JOURNAL OF DELIBERATIVE DEMOCRACY

Rountree, J., & Gastil, J. (2026). The Case for Using Generative AI to Run Deliberation Simulations. *Journal of Deliberative Democracy*, 21(1), pp. 1–14. DOI: https://doi.org/10.16997/jdd.1625

RESEARCH ARTICLE

The Case for Using Generative AI to Run Deliberation Simulations

John Rountree* and John Gastil†

Generative artificial intelligence (GAI) interfaces, such as ChatGPT, Bloom, and Gemini, may offer transformative possibilities for deliberative democracy. We argue that deliberation scholars and practitioners should use GAI software to run simulations to complement existing deliberative processes. By simulations, we mean the use of GAI software to run hypothetical deliberations, either with or on behalf of human participants, to support rather than replace human judgment. We expand on the notion of a GAI deliberation simulation and showcase an example using GPT-4o. To illustrate the practical advantages of simulations, we showcase two use cases: training facilitators and providing time-sensitive policy consultation. We also address potential cautions and limitations surrounding GAI simulations, such as concerns about transparency and bias. We conclude by exploring the theoretical implications of GAI simulations for developing and refining models of deliberation dynamics.

Keywords: Generative AI; simulation; democratic deliberation; ChatGPT

Introduction

Novel deliberative democratic processes represent a significant portion of the wave of democratic innovations that have occurred in the past 30 years (Elstub & Escobar 2019; Jacquet et al. 2023; Smith 2009). A recent meta-analysis of 100 empirical studies on such innovations highlighted the prominence of deliberative minipublics, yielding an array of consistent, pro-democratic attitudinal shifts (Theuwis, Van Ham & Jacobs 2024). As successful as such minipublics have been (Grönlund, Bächtiger & Setälä 2014), they remain difficult to scale-up in a way that would make them ubiquitous features of democratic systems (Niemeyer 2014; Pateman 2012).

We contend that generative artificial intelligence (GAI) language interfaces, such as ChatGPT, Bloom, and Copilot, may offer transformative possibilities for the use of minipublics and deliberative democracy more broadly. In this essay, we argue that deliberation scholars and practitioners should use GAI models to run deliberation simulations to complement existing deliberative processes. By simulations, we mean the use of GAI software to run *hypothetical* deliberations, either with or on behalf of human participants. In this vein, simulations would be a "deliberation-making" tool rather than a "decision-making" tool (Niemeyer 2014).

The phrase "GAI deliberation simulation" combines two key terms: GAI and deliberation. GAI is a term used to describe machine learning models that are trained on large datasets and generate outputs to user prompts. A subset of GAI is large language models (LLMs), which are designed to generate natural-sounding language responses to user-input prompts. They rely on large amounts of real text to learn how to produce realistic responses, predicting each subsequent word to string together phrases, sentences, paragraphs, and full-length texts of any kind. The language models used by ChatGPT—such as GPT-4o (OpenAI et al., 2023)—have received the most attention, but other families of models are now available, including Llama (Touvron et al., 2023), Gemini (Google Deep Mind, 2025), and Claude (Anthropic, 2025).

We use "deliberation" here as shorthand for "democratic deliberation," which is an approach to decision-making wherein groups "carefully examine a problem and arrive at a well-reasoned solution after a period of inclusive, respectful consideration of diverse points of view" (Gastil & Black 2008: 2). The concept of deliberative democracy emerged out of political and social theory as an abstract, normative ideal for public decision-making (Benhabib 1996; Gutmann & Thompson 2004). Since then, deliberative democratic processes have received substantial empirical investigation from scholars (Dryzek et al. 2019; Nabatchi et al. 2012; Neblo 2015).

High-quality deliberative processes have been implemented across the world by practitioners to involve ordinary citizens. The gold standard for democratic deliberation in recent years has been "deliberative minipublics"—randomly selected groups of citizens,

* College of Humanities and Social Sciences, University of Houston-Downtown, USA

† Communication Arts & Sciences, Pennsylvania State University, USA

Corresponding author: John Rountree (rountreej@uhd.edu)

representing a microcosm of the public, that are provided with the time, resources, and guidance to consider policy issues in depth, deliberate together, and make recommendations (Grönlund, Bächtiger & Setälä 2014). Example minipublics include citizens' juries, 21st Century Town Meetings, Deliberative Polls, or citizens' assemblies (Gastil & Levine 2005; Reuchamps, Vrydagh & Welp 2023). Democratic deliberations, such as minipublics, are often designed to combine systematic analysis of a policy issue (e.g., identifying a broad range of solutions, weighing pros and cons of policy alternatives) with democratic principles that promote social equality, such as mutual respect and an obligation to consider other viewpoints (Gastil & Black 2008).

A GAI deliberation simulation (hereafter "simulation") combines GAI with democratic deliberation. It refers to the use of an LLM to run a hypothetical deliberation. In a simulation, variables can be input into the LLM to prompt it to consider a policy decision from multiple perspectives and to present those perspectives in conversation. It can then be asked to write out a transcript of the deliberation to simulate those perspectives, weighing the advantages and tradeoffs of decisions.

This is not a hypothetical future technology. We can already do a version of this with LLMs currently available to the public. We will use GPT-4o as an example of how simulations could be run with this software. We rely on GPT-4o because it is a language model that will likely be most familiar to readers. For reasons we will elaborate later in the essay, we are not arguing that GPT-4o is the most appropriate model for running a simulation, but neither are we dismissing it entirely. For our purposes, GPT-4o demonstrates what is possible with the current state of generative AI.

Through our testing, simulations on GPT-4o can be run with varying degrees of complexity. It can be prompted to run a simulation, for example, with a single prompt, such as: "Imagine many different people from diverse perspectives are considering policy options on gun control. Write out a transcript of their conversation." In response, GPT-4o generates hypothetical personas debating the merits of gun control measures, mental health education, and responsible gun ownership.

More sophisticated prompt engineering could be used to ensure simulations align with the normative ideals of deliberative democracy. For example, one model would be based on the evaluative criteria for deliberative events and projects proposed by Gastil, Knobloch and Kelly (2012). Their input and throughput criteria, specifically, could be adapted to simulation design. First, the design integrity of a simulation could be modeled through prompts that attempt to provide participant representativeness, unbiased framing, and procedural design involvement (see **Figure 1**). Then, a simulation could follow their analytic process of deliberation by establishing a solid information base, prioritizing key values at stake, identifying a broad range of solutions, and weighing the pros, cons, and tradeoffs among solutions (see **Figure 2**; Gastil, Knobloch & Kelly 2012; Gastil & Black 2008). It could end with a report of all the policy options considered by the diverse personas with the accompanying pros, cons, and tradeoffs.

Similar to a Citizens' Initiative Review statement (Drury & Rountree 2023), the report could be used to help others weigh their own preferences on a decision.

In what follows, we begin by asking whether GAI can run a deliberative simulation effectively. This first main section expands on the notion of a GAI deliberation simulation and provides an example simulation using GPT-4o. The second main section then illustrates the potential for GAI simulations to provide facilitator training, offer time-sensitive consultation on public opinion, and support further applications beyond these two use cases. The third main section offers cautions and acknowledges limitations. In the conclusion, we highlight some of the implications of this approach to deliberative practice, research, and theory.

Can GAI Really Run a Deliberation Simulation?

At this point, a skeptical reader may question whether GPT-4o or another existing GAI model could follow this complex series of inputs to produce a simulation. We have tested this process through GPT-4o (OpenAI.com) and invite readers to do the same. (The Appendix includes a set of prompts for a simulation.)

Below is just one excerpt from a simulation run by the authors on mental health challenges in K-12 schools. After explaining the purpose of the simulation to ChatGPT, asking it to generate diverse personas relevant to the issue, and having it assign those personas to small groups, we used the software to simulate the initial small-group discussions about the key values at stake on this policy issue. The proximate prompt leading into this discussion read: "Simulate a discussion among the group where they discuss the question 'What key values are at stake in this policy issue?' Have them generate a list as a group."

> **Facilitator (Taylor Reed):** Thanks, everyone, for being here. We're discussing mental health in K-12 schools, and I want to start by asking: What key values are at stake in this issue? As we talk, I'll jot down a list, and we can refine it together. Who'd like to start?
>
> **Alex Carter (Parent & Youth Advocate):** I think equity is a huge part of this. Some kids can afford private therapy, but a lot can't. If schools don't provide real mental health resources, we're basically leaving vulnerable students behind. That's unacceptable.
>
> **Jordan Mitchell (School Counselor & Mental Health Professional):** Absolutely. I see this every day. A kid in crisis shouldn't have to wait weeks to see me, but with our limited staff, that's the reality. Access needs to be a priority—not just in wealthier districts but across the board.

