# Scaling Dialogue for Democracy: Can Automated Deliberation Create More Deliberative Voters?

James Fishkin, Valentin Bolotnyy, Joshua Lerner, Alice Siu and Norman Bradburn

The theory and practice of what has come to be called "deliberative democracy" have been revived for the modern era with a focus on deliberative microcosms selected through random sampling or "sortition." But might it be possible to spread some of the benefits of deliberation beyond mini-publics to the broader society? Can technology assist with scaling an organized deliberative process? In particular, would those who experience such a process become more deliberative voters? Would their considered judgments from deliberation influence their voting? We draw on a larger than usual experiment with public deliberation and a one-year follow-up in the mid-term U.S. elections to suggest answers to these questions. It has implications for whether spreading an organized deliberative process could, in theory, be used to create more deliberative elections.

Corresponding author: James Fishkin 🆔 is the Janet M. Peck Chair in International Communication, Professor of Communication, and Professor of Political Science (by courtesy) and Senior Fellow of the Freeman Spogli Institute for International Studies at Stanford University where he directs the Deliberative Democracy Lab (jfishkin@stanford.edu). He works primarily on deliberative democracy in both theory and practice.

Valentin Bolotnyy is a Kleinheinz Fellow at Stanford University's Hoover Institution (vbolotnyy@stanford.edu). He is also a Research Affiliate at the Institute of Labor Economics (IZA) and an Affiliated Scholar with the Deliberative Democracy Lab at Stanford's Center on Democracy, Development and the Rule of Law (CDDRL). His research covers public, labor, and political economics, and is often done in partnership with governments.

Joshua Y. Lerner 🆔 is a Research Methodologist at NORC at the University of Chicago in the Department of Methodology and Quantitative Social Sciences (lerner-joshua@norc.org). His research interests include causal inference, natural language processing, machine learning and their implications for American politics.

Alice Siu is Associate Director of the Deliberative Democracy Lab and Senior Research Scholar at the Center for Democracy, Development, and Rule of Law housed at the Freeman Spogli Institute at Stanford University (asiu@stanford.edu). She is a lead collaborator on the Stanford Online Deliberation Platform, a partnership with the Crowdsourced Democracy Team at Stanford. On Deliberative Polling, she has advised policymakers and political leaders around the world, at various levels of government, including in Mongolia, China, and Brazil. Her research interests in deliberative democracy include what happens inside deliberation, such as examining the effects of socio-economic class in deliberation, the quality of deliberation, and the quality of arguments in deliberation

Norman M. Bradburn 🆔 is the Tiffany and Margaret Blake Distinguished Service Professor Emeritus, University of Chicago (Bradburn-Norman@norc.org). He is currently a Distinguished Senior Fellow at NORC at the University of Chicago. Bradburn has been at the forefront in theory and practice in the field of sample survey research. He has focused on psychological well-being and assessing quality of life, particularly through the use of large-scale sample surveys; non-sampling errors in sample surveys; and research on cognitive processes in responses to sample surveys. He is currently the co-Primary Investigator of the Humanities Indicators project at the American Academy of Arts and Sciences.

## The Challenge of Scaling Deliberation

The theory and practice of what has come to be called "deliberative democracy" have been revived for the modern era with a focus on deliberative microcosms selected through random sampling or "sortition." One merit of organized deliberation with a good sample is that it should, in theory, permit inferences about what the entire population would think about an issue if it could deliberate about it under similarly good conditions. However, advocates freely admit that the public will usually not be effectively motivated to consider issues in depth under anything like the "good conditions" of an organized deliberation. What Jane Mansbridge has called "everyday talk" (roughly equivalent to discussion in ordinary life outside of organized deliberative settings) is thought to be deliberative, sometimes, and to some degree, but less so than what could be achieved with an organized design (Mansbridge 1999).

There are many challenges to the deliberative quality of mass discussion in ordinary life. The public is routinely subjected to one-sided advocacy and misinformation, facilitated by social media sharing amongst the like-minded, while, at the same time, other portions of the public often remain inattentive or disengaged from major policy questions. Low information levels and inattention are often explained as an understandable consequence of what Anthony Downs called "rational ignorance," and its companion "rational apathy" (a term commonly applied to shareholder voting in public companies but equally applicable to citizens for many issues). Why should I vote or become involved in thinking about public issues if my individual views are likely at best to have only a miniscule impact on policy (Downs 1957; Hardin 2002)? At the same time, other portions of the public may be intensely motivated, but highly unrepresentative. The voices that are motivated to speak up via self-selected processes are likely to be unrepresentative, and non-deliberative while conventional polling, even when conducted with good samples, will offer results that are some distance from the *overall* public's considered judgments.

The rationale for convening deliberative microcosms such as Deliberative Polling is that they attempt to offer a picture of what the entire public would think under good conditions. The aspiration is to convene a process that gives voice to "the will of the people" (at least on selected issues) through a combination of inclusion and deliberation. In some ways, random sampling offers an admirable form of inclusion compared to most democratic practices. For example, deliberation with a good stratified random sample is likely to be more representative than a low turnout election or referendum from the same population. But even if a deliberating microcosm is highly representative of the views of the entire population (in its pre-deliberation views), it will inevitably be the case that most

people whose "will" is being represented (through the inclusion of random sampling) are not actively involved, and probably not even aware of the process. As Cristina Lafont has pointedly asked, what about those who are left out of the deliberations? If the deliberations produce significant opinion changes then many of them will disagree and not feel represented by the results (Lafont (2020).

One answer to this challenge is the hypothetical claim already mentioned—the results offer a representation of what the public as a whole would think about an issue under stipulated good conditions. And those who are not deliberating are included in the public as a whole from whom the sample is selected. Of course, the claim about "good conditions" must be continually interrogated, both conceptually and empirically. But that is not our focus here (see Fishkin 2018 for a general account and Farrar et al. 2010 and Sandefur et al. 2022 for experiments disaggregating aspects of the Deliberative Polling treatment).

A second response to the problem of those left out might be some robust attempt to bring *the rest* of the public into the dialogue, at least in terms of exposure and awareness. This can be done with media coverage and by advocates of the results invoking the reasons for the post deliberation opinions of the mini-public, or the fact that the results were complied with by decision makers. If the sample is representative, those arguments and the conclusions they support should have some purchase on public opinion of the broader public. Hence broadcasting and print media may amplify the process of deliberation and its conclusions, perhaps with some constructive effect on the viewers (see Rasinski, Bradburn, and Lauen (1999) for an important early study of television viewing of a mini-public deliberation; also see the literature on survey experiments informing people of the results of mini-public deliberations such as Muradova, Culloty, and Suiter 2023, Muradova and Suiter 2022; Germann, Manen, and Murdova 2022.)

Even more effective might be a third response—an attempt to *actually engage* the broader public in organized deliberations. The idea is not merely to expose members of the broader public to deliberations by others in a random sample, but rather, to engage them in a deliberative process organized on the same design. The experience of actually deliberating is likely to be more consequential than mere exposure to some aspect of deliberation by others. Especially on highly polarized issues, mere exposure may have little effect (see the well-established literature on CIE, the "continued influence effect" of misinformation despite exposure to correction, [Ecker et al. 2022, 15]) and it will sometimes backfire (Chong and Druckman 2007; Nyhan and Reifler 2010).[1]

However, *participation* in moderated discussions with diverse others in an appropriately organized design, appears to depolarize opinion and move participants to

their considered judgments even on highly contested and polarized issues (Fishkin et al. 2021). But these conclusions are supported by results from face-to-face deliberations. Because technology offers the promise of more cost-effective and practical scaling of participation in the deliberative process itself (rather than merely the scaling of awareness of deliberations by others) we are interested in whether or not such results can be replicated with new technology. Here we will investigate whether an online, AI Assisted moderator could produce some of the same effects that have been found for those who participate in Deliberative Polls in the face-to-face mode.[2] This is not a merely academic question. If the automated AI-assisted moderator facilitates an effective form of deliberation (as we will inquire later with the hypotheses provided) then scaling participation in the deliberative process becomes a more realistic and attractive possibility.

The idea of organized mass deliberation is proposed in the scheme for "Deliberation Day" (Ackerman and Fishkin 2004). There the idea was to have the whole country deliberate face-to-face in innumerable small groups a few weeks before a national election. A plan for these discussions in every locality of the United States was described and advocated. However, organized face-to-face deliberation is a massive undertaking. It would probably require legislation and an organized institutional support structure.

## The AI-Assisted Online Deliberation Platform

The online version of Deliberative Polling, and indeed, of other deliberative processes, suggests strategies for more cost-effective and realistic implementations for scaling the deliberative process to entire communities, states, or even the nation as a whole. The online deliberations were implemented in this project via the AI-Assisted Stanford Online Deliberation Platform.[3] In theory it can handle any number of small groups (of approximately ten people in each group) in synchronous video-based discussions. The process moderates itself, with a queue for talking time (currently set at 45 seconds per intervention), nudges speakers who are not participating, intervenes for uncivil language, produces nearly instant transcripts of all the discussions to document the arguments for research purposes, orchestrates the movement of the discussions through an agenda of proposals and initial pros and cons of each proposal, etc. and coordinates the group in identifying its two most important questions for the plenary sessions. Evaluations of the platform by participants show ratings as high or higher than projects with human moderators (refer to online appendix table 1 for the evaluations). We call the platform "AI-Assisted" rather than AI facilitated because at crucial points, it draws on the collective intelligence of each small group, rather than having the AI make the decision itself for the group. Has the group sufficiently discussed the first proposal? Has it covered the pros and cons of that proposal? The group is polled on these and similar decisions as it moves through the process. In effect, the platform assists the group in moderating itself.

