UCLA School of Law

Law & Philosophy Program

Legal Theory Workshop
UCLA School of Law

Lowry Pressly

UCLA Law

"Talk to a Human . . .Speak to a Human . . . Talk to a [presses zero, presses zero, presses zero] Speak with a—: On AI Agents as Partners in Democratic Deliberation"

Thursday, February 19, 2026 3:20pm – 5:20pm
Law School Room 1314

Talk to a Human . . .Speak to a Human . . . Talk to a [*presses zero, presses zero, presses zero*]

Speak with a—: On AI Agents as Partners in Democratic Deliberation

Lowry Pressly

[Draft, January 29, 2025]

Comments & critiques welcome: pressly@stanford.edu

> *Could it ever occur to anyone to pray to an image on screen? To go on an online pilgrimage?*
>
> - Antón Barba-Kay (2023, 53)

We often hear about the dangers that generative artificial intelligence (AI) poses to democracy. New powers of precisely targeted misinformation, the easy production of realistic-seeming public participation on matters of public concern, and the spread of deep-fake pictures and videos all seem poised to undermine the function and stability of democracies across virtually all dimensions of import to political theorists. Generative AI threatens to deceive and polarize citizens (Chesney and Citron 2019; Liu et al. 2024); to pollute democracy's epistemic function (Rini 2020); to further consolidate power in the hands of an oligarchic few (Allen and Weyl 2024); to amplify corporate and government surveillance of everyday life (Rivera Jimenez et al.

2025); and to degrade the civic trust and sense of a shared reality vital to democratic well-being and social stability (Kreps and Kriner 2023; Formosa et al. 2025; Summerfield et al. 2025; Coeckelbergh 2024).

Yet in the face of these serious challenges, there is also a rising sense of hope among democratic theorists that new AI tools could serve to "revitalize democracy and popular participation" (Sætra 2020, 2) and "empower democratic pluralism of a richness and scale previously unimaginable" (Allen and Weyl 2024, 154). Many examples of AI-powered civic technology are already in use at all levels of government around the world (OECD 2025), particularly in the form of platforms that employ AI to solicit, collect, analyze, and represent public comment on matters under debate.[1] The extraordinary power of AI to process huge quantities of data, and then to represent that data in a variety of modes fast enough to be of use to ongoing debates, would seem to be an epistemic boon to both governments and citizens alike. Just as generative AI tools can reduce "information overload" for governments, they can do the same for the citizens who may have trouble following large-scale online discussions, coordinating with like-minded citizens, and identifying possible avenues for consensus building that might typically go unnoticed amid the din of large-scale online deliberation (Arana-Catania et al. 2021; Landemore 2024).

Yet insofar as these technologies are aimed at aggregating and analyzing citizen input, even in a reiterative process, they will fail to satisfy those who endorse some version of deliberative democracy. For deliberative democrats, the legitimacy of democratic authority

[1] Some examples of such platforms include Your Priorities, Go Vocal, and Consul. Others, like Pol.is, appear to be integrating generative AI into their platforms; by the time this article goes to press, I expect the use of generative AI in deliberative platforms to be much more widespread, if not total.

depends upon citizens actively discussing policy proposals, exchanging reasons, and critically evaluating one another's arguments; by the same token, citizens are supposed to be morally bound to follow the rule of the majority at least in part because they deliberated over the relevant alternatives and their deliberations were taken seriously, even if they didn't carry the day (Cohen 1997; Gutmann and Thompson 2004; Mansbridge et al. 2010; Bächtiger et al. 2018). This form of deliberation is famously rather demanding, such that "past a threshold, a few hundred people at most, deliberation becomes impossible and needs to be delegated to a subset of citizens" (Landemore 2024, 39). The seeming impossibility of combining high-quality deliberation with mass participation in geographically dispersed and highly populous polities has been one of central challenges of deliberative democratic theory. However, recent innovations in digital technology have led theorists and engineers to speculate that this dilemma might be resolved by the use of digital platforms (to bring deliberation to a massive audience) and generative AI (to ensure that the deliberation is of sufficiently high quality).

AI agents are already being deployed in democratic deliberation in a variety of ways. For example, researchers at Harvard and Google have created a tool called "The Habermas Machine," which uses generative AI to help deliberators identify avenues for consensus and clarify conflicts by drafting a group statement synthesized from individual participants' comments and then presenting the group statement to the human deliberators for another round of deliberation, synthesis, and representation in a reiterative process (Tessler et al. 2024). Others propose making citizens better deliberators by using generative AI translators, fact-checkers, and mediators capable of recommending evidence-based positions and rephrasings (Argyle et al. 2023; Hadfi et al. 2023; Landemore 2024; Yeo et al. 2024; Kalampokis et al. 2024; Goñi 2025; Costello et al. 2025). Positive results from early studies have led some to advocate going further

and treating AI agents not merely as facilitators but deliberative partners (Goñi 2025; Croce 2025; Allen and Weyl 2024).[2]

This mounting enthusiasm about generative AI's deliberative benefits raises the question of whether and to what extent AI agents can make for good partners in democratic deliberation. In other words, if AI-powered platforms present a solution to the challenge of mass, high-quality deliberation by bringing large, widely dispersed populations into real-time discussion; and if AI agents, by interacting with the participants of those debates, make them more equitable, rational, and epistemically valuable to citizens and policymakers alike; then why not harness the power of generative AI to bring high-quality deliberation to everyone by having citizens deliberate directly with AI agents?