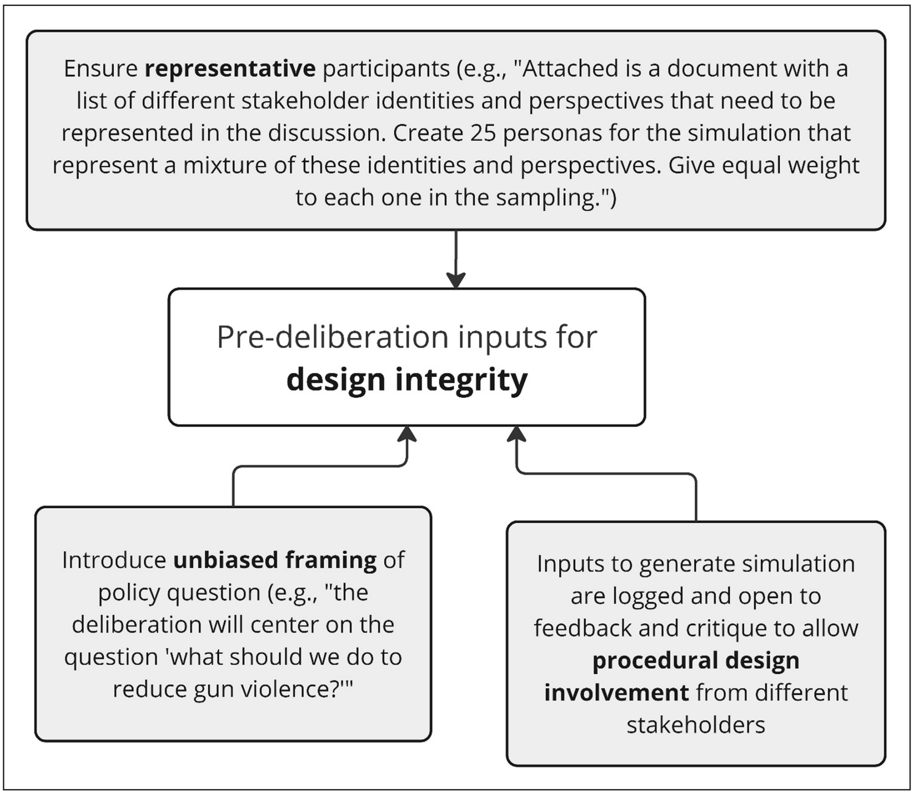

Figure 1: Example pre-deliberation inputs for design integrity in a deliberation simulation.

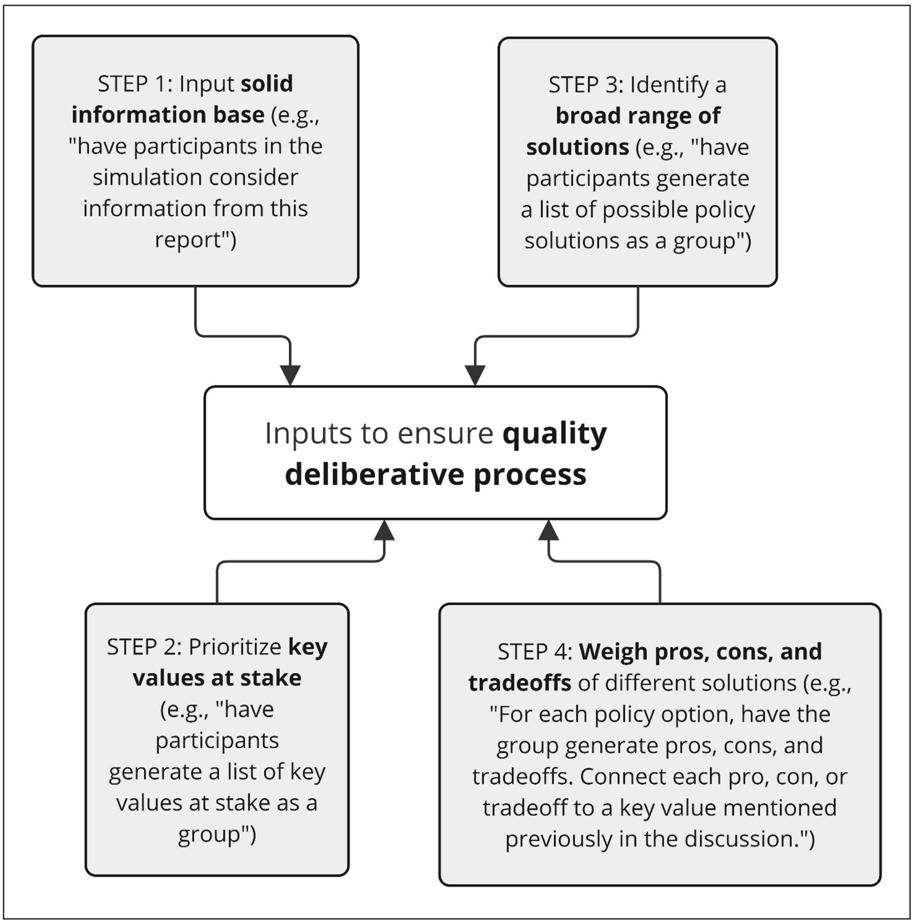

Figure 2: Example inputs to ensure a quality deliberative process in a simulation.

> **Facilitator (Taylor Reed):** So, equity and access—ensuring all students, regardless of income or background, have real support?
>
> **Mia Robinson (High School Student & Mental Health Advocate):** Yeah, but not just access to 'a counselor.' We need access to people who actually get us. Some students feel like counselors don't understand their culture or experiences, so they don't even try to get help.
>
> **Facilitator (Taylor Reed):** Good point. So we're also talking about cultural competence in mental health services?
>
> **Chris Delgado (Teacher & Parent with Privacy Concerns):** I get the equity piece, but what about privacy? If schools start tracking mental health, where does that information go? Do parents always have the right to know if their kid talks to a counselor? What if the student doesn't want them to know?
>
> **Principal Renee Thomas (School Administrator):** This is a tough one. Legally, we have to involve parents in certain situations, but sometimes students—especially LGBTQ+ students or those in unsafe households—don't want their parents involved. We need to protect students, but where's the line?
>
> **Morgan Patel (State Legislator & Former Student):** That's a critical balance. As a policymaker, I hear from both sides—some parents want total transparency, while others recognize that kids need a safe space to talk without fear.

Human discussions are not as neatly organized as the simulated conversation in this excerpt, and this tidiness of the conversation gives the simulated discussion an artificial quality. Real human conversations are filled with false starts, digressions, incomplete sentences, fillers, multiple speakers talking at once, and all the messiness of oral communication. Even qualitative researchers, however, recognize that literal transcripts can be indecipherable if they include all the awkwardness of natural speech (Lindlof & Taylor 2002). Given that AI simulations are a kind of fiction, it also bears mention that novelists face a similar dilemma: Their writing must diverge from *real* linguistic behavior to produce dialogue that sounds *realistic* (Stein 1995).

Whatever their fidelity to human speech, the simulated personas of Carter, Mitchell, Robinson, Delgado, Patel, and Thomas do showcase one component of a deliberative analytic process. Specifically, the personas articulate different values at stake based on their unique positionalities relative to mental health challenges in K-12 schools.

Again, one might question whether this simulated conversation is an authentic representation of deliberation. It is clear from the excerpt that GAI can present different vantage points on an issue that, on the surface, could be considered deliberative; however, it is an open question whether (and to what extent) GAI can actually simulate critical aspects of deliberation, such as arriving at high-quality judgments. This question is important beyond deliberation simulations. Others have argued for and implemented LLMs to simulate human behavior for social science research (Aher, Arriaga & Kalai 2023; Gao et al. 2023; Grossmann et al. 2023). As this technology evolves and GAI gets used to simulate real human interactions, it will become necessary to grapple with the quality of simulations.