The prospect of mass automated deliberation is at hand. What are some key issues to engage the deliberative theory and empirical communities about the potentially realistic possibility of mass deliberation via this kind of technology? What do we need to know to evaluate whether mass applications on this model would usefully contribute to the democratic process? Our initial answers will come from a larger than usual application of Deliberative Polling, but the data also offers a picture that we can use for evaluation of some key elements of what might become organized mass deliberation beyond the confines of random samples.[4] We are proposing Deliberative Scaling on the same model of discussion as Deliberative Polling, with the same (automated) moderator and the same kinds of briefing materials used in the national experiment discussed here.

Given the scale and ambition of mass organized deliberation, why should we take the idea seriously? Because there is evidence that deliberation in organized settings can address three notable problems with the operation of mass democracy. Some of this evidence comes, suggestively, from face-to-face deliberations with random samples. Do we have any reason to believe that similar results might follow if technology were used to scale deliberations on the same model of discussion with larger populations?

Here we will investigate the application of a specific technology which closely models the Deliberative Polling design used in face-to-face deliberations. We will draw inferences from its use in the largest Deliberative Poll conducted up to that time, nearly 1,000 participants in 104 small groups; plus, nearly 700 in a pre-post control group that only answered the same questionnaire at the time of recruitment and at the end of the process).[5]

Deliberations by random samples, mostly in face-to-face discussions, seem to produce at least three notable effects, which if scaled to very large numbers in the broader society, would likely help with three major challenges facing democracy. First, competitive democracies often seem crippled by extreme partisan polarization. The divisions between parties at the mass level are not just extreme, they are alleged to be "calcified" or intractable (see Sides, Tausanovitch, and Vavreck 2022), fueling not only division but deadlock (Hetherington and Rudolph 2015). Yet there is evidence that face-to-face deliberation can produce significant partisan depolarization. Can the same be true of deliberations with the AI-assisted online moderator? Second, democracies are supposed to make a connection between the will of the people and what is actually done. If deliberative democracy applications are intended to clarify the "public will" they must identify the

considerations that support the public's considered judgments. Face-to-face deliberations have a record of doing so. Will the automated online process be equally successful? Third, the predominant accounts of voting behavior in our competitive democracies are that the tribalism of party loyalties mostly determines who turns out and who gets elected (Green, Palmquist, and Schickler 2004; Achen and Bartels 2016). The idea that voters evaluate policies on the merits and cast their votes accordingly is commonly treated as a quaint leftover from the now widely abandoned "classical theory" of democracy (Schumpeter 1942; Achen and Bartels 2016; Posner 2003; Shapiro 2003). But if voters do not really consider policy positions in casting their votes, then how are they to exercise popular control? On the other hand, there is evidence that face-to-face deliberation, whether in juries or deliberative mini-publics such as Deliberative Polling creates a latent variable of civic engagement that has a significant effect on voting (Gastil, Deess, and Weiner 2002, Fishkin et al. 2024). Do these effects occur with the automated online process?

We propose three hypotheses corresponding to these three notable effects. The hypotheses are based on results from face-to-face deliberations as well as related literature.

HYPOTHESIS 1: The opinion changes from the automated online deliberative process will show evidence of partisan depolarization.

One of the possible benefits of scaling mass deliberation might be the diminution of extreme partisan polarization. A national face-to-face Deliberative Poll produced some striking results along these lines (Fishkin et al. 2021). Do we see similar evidence for the online process with the automated platform? Following our previous work (Fishkin et al. 2021) we define depolarization in terms of whether or not the mean positions of the two parties move closer together. This can result from one-sided or two-sided movement (Fishkin et al. 2021, 4).[6] Why might this happen? As we argued in Fishkin et al. 2021, "Kunda hypothesized … that there are two kinds of "motivated reasoning"—"directional" and "accuracy based" (Kunda 1990). There is an obvious argument that discussion will increase partisan polarization on issues that are already polarized because of directional motivated reasoning. But if an experimental treatment could encourage the second kind of motivated reasoning—accuracy-based—then reasoning on the merits would likely result. Further, if on issues of partisan polarization, people have arrived at positions without seriously considering (or even encountering) arguments on the other side, then we might well expect accuracy-based motivated reasoning to reduce partisan polarization by overcoming the legacy of previously one-sided reasoning. The expectation here is not that deliberation will always depolarize but that it will likely depolarize on issues where participants begin the exercise

with high levels of partisan polarization. Generally, Kunda concluded that subjects will be motivated to be accurate when they think their views will matter, and when they have to share the reasons supporting them" (Kunda 1990, 481).[7] This is the framework we applied in Fishkin et al. 2021, 1468, and we employ it again here.

HYPOTHESIS 2: The opinion changes from the automated online deliberative process will show evidence of the effect of identifiable reasons.

Deliberative mini-publics are designed to foster the weighing of competing reasons by the participants in an evidence-based environment. But it is an empirical question whether they succeed in fostering opinions and opinion movements that show the effect of identifiable reasons. For evidence of the role of identifiable reasons in face-to-face Deliberative Polls, see Fishkin 2018, especially Part III. Given that the topic in these discussions is climate change, we also rely on Krosnick et al. 2006, who identified questions that affect the importance of climate in U.S. national public opinion.

Lastly, does deliberation with the online automated moderator platform help facilitate deliberative voting in actual elections?

HYPOTHESIS 3. The experience of online deliberation with the automated moderator will produce a latent variable of civic engagement that increases deliberative voting (so that voters connect their post-deliberation policy priorities with how they vote in actual elections). This latent variable can be investigated via causal mediation analysis.

The very idea of a deliberative voter, one who thoughtfully considers the arguments for and against candidates (or parties or policies) and then votes to a considerable degree in accord with those considered judgments has been largely dismissed in democratic theory since Schumpeter relegated it to the "classical theory" of democracy, to be superseded by the modern "competitive theory" (which makes no mention of the public will, just the "competitive struggle for the people's vote"). The most systematic, empirically based reprisal of Schumpeter's account, the "realist theory" of Achen and Bartels (2016) dismisses anything like the deliberative voter as a "folk theory" never to be realized. Yet despite Schumpeter and his followers, the aspiration has a continuing hold on normative theories of democracy even beyond the literature on deliberative democracy. Robert Dahl, for example, included the requirement of "enlightened understanding" on "the issues to be decided", as one of the essential criteria for a system to be considered democratic. However, he offered little basis for concluding it might be practical (apart from his endorsement of what we would now call a

deliberative minipublic, the "mini-populous" in his concluding reflections on an advanced democratic society (Dahl 1989, 126-128 and 340-341).

Can deliberation in an organized setting bring the deliberative voter to life? In principle, might we use technology to produce this effect? If so, this might make scaling to large numbers a more practical aspiration, particularly if the effect had staying power.

Gastil et al. (2002) found indirect effects on voting from serving on a jury that reached a verdict. They found that depth of deliberation (which they measured by the number of counts considered by a jury) was one of the mediators for the likelihood of voting (see also Gastil et al. 2010, where they extend the analysis of jury participation's effect on turnout with additional data). Fishkin et al. (2024) found that indirect effects from face-to-face deliberation affected voting a year later in the presidential election. In that case the dependent variable was not just turnout but voting intention about whom to vote for. Once again the analysis focused on a latent variable of civic engagement analyzed via causal mediation analysis.

## America in One Room: Climate and Energy

What would the American public really think about our climate and energy challenges if it had the chance to deliberate about them in depth, with good and balanced information? If the American people—or in this case, a representative sample of them—could consider the pros and cons of our different energy options, which would they support? Which would they cut back on? What possible paths to Net Zero (the point at which we are not adding to the total of greenhouse gases in the atmosphere) would seem plausible to them? Which proposals would they resist? Can the public's conclusions chart a way forward on our climate and energy dilemmas—a way forward that transcends our great divisions, especially our deep partisan differences? The aim of the project was not to persuade the participants to move in any particular direction. There were balanced briefing materials with suggested pros and cons for each proposal, balanced plenary panels, and efforts to equalize the participation of all the participants.

NORC at the University of Chicago selected a nationally representative sample of 962 respondents to deliberate on 72 substantive questions, and a second representative sample (a pre- and post-control group of 661) that did not deliberate but took essentially the same questionnaire in the same time period. Each group completed the survey upon being recruited (in early August) and again at the end of the experiment (in late September 2021) as well as a year later (close to the mid-term Congressional elections). The treatment and control groups were two independent samples both drawn from the same probability-based panel, selected so that there is no overlap between the two samples and weighted so that they are comparable. See the online appendix section on "Sampling and Weighting" for a detailed description as well as the Balance Table (A2) in the online appendix.

On 66 of the 72 issue propositions in the survey, the participants changed significantly over the course of the deliberation. Almost all the changes were toward doing more to combat climate change, and these changes were generally in the same direction across party and demographic divides. Democrats were initially more supportive of ambitious policies to address climate change, and Republicans were initially more skeptical (with Independents falling in between the two). However, by the end of the deliberations, majorities of Republicans had come to support the general principle of "serious action to reduce greenhouse gases," along with a number of specific proposals to "dramatically accelerate" adoption of renewable sources of energy and to slow deforestation. At the same time, Democrats became more supportive of including a new generation of nuclear power plants in the future energy mix, and a majority of Republicans remained wary of a hard deadline of 2050 for phasing out oil and natural gas.