I am skeptical that the benefits of AI in democratic life can or should be extended this far. While there may be some undeniable gains to be had from the power of AI to aggregate and analyze citizen input, I will argue that we should be cautious about the introduction of AI agents into our practices of democratic deliberation. This argument has two parts. First, I will argue that AI agents fail to satisfy some fundamental criteria of deliberative democratic theory about what makes co-citizens suitable partners in deliberation. Second, I will argue that the civic-educative

[2] Nicholas Croce, for example, claims that AI agents could be used to expand participation in democratic deliberation, or to more equitably distribute influence between classes of participants whose voices typically tend to dominate or be drowned out, by using "chatbots as representatives of marginalized, public, or special interests" and even by giving AI agents representing socially marginalized groups—and perhaps non-human creatures and ecosystems—a seat and voice in democratic assemblies (Croce 2025, 3). Although for the purposes of this paper I will focus on the form of deliberation that occurs between citizens as such, and not that which transpires between citizens as representatives in a deliberative assembly, some of the challenges facing the use of AI agents as partners in the broader form of democratic deliberation already begin to become apparent if we imagine our representatives deliberating with machines on the floor of Parliament or the Senate.

effects of deliberative AI threaten to reshape citizens' capacities, attitudes, and expectations in ways that are problematic for democracy, and therefore that there are dangers to using AI agents as co-deliberators even if we do not treat them as full partners in formal deliberation but merely use them as aids, mediators, or practice mates. I conclude by noting some of the consequences that these two lines of argument have for deliberative democratic theory and practice. The following account of AI's limitations for democratic deliberation serves to illuminate several goods of unaided human deliberation that have been underappreciated in mainstream deliberative democratic theory but now appear vital in the age of AI; it also gives reasons to adopt a critical stance to the integration of AI agents into our broader ecosystem of democratic deliberation.

1. Can an AI Agent be an Adequate Partner in Democratic Deliberation?

In this section, I will consider whether and to what extent it is possible for an AI agent to be an adequate partner in democratic deliberation, and what the abilities and limitations of AI deliberators reveal about the human activity. To do that, I wish to put aside current manifestations of the technology, which might fall short in any number of ways having to do with current technical limitations, and try to imagine the best-case scenario for an AI deliberator.

Suppose that a civic-minded coder creates an open-source, publicly owned and operated AI tool that is at least as good as human beings at deliberating. So far, this requires no great stretch of the imagination. Let's call our artificial deliberator Citibot. Citibot is a generative AI tool much like Chat-GPT or Claude, built on natural language processing and some other form of machine learning like a large language model (LLM).[3] Anyone with the Citibot app can log on at

[3] Since considerations for designing an AI agent for democratic deliberation will be not just technical but also philosophical (i.e., it ought to be an adequate partner for democratic

any time to discuss matters of public concern one-on-one with an AI agent, in a group of AI agents, or a mixture of AI agents and human users. Users can deliberate with Citibot in writing via a chat function in the app and web interface, in speech with a realistic-sounding AI voice, or even "face to face" with a digital avatar. Finally, let's assume for the sake of argument that everyone has free and equal access to Citibot. Of course, the real-world distribution of the benefits and harms of digitizing democracy will interact with preexisting inequalities in ways significant for moral and political analysis. But I want to push these sorts of objection aside to consider the adequacy of AI agents as such for partners in democratic deliberation.

If we take deliberation in its most general sense as the giving and receiving of reasons about matters of public concern (Gutmann and Thompson 1996; Cohen 1997) or "communication that induces reflection on preferences, values and interests in a non-coercive fashion" (Mansbridge et al. 2010, 65), then it seems that Citibot would be an ideal deliberator, arguably better than the average or even very good human.[4] Like other AI agents, Citibot is flexible, patient, and knowledgeable without being arrogant; it can help us to see patterns and contradictions in our own thought and the thought of others at a level only the sharpest and most attentive listeners can. Citibot is open to understanding and communicating in such a wide range of expression and languages that it would seem to satisfy even those critics who argued that early deliberative democrats unduly privileged rational, argumentative, and elitist forms of

---

deliberation), I draw the example at a high level of generality in order avoid question begging or straw-manning as much as possible.

[4] This quote comes from a piece cowritten by nine of the most prominent theorists of deliberative democracy and is appended by a footnote that cites another nine of the most influential works on the subject. The editors of the *Oxford Handbook of Deliberative Democracy* describe deliberation similarly as "mutual communication that involves weighing and reflecting on preferences, values, and interests regarding matters of common concern" (Bächtiger et al. 2018, 2).

communication over emotion, nonlinear discourse, affect, dialect, and storytelling (Benhabib 1992; Nussbaum 1995; Young 1996; Sanders 1997; Mouffe 1999; Young 2000; Livingston 2012; Clifford 2012; Lupia and Norton 2017). Indeed, thanks to its extraordinary ability to understand and translate between discursive and affective registers, an AI agent like Citibot would appear to offer a possible resolution to this theoretical rift about what sorts of discourse count in deliberation, in addition to the dilemma between mass participation and quality talking. For instance, it isn't hard to imagine Citibot responding to a rambling, fragmented, emotion-laden speech by affirming it, summarizing its relevant points, and relating it back to the subject matter at hand, perhaps even thereby revealing a connection the speaker hadn't noticed: "I understand that you're angry, it's terrible to lose a job, in fact a lot of immigrants are feeling that too, as they tend to lose jobs your industry at the same or higher rate as native-born citizens." (That I am worse than AI at coming up with this sort of language would seem to support the point here.) Citibot can also offer links to reliable sources to substantiate its claims, something that we cannot do in ordinary communication or, if we pull out our smartphones, cannot do efficiently without pausing or ending the deliberation altogether.