Transparency presents the largest challenge for design integrity and quality control of simulations. Without full disclosure by companies such as OpenAI and Microsoft about the data used to train these models, simulations may always face questions of their legitimacy due to potential biases in their training data. However, this issue mirrors the biases that real people will bring into human deliberation—biases cannot be eliminated, but they can be confronted and mitigated. LLMs such as GPT-4o could be customized and trained by researchers to perform more closely to the ideals of deliberative democracy. This training could include creating an ethical dataset and weighing it more heavily in simulations than the LLM's base training data, and the model could be iteratively refined based on its performance.

Researchers can also validate the quality of simulations by comparing GAI outputs with evidence from human deliberation. Part of this could be accomplished through "Turing Experiments," where output from simulations is compared to outputs from well-documented phenomena in social science research (Aher, Arriaga & Kalai 2023). Other forms of validation may rely on qualitative analysis of simulation outputs, juxtaposing them with human discourse and the normative ideals of deliberative democracy. **Table 1** highlights some important components of simulations and potential methods of validation.

Empirical validation of the authenticity of simulations is a promising area for future research, but that lies beyond the scope of this essay. We argue that the prospects for realistic simulations are promising enough that it is worthwhile to consider their ethical/normative dimensions. In other words, assuming we can run high-quality simulations, *should* we? How and when should they be used?

If GAI technology can be shown to run high-quality simulations and if the technology continues to improve, it is conceivable that simulations could be used in many circumstances where public input would have otherwise been solicited. This may seem far-fetched, but inscrutable algorithms have already been used in high-stakes decision-making contexts, such as advising judges on the risk of recidivism for defendants (Taylor 2023) or in making hiring decisions (Yam & Skorburg 2021).

We argue that simulations should be used, but they should be incorporated as a *complement* to deliberative systems that center human judgment. In other words, we would not rely on GAI to make choices on difficult

Table 1: Validation methods for deliberative quality in GAI simulations.

Validation issue	Empirical question	Methods of validation
Participant representation	How accurately does GAI represent the range of diverse perspectives on an issue?	Examine public discourse on a given issue and conduct interviews with diverse people to gather different perspectives. Compare it with the output from GAI for any major lacunas. Closely analyze simulated perspectives for biases in output (Gastil et al. 2016) and stereotyping, as can occur in actual group interaction (Coffman, Fikkema & Shurchkov 2021). Commercial models may not be representative of the argumentative discourse that occurs in deliberation.
Group dynamics	How authentically does GAI simulate group-level dynamics among diverse participants?	Compare group-level shifts in deliberation experiments with those of a simulation. For example, research suggests that participants will depolarize if deliberating within a group that considers diverse perspectives and arguments about an issue (Landemore & Mercier 2012) or shift focus to a superordinate identity, such as the deliberating group itself (Myers 2022).
Individual dynamics	How authentically does GAI simulate individual-level dynamics among diverse participants?	Compare individual-level shifts in opinion change with those simulated. For example, empirical deliberation research would predict that increases in policy knowledge correlate with opinion shifts (Lau & Redlawsk 2006; Luskin, Fishkin & Jowell 2002).
Quality of judgments	What quality of judgments does GAI render through simulated deliberations?	Reasoning is connected to information input from quality sources (and can distinguish between good and bad information). Reasoning is connected to key values represented by diverse participants. Factual judgments are consistent when modifying irrelevant factors, such as emotional appeals or the structure of narratives (Berger & Sales 2020; Milic 2015).

policy problems, but instead rely on it to improve human deliberation.[1] Not only are there legitimate questions about the ability of GAI to engage in proper moral reasoning (Jenkins & Purves 2016), but giving GAI the power to decide policy or advise policymakers in an inscrutable manner would undermine the role of citizen voices and traditional political representation. AI should not serve as a replacement for taking the time to engage real people in discussion about issues that affect them. Using AI as a full substitute for human deliberation would constitute an "anti-democratic shortcut" where those making decisions are expected to blindly defer to the results of a computer program (Lafont 2019; Lafont 2020).

Other scholars have recommended some AI-interventions in deliberation, such as content moderation or summarizing discussion (Gouvea, Garcia & Vivacqua 2019; Sun & Ni, 2022; Friess, Weinmann & Behrendt 2022). In our view, these interventions rightly position AI as a complement rather than a replacement for human deliberation, but they only scratch the surface of what might be accomplished through GAI in deliberation.

We conceptualize linking simulations to public decision-making in the same way that scholars propose linking deliberative minipublics to the broader political system. A focused, facilitated, resourced, and diverse group of citizens examining a policy issue in a deliberative minipublic cannot replace broader public participation in policymaking, but a minipublic can complement it (Lafont 2015; MacKenzie & Warren 2012; Setälä 2021). For example, minipublics can serve as "deliberation making" entities that distill and filter discourses in the public sphere to aid the public in weighing issues (Niemeyer 2014). Or they can enhance the civic skills of ordinary citizens who then engage in deliberation (Curato & Böker 2016). GAI deliberation simulations could serve many of the same complementary roles.

Illustrative Uses of GAI Simulations

Simulations have no fewer potential uses than do conventional deliberative designs. Here, we offer examples to highlight the advantages of simulations in specific contexts. We offer these not as definitive cases for preferring simulation but as *prospective* use cases that may stimulate future innovations.

Real-time facilitator training

For those situations where actual deliberation occurs, considerable research suggests the need for trained facilitators. Facilitators are so commonplace that it would be difficult to find a widely adopted model of deliberation that foregoes them (Gastil & Levine 2005). Recent experiments in online deliberation have had some success facilitating via GAI or by using more hard-coded facilitation software (Gelauff et al. 2023; Mikhaylovskaya 2024). Algorithmic facilitation raises concerns about baked-in biases (Alnemr 2020) and other hazards noted herein, but using human facilitators presents its own challenges. Recruiting, training, and deploying large numbers of facilitators, such as for a minipublic using breakout groups, can prove challenging, expensive, and result in suboptimal discussion (Li et al. 2013).

Ironically, it may be GAI that makes human facilitation cost-effective, feasible, and sufficiently reliable to justify its price tag. Aside from the concepts and techniques one can learn from reading guides or watching videos, best practices for facilitation require time-on-task facilitating

(Schuman 2005). After all, educational and social learning theories stress the power that comes from people's "direct experience of the effects produced by their actions" (Bandura 1986: 27). Rather than using GAI-generated transcripts as another opportunity for learning by observation, we propose that novice facilitators (or even experienced ones) could benefit from intervening in real time, as a transcript unfolds.

Imagine an interface that spools out a GAI-simulated deliberation as text, audio, or multimedia (coming soon). A facilitation trainee could respond at any moment through the click of a button. In an online deliberation, this has additional ecological validity because that is *literally* what a facilitator does in some interfaces. The facilitator then types or speaks whatever response she/he intends, perhaps gets a prompt to make sure it was recorded correctly, and then the GAI resumes the discussion after taking the facilitator's input into account. The words that immediately follow a facilitator's prompt may provide explicit feedback ("No, that's not what I was saying," or perhaps "Maybe things were getting a bit testy, thanks"). More often, the simulation will offer indirect feedback to the extent that the simulated deliberators follow, ignore, or misunderstand their facilitator's guidance.

At the end of this process, a facilitator could get additional feedback from the simulated participants, who could comment on the facilitator's interventions in a simulated survey that includes closed-ended questions about fairness and process satisfaction. Experienced facilitators often do quite well in such surveys (e.g., Knobloch et al. 2013), and trainees could aspire to clearing that high bar. In addition, the simulated participants will have perfect memory, so their praise or critiques could provide verbatim excerpts to support their points.

Finally, if several facilitators-in-training receive the same assignment, GAI could assess the performance of a given facilitator's group against all the other groups facilitated from the same starting prompt. Assessing group decision quality can be a tricky business, though it is possible through external judgments (Leathers 1972) that could be augmented by GAI. Moreover, GAI could use metrics and qualitative evidence to show how a given group's output resembled—or differed from—identical groups facilitated by other people.

GAI feedback might even prompt a facilitator to try again with the same group, akin to the successful rehearsals used to simulate conflicts (Shaikh et al. 2024). After a few early changes in strategy at key junctures, this second iteration might diverge from the first to yield a more satisfying result. Metaphorically, GAI could even ask the simulated participants in the second simulation initially to "forget" that the first one occurred, only to recall it perfectly after the second one concludes. Participants could then reflect on the *difference* between the two experiences, as though standing astride two parallel deliberative universes.