Following our work in Fishkin et al. (2021) and Fishkin et al. (2024) we focus our attention on the most extremely polarized issues and then use those items to construct a policy-based score (PBS).[8] We focus on these particular issues out of the 72 total policy questions posed to study participants, because they enable us to clearly classify an individual on a left-right policy dimension. Only for the polarized issues are Democrat respondents concentrated on the left and Republican respondents concentrated on the right, allowing for a unidimensional scale. These are also the issues where beliefs are strongest and most passionate, making any sort of change especially difficult and unlikely. We thus view any changes we measure to an individual's PBS as a lower-bound on the capacity of deliberation to change minds and attitudes on policy issues. Following the methodology in Fishkin et al. (2024), which was an experiment involving in-person deliberation, has the additional benefit of allowing us to test whether deliberation on the automated platform can generate similar depolarization as measured for the most extremely polarized issues from the face-to-face project.

While Fishkin et al. (2024) from the face-to-face study, used 26 questions across five different issue areas to construct its PBS, our questions here only relate to climate. When we follow the same criteria for question selection, we identify only 9 questions that can be classified as extreme partisan polarization. These criteria require 1) 50% or more of Democrats and Republicans to be at opposite sides of an issue and for 2) 15% or more of Democrats and Republicans to be at opposite extremes of an issue.[9] While our results are robust to using a PBS constructed in this way out of nine questions, they are

noisier than if we use a measure of extreme polarization that is constructed out of a larger set of questions. Thus, the results we present throughout the main body of this paper use a twenty-seven question PBS, with these questions identified only using criteria 2) provided earlier (Alpha = 0.98).[10]

The PBS ranges from 0 to 10, with 0 denoting most favorable to climate action and 10 denoting least favorable. The T1 score (using the pre-period survey) and T2 score (using the post-period survey) for each individual are an average over their responses to each of the twenty-seven questions. Before averaging, we make sure that the response to each question is converted to the PBS rubric (e.g., if the extreme Republican response to the question was 0, we flip *everyone's* scores, turning *0*s into 10, *1*s into *9*s, etc. before averaging). We do this because of the way some questions were worded.

## Results

Changes in the overall PBS between T1 and T2 are shown in a binned scatterplot in figure 1.[11] That figure and those that follow are intended to be descriptive in assisting the reader to visualize the data. The statistical significance of the changes in the PBS between treatment and control groups are presented in table 1 below. While the control group sees no change in its PBS, except for what is likely mean reversion at the extremes, Climate Deliberation participants see significant changes. Those who go into the deliberation just left of center come out of it with their

positions on polarizing questions moving further to the left. The movement to the left—in our case movement towards acting to address climate change—gets more pronounced for those who start off with more conservative attitudes towards climate change issues. For a difference-in-difference analysis comparing the movement in treatment and control groups on all the substantive questions, refer to online appendix table A5. Of the twenty-seven extremely polarized questions, twenty-two changed significantly in comparison to the control group.

Figure 1 visualizes the contrast between the treatment and control group. It is useful for visualizing the depolarization in HYPOTHESIS 1. With a majority of Republicans starting on the right side (6 to 10 on the PBS scale) one can see the magnitude of movement to the left (negative numbers for the vertical axis) among the deliberators but not for the control group. See table 1 for regressions showing the significant differences between treatment and control for figures 1, 4, and 5.

More specifically, when the movements from Republicans and Democrats are pictured separately for both treatment and control groups, we can clearly see substantial partisan depolarization.

For the control group we see little overall movement in figure 2 (although we see some Republicans actually moving upwards and hence to a more conservative position on climate change if they started left of center). However for the treatment group as pictured in figure 3 we see substantial movement from the Republicans on the right side, with negative movements (hence movements left toward more

**Figure 1**
**Policy-based score (PBS) changes (T2–T1)**
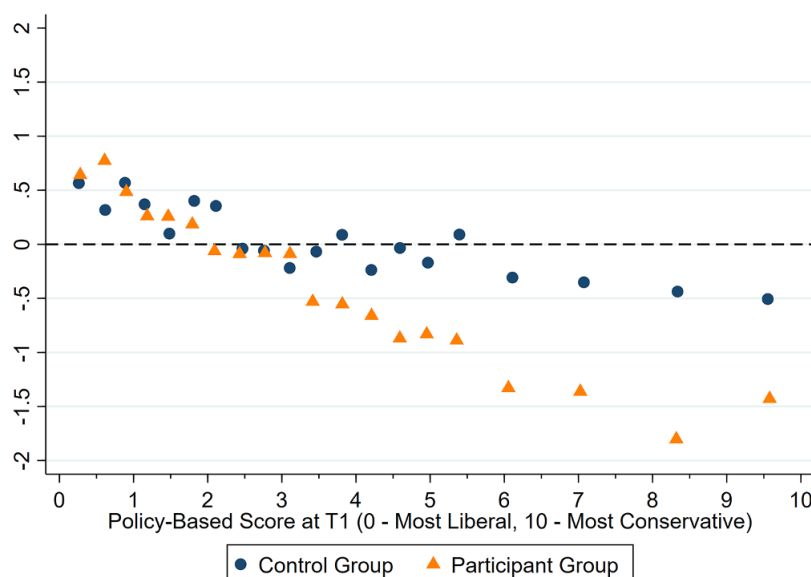


*Note:* Policy-Based Score is constructed for each individual based on responses to 27 questions identified as the most polarizing: 15%+ of each party are at opposite extremes (0 or 10).

**Table 1**
**Differences between participant and control groups between Time 1 and Time 2**

| | (1) | (2) | (3) | (3) | (5) | (6) |
|---|---|---|---|---|---|---|
| | ΔPBS | ΔPBS | ΔWorry | ΔWorry | ΔBelief | ΔBelief |
| Participant=1 | 0.226** | 0.297*** | 0.293* | 0.240 | 0.100 | 0.000 |
| | (0.104) | (0.097) | (0.163) | (0.150) | (0.144) | (0.122) |
| Policy-Based Score (PBS) at T1 | −0.105*** | −0.121*** | 0.042 | 0.037 | 0.041* | 0.012 |
| | (0.018) | (0.020) | (0.028) | (0.030) | (0.024) | (0.028) |
| Participant*Policy-Based Score (PBS) at T1 | −0.170*** | −0.174*** | 0.049 | 0.066* | 0.132*** | 0.153*** |
| | (0.023) | (0.023) | (0.036) | (0.039) | (0.031) | (0.034) |
| Constant | 0.408*** | −0.917 | −0.111 | −0.890 | −0.186 | −0.651* |
| | (0.084) | (0.613) | (0.132) | (0.893) | (0.116) | (0.378) |
| Observations | 1,403 | 1,394 | 1,401 | 1,393 | 1,393 | 1,384 |
| Adj. R^2 | 0.235 | 0.250 | 0.028 | 0.020 | 0.080 | 0.069 |
| R^2 | 0.237 | 0.284 | 0.030 | 0.063 | 0.082 | 0.111 |
| Robust Standard Errors | No | Yes | No | Yes | No | Yes |
| Controls | No | Yes | No | Yes | No | Yes |

*Note:* Regressions correspond to figures 1, 4, and 5. Figures 1, 4, and 5 show raw relationships, without any controls in the background and correspond to columns 1, 3, and 5 in this table. Columns 2, 4, and 6 in this table include controls and report robust standard errors. Controls include dummies for high school or less education, female, over 60 in age, white, married, employed, low income, rural resident, internet connection at home, lean Republican, and state of residence.

**Figure 2**
**Party differences in policy-based score changes for control group**



*Note:* Policy-Based Score is constructed for each individual based on responses to 27 questions identified as the most polarizing: 15%+ of each party are at opposite extremes (0 or 10).

concerted climate action). Expressed another way, on many issues the Republicans started out with minority support for climate action in the 35% range and moved to majority support of 55% or more, supporting the position of Democrats.

We take these depolarizing movements between Republicans and Democrats as a summary answer for HYPOTHESIS 1. There was systematic movement in a de-polarizing direction, summing over all of the initial polarizing issues.