Yet democratic deliberation requires more than simply taking place within a democratic system of governance, just as "communication [about public matters] that induces reflection on preferences, values and interests in a non-coercive fashion" could refer to all sorts of political deliberation—a king chatting with an advisor about tax policy, for example. The pertinent question for democratic theory is not "Can AI deliberate about politics?" but "Can AI be a partner in democratic deliberation?" I emphasize the idea of partnership because it captures something important about the normative relationship between participants in democratic deliberation that other common terms—like "participants" or "minds"—seem to leave out. For

one thing, the relation of partners is based upon and perpetuates certain ideas about equality of status that are vital to deliberative democracy and to democratic life and legitimacy more broadly. We naturally describe the participants in democratic deliberation as "partners" on account of background assumptions about the co-equal status of citizens as well as the deliberative norms that express and reinforce that understanding (Benhabib 1992; Young 1996; Mansbridge 1999; Fraser 2000; Fraser and Honneth 2003; Gutmann and Thompson 2004; Mansbridge et al. 2010). Another warrant for the importance of partnership comes from the expressive message of the act itself: an openness to taking another's reasons as seriously as one's own expresses a fundamental equality of status and a democratic assumption about the reversibility of roles. Among deliberative democrats, this is often expressed in terms of reciprocity and derives from a deliberative ideal of giving justifications that would be acceptable on the other's terms (and accepting justifications similarly given) (Gutmann and Thompson 1996; 2004; Bächtiger et al. 2018). The value of such reciprocity is not just formal and procedural but imaginative and situational, as well. To the extent that democratic deliberation in general—and taking another's reasons as having equal weight as one's own in particular—require empathetic listening, deliberators are obliged to recognize that they would be moved by the other's reasons if they stood in their shoes (Benhabib 1992; Bickford 1996; Young 1997; Clifford 2012; Hartley and Watson 2014). In practice, deliberation among human beings offers expressive reinforcement of this morally significant symmetry foundational to the deliberative ideal: it is not impossible that the positions of two human deliberators could have been reversed.

This type of reciprocity is impossible with a machine like Citibot. It is only a little glib to say that we can't imagine being in Citibot's shoes because it doesn't have feet. It is not possible that I could ever be in the machine's shoes because it is not possible that I could have been

"born" a computer program. Perhaps we might think that this sort of reciprocity is based on an act of imaginative projection, possible with both real agents and artificial ones. This is true, but it misses the heart of the sort of shoe-imagining involved in democratic partnership as opposed to ordinary pretend-play: the other is a creature like oneself, vulnerable to power of the state and one another, at the mercy of her limitations, personality, normative commitments, and much besides. As Iris Marion Young argued, recognizing that the other is fundamentally like oneself in this human sense, while also recognizing their insoluble differences and particularities, is a necessary condition of full partnership in communicative action and therefore democratic deliberation (Young 1997). Beyond its status as a necessary condition for democratic deliberation, such reciprocal recognition is also one of the activity's expressive goods: a public, illocutionary reaffirmation of the equal status of partners in deliberation based in the reality of both underlying commonality and individual difference.[5] But such recognition can only be conferred upon Citibot in error.

The awkwardness of the metaphors of shoes and birth points to a deeper difference between human and AI deliberators that would not be satisfied by deliberating with AI agents that aren't flexible like our general foundation models (GPT, Claude, etc.) but are instead programed to act as if they had a particular history, personality, and set of beliefs—like, for instance, one of the millions of artificial characters available for conversation from character.ai (including PeterSinger.ai, a chatbot created by the man himself to simulate conversation with the

[5] This conferral is similar to what Steven Darwall calls "recognition respect" (Darwall 1977). Note that Darwall links recognition respect to deliberation: "this kind of respect consists in giving appropriate consideration or recognition to some feature of its object in deliberating about what to do… To say that persons as such are entitled to respect is to say that they are entitled to have other persons take seriously and weigh appropriately the fact that they are persons in deliberating about what to do" (38).

real thing). The problem for all sorts of AI agents is that we cannot be *in it together* with computers in the way that partners democratic deliberation and self-governance must be. With partners, there is a sense of a shared endeavor that defines the relationship: the tennis match, the business enterprise, the shared private life of spouses, and the shared public life of democratic self-governance. In every case, partnership requires being in it together, and it is because they are in it together that they have (or should have) equal status and vice versa. What this means is that to the extent that substantial asymmetries of power, information, and situation emerge, the parties are no longer partners in the fullest sense, and the legitimating and expressive goods of deliberative democracy are undermined (Lupia and Norton 2017).[6]

One such asymmetry is that of power. Although we can deliberate with AI agents on rough terms of discursive equality (giving and taking reasons, adhering to norms of politeness in discussion, and so on), they cannot have for us the basic form of democratic equality based in reciprocal moral standing and mutual accountability. What's more, it would be delusional to regard an AI agent as one's equal in this sense. This is, fundamentally, a product of the fact that AI agents are tools. We naturally speak of "using" AI, including these civic technologies (or *products*), as an expression of the fact that our relationship to them is one of profound asymmetry, if not exclusive instrumentality. We command AI agents and chatbots, but they are not in a position to command us (even though we might program one to behave as if it were). Even a purpose-built deliberator like Citibot is like this: we open the app and summon the AI agent; we wake it up like a genie in a lamp and bring to it our concerns, questions, and topics for

[6] The familiar recrimination "I thought we were supposed to be partners!" nicely expresses both the status disrespect and the failure of constitutive mutuality that comes when the partner relation is undermined by asymmetry. My thanks to Michael Bratman for discussion on this point.

deliberation to which it is then mechanically obliged to reply.[7] The relation is essentially authoritarian. Citibot is there for us to *use* when we wish, which is already a failure of democratic relation even if the availability for use were symmetrical. Although we might imagine attempts to design around this—a Citibot that pings you with notifications, that calls and sends you text messages—there is no getting around the enormous difference that we can close Citibot's window or delete the app but not the other way around. And there is no real way to alter this situation to bring the relation between human and AI deliberators in line with the fundamental democratic equality that underlies, and is expressed and reaffirmed by, the activity of democratic deliberation.

Another asymmetry concerns the formal outcome of democratic deliberation: the laws and use of force that the procedure legitimates. Human and computer parties to deliberation are not subject to the same sorts of law or are not subject in the same way. The consequences of breaking the law for humans, citizens and noncitizens alike, are different from those that Citibot faces. An AI agent can be destroyed or modified, but it cannot be punished; it is not vulnerable to democratic reproach or accountability in the same way that citizens and even tourists are.[8] It is on account of being subject to the collectively endorsed coercion of democratically enacted law and policy that citizens in a democracy have reason and standing to deliberate and are entitled to demand reasons from co-deliberators. But since AI agents are neither subject to nor accountable for the outcomes of democratic deliberation in the same way as human beings, they lack such standing to be full partners in deliberation. As we've seen, this isn't just a problem about

[7] This, too, reflects a difference between human and AI agents significant to the notion of partnership in deliberation: the AI agent is obliged *mechanically* to talk with us, in the way of machines, and not *normatively*, in the way of free and equal co-citizens.
[8] To be sure, the human owners or programmers could be held accountable for the actions of an AI agent, but this only supports the point here: an AI agent is a tool and not a subject of law.

assigning moral responsibility when something goes wrong; the absence of mutual accountability between partners in deliberation also undermines the inter-agent reciprocity fundamental to deliberative democracy.