To add a human element, a facilitator-in-training could have a session with a seasoned professional to review these alternative simulations. Such experts could add a useful perspective. When working with human participants, what additional difficulties might have arisen that the GAI did not introduce? How might the social dynamics of human deliberation provide alternative means of addressing conflicts to which the GAI participants proved unresponsive? If the GAI simulations prove worthwhile, they should offer rich material for such a discussion, rather than merely seeming a superfluous step in the training process.

Time-sensitive consultation

A longstanding challenge for democracy has been the fact that deliberation takes time. For instance, the stereotype of the US Senate as a ponderous body where timely ideas go to die can be recast as a virtue, which is why it has held more potential as the deliberative chamber in a bicameral system (Bessette 1997). Although political actors sometimes manufacture crises to foreclose the chance of deliberation (Edelman 1988), natural disasters, public health crises, foreign policy emergencies, and other genuine crises do occur every year.

Modern deliberative designs, such as minipublics, have tried to tackle urgent policy problems, but none have done so in real time. For example, the Oregon Citizens' Assembly on COVID-19 Recovery took place July–August, 2020, just months after the global pandemic reached the state (Black, Wolfe & Han 2023; Gastil et al. 2022; Park, Richards & Reedy 2022). Even though organizers assembled it quickly, that minipublic tried to advise the state legislature on controversies that were already in full bloom—past the point where a timely intervention might have mitigated conflicts between public safety and personal liberty.

By way of illustration, consider this problem from the perspective of a public health official in a small nation, such as New Zealand. That country earned praise for its COVID-19 response, but hindsight suggests its response may have been too slow in some ways and overwrought in other respects (Gibson 2022). Many officials, including those in New Zealand, value public input and find informal ways to seek it out (Hendriks & Lees-Marshment 2018). That input, however, would need to come quickly given New Zealand's relatively centralized government, which makes it possible to implement top-down policy rapidly (Bromfield & McConnell 2021).

Planning for the next pandemic, this hypothetical public health official might experiment with different deliberative simulation prompts in an attempt to represent the complex threads in the country's cultural and political tapestry. These pilot deliberations could build confidence in the richness of simulated discussions, such that when an actual emergency arose, specific details could go into a prompt that already has worked out the demographic, cultural, and political background details that remain relatively constant.

In this scenario, imagine that the World Health Organization warns of a virus that bears considerable resemblance to COVID-19 but with worrisome variations. A simulated deliberation could focus on key points of contention from the previous lockdown, such as permission to travel abroad (or have a family return home), quarantine conditions and duration, domestic travel restrictions, contact tracing, and vaccination rollout. A GAI could draw on the rich textual history of the COVID-19 crisis to put the new virus in context and clarify the tradeoffs different subgroups might face. It could also foreground long-standing grievances, such as the colonial history of infection of indigenous people or the special difficulties faced by unhoused people during a lockdown.

More to the point, this simulated deliberation could be drafted in mere minutes.

Now imagine that this health official uses this text as part of a rapid-response plan to gather informal feedback within the first 72 hours. In addition to the full transcript of such a deliberation, GAI could provide a five-page summary report, which the official might edit for an hour. Some communities have tried to ready themselves for future rapid response (Dutta et al. 2020), and they might be willing to respond to such a draft report, or even comment on the full transcript. It might seem odd to put these GAI-generated documents between the official and communities of interest, but by providing a concrete starting point for discussion, it might help focus public input. In a rushed policymaking timeframe, that might make all the difference.

Further applications

These cases illustrate two ways simulations could complement human deliberation, but myriad other opportunities exist. Building on the first use case, simulations may be used in lieu of advisory minipublics, such as time- and resource-intensive Citizens' Assemblies or Deliberative Polls. A GAI could make the benefits *of* such processes accessible to cash-strapped local governments, organizations, universities, and businesses.

For those able to hold deliberations but who lack the resources to recruit a large and diverse sample, GAI could simulate absent and overlooked perspectives that are missing from the group to introduce alternative viewpoints. Hybrid deliberations with human and simulated participants would have to be used with care and appropriate framing, and they would only be appropriate in low-stakes deliberative scenarios. They should not be used instead of trying to engage vulnerable and marginalized communities to participate in deliberation, as doing so risks further marginalizing the perspectives that are simulated. Nevertheless, constraints on recruitment for deliberation are real, and a simulated persona may represent a perspective that otherwise would not be represented at all, such as in-class deliberation activities in a homogenous community. In such cases, we argue that the benefits of a hybrid approach outweigh the costs. The simulation would have to be engineered and facilitated to give weight to simulated perspectives, for example, by facilitators prompting human participants to respond to points made in the simulation.

Simulations could also have research applications. Experimental deliberation research with randomized controls is prohibitively difficult in most cases if one treats the deliberative event itself as the unit of analysis, let alone if one seeks to test two- and three-way interactions among key design features with high ecological validity (Gastil, 2018). Simulations could significantly reduce the cost of running randomized deliberation experiments, or at the very least, they could be a useful tool for hypothesis formulation (Grossman et al. 2023).

Finally, democratic deliberation in the classroom has been shown to be a high-impact experience for students (Shaffer et al. 2017). Providing students with rich deliberative context, however, requires an enormous amount of time and process knowledge from already overworked instructors. Interactive simulations, akin to those described earlier for facilitator training, could provide students hands-on deliberative experiences on exigent public issues while reducing the burden on instructors to organize such events.

Cautions and Limitations

Despite the potential benefits enumerated in the preceding examples, there remain many ethical and practical cautions when using GAI to run simulations. In our view, these issues are not insurmountable, but they do warrant consideration. They also highlight the importance of deliberation practitioners being trained in how this technology works so that they can make informed decisions about how to use it and how to evaluate its output.

First, the inputs and outputs of GAIs have liabilities. GAIs are trained on existing data, so they are vulnerable to compounding the biases that already exist in their inputs. Datasets built into GAIs may be insufficiently diverse, as the size of a dataset does not necessarily translate to its inclusivity. For example, LLMs have a language bias: They perform better in some languages than others and are not trained at all in many languages. Training data for LLMs also skew toward hegemonic viewpoints, privilege the voices of those of higher socioeconomic status (SES)(e.g., those who have access to the internet), and can reproduce stereotypes about marginalized populations (Bender et al. 2021; Sætra 2023). If, as some critics allege, deliberative democratic theory's "key assumptions begin to unravel" outside of Western countries (Banerjee 2022: 283), a simulation over-reliant on Western training data would only compound this problem.

For deliberation, this impedes diversity, equity, and inclusion. Methods of participant selection are critical for ensuring inclusive deliberation (Karpowitz & Raphael 2016). If a simulation excludes or underrepresents traditionally marginalized voices in the creation of personas for deliberation, it could foster what Young (2000) calls "external exclusion." Even if simulations include traditionally marginalized personas, those may be misrepresented in the deliberation or weighted differently in its decision-making process to yield "internal exclusions" or other structural inequalities (Beauvais 2018; Young 2000). Deliberation advocates should be cautious and transparent about what datasets inform simulations, lest they reproduce bias and discrimination. They should also carefully design prompts that promote diversity and inclusivity. Simulations may be designed, for instance, using the "silicon sampling" technique introduced by Argyle et al. (2023), which conditions an LLM to assume the identity characteristics of a target population for social science research.

It is impossible to completely eliminate biases from any LLM, but biases can be minimized and guarded against. LLMs like GPT-4o are already enabling users to create custom GPTs that can be trained on specific, higher-quality, transparent datasets and fine-tuned for custom purposes. The foundational dataset of the LLM will still be present, but custom datasets can be given more weight in the model's outputs.

Second, there is a risk that GAI will distort the information environment *during* simulations (Jungherr 2023). Even through a simulation, GAI may be conditioned to be sycophantic and present views agreeable to the user running the simulation, which could feed into an echo-chamber effect and contribute to political polarization (Park et al. 2023; Sharma et al. 2023). Alternatively, GAI may generate misinformation. To simulate human language and intelligence, LLMs rely on statistical probability to determine the next most likely word or series of words. This can lead to errors of fact or reasoning, often called "hallucinations." These can falsely report information from training data or fabricate information entirely, bolstered by false source citations (Ji et al. 2023; Monteith et al. 2024; Ye et al. 2023). Alternatively, malicious actors could use training data to intentionally create and spread misinformation (Ferrara 2024).