What drives these significant movements, moderating the views of those most skeptical about acting to address climate change? In the spirit of HYPOTHESIS 1, what appears to be motivating these changes? Our first clues arise from a look at how beliefs about climate change and worry about its consequences change as a result of the

**Figure 3**
**Policy-Based Score Changes (T2–T1), Participant Group**
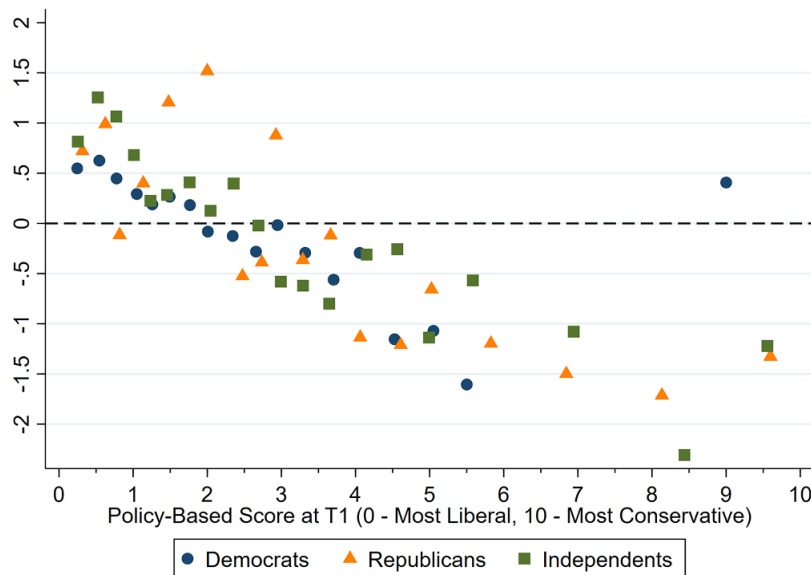


*Note:* Policy-Based Score is constructed for each individual based on responses to 27 questions identified as the most polarizing: 15%+ of each party are at opposite extremes (0 or 10).

deliberations. Figure 4 is a binned scatter plot showing how agreement with four factual statements about climate changes for individuals who start off at different points on the ideological spectrum.[12] Here, as with the overall PBS (which does not include these questions), change occurs only for the deliberators and grows in magnitude as one moves from left-of-center to the right on the T1 PBS. Those who start off most conservative on climate policies see the biggest increase in their agreement with factual statements about climate change as a result of the deliberation.

Figure 5 provides a similar picture for a measure of worry about the consequences of climate change. The y-axis in this binned scatter plot shows changes, between T1 and T2, in a worry score constructed as an average for each individual over three questions.[13] Here, too, worry increases for the deliberators, with change in worry levels increasing the most for those who start off more conservative on climate change actions. In contrast to the linear increase in belief of statements concerning the dangers posed by climate change, however, the increase in worry is more of a level shift upwards.

Figures 1, 4, and 5 are descriptive, to give the reader a picture of the differences between the treatment control groups. The significance of these differences is detailed in table 1.

As in other analyses in our study, using these controls enables us to not have to do any additional re-weighting to make the Control group and the Participant group comparable on initial observable characteristics. Furthermore,

our analyses suggest that the controls do not meaningfully affect our findings.

How do we explain these differences between treatment and control? We are especially interested in the reasoning that seems to motivate the differences. The reasons are causally proximate causes for the positions, reasons that are presumably affected by the deliberative process. Hence our interest in HYPOTHESIS 2.

Informed by these visual inspections of changes in key variables and potential mechanisms between T2 and T1 for the participant and control groups, we turn to capturing what we're seeing in regression form. Table 2 provides estimated coefficients for a model that hypothesizes that changes in worry and belief about the seriousness of climate change's impacts during the course of the deliberations might be connected to the change in PBS we documented in figure 1. Column 1 presents coefficients from a regression of change in PBS between T1 and T2 at the individual level on a dummy variable for whether the individual was in the participant or control group, the change for the individual in his/her worry and belief scores between T1 and T2, and the interaction between the dummy variable and his/her worry and belief score changes. The estimates in column 1 do not include controls for demographics. As we see from the coefficients on the interactions between the Participant dummy variable and the Change in Worry Score, an increase in worry about climate moves individuals to the left in their views on climate policies, disproportionately so for the participant group. A similar picture emerges for the relationship

**Figure 4**
**Climate belief score changes (T2–T1)**
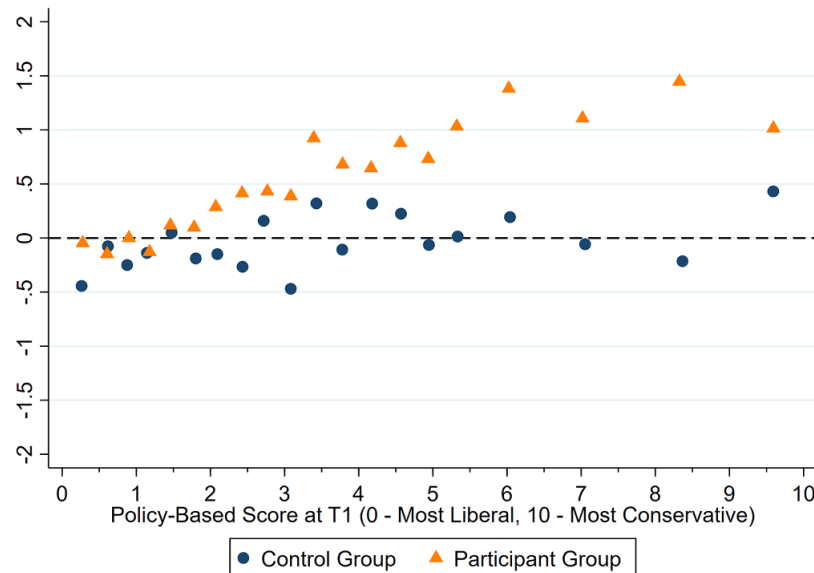


*Note:* Policy-Based Score is constructed for each individual based on responses to 27 questions identified as the most polarizing. Climate Belief Score is constructed using 4 questions on one's level of belief about climate change being a problem (10 is highest level of agreement with statements).

**Figure 5**
**Climate worry score changes (T2–T1)**
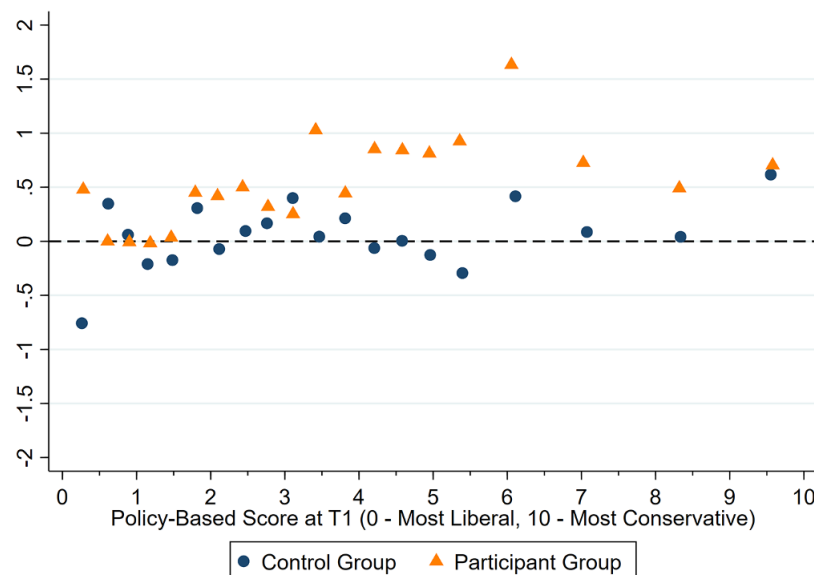


*Note:* Policy-Based Score is constructed for each individual based on responses to 27 questions identified as the most polarizing. Climate Worry Score is constructed using 3 questions on one's worry about the current condition of the natural environment (10 is "as worried as can be").

between the change in belief about climate change impacts: those who see an increase in such beliefs also see more movement to the left on the overall PBS (again, disproportionately so for the participant group). In column 2, the estimated model also includes control variables that capture the individual's level of education, gender,

## Table 2
## Potential drivers of changes in policy-based score (PBS) between Time 1 and Time 2

|  | (1) | (2) |
|---|---|---|
|  | ΔPBS | ΔPBS |
| Participant=1 | −0.033 | −0.045 |
|  | (0.056) | (0.061) |
| Change in Worry Score (T2–T1) | −0.050 | −0.040 |
|  | (0.032) | (0.032) |
| Participant*Change in Worry Score (T2 − T1) | −0.064 | −0.074* |
|  | (0.042) | (0.042) |
| Change in Belief Score (T2–T1) | −0.212*** | −0.190*** |
|  | (0.040) | (0.038) |
| Participant*Change in Belief Score (T2–T1) | −0.135*** | −0.129*** |
|  | (0.050) | (0.049) |
| Constant | −0.042 | −1.691* |
|  | (0.041) | (0.955) |
| Observations | 1,391 | 1,383 |
| Adj. R^2 | 0.226 | 0.257 |
| R^2 | 0.229 | 0.292 |
| Controls | No | Yes |

\* $p < 0.10$
\*\* $p < 0.05$
\*\*\* $p < 0.01$.
*Note:* Standard errors are robust and reported in parentheses. Controls include dummies for high school or less education, female, over 60 in age, white, married, employed, low income, rural resident, internet connection at home, lean Republican, and state of residence. Worry Score and Belief Score are averages over three questions and four questions, respectively, where 0 is no worry/belief and 10 is highest worry/belief.

