

35. Fesenfeld LP, Rinscheid A. Emphasizing Urgency of Climate Change is Insufficient to Increase Policy Support. One Earth. 2021; 4: 411–424. https://doi.org/10.1016/j.oneear.2021.02.010

36. Fesenfeld L, Sun Y, Wicki M, Bernauer T. The role and limits of strategic framing for promoting sustainable consumption and policy. Glob Environ Chang. 2021; 68: 102266. https://doi.org/10.1016/j.gloenvcha.2021.102266

37. Grimmer J, Messing S, Westwood SJ. Estimating heterogeneous treatment effects and the effects of heterogeneous treatments with ensemble methods. Polit Anal. 2017; 25: 413–434.

38. Ratkovic M, Tingley D. Sparse estimation and uncertainty with application to subgroup analysis. Polit Anal. 2017; 25: 1–40.

39. Kahan DM, Carpenter K. Out of the lab and into the field. Nat Clim Chang. 2017; 7: 309. Available: https://doi.org/10.1038/nclimate3283

40. James KL, Randall NP, Haddaway NR. A methodology for systematic mapping in environmental sciences. Environ Evid. 2016; 5: 1–13.

41. Kidd LR, Garrard GE, Bekessy SA, Mills M, Camilleri AR, Fidler F, et al. Messaging matters: A systematic review of the conservation messaging literature. Biol Conserv. 2019; 236: 92–99.

42. Comfort SE, Park YE. On the field of environmental communication: A systematic review of the peer-reviewed literature. Environ Commun. 2018; 12: 862–875.

43. Li N, Su LY-F. Message framing and climate change communication: A meta-analytical review. J Appl Commun. 2018; 102: 1c–1c.

44. Blackwell M, Olson M. Reducing model misspecification and bias in the estimation of interactions. Working Paper available at https://mattblackwell.org/files/papers/lasso . . .; 2020.

45. Beiser-McGrath J, Beiser-McGrath LF. Problems with products? Control strategies for models with interaction and quadratic effects. Polit Sci Res Methods. 2020; 8: 707–730.

46. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009; 151: 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135 PMID: 19622511

47. Collaboration for Environmental Evidence. Guidelines and Standards for Evidence synthesis in Environmental Management. Springer; 2018. Available: https://environmentalevidence.org/information-for-authors/

48. Oxley Z. Framing and political decision making: An overview. Oxford Res Encycl Polit. 2020.

49. Olsen A. Equivalency framing in political decision making. Oxford Research Encyclopedia of Politics. 2020.

50. Spence A, Pidgeon N. Framing and Communicating Climate Change: The Effects of Distance and Outcome Frame Manipulations. Glob Environ. 2010; 20: 656–667.

51. Brügger A, Dessai S, Devine-Wright P, Morton TA, Pidgeon NF. Psychological responses to the proximity of climate change. Nat Clim Chang. 2015; 5: 1031–1037. https://doi.org/10.1038/nclimate2760

52. Singh SP, Swanson M. How issue frames shape beliefs about the importance of climate change policy across ideological and partisan groups. PLoS One. 2017; 12: e0181401. https://doi.org/10.1371/journal.pone.0181401 PMID: 28727842

53. Stokes LC, Warshaw C. Renewable energy policy design and framing influence public support in the United States. Nat Energy. 2017; 2: 17107.

54. Bansak K. Estimating causal moderation effects with randomized treatments and non-randomized moderators. J R Stat Soc Ser A Stat Soc. 2021; 184: 65–86. https://doi.org/10.1111/rssa.12614

55. Beiser-Mcgrath J, Beiser-Mcgrath LF. The Consequences of Model Misspecification for the Estimation of Nonlinear Interaction Effects. Polit Anal. 2023; 31: 278–287. https://doi.org/10.1017/pan.2022.25

56. Imai K, Ratkovic M. Estimating treatment effect heterogeneity in randomized program evaluation. Ann Appl Stat. 2013; 7: 443–470.

57. Hainmueller J, Hazlett C. Kernel regularized least squares: Reducing misspecification bias with a flexible and interpretable machine learning approach. Polit Anal. 2014; 22: 143–168.

58. Imai K, Strauss A. Estimation of heterogeneous treatment effects from randomized experiments, with application to the optimal planning of the get-out-the-vote campaign. Polit Anal. 2011; 19: 1–19.