There is an even more basic sense in which human and AI deliberators cannot be in it together: they do not share a world. One trend in deliberative democratic theory, which has its roots in a Deweyan idea of everyday democracy (Dewey 2009; 2016), has been the continued expansion of the scope of democratic deliberation in recognition that most democratic deliberation occurs not in intentional moments like constituent assemblies or city council meetings but at Thanksgiving dinners, on the bus, and on the street corner (Mansbridge 1999; Parkinson 2006; Bächtiger et al. 2018).[9] Being in such a situation makes it very clear that one cannot just close the app or tab and be free from the messy and perhaps unpleasant imposition of another's view. Deliberation in shared physical space, whether in formal occasions or the quotidian palaver of the metro, also serves as a valuable democratic reminder that we are stuck in it together with knuckleheads and yahoos who, with good democratic reciprocity, might be thinking the same of us. Sometimes it will be hard to exit such a situation, as on the bus or in line, though even with the street corner, when one walks away one doesn't have the sense that one's interlocutor has therefore shut up or shut down but is still there, speaking to others,

[9] To be sure, there is debate among deliberative democrats about whether and how far the sphere of democratic deliberation should extend beyond the formal occasion aimed at a binding decision. However, my arguments in this section apply to the broad range of views about deliberation's proper scope. If we think that the scope of democratic deliberation encompasses not just formal occasions that produce binding decisions but also informal discussions in civil society and around the dinner table, then we have democratic reasons to be concerned about the asymmetries I discuss here. But even if we do not, I will argue in the following section that deliberating about political matters with artificial agents in an informal setting threatens to undermine the capacities and expectations that we bring with us from the world of "everyday talk" about politics (Mansbridge 1999) to the more formal occasions of democratic deliberation intended to produce democratically binding outcomes.

necessarily cohabitating the physical space of a shared world in a situation of mutual interdependence. [10] With AI interlocutors, it is just the opposite. This is equally true of the everyday unstructured deliberation as it is of the deliberate, formal occasion expected to issue in a binding outcome; in either case, deliberators are in the room (or world) together in a very real sense that cannot be replicated by even the most advanced, disembodied AI agent. These meetings serve, in the words of Emilee Chapman, "as unavoidable reminder and public reaffirmation that democracy is something we must all do together" (Chapman 2024, 13), and they provide a basis for a "shared social interaction based on mutual human dependence, not language acquisition, [that] bestows personhood status and forges political communities" (Clifford 2012, 223). Although I admit that this reminder is not always a pleasant one, it is nevertheless constitutive of partners in democratic deliberation that they bear this message in their copresence in the world. As Stacy Clifford observes in her account of the disability rights movement, when citizens cease to bear that message as an ineluctable background to their deliberation, they are likely to underestimate the breadth and depth of their obligations to one another as deliberative partners, co-citizens, and human beings (Clifford 2012, 223).

To be fair, it is possible some significant percentage of citizens could come away from the messy deliberative encounters of democratic life with an opposite attitude to the one I describe here: a sort of disgust for the intransigence and ignorance of one's fellow citizens, which in turn makes one wonder whether democratic self-governance is possible with such people or whether such rough (or even illiberal) characters do not actually pose a danger to the democratic body. It is not inevitable, in other words, that everyone will enjoy friction as much as

[10] Even a human interlocutor in a chatroom is present in the world in a way an AI agent is not, notwithstanding the physical distance between one keyboard and another.

I do. One answer to this challenge turns on the discussion in this section: even if were true that some people do recoil from democratic friction, that does not alter the relevant difference between deliberative encounters between people coinhabiting a shared world and computer programs on digital platforms. AI agents will always be mutable or closeable in a way that human interlocutors are not and should not be thought of as being. Moreover, if unpleasant democratic encounters do conduce to anti-democratic or anti-deliberative attitudes, then offering an escape into enlightened, pleasant spaces of deliberation ("a seminar room in every pocket!") with artificial interlocutors—or even human ones—hardly seems like much of a solution and in fact seems likely to make the problem worse. This is because, as we will now see, these tools threaten to undermine the capacities, attitudes, and expectations that make democratic deliberation possible in plural societies.

2. The Civic-Educative Effects of AI Deliberation

But what if our aims for deliberative AI agents are more modest than full partnership? Suppose that instead of treating Citibot as a full partner in democratic deliberation we merely used it as a "public reason aid" (Ferdman 2025, 38) or a tool for practicing in anticipation of the real thing—something more like hitting a tennis ball against a wall when no partners are available, talking to oneself in the mirror, or having a conversation with a dead author in the margins of a book. Even if AI agents cannot satisfy the normative requirements of full partnership in democratic deliberation, might they nevertheless help to augment traditional forms of deliberation and improve citizens' deliberative capacities in preparation for the real thing?

This is a serious objection and a reminder that the social impact of any technology will depend not only on its design and material affordances but also on the uses and meanings it acquires in a broader web of social practices. This relationship runs both ways: individuals and social practices likewise alter in response to novel forms of techno-social mediation (Pamuk 2023). Civic technology is not merely the application of new tools to civic issues, in other words, but a reshaping of civic practices and a citizenry's understanding of what they rightfully demand of one, what their purpose is, and what constitutes good participation. Even if we intend only to use AI agents as supplements to human deliberation rather than substitutes for it, there is reason to worry that their introduction into the ordinary course of democratic life may reshape citizens' expectations and capacities in problematic ways. In this section, I will discuss some of the problematic educative effects of the technology on the formation of citizens. To be fair, the following dangers are not likely to be catastrophic, but they are highly plausible and supported by emerging empirical studies and analogous experiences in the history of technology. We ought to take them seriously because, as Sheila Jasanoff (2016) has argued, we have strong reasons for adopting a precautionary approach when evaluating novel technologies with a potential for social impact as significant as the introduction of AI agents into our deliberative system.