## Table 3
## DV: Support democratic control of congress (Time 3)

|  | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Deliberation (participation in the treatment) | −1.408* | −0.628 | −0.534 |
|  | (0.707) | (0.783) | (0.696) |
| Importance Climate for Vote | −0.049 | — | — |
|  | (0.066) |  |  |
| Deliberation* Importance Climate | 0.717* | — | — |
|  | (0.079) |  |  |
| Importance Crime for Vote | — | −0.285*** | — |
|  |  | (0.077) |  |
| Deliberation* Importance Crime | — | 0.075 | — |
|  |  | (0.098) |  |
| Importance Democracy for Vote | — | — | 0.099 |
|  |  |  | (0.066) |
| Deliberation* Importance Democracy | — | — | 0.057 |
|  |  |  | (0.087) |
| Num.Obs. | 1,213 | 1,215 | 1,191 |
| Adjusted R2 | 0.769 | 0.779 | 0.761 |

+ $p<0.1$
\* $p<0.05$
\*\* $p<0.01$
\*\*\* $p<0.0.01$
*Note:* The models are logit regressions with random intercepts for state. Models include demographic controls for respondent age, gender, race, income, education, and region, respondent party ID, as well as respondent T1 PBS. We include survey weights to ensure the balance between the treated and control groups. The results in table 3 hold when we run OLS regressions instead of logistic regressions (refer to online appendix table A6), to address concerns raised by Berry, DeMerrit, and Escarey (2010) about logistic regression and interactions. Because we are estimating an average treatment effect, using a linear probability model should produce consistent results (Angrist and Pischke 2009). We also show the marginal effects plot to demonstrate how different the two sets of equations are.

age, race, marital status, employment status, income level, rural or metropolitan area of residence, existence of internet connection at home, political party preference, and current state of residence. The coefficients on our interactions of interest are similar with this model specification to those in column 1. Given the patterns we observe in figures 4 and 5, these are exactly the findings we would expect to see in regression form.[14]

## Did Deliberation Produce More Deliberative Voters?

HYPOTHESIS 3 posits that deliberation will produce a latent variable of civic engagement that will yield more deliberative voters. Recall that the national deliberation in this experiment occurred in late summer /early autumn 2021 (T1 at time of recruitment in early August, T2 after deliberation in late September) and the third questionnaire (T3) was fielded in October 2022 before the mid-term elections. Did the deliberations in 2021 affect vote choice a year later in the midterms in 2022? Because of the time lapse of a year, because the intervention was via an online discussion, and because the election was a mid-term for

control of Congress (rather than a presidential election) the hypothesis sets a high bar.

We will look at how much deliberation predicted support for Democratic control of Congress in the mid-term elections through the prism of how much it changed the relative importance of climate. Because the Republicans are generally more skeptical of climate change and the Democrats more in favor of actions to address climate change,[15] we hypothesize that support for Democratic control for Congress will be correlated with an increase in concern for climate change resulting from participation in the deliberative groups.[16]

To test this proposition, we fit a logistic regression model on support for Democratic control of Congress against deliberation. We also interacted deliberation with respondent prioritization of climate as an issue that drove their vote. Admittedly, we are assuming that those who

**Figure 6**
**Predicted support of democratic congress: Climate importance**
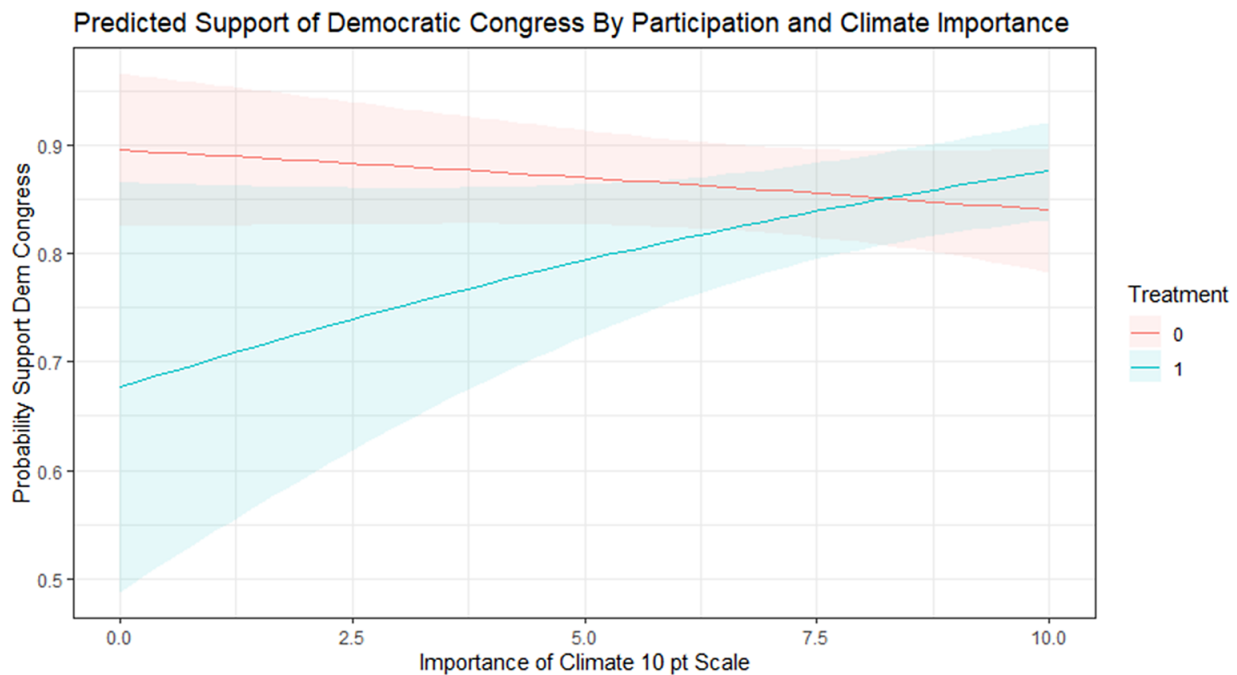


*Notes:* Marginal Effects plot based on Model 1 from table 3. All control variables are held at either their median for continuous variables or their mode for categorical variables.

support action on climate change will see that as a reason for supporting democratic control of Congress, given the evident party differences on the issue. Obviously, depending on the issue, the impact on deliberative voting could yield support for either party (depending on the party positions on the issue in question).

Model 1 in table 3 tests this directly near the time of the election by interacting participation with respondent reported climate importance on preference for control of Congress. If participation in the deliberation treatment has the effect we predict, we should see a significant and positive effect in this interaction, meaning that the linking of climate salience with preference for control of Congress is only true for participants in the deliberation treatment. We find exactly that—while also finding no strong direct impact of participation independent of climate importance and no impact of climate importance for the non-participants on preference for control of Congress.

We use Models 2 and 3 to test placebo importance measures to see if the effect exists for issues that were not discussed during the deliberation. Instead of climate importance, Models 2 and 3 use crime and democracy importance as plausible substitutes that should be independent of deliberation, since neither issue was discussed. And for those two issues, we find the opposite relationship than we did with climate: the overall issue importance for crime or democracy strongly predicts respondent preference for

control of Congress and there is no moderating effect by participation in the treatment at all. That is to say, it appears unlikely that respondent participation had any discernible impact on the overall importance respondents cited for the issues of crime and democracy, and to that effect, it is only through affecting overall climate importance that the treatment drove changes in respondent preferences. We see this illustrated more clearly in figures 6 and 7, which are marginal effects plots derived from Models 1 and 2 from table 3.

We see in figure 6 that, as climate increases in overall issue importance, the probability that a respondent will support Democratic control of Congress increases hardly at all for members of the control group, but quickly for members of the treatment group. And, among those who listed climate as the most important issue, there was a 10% difference in their likelihood of supporting Democratic control of Congress by treatment status. Indeed, the opposite relationship also holds true: if, after the deliberation and discussion over climate issues you still did not think that climate was an important issue weighing on your vote, you were 23% less likely to support Democratic control of Congress than if you had not participated at all in the first place.

Figure 7 provides a revealing counterfactual to figure 6 by looking at crime as opposed to climate, an issue that should not have been affected by participation in the deliberation. And, indeed, we see that borne out when comparing the treatment and control groups on the figure: there is virtually

**Figure 7**
**Predicted support of democratic congress: Crime importance**



Predicted Support of Democratic Congress By Participation and Crime Importance
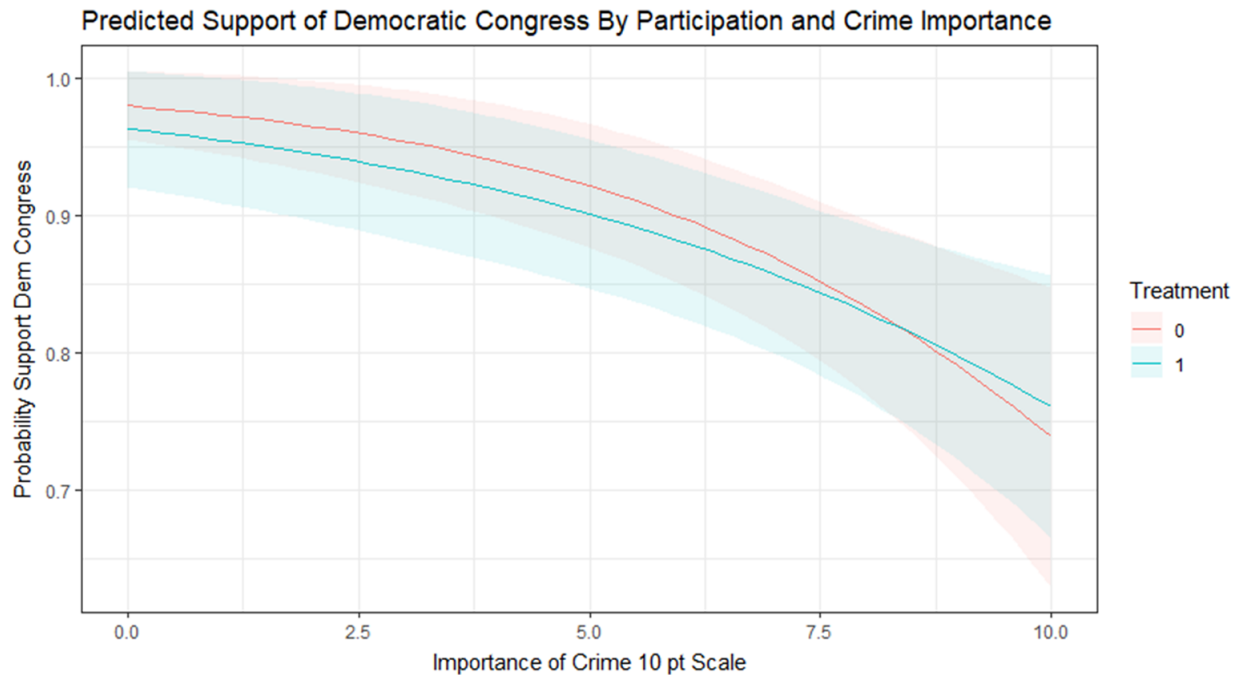
*Notes:* Marginal Effects plot based on Model 2 from table 3. All control variables are held at either their median for continuous variables or their mode for categorical variables.