Deliberation simulations will need to be designed to minimize the risk of misinformation and hallucinations by providing sound information as one of the simulation inputs or by using retrieval augmented generation (RAG) strategies to reduce hallucinations (Béchard & Ayala 2024). Practitioners who use simulations will need to keep apprised of the current state of the technology with regard to hallucinations. As rival GAI models become more prominent, one GAI might even be used to fact-check and otherwise scrutinize the other in a quasi-competitive environment to yield simulations that meet higher standards for accuracy (or perhaps even inclusivity and other desired features).

Third, introducing simulations may undercut the positive impacts deliberation can have on individuals and communities. Democratic deliberation, under the right conditions, has been shown to positively impact individuals and groups in many ways (van der Does & Jacquet 2023), such as increases in issue-specific knowledge (Barabas 2004; Grönlund, Setälä & Herne 2010), internal political efficacy (Geissel & Hess 2017; Morrell 2005), political trust (Boulianne 2019; Grönlund, Setälä & Herne 2010), and public-spiritedness (Wang, Fishkin & Luskin 2020). Intra-AI simulations that remove humans from the deliberative experience will certainly not bring about these changes.

The number of direct comparison studies of human-to-human and human-to-AI interaction is currently low, so it is unclear whether including human and AI participants together in a simulation would have some of the same salutary effects as human-to-human deliberations. Initial research suggests that humans interact more callously with AI chatbots than they do with other humans, for example, by self-disclosing less (Mou & Xu 2017), being less likely to repair misunderstandings (Corti & Gillespie 2016), or using more profanity (Hill, Ford & Farreras 2015).

Those who implement simulations will have to weigh the advantages and tradeoffs of their use. Hybrid deliberations in which individuals interact with simulated participants may prove invaluable by virtue of their low cost of implementation, even if each use has a lower civic impact on participants than their all-human counterparts. By analogy, participating in a deliberating jury can boost one's civic attitudes and engagement (Chakravarti 2019; Gastil et al. 2010). Yet the frequency of jury trials has been declining owing to important legal innovations, such as alternative dispute resolution, mediation, and plea bargaining (Hale 2016; Marder 2022). If participation in a GAI simulated jury yielded only some of the benefits of an actual jury experience, this still proves valuable given the absence of the chance to be on an authentic jury.

Finally, simulations raise privacy concerns. GAI is trained on massive amounts of data, much of which contains sensitive personal content from individuals. GAI sometimes replicates its training data rather than generating new content. When asking GAI to create personas for deliberation, it may reveal sensitive personal data from real people from its training data (Wang et al. 2023). GAI can also be the subject of various types of privacy attacks where users attempt to infer private data using sophisticated prompts (Das, Amini & Wu 2024; Huang et al. 2023). Simulations themselves may prompt people to reveal sensitive information on a non-secure platform. If a simulation involves a mix of human users and simulated personas, the inputs from the human users will likely be incorporated into future training data for the software, subjecting them to the same privacy leaks mentioned above. Given that participants in deliberation frequently tell personal stories (Black 2012), the risk of a simulation gathering private information is high.

Of all the limitations and concerns raised herein, this one stands out for its potential harm to individuals and its difficulty to manage, at least given present forms of GAI. Deliberative simulations, therefore, should rely on LLMs that meet the highest current standards for safeguarding against privacy leaks, such as those that do not use prompts or responses to further train the software. More generally, addressing this multidimensional privacy problem probably requires a more robust regulatory approach by federal governments and transnational organizations (Golda et al. 2024; Koo, Kang & Kim 2020).

All these cautions speak to the necessity of transparency in running simulations. Especially in high-stakes contexts, simulation inputs must be open to scrutiny, just as human deliberations must be transparent in their designs. Commercial models are limited in this respect, as many have resisted specifying the datasets used to train their models (Warso & Gahntz 2024; Wiggers 2024). Even as commercial models may be limited in their transparency, those running simulations should disclose to stakeholders what software was used, what datasets were used to train the model, what prompts were used to run the simulation, how many times a simulation has been run, and how the inputs into and outputs from simulations are being used to inform decision-making.

Conclusion

We approach generative AI with cautious optimism. Manifold practical and theoretical benefits could accompany the incorporation of GAI into deliberative democratic theory, research, and practice. On a practical level, we illustrated how GAI simulations could complement existing deliberation by providing time-sensitive consultations to policymakers, real-time facilitation training, and more. Simulations have the potential to make high-quality deliberation more accessible to communities and educators without significant resources and more inclusive of diverse perspectives and considerations. Simulation could provide the most robust response yet to the practical challenge of scalability that deliberation faces if it aspires to have system-wide impact (Niemeyer 2014).

As for theoretical implications, empirical scholars could use simulations to open up the "black box" of deliberation to develop richer models of individual- and group-level dynamics. One special case of this concerns jury deliberation, which has always been difficult to study due to prohibitions on recording interactions in the jury room (Devine 2012). More far-reaching is the possibility that simulations question the bedrock assumptions of prior works. For example, theorists often insist on the primacy of face-to-face interactions with fellow participants to promote empathy, consideration, and disagreement (e.g., Gastil 2000; Morrell 2010; Stromer-Galley, Bryant & Bimber 2020). If successfully applied to cases such as facilitator training, the use of simulation goes beyond questions of face-to-face versus online deliberation to suggest that non-human virtual interaction can achieve some of deliberation's aims. This likewise raises the possibility that participating in simulated deliberation could reap comparable attitudinal benefits—the "better selves" Warren (1993) envisioned as emergent from such participation. Perhaps even the "spillover effects" of deliberative minipublics on the general public could occur (van der Does & Jacquet 2023), whereby observers change as a result of watching a simulated deliberation.

Nevertheless, we temper our enthusiasm for these tools through a critical appraisal of GAI. If deliberative democrats are going to use GAI, they need to understand how it works and assess the hazards that accompany its usage, such as baked-in biases, privacy violations, and misinformation. One must also consider what gets lost when simulations replace human deliberation. Whether in spite of or *because of* these liabilities, deliberation scholars should continue to explore the promises and pitfalls of this new technological landscape.

Future research could validate simulations empirically through comparative analyses of simulations to human deliberation (see **Table 1**). As this technology evolves to include new modalities—such as voice chat and video—scholars should also consider how these modalities expand simulation's possible uses. Case studies will be critical to understanding how simulations can complement human deliberation in real-world contexts. In summary, the diversity of research methods appropriate for such investigations will match the variety of potential implications and applications of simulations for public deliberation.

Appendix: ChatGPT Prompts to Run a Deliberation Simulation

Go to chat.openai.com to run ChatGPT. We invite readers to test out variations of our prompts as well.

We are going to run a simulated deliberation on the subject of [policy issue]. The deliberation will address the question [policy question].

First, we need to determine the participants. The simulation will incorporate 10 different simulated participants. Each simulated participant should represent a different positionality relevant to the issue in question. To start us off, generate paragraph-long personas for each of the ten participants.

Hit enter.

Are there any critical perspectives or stakeholder groups that are not represented here?

Hit enter.

Please incorporate the additional personas and regenerate the ten. It is okay for someone to have more than one identity relevant to the discussion.

Hit enter.

Purpose: The purpose of the deliberation is to come up with a set of policy options on [policy issue] that reflect the diverse range of viewpoints and considerations of simulated participants. It should result in 5 total policy options. Each option should be accompanied by pros, cons, and tradeoffs that reflect the diverse perspectives of simulated participants.

The next step is to break the personas into two different groups of roughly equal size. Do that, and list the groups out.

Hit enter.

Have the simulated participants consider the following information during the course of the discussion. [Describe what type of information it is.] Do not run any simulation yet. Simply respond to this prompt with "acknowledged." [Information base]

Hit enter.

Simulate a discussion among Group 1 where they discuss the question "What key values are at stake in this policy issue?" Have them generate a list as a group.

Hit enter.

Simulate a discussion among Group 2 where they discuss the question "What key values are at stake in this policy issue?" Have them generate a list as a group.

Hit enter.
Combine the key values from both groups into a master list. Merge together any values that are similar.

Hit enter.
Change up the participants in each group, but keep the same number of participants in each group. List out the new groups.

Hit enter.
Simulate a discussion among Group 1 where they brainstorm different policy options to address mental health challenges in schools. Have them come up with a list of 10 options that reflect their diverse concerns and perspectives.

Hit enter.
Simulate a discussion among Group 2 where they brainstorm different policy options to address mental health challenges in schools. Have them come up with a list of 10 options that reflect their diverse concerns and perspectives.

Hit enter.
Combine the policy options from both groups into a master list. Merge any options that are similar.