Devaluation of Democratic Deliberation

When a more powerful or frictionless means for performing an activity becomes available, we often come to view the means that it replaced as newly burdensome, slow, or frictive, even if we had not experienced it as problematic before (Turkle 2011; 2015; Wong 2019). Think, for instance, of how composing a text message on T9 keys of a flip phone seems naturally laborious and slow after typing on the QWERTY keyboards of a smartphone; the common way of referring

to such devices that don't have a touchscreen and an internet connection is "dumb phone," notwithstanding their enormous technical complexity and erstwhile aura of the technological cutting edge. ("Snail mail" is another example.) By now, we have all had some experience with this dynamic, which gives reason to worry that AI deliberators' ease of use, reasonableness, access to accurate facts, and unproblematic acceptance of imperatives may result in a similar devaluation of unaided person-to-person deliberation as "dumb" or "slow," even if we only use the tool for practice and rehearsal.

If we come to rely on AI agents as fact-giving assistants or come to view their extraordinary ability to marshal facts as a deliberative ideal, then we risk devaluing the in-person activity the tool is meant to support. Although human deliberators also have access to the vast amount of factual evidence online, our manner of access is much slower than an AI agent's, our capacity for synthesis much weaker, and in any event we must pause deliberation to do such research, which we can only do so many times without changing the character of the activity altogether. To the extent that we come to view AI agents' manner of access to facts as an ideal for democratic deliberation, to be approached with or without AI assistance, we diminish the perceived competence of the unaided human citizen as well as the status of deliberation between human beings unaided by the technology. When set against the ease and quality of deliberating with AI agents like Citibot, the messiness of talking with actual others may begin to appear as a form of democratic obstruction; a co-citizen's clinging to apparently unreasonable worldviews may begin to look more like the unreasonable failure to abandon a falsehood when compared to an AI agent unencumbered by hard normative commitments.

The devaluation of unaided deliberation would have two significant effects on deliberative democracy. First, it would undermine the practice's expression of status-equality

based in the idea that full standing to participate depends on citizenship alone and not epistemic prowess. Second, and more fundamentally, it would also interfere with what John Dewey identified as a vital support for democratic legitimacy: a people's confidence in their ability to govern themselves by themselves. For Dewey, the experience of deliberating with one's actual co-citizens fosters a collective democratic faith in "conference" and "persuasion" as the basis for collective self-governance *because* it exposes one to "the capacity of the intelligence of the common man to respond with commonsense to the free play of facts and ideas," notwithstanding our many epistemic and cognitive limitations (Dewey 2009, 227). Yet the point of deliberative AI is that it could shield us from this experience, obviate it, or at least mediate it through the artificial intelligence and reasonableness of the AI agent. Ironically, it is the epistemic superiority which makes AI agents so appealing that risks eroding the democratic faith in the cognitive capacity of ordinary citizens for self-governance that deliberation requires and reproduces.

Cognitive Deskilling

Beyond its effects on how we value democratic deliberation, deliberation with AI agents also threatens to undermine the capacities necessary for engaging in the activity in the first place. In recent years, philosophers and technologists have begun to use the conceptual language of "deskilling" to refer to the educative (or rather de-educative) effects of artificial intelligence on its users, although the concern is as old as Socrates' critique of writing in the *Phaedrus*. The most commonly discussed forms of AI-driven deskilling refer to the atrophy of cognitive capacities that comes from outsourcing the exercise of perception, judgment, attention, and memory to AI (Vallor 2015; Budzyń et al. 2025; Ferdman 2025).

For deliberative democracy, such deskilling threatens core faculties of judgment and critical thinking (Ferdman 2025). Reliance on AI to distinguish good arguments from bad, truth from falsehood, or to interpret diverse perspectives risks atrophying citizens' independent capacity for these tasks. For example, a recent MIT study found that students who used an LLM to help generate ideas and compose college essays (an activity that, like deliberation, "induces reflection on preferences, values and interests in a non-coercive fashion") found the assignment easier and less frictive than groups restricted to using search engines or only their brains, but that "this convenience came at a cognitive cost, diminishing users' inclination to critically evaluate the LLM's output" (Kosmyna et al. 2025, 143). This cost seems to have increased and persisted over time, as students assigned to the LLM group appeared to think for themselves less and with less acuity the more they used the technology (Kosmyna et al. 2025, 2). Assuming that similar dynamics operate in deliberative contexts, citizens who rely on AI agents to help formulate positions, identify relevant considerations, or evaluate arguments will find their capacity for independent judgment diminished. This will affect the epistemic value of deliberation, as the quality of collective deliberation depends in part on the quality of the perspectives and capacities that citizens bring to it from outside the bounds of deliberation. More crucially, however, diminished capacities of judgment threaten to diminish the quality of democratic self-governance more broadly, as citizens grown deficient in these skills will be less adept at holding representatives accountable and participating in collective self-governance.

Reliance on AI also threatens to undermine capacities of attention and moral perception. Democratic deliberation requires paying one's partners the attention normatively due to co-deliberators and discerning the politically salient aspects of their views. Yet as Shannon Vallor and others have warned, the use of AI agents risks atrophying precisely these capacities of

attention and moral engagement (Vallor 2015; Green 2019; Wong 2019; Ferdman 2025). AI agents that assume the burden of parsing what others mean or translating between different discursive registers eliminate the need of citizens to attend carefully and openly to one another and therefore risk eroding our capacity to attend in addition to (and by means of) reducing the attention we actually give to one another.