no difference in the relationship between the importance of crime on vote and on preference for Democratic control of Congress. There is a consistently strong connection between crime importance and lessened support for Democratic control of Congress – indeed stronger than climate issues —but one that is unaffected by the treatment.[17]

## Causal Mediation Analysis and Deliberative Voting

It may seem surprising that a weekend of deliberation could have an effect on voting preferences a year later. Most proposals to insert organized deliberations into an election context look only at short-term effects. What are the opinions (and reasons) of the sample soon before primary elections (Fishkin 1991; McCombs and Reynolds 1999), before ballot propositions in the United States (as with the Citizens Initiative Review, see Gastil 2000 and Gastil and Noblock 2020) or before national referendums (Andersen and Hansen 2007). Few have contemplated long-term effects of deliberation (or indeed any other intervention) on voting behavior for periods as long as a year preceding an election.

Unlike the two cases cited earlier, Gastil et al. (2002) and our own work in Fishkin et al 2024, the intervention here was online, not face to face. The online intervention is of special interest because of its potential for scaling. Furthermore, because the deliberations were

about a specific substantive topic, we can compare the salience of that topic at the time of the election with the salience of other topics as a voting issue. In this case, our results show that deliberation in an organized process, even through an automated online discussion, can create more deliberative voters in a national election a year later.

But what accounts for the persistence of climate salience when there are so many other issues in a highly contentious national election? We turn to causal mediation analysis.

## Causal Mediation Analysis: Estimating Direct and Indirect Effects of Deliberation

This section reprises the framework we used in Fishkin et al. (2024) because we are exploring whether the effects from the face-to-face case can be replicated with the online AI-assisted deliberations. The traditional method of exploring relationships between a treatment and outcomes is by using a regression model. However, this method fails to disentangle underlying causes and effects that are indirect rather than direct. In our case, we know that there is an effect of participating in the deliberations on an individual's climate preferences and it seems to have an impact on electoral preferences over a year out. It is, however, unsatisfactory to state that participating in the deliberations is the direct cause of these electoral

**Table 4**
**Average causal mediated effect (ACME) of deliberation (95% CI)**

| | | Dependent Variable | |
| --- | --- | --- | --- |
| | | Climate Importance on Vote Choice? | Congress Control Vote Preference (Dem) |
| Mediation | Worry Climate Scale | **0.0458**\*\* | 0.0026 |
| | $\Delta$ T1 $\rightarrow$ T2 | **(0.0148, 0.081)** | (−0.005, 0.01) |
| | Worry Climate Scale | 0.0263 | 0.0003 |
| | $\Delta$ T1 $\rightarrow$ T3 | (−0.0195, 0.080) | (−0.001, 0.001) |
| | Climate Belief Scale | **0.0708**\*\*\* | −0.0019 |
| | $\Delta$ T1 $\rightarrow$ T2 | **(0.0246, 0.121)** | (−0.011, 0.011) |
| | Climate Belief Scale | **0.1030**\*\* | 0.0018 |
| | $\Delta$ T1 $\rightarrow$ T3 | **(0.0371, 0.172)** | (0.0028, 0.011) |
| | Climate Knowledge Scale | 0.0521\* | **0.0016**\*\* |
| | $\Delta$ T1 $\rightarrow$ T2 | (−0.0013, 0.121) | **(0.001, 0310)** |
| | Climate Knowledge Scale | **0.0249**\*\* | **0.0052**\*\* |
| | $\Delta$ T1 $\rightarrow$ T3 | **(0.0044, 0.052)** | **(0.011, 0.010)** |
| | Observations | 1,361 | 1,391 |

\* $p<.1$
\*\* $p<.05$
\*\*\* $p<.01$
*Note:* Each model is fit using a generalized linear mixed effects model for both the mediators and the dependent variables—linear models for each of the mediators and for the dependent variables. Random intercepts were fit at the state level. Models include demographic controls for age, gender, race, income, education, and region, as well as respondent T1 PBS. We include survey weights to ensure the balance between the treated and control groups. Observations include participants and control groups members. Climate importance on vote choice is a 0–10 scaled variable, with 0 being of no importance at all, and with 10 being extremely important for respondent vote choice. Congress Control Vote Preference for Democrats is a binary variable (1 if support Democratic control of both chambers, 0 otherwise). Models are fit using the "mediation" package in R with 95% CI included in the parenthesis. This approach is generally parallel to our work in Fishkin et al. 2024.

preferences: surely there were intermediate steps caused by the deliberations that, when taken together, affect these outcomes.

When faced with the possibility of indirect effects, investigators may have prior knowledge that an explanatory variable plausibly exerts its effect on an outcome via direct and indirect pathways. In the indirect pathway, there exists a *mediator* that transmits the causal effect. As noted, we outline mediation analysis later following along closely with the discussion of mediation analysis in our previous work (Fishkin et al. 2024, 13-16). We follow Imai, Keele, and Tingley (2010) and Imai et al. (2011) (see also Zhang et al. 2016 building on the same framework).

Suppose we have variables T and Y indicating the treatment variable and outcome variable, respectively. Mediation in its simplest form involves adding a mediator M between T and Y. The sequential ignorability assumption, critical to causal mediation analysis, states that the treatment (explanatory variable T) is first assumed to be ignorable given the pre-treatment covariates, and then the mediator variable (M) is assumed to be ignorable given the observed value of the treatment as well as the pre-treatment covariates (Imai, Keele, and Tingley 2010; Imai et al. 2011).

The first part is often satisfied by randomization, while the second part implies that there are no unmeasured confounding variables between mediator and outcome.

The standard mediation analysis starts with three equations, usually modeled with continuous outcomes (though advances in methods now allow for most parametric modeling approaches for stage of the mediation):

$$Y = i_1 + cT + e_1 \qquad [1]$$

$$Y = i_2 + c'T + bM + e_2 \qquad [2]$$

$$M = i_3 + aT + e_3 \qquad [3]$$

As Zang et al. explicate this standard approach:

where $i_1$, $i_2$, and $i_3$ denote intercepts, Y is the outcome variable, T is the treatment variable, M is the mediator, c is the coefficient linking T and Y (total causal effect), c' is the coefficient for the effect of T on Y adjusting for M (direct effect), b is the effect of M on Y adjusting for explanatory variables, and a is the coefficient relating to the effect of T on M. $e_1$, $e_2$, and $e_3$ are residuals that are uncorrelated with the variables in the right-hand side of the equation and are independent of each other. Under this specific model, the causal mediation effect (CME) is represented by the product coefficient of ab. Of note, Eq. [3] can be substituted into Eq. [2] to eliminate the term M:

$$Y = i_2 + bi_3 + (c' + ab)T + e_2 + be_3 \qquad [4]$$

It appears that the parameters related to direct (c') and indirect effect (ab) of T on Y are different from that of their total effect. That is, testing the null hypothesis c=0 is unnecessary since CME

can be nonzero even when the total causal effect is zero (i.e., direct and indirect effects can be opposite), which reflects the effect cancellation from different pathways. (Zang et al. 2016, 2)

This standard setting for mediation analysis was refined and brought into the potential outcomes framework in Imai et al. (2011). The authors propose a set of methods that unifies the approach to identifying direct and indirect effects, relying on a set of assumptions that are more readily testable than classical mediation analysis provides.

## Causal Mediation Analysis: Climate Change and Voting

We follow the procedures for using causal mediation analysis to study indirect effects of deliberation laid out in Fishkin et al. (2024), where we looked at the effects of deliberation on voting behavior through an activation of a latent civic awakening. We believe our research design lends itself to a similar analytic framework for this online case. We ran causal mediation analysis to determine the indirect effect of deliberation (from T1 to T2) on respondents prioritizing climate as a consideration in voting behavior and on respondent preferences for which party controls Congress at T3. We believe the mediating effects of deliberation on these two respective outcomes are going to be changes in how much the respondent worried about climate (as represented in the "Worry Climate Scale"), how many true things they believe about the climate (as represented in the "Climate Belief Scale"), and in how knowledgeable about the climate they are (as represented in the "Climate Knowledge Scale"). We view these three mediators as an overall latent climate engagement dimension, induced by deliberation at the event a year earlier. For more details on the model, refer to the "Addendum on Causal Mediation Analysis" in the online appendix.