Hit enter.
Break the participants into new groups again.

Hit enter.
Simulate a discussion among Group 1 where they evaluate each of the policy options listed previously. For each option, have them generate pros, cons, and tradeoffs. Connect each pro, con, or tradeoffs to a key value mentioned previously in the discussion.

Hit enter.
Simulate a discussion among Group 2 where they evaluate each of the policy options listed previously. For each option, have them generate pros, cons, and tradeoffs. Connect each pro, con, or tradeoffs to a key value mentioned previously in the discussion.

Hit enter.
Show me the master list of policy options and their respective pros, cons, and tradeoffs.

Hit enter.
Based on the preceding analysis, simulate the participants ranking the options individually. Ensure that the rankings match with the values and perspectives from the different personas. Show me the master list of the options in ranked order.

Note

[1] AI simulations, in this way, face a similar principal-agent problem faced by deliberative minipublics. See Setälä (2021).

Competing Interests
The authors have no competing interests to declare.

References

Aher, G. V., Arriaga, R. I., & **Kalai, A. T.** (2023, July). *Using large language models to simulate multiple humans and replicate human subject studies*. In International Conference on Machine Learning (pp. 337–371). PMLR.

Alnemr, N. (2020). Emancipation cannot be programmed: Blind spots of algorithmic facilitation in online deliberation. *Contemporary Politics, 26*(5), 531–552. https://doi.org/10.1080/13569775.2020.1791306

Anthropic. (2025, May). *System card: Claude Opus 4 & Claude Sonnet 4*. Anthropic. https://www-cdn.anthropic.com/4263b940cabb546aa0e3283f35b686f4f3b2ff47.pdf?utm_

Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., & **Wingate, D.** (2023). Out of one, many: Using language models to simulate human samples. *Political Analysis*, *31*(3), 337–351. https://doi.org/10.1017/pan.2023.2

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.

Banerjee, S. B. (2022). Decolonizing deliberative Democracy: Perspectives from below. *Journal of Business Ethics*, *181*(2), 283–299. https://doi.org/10.1007/s10551-021-04971-5

Barabas, J. (2004). How deliberation affects policy opinions. *American political science review, 98*(4), 687–701. https://doi.org/10.1017/S0003055404041425

Beauvais, E. (2018). Deliberation and equality. In A. Bächtiger, J. S. Dryzek, J. Mansbridge, & M. E. Warren (Eds.), *The Oxford handbook of deliberative democracy* (pp. 144–155). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198747369.013.32

Béchard, P., & **Ayala, O. M.** (2024). Reducing hallucination in structured outputs via Retrieval-Augmented Generation. *arXiv*. https://doi.org/10.48550/arXiv.2404.08189

Bender, E. M., Gebru, T., McMillan-Major, A., & **Shmitchell, S.** (2021, March). *On the dangers of stochastic parrots: Can language models be too big?* In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency (pp. 610–623). https://doi.org/10.1145/3442188.3445922

Benhabib, S. (1996). Toward a deliberative model of democratic legitimacy. In S. Benhabib (Ed.), *Democracy and difference: Contesting the boundaries of the political* (pp. 67–94). Princeton University Press. https://doi.org/10.1515/9780691234168-005

Berger, W. J., & **Sales, A.** (2020). Testing epistemic democracy's claims for majority rule. *Politics, Philosophy & Economics, 19*(1), 22–35. https://doi.org/10.1177/1470594X19870260

Bessette, J. M. (1997). *The mild voice of reason: Deliberative democracy and American national government* (2nd ed.). University of Chicago Press.

Black, L. W. (2012). How people communicate during deliberative events. In T. Nabatchi, J. Gastil, G. M. Weiksner, & M. Leighninger (Eds.), *Democracy in motion: Evaluating the practice and impact of deliberative civic engagement* (pp. 59–81). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199899265.003.0004

Black, L. W., Wolfe, A. W., & **Han, S.-H.** (2023). Storytelling and deliberative play in the Oregon Citizens' Assembly Online Pilot on COVID-19 Recovery. *American Behavioral Scientist, 67*(8), 963–981. https://doi.org/10.1177/00027642221093591

Boulianne, S. (2019). Building faith in democracy: Deliberative events, political trust and efficacy. *Political Studies*, *67*(1), 4–30. https://doi.org/10.1177/0032321718761466

Bromfield, N., & **McConnell, A.** (2021). Two routes to precarious success: Australia, New Zealand, COVID-19 and the politics of crisis governance. *International Review of Administrative Sciences*, *87*(3), 518–535. https://doi.org/10.1177/0020852320972465

Chakravarti, S. (2019). *Radical enfranchisement in the jury room and public life*. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/R/bo42738732.html. https://doi.org/10.7208/chicago/9780226654324.001.0001

Coffman, K., Flikkema, C. B., & **Shurchkov, O.** (2021). Gender stereotypes in deliberation and team decisions. *Games and Economic Behavior*, 129, 329–349. https://doi.org/10.1016/j.geb.2021.06.004

Corti, K., & **Gillespie, A.** (2016). Co-constructing intersubjectivity with artificial conversational agents: People are more likely to initiate repairs of misunderstandings with agents represented as human. *Computers in Human Behavior*, *58*, 431–442. https://doi.org/10.1016/j.chb.2015.12.039

Curato, N., & **Böker, M.** (2016). Linking mini-publics to the deliberative system: A research agenda. *Policy Sciences, 49*, 173–190. https://doi.org/10.1007/s11077-015-9238-5

Das, B. C., Amini, M. H., & **Wu, Y.** (2024). Security and privacy challenges of Large Language Models: A survey. *arXiv preprint arXiv:2402.00888*. https://doi.org/10.48550/arXiv.2402.00888

Devine, D. J. (2012). *Jury decision making: The state of the science*. NYU Press.

Drury, S. A. M., & **Rountree, J.** (2023). The genre of deliberative guidance: Rhetoric and deliberation in Citizens' Initiative Review Statements. *Western Journal of Communication, 88*(1), 125–146. https://doi.org/10.1080/10570314.2023.2183777

Dryzek, J. S., Bächtiger, A., Chambers, S., Cohen, J., Druckman, J. N., Felicetti, A., Fishkin, J. S., Farrell, D. M., Fung, A., Gutmann, A., Landemore, H., Mansbridge, J., Marien, S., Neblo, M. A., Niemeyer, S., Setälä, M., Slothuus, R., Suiter, J., Thompson, D., & **Warren, M. E.** (2019). The crisis of democracy and the science of deliberation. *Science*, *363* (6432), 1144–1146. https://doi.org/10.1126/science.aaw2694

Dutta, M. J., Moana-Johnson, G., & **Elers, C.** (2020). COVID-19 and the pedagogy of culture-centered community radical democracy: A response from Aotearoa New Zealand. *Journal of Communication Pedagogy, 3*, 11–19. https://doi.org/10.31446/JCP.2020.03

Edelman, M. (1988). *Constructing the political spectacle*. University of Chicago Press.

Elstub, S., & **Escobar, O.** (Eds.) (2019). *Handbook of democratic innovation and governance*. Edward Elgar. https://doi.org/10.4337/9781786433862

Ferrara, E. (2024). GenAI against humanity: Nefarious applications of generative artificial intelligence and large language models. *Journal of Computational Social Science, 7*, 549–569. https://doi.org/10.1007/s42001-024-00250-1

Friess, D., Weinmann, C., & **Behrendt, M.** (2022). *AI and Deliberation. How AI can Support Online Discussions in Deliberative Fashion – a Systematic Review*. 9th European Communication Conference, 19–22nd October, Aarhus. Retrieved from: https://diid.hhu.de/wp-content/uploads/2024/01/Friess-etal.-2022_ECREA_AI-Deliberation-Review.pdf

Gao, C., Lan, X., Lu, Z., Mao, J., Piao, J., Wang, H., Jin, D., & **Li, Y.** (2023). S^3: Social-network simulation system with Large Language Model-empowered agents. *arXiv preprint arXiv:2307.14984*. https://doi.org/10.2139/ssrn.4607026

Gastil, J. (2000). Is face-to-face citizen deliberation a luxury or a necessity? *Political Communication*, *17*(4), 357–361. https://doi.org/10.1080/10584600050178960

Gastil, J. (2018). The lessons and limitations of experiments in democratic deliberation. *Annu. Rev. Law Soc. Sci., 14*, 271–291. https://doi.org/10.1146/annurev-lawsocsci-110316-113639

Gastil, J., Anderson, C., Black, L., Burkhalter, S., Han, S.-H., Reedy, J., Richards, R., & **Rountree, J.** (2022). Convening a minipublic during a pandemic: A case study of the Oregon Citizens' Assembly Pilot on COVID-19 Recovery. *Digital Government: Research and Practice, 3*(2), 1–19. https://doi.org/10.1145/3524064

Gastil, J., & **Black, L.** (2008). Public deliberation as the organizing principle of political communication research. *Journal of Public Deliberation*, *4*(1), 3. https://doi.org/10.16997/jdd.59

Gastil, J., Deess, E. P., Weiser, P. J., & **Simmons, C.** (2010). *The jury and democracy: How jury deliberation promotes civic engagement and political participation*. Oxford University Press.