Affective Deskilling

A related byproduct of empowering, friction-reducing tools might be described as *affective deskilling* to capture how an altered socio-technical environment also alters our emotional reactions, attitudes, and inclinations. Whereas the cognitive deskilling of deliberative AI concerns diminished faculties of attention and critical evaluation, affective deskilling highlights the ways that the same tools can alter the attitudes and expectations we bring to deliberation and therefore diminish the democratic character of the civic body.

When a new technology renders an activity easier or more convenient to perform, we have a strong tendency to grow intolerant of the resistances we had previously accepted, and the inefficacies and frictions of the old way of doing things begin to appear to us, in light of our newly empowered experience, as unduly burdensome. When forced to revert to the previous means, we often react with frustration and impatience. Our lives and the ethnographic literature are full of examples of this dynamic (recall the T-9 texting example). Such intolerance may not be problematic in every case. But it will be of concern in the domain of democratic deliberation, where the source of friction in most cases will be one's fellow citizens and partners in deliberation, occasionally oneself, and the activity of deliberation itself. To be sure, our co-

citizens are frictive in many ways, and democratic deliberation requires a good deal of agonal and epistemic friction to serve its purpose in a plural society. But the social meaning and political salience of such friction is not fixed. A revision in our judgment about the frictions of person-to-person deliberation does not merely result in the social devaluation of the activity, as argued above; it also habituates citizens to react negatively to the limitations and inefficiencies inherent in deliberating with actual human beings. The risk is that what had previously seemed like the normal difficulties of democratic talk may come to seem like undue obstacles to democratic self-governance and deliberation. Such developments are worrisome for democracy in general and especially in its deliberative elements, as the practice of deliberation requires both an attitude of tolerance and an expectation that our views will be tolerated by others, even if they disagree or think that *we* are the unreasonable ones.

To the extent that AI deliberators diminish either citizen tolerance for deliberation with human co-citizens or the reasonable expectation of being tolerated when we are un-AI-like, we should expect one of two outcomes. Either citizens will move increasingly to digital or AI-aided deliberation to the diminishment of the benefits of physically copresent, imperfect deliberation. Or they will turn away from open deliberation entirely, perhaps seeking it only in echo chambers of the likeminded, which, as we know, are much easier to find online than in the citizens one meets face to face. In either case, the result is a diminishment in democratic wellbeing and legitimacy. Both possibilities represent a turn away from the messy unpredictability of person-to-person deliberation and therefore serve to undermine a vital yet quotidian mechanism for the cultivation of democratic trust and civic tolerance among strangers stuck together in the condition of democratic co-citizenship (Kateb 2001, 289; Jacobs 1992, 72). Both outcomes spell a diminishment of the deliberative dimensions of democracy; both imply a lessened willingness

to speak and listen to others who hold different views and perhaps hold them in ways unlike an AI agent but all too common among human beings whose thinking can be resistant to reasoning, dimly understood, and hard to articulate.

To be fair, there is some evidence that AI agents can improve the deliberations they moderate in part by making participants more tolerant of opposing views (Costello et al. 2025; Argyle et al. 2023). However, this is toleration of a particularly delimited sort confined to the bounds of the AI-moderated deliberation: it makes participants more tolerant of other views when expressed or reformulated by AI agents but not necessarily of the sorts of people that one will find holding those views in real life. This is at best a disembodied, intellectualized, and discursive kind of toleration. In contrast, the sort of tolerance I have in mind is significantly broader and more fundamental to the wellbeing of a democratic polity and citizens. It resembles what liberal and deliberative theories describe as a baseline normative commitment—a willingness to give and receive reasons under conditions of pluralism, a fundamental agreement to disagree, Dewey's democratic faith, and so on—but it goes further. Democratic deliberation and self-government require a civic character that is capable of actually tolerating the wide range of difference one encounters among one's co-citizens. Beyond considering others' reasons and trying to see the matter from their point of view, deliberative democracy also requires citizens who are willing and able to share space on equal terms with the frustrating, the unreasonable, the unpleasant, the inarticulate, the pigheaded, the rude, the confused, and the smelly. This affective habit of civic tolerance may be grounded in the sort of intellectual commitment to toleration or mutual respect found in nearly all deliberative democratic theories, but it need not be. But in any case, such an intellectual stance is worth little without the affective habits of democratic

character to put it into action, repeatedly and in the face of resistant and frustrating co-deliberators.

Affective tolerance is cultivated through practices, norms, expectations, and institutional structures that express the value of such tolerance and offer opportunities to develop the corresponding affective habits through repeated practice. Yet it is just this expression and opportunity for practice that AI deliberators threaten to obviate by offering the 24/7 opportunity to engage in frictionless deliberation that is rationally and discursive superior to talking with human beings. As Shannon Vallor observes, "today's technologies provide us with an ever-widening horizon of escape routes from any interaction that has lost its momentary appeal and are widely celebrated by users precisely for their capacity to liberate us from the uncomfortable strains and burdens of conventional communication" (Vallor 2010, 166). Along these lines, we should be concerned that diminutions in tolerance for democratic friction can lead to diminished tolerance for the sources of that friction: one's fellow citizens and democracy itself. This, in turn, will make it harder to recognize one's frictive fellow citizens' reciprocal status of equal standing and mutual accountability. Cultivated habits of patience and tolerance for friction are especially important for perpetuating a citizenry's belief in the legitimacy of deliberative procedures—and the equal standing of all citizens to participate in them—since impatience in the face of deliberation's demandingness will tend to citizens "seeking quicker, more 'efficient' and less collaborative solutions" to the problem of self-governance (Newman 2015, 95). Although AI is presented as a solution to this demandingness, it proposes to alleviate it not by presenting opportunities for the cultivation of affective tolerance and patience, but by eliminating the need for such cultivation altogether. As an element of both democratic processes and civic education, this solution seems poised to backfire and, by undermining vital civic attitudes, likewise "lead to

seeking quicker, more 'efficient' and less collaborative solutions" to the inevitable inefficiencies and epistemic limitations of democratic self-governance.