In this analysis we adapt our approach to the indirect effects of deliberation on voting in Fishkin et al. 2024, 17. Participation in the deliberations significantly increased worry about climate, the number of correctly held climate beliefs, and the overall knowledge about climate, as demonstrated in pervious sections. Because of this, we know that it is possible that these three mediators will have significant indirect effects on outcomes, even if there is weaker evidence for a direct effect of these mediators. Table 4 shows the indirect effects of deliberation on climate worry, beliefs, and knowledge on how important climate is to vote choice, and respondent preference that Democrats control Congress. The values are the average causal mediated effect, with bootstrapped 95% confidence intervals which follow.

We find consistent effects of deliberation on the aggregate latent climate dimension—as measured through changes in climate worry, belief, and knowledge—on

respondent valuation of the importance of climate on their vote choice. Note that we do not get to see their specific vote choice, so this is only a self-reported measure of what was important to their vote. The fact that the effect is consistent across five of the six indicators of latent climate engagement is indicative of the robustness of the effect. We do not find an indirect effect on respondent preferences for who controls Congress—with the exception of the Climate Belief Scale in both formulations (time 1 to time 2 and time 1 to time 3) and the Climate Knowledge Scale from time 1 to time 3. The presence of these indirect effects, as measured by the causal mediation analysis, suggests that the effects of participating persisted in ways that raised climate salience even a year after the event. And as we have shown, climate salience at the time of the election has an effect on deliberative voting. Voters connect the salience they attribute to climate with voting for a Democratic congress. By contrast the control group did not prioritize climate but instead prioritized crime or other issues in their voting preferences for who should control Congress.

## Conclusion

We attempt to connect discussions in political theory about a more deliberative society with empirical issues that might confront efforts to actually achieve one through the spread of organized deliberation. Technology offers new deliberative possibilities. This paper is an exploration of what light a national controlled experiment can shed on the challenge of scaling. If, as suggested by these results, the online process with the AI-assisted moderator can be conducted at a high level of deliberative quality, then the road is open to new designs for deliberative systems, at some points employing mini-publics, at some points employing mass deliberation with this kind of technology.[18] For example, the mini-publics can be used for agenda setting and the mass deliberation could be used to help prepare citizens for voting in referenda or elections, as sketched in Deliberation Day (Ackerman and Fishkin 2004; Fishkin 2018). Whatever the design, the prospect of viable forms of mass organized deliberation can close the gap between mini-publics and the mass public and provide a more solid basis for deliberative policy-making and more deliberative elections.

Some deliberative theorists have argued that deliberation and competitive elections are sufficiently incompatible that connecting them is not an appropriate aspiration for deliberative designs (see, for example, Thompson 2013, who supports "deliberation about elections" but opposes "deliberation in elections"). We already have elections saturated with political tribalism. Party loyalties, and partisanship seem to determine elections (see for example, Green, Palmquist, and Schickler 2004 and Achen and

Bartels 2016), and the more competitive the election, the more it seems likely that partisanship will cloud whatever deliberative capacities voters might have. From this perspective, some theorists argue that we need to give up on competitive elections entirely and replace them with legislatures selected by random sampling or "sortition" (Van Reybrouck 2016; Gastil and Wright 2019; Landemore 2020). A key point from this sortition perspective is that random samples of the people are not running for re-election so they can make sincere judgments on the merits (outside the hothouse environment of political campaigns). Other ("realist") theorists say we need to give up on deliberation affecting elections and accept that competitive democracy will not connect meaningfully to the "will of the people." That is just a "pipe dream hardly worthy of the attention of a serious person" (Posner 2003, 163), a "fairy tale" of what Achen and Bartels call the "folk theory of democracy" (Achen and Bartels 2016, 7). On this view "voting behavior primarily reflects and reinforces voter's social loyalties" hence "it is a mistake to suppose that elections result in popular control" (Achen and Bartels 2016, 4).

In between these two extremes of giving up on elections (by having sortition assemblies) and giving up on deliberation (by having pure Schumpeterian competition without any meaningful "will of the people") we can aspire to foster deliberative voters at scale. We can aspire to create enough deliberative voters so that we will have more deliberative elections. This paper offers experimental evidence that sheds some light on these long-term deliberative possibilities. First, the automated online process (and its future improved successors) makes this aspiration substantially more practical, less expensive and less utopian. Second, an unexpected indication that it could work is the kind of result uncovered here—deliberative voting fostered by an online process and for voting in a mid-term election, no less. Presumably, if there were lasting effects on deliberative voting a full year later, there would also have been such effects, probably stronger, if the time gap had been shorter. It seems plausible to confirm that the automated online deliberative process has an effect on producing more deliberative voters. This effect is crucial for designing a system where thoughtful popular control can be restored through deliberation.

## Supplementary material

To view supplementary material for this article, please visit http://doi.org/10.1017/S1537592724001749.

## Data replication

Data replication sets are available in Harvard Dataverse (Bolotnyy et al. 2024) at: https://doi.org/10.7910/DVN/IIOG1S

## Notes

1 In a later study, Nyhan et al. and others found the backfire effect difficult to replicate (see Nyhan et al. 2019). However, they found that misinformation persists despite exposure to correction, a result consistent with the continued influence effect (CIE).

2 In a separate collaboration in Finland we will report on a controlled experiment comparing opinion change with this automated platform with results from small groups with human moderators (both online with Zoom and face-to-face conditions) all on the same topic (see Grönlund et al. 2024).

3 https://deliberation.stanford.edu/tools-and-resources/online-deliberation-platform

4 Landemore (2022) distinguishes three ways that deliberative democracy could be brought to the masses: a) some mechanism whereby each citizen could talk with every other; b) division of the population into large numbers of small groups; and c) proliferation of random samples to deliberate as minipublics. In this paper we are exploring b) and c) as strategies for spreading deliberation with the ultimate aim of fostering more deliberative voters.

5 The deliberative condition also included an oversample from Texas and California to permit separate inferences about those two states. Refer to the online appendix discussion of methodology.

6 The definition includes provision for a further complexity: when the two parties move from opposite sides to the same side of the divide between opposition and support, that is considered depolarization even if the means are not closer, because the other party moved

even further on that side (see page 4 of our earlier work, Fishkin et al. 2021)

7 As we note in Fishkin et al. 2021, 1468, the Deliberative Poll appears to fit this design. The task of the small groups is to discuss the arguments for and against each policy proposal and then formulate group questions for the balanced panels of competing experts in the plenary sessions, all so that the participants can come to their own, individually considered judgments.

8 The methodological approach is also similar in spirit to the weighted averaging of issues scale method advanced in Ansolabehere, Rodden, and Snyder (2008).

9 Answer choices for our questions range from 0 to 10. Having Democrats and Republicans at opposite sides of an issue means having the majority of respondents of one party locate themselves in the 0–4 range and the majority of respondents of the other party locate themselves in the 6–10 range. Extremes are defined as 0 or 10. Additionally, the set of questions we deem extremely polarizing is robust to using a definition of party identification that allows people to specify whether they lean towards a particular party rather than firmly identify with one.

10 We also include as a robustness check a factor analytic-based measure of ideology that relies on Poole's (1998) BasicSpace scaling method. We show a comparison between this measure and time 1 PBS in online appendix figure A1. We find a 0.97 correlation between these two measures, indicating that they are functionally equivalent.

11 A binned scatterplot splits the sample up into groups based on the x-axis variable (in our case Time 1 PBS) such that each group (bin) has an equal number of individuals. Each circle, triangle, or square depicted in our figures shows the average y-axis value for the individuals in each x-axis group.

12 The four questions being averaged for each individual to obtain their T1 and T2 belief scores are the following. Answers ranged from 0 (Strongly disagree) to 10 (Strongly agree):

  (1) Our planet is experiencing an increase in global temperatures that will greatly harm our quality of life.
  (2) Rising temperatures are caused by human activities that emit greenhouse gases, like carbon dioxide and methane, which trap heat in the atmosphere and warm the earth's climate.
  (3) Failure to address these issues (rising temperatures) will threaten human life on earth within the next century.

  (4) In order to stop the increase in global temperatures, humans must stop adding to the total amount of climate-heating gases in the atmosphere (reach Net Zero).

13 The three questions being averaged for each individual to obtain their T1 and T2 worry scores are the following. Answers ranged from 0 (not at all worried) to 10 (as worried as can be):

  (1) How worried are you about the current condition of the natural environment in your local area?
  (2) How worried are you about the current condition of the natural environment in the United States?
  (3) How worried are you about the current condition of the natural environment of the earth as a whole?

14 Note that we do not present or discuss the coefficients on our control variables, because this is an experimental design and those coefficients would be meaningless or misleading (Keele, Stevenson, and Elwert 2020).

15 See https://climatecommunication.yale.edu/visualizations-data/partisan-maps-2018/ for visualizations of party positions on climate change constructed from survey data.

16 We asked both the treatment and control group to rate the importance of nine issues in the upcoming election: threats to democracy, cost of living, jobs and the economy, immigration and the situation at the border, climate change, guns, abortion, crime, and Covid 19.

17 For an alternative version of figures 6 and 7 based on second differences, see figures A5 and A6 in the online appendix.