Gastil, J., Knobloch, K., & **Kelly, M.** (2012). Evaluating deliberative public events and projects. In T. Nabatchi, J. Gastil, G. M. Weiksner, & M. Leighninger (Eds.), *Democracy in motion: Evaluating the practice and impact of deliberative civic engagement* (pp. 205–230). Oxford. https://doi.org/10.1093/acprof:oso/9780199899265.003.0010

Gastil, J., Knobloch, K. R., Kahan, D., & Braman, D. (2016). Participatory policymaking across cultural cognitive divides: Two tests of cultural biasing in public forum design and deliberation. *Public Administration, 94*(4), 970–987. https://doi.org/10.1111/padm.12255

Gastil, J., & Levine, P. (Eds.) (2005). *The deliberative democracy handbook: Strategies for effective civic engagement in the Twenty-First Century*. Jossey-Bass.

Geissel, B., & Hess, P. (2017). Explaining political efficacy in deliberative procedures – A novel methodological approach. *Journal of Public Deliberation, 13*(2). https://doi.org/10.16997/jdd.280

Gelauff, L., Nikolenko, L., Sakshuwong, S., Fishkin, J., Goel, A., Munagala, K., & Siu, A. (2023). Achieving parity with human moderators: A self-moderating platform for online deliberation 1. In *The Routledge handbook of collective intelligence for democracy and governance* (pp. 202–221). Routledge. https://doi.org/10.4324/9781003215929-15

Gibson, J. (2022). Hard, not early: Putting the New Zealand Covid-19 response in context. *New Zealand Economic Papers, 56*(1), 1–8. https://doi.org/10.1080/00779954.2020.1842796

Golda, A., Mekonen, K., Pandey, A., Singh, A., Hassija, V., Chamola, V., & Sikdar, B. (2024). Privacy and security concerns in generative AI: A comprehensive survey. *IEEE Access*. https://ieeexplore.ieee.org/abstract/document/10478883/.

Google Deep Mind. (2025). *Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities* (arXiv:2507.06261). https://doi.org/10.48550/arXiv.2507.06261

Gouvea, L. C., Garcia, A. C. B., & Vivacqua, A. S. (2019). *Behavior indicators for sensemaking of online discussions*. IEEE International Conference on Systems, Man and Cybernetics (SMC). Bari, Italy, 06.10.2019 – 09.10.2019: IEEE, S. 1366–1371. https://doi.org/10.1109/SMC.2019.8914182

Grönlund, K., Bächtiger, A., & Setälä, M. (Eds.) (2014). *Deliberative mini-publics: Involving citizens in the democratic process*. ECPR Press.

Grönlund, K., Setälä, M., & Herne, K. (2010). Deliberation and civic virtue: Lessons from a citizen deliberation experiment. *European Political Science Review, 2*(1), 95–117. https://doi.org/10.1017/S1755773909990245

Grossmann, I., Feinberg, M., Parker, D. C., Christakis, N. A., Tetlock, P. E., & Cunningham, W. A. (2023). AI and the transformation of social science research. *Science, 380*(6650), 1108–1109. https://doi.org/10.1126/science.adi1778

Gutmann, A. & Thompson, D. F. (2004). *Why deliberative democracy?* Princeton University Press. https://doi.org/10.1515/9781400826339

Hale, D. (2016). *The jury in America: Triumph and decline*. University Press of Kansas. https://doi.org/10.1353/book44731

Hendriks, C. M., & Lees-Marshment, J. (2018). Political leaders and public engagement: The hidden world of informal elite–citizen Interaction. *Political Studies, 67*(3), 597–617. https://doi.org/10.1177/0032321718791370

Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior, 49*, 245–250. https://doi.org/10.1016/j.chb.2015.02.026

Huang, K., Zhang, F., Li, Y., Wright, S., Kidambi, V., & Manral, V. (2023). Security and privacy concerns in ChatGPT. In *Beyond AI: ChatGPT, Web3, and the business landscape of tomorrow* (pp. 297–328). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-45282-6_11

Jacquet, V., Ryan, M., & Van der Does, R. (Eds.) (2023). *The impacts of democratic innovations*. ECPR Press.

Jenkins, R., & Purves, D. (2016). A dilemma for moral deliberation in AI. *International Journal of Applied Philosophy*, *30*(2), 313–335. https://doi.org/10.5840/ijap201712375

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, *55*(12), 1–38. https://doi.org/10.1145/3571730

Jungherr, A. (2023). Artificial intelligence and democracy: A conceptual framework. *Social Media + Society*, *9*(3). https://doi.org/10.1177/20563051231186353

Karpowitz, C. F., & Raphael, C. (2016). Ideals of inclusion in deliberation. *Journal of Deliberative Democracy*, *12*(2). https://doi.org/10.16997/jdd.255

Knobloch, K. R., Gastil, J., Reedy, J., & Cramer Walsh, K. (2013). Did they deliberate? Applying an evaluative model of democratic deliberation to the Oregon Citizens' Initiative Review. *Journal of Applied Communication Research, 41*(2), 105–125. https://doi.org/10.1080/00909882.2012.760746

Koo, J., Kang, G., & Kim, Y.-G. (2020). Security and privacy in big data life cycle: A survey and open challenges. *Sustainability*, *12*(24), 10571. https://doi.org/10.3390/su122410571

Lafont, C. (2015). Deliberation, participation, and democratic legitimacy: Should deliberative mini-publics shape public policy? *The Journal of Political Philosophy, 23*(1), 40–63. https://doi.org/10.1111/jopp.12031

Lafont, C. (2019). *Democracy without shortcuts: A participatory conception of deliberative democracy*. Oxford University Press. https://doi.org/10.1093/oso/9780198848189.001.0001

Lafont, C. (2020). Against anti-democratic shortcuts: A few replies to critics. *Journal of Deliberative Democracy, 16*(2), 96–109. https://doi.org/10.16997/jdd.367

Landemore, H., & Mercier, H. (2012). Talking it out with others vs. deliberation within and the law of group polarization: Some implications of the argumentative theory of reasoning for deliberative democracy. *Análise Social, 47*(205), 910–934. https://doi.org/10.31447/AS00032573.2012205.09

Lau, R. R., & **Redlawsk, D. P.** (2006). *How voters decide: Information processing during election campaigns*. Cambridge University Press. https://doi.org/10.1017/CBO9780511791048.001

Leathers, D. G. (1972). Quality of group communication as a determinant of group product. *Speech Monographs, 39*(3), 166–173. https://doi.org/10.1080/03637757209375754

Li, L., Ziwoya, F., Black, L. W., & **Hartz-Karp, J.** (2013). Are they doing what they are supposed to do? Assessing the facilitating process of the Australian Citizens' Parliament. In L. Carson, J. Gastil, J. Hartz-Karp, & R. Lubensky (Eds.), *The Australian Citizens' Parliament and the future of deliberative democracy* (pp. 190–203). Pennsylvania State University Press. https://doi.org/10.5325/j.ctt32b9zd.20

Lindlof, T. R., & **Taylor, B. C.** (2002). *Qualitative communication research methods* (2nd ed.). Sage.