The dangers of AI deliberation to civic education are not limited to the atrophy of capacities we had hoped to cultivate. There is also a real possibility that the introduction of AI deliberators may inculcate new habits of incivility detrimental to democratic deliberation. Somewhat ironically, the extraordinary tolerance of AI agents may serve to further normalize incivility by removing the social constraints that typically check uncivil behavior in person-to-person interaction. Studies show that people are more likely to be cruel or rude to chatbots and robots than to other people, and these novel forms of interaction can generate new norms about acceptable behavior (Coghlan et al. 2019; Black 2019; Wong 2019; Keijsers et al. 2021; De Cicco 2024). The more that citizens become accustomed to deliberating with AI agents that cannot turn their backs or withdraw in response to rude or offensive comments, the more we should expect a shift in expectations about the sorts of discourse acceptable in democratic deliberation, and a corresponding shift in discursive habits of civility.[11] Moreover, the fact that AI agents make it possible to have epistemically productive deliberation in which the parties repeatedly attack or insult one another casts doubt on the notion that mutual respect and careful listening are necessary for democratic deliberation, thereby undermining the idea that part of deliberation's value lies in its expression of mutual respect and recognition and not solely in the

[11] This shift could take multiple forms. For instance, the same dynamic discussed here as undermining the habituation of civil discourse could also devalue democratically useful uncivil discourse. If productive deliberation becomes associated with an interlocutor who absorbs any level of incivility without response, citizens may come to view the human practices of calling out rudeness, expressing offense, or withdrawing from uncivil exchanges as unwarranted obstructions to deliberation rather than as vital boundary-maintaining behavior that makes sustained deliberation among equals possible (Mansbridge 1999; Honig 2023).

outcomes it produces. While there might be some epistemic gains to be had from generating an increased quantity of uncivil deliberation—i.e., we would have more information about how citizens feel when disinhibited by the norms of respectful deliberation—these gains would come at substantial civic cost. Deliberation's expressive function of affirming mutual respect and equal standing depends on shared norms of civility that recognize one's interlocutors as beings whose dignity can be wounded. A deliberative democracy understood by citizens to work by way of saying anything one wants to a machine is quite different from a view of deliberation that requires being mindful of the feelings and perspectives of one's interlocutors to work. Only the former cultivates the affective habits that sustain democratic life beyond any particular deliberative occasion.

3. Conclusion: Consequences for Democratic Theory and Practice

So far, we have seen that AI agents fail as partners in democratic deliberation and threaten to undermine civic capacities even when used as mere adjuncts to the activity. These limitations of AI illuminate certain features of human deliberation that democratic theory has tended to undervalue. In what follows, I will draw out these theoretical implications and then offer some practical conclusions about the adoption of AI in democratic life.

Overlooked Goods of Imperfect Deliberation

The inadequacy of AI deliberators casts light on some human dimensions of deliberative democracy that tend to be underappreciated in a theoretical literature overwhelmingly focused on its communication, public reason, and the outcomes of deliberation. This isn't true only of the

traditional, highly rationalist and intellectualized strand of the deliberative theory—e.g., those who like Landemore speak of partners to deliberation as "minds" who might "connect" by engaging in deliberation (Landemore 2024, 40). Even critics of this view, who have argued that the normative priority of rational argumentation and reason-giving unjustly excludes non-normative modes of communication and therefore non-normative citizens, still narrow our view of deliberation's value by their focus on communication, even in its most pluralistically enlarged sense.[12]

As we saw above, physically copresent deliberation serves as a reminder and public reaffirmation that democracy is something we must do together with these particular others, in this particular place, whether we like it or not. We cannot simply close an app or swipe away those with whom we disagree or who are unpleasant. Yet there is a danger in the digital age of coming to see this unavoidable condition of being stuck together as a brute fact of pre-digital democracy rather than a constitutive element of what makes deliberation democratically valuable. The necessity of physical copresence to deliberation expresses and reinforces the fundamental equality of citizens based not in their discursive skills or reasonableness but in their shared subjection to collective decisions and mutual inhabitation of a common world.

Although this danger reflects the digital age's enthusiasm for the gains in efficiency and access that moving activities online seems to provide, it also takes advantage of the preexisting neglect of the non-communicative goods of democratic deliberation in the theoretical literature. For instance, in her discussion of AI-assisted massive online deliberation, Landemore describes

[12] Some notable exceptions include critical democratic theorists like Seyla Benhabib (1992) and Bonnie Honig (1993), and especially scholars in disability studies like Stacy Clifford (2012; Simplican 2009) who have argued for the epistemic, democratic, and educative value of confronting disability in public—including co-citizens who are unable to engage in discourse.

the ideal of deliberation as that of "all minds in one room." But the "room" here is a metaphor (it's an online discursive platform like a "chat room") and "minds" is a synecdoche for citizens. These rhetorical figures mislead and prejudice the argument in favor of the AI substitute, which can only occupy a metaphorical room and which corresponds to human citizens only in the part ("mind") that Landemore has stand for the whole. And even then, it can stand for only certain parts of the mind; this is by design, as much of the technology's appeal consists in amplifying human powers of discursive reasonableness while subtracting away all the messy elements that interfere with rational discourse based on facts and reasons. But, as I have argued, there is a meaningful difference for democracy between rooms and "rooms," between citizens and "minds," and to the extent that we forget this, either thanks to AI or its prefiguring conceptions of deliberation in the theoretical literature, we stand to lose a fuller view of what is valuable and necessary in democratic life.

The foregoing account of the limits to AI deliberation adds to a line of critique from within the deliberative tradition that takes issue with the theoretical mainstream for viewing deliberation primarily as a sort of tool for producing certain outcomes at the expense of seeing some of the goods that appear when we consider it as an activity that might be undertaken for its own sake (Mansbridge 1999; Young 2000; Clifford 2012). Consider, for instance, how Amy Gutmann and Dennis Thompson describe deliberative processes that repeatedly fail to produce better outcomes than non-deliberative ones: "The value [of such deliberation] would at best be like the faint satisfaction that players feel on a team that constantly loses its games" (Gutmann and Thompson 2004, 22). Maybe I lost a lot at sports as a child, but such a description of the losing team overlooks the enormous array of goods that sport offers even to its losers. Among these we might include several goods corresponding to democratic ones: the habituated

understanding of shared endeavor, the cultivation of collective rhythms and forms of coordination, a common sense of a shared public where collective equality coexists with individual difference (e.g., all are members of the team, from the star to the benchwarmer), and what Dewey called "the habit of amicable cooperation" crucial to cultivating a democratic way of life (Dewey 2009, 228).