59. Goplerud M, Imai K, Pashley NE. Estimating heterogeneous causal effects of high-dimensional treatments: Application to conjoint analysis. arXiv Prepr arXiv220101357. 2022.

60. Goplerud M. Modelling Heterogeneity Using Bayesian Structured Sparsity. arXiv Prepr arXiv210315919. 2021.

61. Schuldt JP, Konrath SH, Schwarz N. "Global warming" or "climate change"? Whether the planet is warming depends on question wording. Public Opin Q. 2011; 75: 115–124.

62. Schuldt JP, Enns PK, Cavaliere V. Does the label really matter? Evidence that the US public continues to doubt "global warming" more than "climate change." Clim Change. 2017; 143: 271–280.

63. Soutter ARB, Mottus R. "Global warming" versus "climate change": A replication on the association between political self-identification, question wording, and environmental beliefs. J Environ Psychol. 2020; 69: 101413.

64. Schuldt JP, Enns PK, Konrath S, Schwarz N. Shifting views on "global warming" and "climate change" in the United States. J Environ Psychol. 2020; 69: 25–27. https://doi.org/10.1016/j.jenvp.2020.101414

65. Westlake WJ. Statistical aspects of comparative bioavailability trials. Biometrics. 1979; 273–280. PMID: 583027

66. Berger RL, Hsu JC. Bioequivalence trials, intersection-union tests and equivalence confidence sets. Stat Sci. 1996; 11: 283–319.

67. Wellek S. Testing statistical hypotheses of equivalence and noninferiority. CRC Press; 2010.

68. Chong D, Druckman J. Dynamic public opinion: Communication effects over time. Am Polit Sci Rev. 2010; 104: 663–680. https://doi.org/10.1017/S0003055410000493

69. Chong D, Druckman J. Framing public opinion in competitive democracies. Am Polit Sci Rev. 2007; 101: 637–655. Available: http://journals.cambridge.org/article_S0003055407070554

70. Druckman J, Nelson K. Framing and deliberation: How citizens' conversations limit elite influence. Am J Polit. 2003 [cited 7 Jan 2017]. Available: http://onlinelibrary.wiley.com/doi/10.1111/1540-5907.00051/full

71. Druckman J. Public opinion: Stunted policy support. Nat Clim Chang. 2013; 3: 617–617. https://doi.org/10.1038/nclimate1939

72. Kinder DR. Curmudgeonly advice. J Commun. 2007; 57: 155–162. https://doi.org/10.1111/j.1460-2466.2006.00335.x

73. Levine AS, Kline R. A new approach for evaluating climate change communication. Clim Change. 2017; 142: 301–309.

74. Roberts ME, Stewart BM, Tingley D, Lucas C, Leder-Luis J, Gadarian SK, et al. Structural topic models for open-ended survey responses. Am J Pol Sci. 2014; 58: 1064–1082.

75. Chapman DA, Lickel B, Markowitz EM. Reassessing emotion in climate change communication. Nat Clim Chang. 2017; 7: 850–852. https://doi.org/10.1038/s41558-017-0021-9

76. Druckman J. The Implications of Framing Effects Implications for Citizen Competence. Polit Behav. 2001; 23: 225–256. https://doi.org/10.1023/A:1015006907312

77. Slothuus R. More Than Weighting Cognitive Importance: A Dual-Process Model of Issue Framing Effects. 2008; 29: 1–28.

78. Sniderman PM, Theriault SM. The structure of political argument and the logic of issue framing. Studies in public opinion: Attitudes, nonattitudes, measurement error, and change. Princeton University Press Princeton; 2018. pp. 133–165.

79. Morton TA, Rabinovich A, Marshall D, Bretschneider P. The future that may (or may not) come: How framing changes responses to uncertainty in climate change communications. Glob Environ Chang. 2011; 21: 103–109.

80. Nabi RL, Gustafson A, Jensen R. Framing climate change: Exploring the role of emotion in generating advocacy behavior. Sci Commun. 2018; 40: 442–468.

81. Schifman B. Equivalence and Issue Framing Effects in the News Media and their Effect on Preferences Regarding Climate Change. Citeseer; 2007.

82. Green DP, Kern HL. Modeling heterogeneous treatment effects in survey experiments with Bayesian additive regression trees. Public Opin Q. 2012; 76: 491–511.