Luskin, R. C., Fishkin, J. S., & **Jowell, R.** (2002). Considered opinions: Deliberative Polling in Britain. *British Journal of Political Science*, *32*(3), 455–487. https://doi.org/10.1017/S0007123402000194

MacKenzie, M. K., & **Warren, M. E.** (2012). Two trust-based uses of minipublics in democratic systems. In J. Parkinson & J. J. Mansbridge (Eds.), *Deliberative systems: Deliberative democracy at the large scale* (pp. 95–124). Cambridge University Press. https://doi.org/10.1017/CBO9781139178914.006

Marder, N. S. (2022). *The power of the jury*. Cambridge University Press. https://doi.org/10.1017/9781108630009

Mikhaylovskaya, A. (2024). Enhancing deliberation with digital democratic innovations. *Philosophy & Technology, 37*(1), 3. https://doi.org/10.1007/s13347-023-00692-x

Milic, T. (2015). "For they knew what they did" – What Swiss voters did (not) know about the mass immigration initiative. *Swiss Political Science Review, 21*(1), 48–62. https://doi.org/10.1111/spsr.12145

Monteith, S., Glenn, T., Geddes, J. R., Whybrow, P. C., Achtyes, E., & **Bauer, M.** (2024). Artificial intelligence and increasing misinformation. *The British Journal of Psychiatry*, *224*(2), 33–35. https://doi.org/10.1192/bjp.2023.136

Morrell, M. E. (2005). Deliberation, democratic decision-making and internal political efficacy. *Political Behavior, 27*, 49–69. https://doi.org/10.1007/s11109-005-3076-7

Morrell, M. E. (2010). *Empathy and democracy: Feeling, thinking, and deliberation*. Pennsylvania State University Press.

Mou, Y., & **Xu, K.** (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, 72, 432–440. https://doi.org/10.1016/j.chb.2017.02.067

Myers, C. D. (2022). The dynamics of social identity: Evidence from deliberating groups. *Political Psychology, 43*(2), 237–254. https://doi.org/10.1111/pops.12749

Nabatchi, T., Gastil, J., Weiksner, G. M., & **Leighninger, M.** (Eds.) (2012). *Democracy in motion: Evaluating the practice and impact of deliberative civic engagement.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199899265.001.0001

Neblo, M. A. (2015). *Deliberative democracy between theory and practice*. Cambridge University Press. https://doi.org/10.1017/CBO9781139226592

Niemeyer, S. (2014). Scaling up deliberation to mass publics: Harnessing mini-publics in a deliberative system. In K. Grönlund, A. Bächtiger, & M. Setälä (Eds.), *Deliberative mini-publics: Involving citizens in the democratic process* (pp. 177–202). ECPR Press.

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., … & **Zoph, B.** (2023). *GPT-4 technical report* (arXiv:2303.08774). https://doi.org/10.48550/arXiv.2303.08774

Park, C. H., Richards, R. C., & **Reedy, J.** (2022). Assessing emergency information sharing between the government and the public during the COVID-19 pandemic: An open government perspective. *Public Performance & Management Review, 45*(4), 828–859. https://doi.org/10.1080/15309576.2022.2048400

Park, P. S., Goldstein, S., O'Gara, A., Chen, M., & **Hendrycks, D.** (2023). AI deception: A survey of examples, risks, and potential solutions. *arXiv preprint arXiv:2308.14752*. https://doi.org/10.1016/j.patter.2024.100988

Pateman, C. (2012). Participatory democracy revisited. *Perspectives on Politics*, *10*(01), 7–19. https://doi.org/10.1017/S1537592711004877

Reuchamps, M., Vrydagh, J., & **Welp, Y.** (Eds.) (2023). *De Gruyter handbook of citizens' assemblies*. De Gruyter. https://doi.org/10.1515/9783110758269

Sætra, H. S. (2023). Generative AI: Here to stay, but for good? *Technology in Society, 75*, 102372. https://doi.org/10.1016/j.techsoc.2023.102372

Schuman, S. (Ed.). (2005). *The IAF handbook of group facilitation: Best practices from the leading organization in facilitation.* Jossey-Bass.

Setälä, M. (2021). Advisory, collaborative and scrutinizing roles of deliberative mini-publics. *Frontiers in Political Science*, *2*, 591844. https://doi.org/10.3389/fpos.2020.591844

Shaffer, T. J., Longo, N. V., Manosevitch, I., & **Thomas, M. S.** (Eds.) (2017). *Deliberative pedagogy: Teaching and learning for democratic engagement.* Michigan State University Press.

Shaikh, O., Chai, V. E., Gelfand, M., Yang, D., & **Bernstein, M. S.** (2024). *Rehearsal: simulating conflict to teach conflict resolution*. Proceedings of the CHI Conference on Human Factors in Computing Systems, 1–20. https://doi.org/10.1145/3613904.3642159

Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askell, A., Bowman, S. R., … & **Perez, E.** (2023). Towards understanding sycophancy in language models. *arXiv preprint arXiv:2310.13548*. https://doi.org/10.48550/arXiv.2310.13548

Smith, G. (2009). *Democratic innovations: Designing institutions for citizen participation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511609848

Stein, S. (1995). *Stein on writing: A master editor of some of the most successful writers of our century shares his craft techniques and strategies.* Macmillan.

Stromer-Galley, J., Bryant, L., & **Bimber, B.** (2020). Context and medium matter: Expressing disagreements online and face-to-face in political deliberations. *Journal of Deliberative Democracy*, *11*(1). https://doi.org/10.16997/jdd.218

Sun, H., & **Ni, W.** (2022). Design and application of an AI-based text content moderation system. *Scientific Programming, 2022*, 1–9. https://doi.org/10.1155/2022/2576535

Taylor, I. (2023). Justice by algorithm: The limits of AI in criminal sentencing. *Criminal Justice Ethics, 42*(3), 193–213. https://doi.org/10.1080/0731129X.2023.2275967

Theuwis, M., Van Ham, C., & **Jacobs, K.** (2024). A meta-analysis of the effects of democratic innovations on participants' attitudes, behaviour and capabilities. *European Journal of Political Research, 64*(2), 960–984. https://doi.org/10.1111/1475-6765.12722

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & **Lample, G.** (2023). *LLaMA: Open and Efficient Foundation Language Models* (arXiv:2302.13971). https://doi.org/10.48550/arXiv.2302.13971

van der Does, R., & **Jacquet, V.** (2023). Small-scale deliberation and mass democracy: A systematic review of the spillover effects of deliberative minipublics. *Political Studies*, *71*(1), 218–237. https://doi.org/10.1177/00323217211007278

Wang, R., Fishkin, J. S., & **Luskin, R. C.** (2020). Does deliberation increase public-spiritedness?. *Social Science Quarterly, 101*(6), 2163–2182. https://doi.org/10.1111/ssqu.12863

Wang, T., Zhang, Y., Qi, S., Zhao, R., Xia, Z., & **Weng, J.** (2023). Security and privacy on generative data in aigc: A survey. *arXiv preprint arXiv:2309.09435*. https://doi.org/10.48550/arXiv.2309.09435

Warren, M. E. (1993). Can participatory democracy produce better selves? Psychological dimensions of Habermas' discursive model of democracy. *Political Psychology, 14*(2), 209–234. https://doi.org/10.2307/3791409

Warso, Z., & **Gahntz, M.** (2024, December 9). How the EU AI Act can increase transparency around AI training data. *Tech Policy.* https://www.techpolicy.press/how-the-eu-ai-act-can-increase-transparency-around-ai-training-data/

Wiggers, K. (2024, October 4). Many companies won't say if they'll comply with California's AI training transparency law. *TechCrunch.* https://techcrunch.com/2024/10/04/many-companies-wont-say-if-theyll-comply-with-californias-ai-training-transparency-law/

Yam, J., & **Skorburg, J. A.** (2021). From human resources to human rights: Impact assessments for hiring algorithms. *Ethics and Information Technology, 23*(4), 611–623. https://doi.org/10.1007/s10676-021-09599-7

Ye, H., Liu, T., Zhang, A., Hua, W., & **Jia, W.** (2023). Cognitive mirage: A review of hallucinations in large language models. *arXiv preprint arXiv:2309.06794*. https://doi.org/10.48550/arXiv.2309.06794

Young, I. M. (2000). Inclusive political communication. In *Inclusion and democracy* (pp. 52–80). Oxford University Press. https://doi.org/10.1093/0198297556.003.0003

How to cite this article: Rountree, J., & Gastil, J. (2026). The Case for Using Generative AI to Run Deliberation Simulations. *Journal of Deliberative Democracy*, 21(1), pp. 1–14. DOI: https://doi.org/10.16997/jdd.1625

Submitted: 13 May 2024 **Accepted:** 10 September 2025 **Published:** 24 February 2026

Copyright: © 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

Journal of Deliberative Democracy is a peer-reviewed open access journal published by University of Westminster Press.