18 There is still an important practical consideration to be confronted: how much diversity is needed for the deliberations in the small groups. For an analysis based on the data from this project, see the online appendix for an "Addendum on Group Level Diversity and Individual Change".

## References

Achen, Christopher, and Larry Bartels. 2016. *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton, NJ: University Press.

Ackerman, Bruce, and James S. Fishkin. 2004. *Deliberation Day*. New Haven, CT: Yale University Press.

Andersen, Vibeke Normann, and Kasper M. Hansen. 2007. "How Deliberation Makes Better Citizens: The Danish Deliberative Poll on the Euro." *European Journal of Political Research* 46:531–56. doi:10.1111/j.1475-6765.2007.00699.x

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder. 2008 "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting." *American Political Science Review* 102(2): 215–32. doi:10.1017/S0003055408080210

Berry, William, Jacqueline H.R. DeMeritt, and Justin Esarey. 2010. "Testing for Interaction in Binary Logit and Probit Models: Is a Product Term Essential?" 54(1): 248–66 (https://www.jstor.org/stable/20647982).

Bolotnyy, Valentin, James Fishkin, Alice Siu, Joshua Lerner, Norman Bradburn. 2024. "Replication Data for: Scaling Dialogue for Democracy: Can Automated Deliberation Create More Deliberative Voters?" *Harvard Dataverse*, doi:10.7910/DVN/IIOG1S.

Chong, Dennis, and James Druckman. 2007. "Framing Public Opinion in Competitive Democracies." *American Political Science Review* 101(4): 637–55. doi:10.1017/S0003055407070554

Dahl, Robert A. 1989. *Democracy and Its Critics*. New Haven: Yale University Press.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.

Ecker, Ullrich K.H., Stephan Lewandowsky, John Cook, Philipp Schmid, Lisa K. Fazio, Nadia Brashier, Panayiota Kendeou, Emily K Vraga, and Michelle A. Amazeen. 2022. "The Psychological Drivers of Misinformation Belief and Its Resistance to Correction." *Nature Reviews: Psychology* 1:13–29. http://doi.org/10.1038/s44159-021-00006-y

Farrar, Cynthia, James S. Fishkin, Donald P. Green, Christian List, Robert C. Luskin, and Elizabeth Levy Paluck. 2010. "Disaggregating Deliberation's Effects: An Experiment within a Deliberative Poll." *British Journal of Political Science* 40(2): 333–47.

Fishkin, James. 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. New Haven, CT: Yale University Press.

———. 2018. *Democracy When the People Are Thinking*. Oxford: Oxford University Press.

Fishkin, J., A, Siu, L. Diamond, and N. Bradburn. 2021. "Is Deliberation an Antidote to Extreme Partisan Polarization? Reflections on "America in One Room." *American Political Science Review* 115(4): 1464–81. doi:10.1017/S0003055421000642

Fishkin, James, Valentin Bolotnyy, Joshua Lerner, Alice Siu, and Norman Bradburn. 2024. "Can Deliberation Have Lasting Effects?" *American Political Science Review*, FirstView, 1–24. doi:10.1017/S0003055423001363

Gastil, J. 2000. *By Popular Demand: Revitalizing Representative Democracy through Deliberative Elections*. Berkeley, CA: University of California Press.

Gastil, J., E.P. Deess, and P.J. Weiser. 2002. "Civic Awakening in the Jury Room: A Test of the Connection between Jury Deliberation and Political Participation." *Journal of Politics* 64(2): 585–95. doi:10.1111/1468-2508.00141

Gastil, J., E. Pierre Dees, Philip J. Weiser, and Cindy Simmons 2010. *The Jury and Democracy: How Jury Deliberation Promotes Civic Engagement and Political Participation*. Oxford: Oxford University Press.

Gastil, J., and Katherine R. Noblock 2020. *Hope for Democracy: How Citizens Can Bring Reason Back into Politics*. Oxford: Oxford University Press.

Gastil J., and Wright, E.O. 2019. *Legislature by Lot: Transformative Designs for Deliberative Governance*. New York: Verso Books.

Germann, M., S. Marien, and L. Muradova. 2022. "Scaling Up? Unpacking the Effect of Deliberative Mini-Publics" *Political Studies*. doi:10.1177/00323217221137444

Green, Donald, Bradley Palmquist, and Eric Schickler. 2004. *Partisan Hearts and Minds*. New Haven, CT: Yale University Press.

Grönlund, Kimmo, Kaisa Herne, James Fishkin, Janette Huttunen, Maija Jäske, Marina Lindell, Isak Vento, Kim Backström, Alice Siu. and Lodewijk Gelauff. 2024. "Good Deliberation Regardless of Mode? A Comparison of Automated Online, Human Online and Face-to-Face Moderations in a Deliberative Mini-Public." Presented at the annual meeting of the American Political Science Association, Philadelphia, September 2024.

Hardin, Russell. 2002. "Street Level Epistemology and Democratic Participation" *Journal of Political Philosophy* 10(2): 212–22. doi:10.1111/1467-9760.00150

Hetherington, Marc, and Thomas Rudolph. 2015. *Why Washington Won't Work*. Chicago: University of Chicago Press.

Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. "A General Approach to Causal Mediation Analysis." *Psychological Methods* 15(4): 309–34. doi:10.1037/a0020761

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105(4): 765–89. doi:10.1017/S0003055411000414

Keele, Luke, Randolph T. Stevenson, and Felix Elwert. 2020. "The Causal Interpretation of Estimated Associations in Regression Models." *Political Science Research and Methods* 8(1): 1–13. doi:10.1017/psrm.2019.31

Krosnick, Jon A., Allyson L. Holbrook, Laura Lowe, and Penny S. Visser. 2006. "The Origins and Consequences of Democratic Citizens' Policy

Agendas: A Study of Popular Concern about Global Warming." *Climatic Change* 77:7–43 doi:10.1007/s10584-006-9068-8

Kunda, Z. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108(3): 480–98.

Lafont, Cristina. 2020. *Democracy without Shortcuts: A Participatory Conception of Deliberative Democracy.* Oxford: Oxford University Press.

Landemore, H. 2022. "Can AI Bring Deliberative Democracy to the Masses?" Presented at Human Centered Artificial Intelligence Seminar, Stanford University, October 2022.

Mansbridge, Jane J. 1999. "Everyday Talk in the Deliberative System. In *Deliberative Politics: Essays on Democracy and Disagreement*, ed. S. Macedo, 211–242. Oxford: Oxford University Press.

McCombs, Max, and Amy Reynolds, eds. 1999. "*The Poll with a Human Face: The National Issues Convention Experiment in Political Communication.*" Mahwah, NJ: Lawrence Erlbaum Associates.

Muradova, Lala, Eileen Culloty, and Jane Suiter. 2023. "Misperceptions and Minipublics: Does Endorsement of Expert Information by a Minipublic Influence Misperceptions in the Wider Public?" *Political Communication* 40(5): 555–75. doi:10.1080/10584609.2023.2200735

Muradova, Lala, and Jane Suiter. 2022. "Public Compliance with Difficult Political Decisions in Times of a Pandemic: Does Citizen Deliberation Help?" *International Journal of Public Opinion Research* 34(3): edac026. doi:10.1093/ijpor/edac026

Nyhan, Brendan, and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32(2): 303–30. doi:10.1007/s11109-010-9112-2

Nyhan, Brendan, E. Porter, J. Reifler, T.J. Wood. 2019. "Taking Fact-Checks Literally but Not Seriously? The Effects of Journalistic Fact-Checking on Factual Beliefs and Candidate Favorability." *Political Behavior*. doi: 10.1007/s11109-019-09528-x

Poole, Keith T. 1998. "Recovering a Basic Space from a Set of Issue Scales." *American Journal of Political Science* 42(3): 954–93. doi:10.2307/2991737

Posner, Richard. 2003. *Law, Pragmatism and Democracy.* Cambridge: Harvard University Press.

Rasinski, Kenneth, N. Bradburn, and D. Lauen 1999. "Effects of NIC Media Coverage Among the Public." In *The Poll with a Human Face*, ed. Max McCombs and Amy Reynolds. Mahwah, NJ: Lawrence Erlbaum Associates.

Sandefur, Justin, Nancy Birdsall, James Fishkin, and Mujobu Moyo. 2022. "Democratic Deliberation and the Resource Curse: A Nationwide Experiment in Tanzania." *World Politics* 74(4): 564–609. doi:10.1017/S0043887122000090

Schumpeter, Joseph A. 1942. *Capitalism, Socialism and Democracy.* New York: Harper and Row.

Shapiro, Ian 2003. *The State of Democratic Theory* Princeton, NJ: Princeton University Press.

John Sides, Chris Tausanovitch, and Lynn Vavreck. 2022. "*The Bitter End: The 2020 Presidential Campaign and the Challenge to American Democracy.*" Princeton, NJ: Princeton University Press.

Thompson, Dennis 2013. "Deliberate about, Not in, Elections." *Election Law Journal* 12(4) 372–85. doi:10.1089/elj.2013.0194

Van Reybrouck, David. 2016. *Against Elections.* London: Bodley Head.

Zhang Z., C. Zheng, C. Kim, S. Van Poucke, S. Lin, and P. Lan. 2016. "Causal Mediation Analysis in the Context of Clinical Research." *Annals of Translational Medicine* 4(21): 425. doi:10.21037/atm.2016.11.11.