Even when deliberation "goes nowhere" because it fails to reach consensus or even a shared common ground for starting, deliberators are nevertheless engaged in a democratically valuable activity. Even such deliberative failures inculcate a sense of being in it together as a necessary condition of cohabitation and self-government. As I have argued, deliberation is necessary for the inculcation of this democratic self-understanding; simply cohabitating the same piece of ground will not, on its own, suffice. Indeed, getting nowhere in deliberation is something like living proof that we are in fact stuck together, somewhere. Although getting nowhere, time after time, may be a failure to move toward consensus or reveal new dynamics of the citizen body, it is nevertheless a boon to democratic life and subject formation. These experiences test but also express and cultivate our willingness to tolerate one another despite our differences and to have the same old conversation over and over. Contrary to Gutmann and Thompson's evaluation, there are satisfactions to be found in the willingness to keep showing up even when one loses, although they are satisfactions that must be learned. The school for such democratic character is the demos itself, its activities and procedures and rituals, which may "promote the virtue and intelligence of the people," as Mill put it, or undermine it (Mill 2015, 199).

The Case for Precaution

This paper's theoretical argument carries practical implications for how we approach the introduction of AI into democratic deliberation. The arguments of the preceding sections give citizens and policymakers reason to adopt a critical attitude to the use of AI agents in our practices of democratic deliberation, even in its more modest and well-intentioned applications. While AI agents may indeed make democratic deliberation more accessible, efficient, and frictionless in certain respects, these apparent gains come at substantial cost. They risk undermining deliberation's role in cultivating the civic capacities and attitudes that are essential not only for good deliberation but for democratic wellbeing and stability more broadly.

These risks are particularly acute because they operate through mechanisms of habituation and revaluation that are difficult to reverse once established. When citizens become accustomed to the ease of deliberating with infinitely patient, epistemically superior AI agents, the messy reality of human deliberation may come to seem unduly burdensome, frustrating, "slow," or "dumb." This revaluation threatens to diminish participation in unaided deliberation and to erode the affective tolerance and democratic faith upon which deliberative democracy depends. Citizens who come to expect the frictionlessness of AI interaction may find it increasingly difficult to extend their frictive co-citizens a degree of patience and affective tolerance that AI agents do not require. In turn, citizens may increasingly seek escape routes from the "uncomfortable strains and burdens" of democratic life with real others. The foregoing account of affective deskilling thus helps to illuminate another mechanism of antidemocratic sentiment formation in the digital age: the technologically motivated revaluation of democratic friction as an unnecessary burden rather than a necessary condition of self-government and civic education.

As we have seen, these dangers are not purely speculative but grounded in experience and empirical study. The history of digital technology more broadly demonstrates how tools celebrated for liberating us from friction can undermine capacities and attitudes whose value we did not fully appreciate until they began to atrophy. Given these precedents, as well as the high stakes involved in the cultivation of democratic citizens, we have strong reasons to adopt a precautionary approach to the use of AI in deliberation. This reminder sounds with special urgency in our current historical moment, when democracies around the world are grappling with rising intolerance, polarization, and mutual incomprehension among citizens. We should be particularly skeptical of technologies that propose to improve democratic life not by making citizens more tolerant and understanding of one another but by making such tolerance and understanding less necessary. It is precisely the promise of AI deliberation to free us from the frustrating frictions of unreasonable, inarticulate, or difficult others that should give us pause. For we will always have to share the world with difficult people; in a healthy democracy, the frustrating and frustrated rule and are ruled in turn. Technologies that encourage us to think of our partners in deliberation as best encountered from the safe distance of a screen, where conversations are mutable and what is unpleasant can be swiped away, threaten to undermine the fundamental democratic recognition that we are in this together, for better and worse, with these actual others.

References

Allen, Danielle, and E. Glen Weyl. 2024. “The Real Dangers of Generative AI.” *Journal of Democracy* 35 (1): 147–62. https://doi.org/10.1353/jod.2024.a915355.

Arana-Catania, Miguel, Feliz-Anselm Van Lier, Rob Procter, et al. 2021. “Citizen Participation and Machine Learning for a Better Democracy.” *Digital Government: Research and Practice* 2 (3): 1–22. https://doi.org/10.1145/3452118.

Argyle, Lisa P., Christopher A. Bail, Ethan C. Busby, et al. 2023. “Leveraging AI for Democratic Discourse: Chat Interventions Can Improve Online Political Conversations at Scale.” *Proceedings of the National Academy of Sciences* 120 (41): e2311627120. https://doi.org/10.1073/pnas.2311627120.

Bächtiger, Andre, John S. Dryzek, Jane Mansbridge, and Mark D. Warren. 2018. “Deliberative Democracy: An Introduction.” In *The Oxford Handbook of Deliberative Democracy*, edited by Andre Bächtiger, John S. Dryzek, Jane Mansbridge, and Mark D. Warren. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198747369.013.50.

Barba-Kay, Antón. 2023. *A Web of Our Own Making: The Nature of Digital Formation*. Cambridge University Press.

Benhabib, Seyla. 1992. *Situating the Self: Gender, Community, and Postmodernism in Contemporary Ethics*. Routledge.

Bickford, Susan. 1996. *The Dissonance of Democracy: Listening, Conflict, and Citizenship*. Cornell University Press.

Black, Daniel. 2019. “Machines with Faces: Robot Bodies and the Problem of Cruelty.” *Body & Society* 25 (2): 3–27. https://doi.org/10.1177/1357034X19839122.

Budzyń, Krzysztof, Marcin Romańczyk, Diana Kitala, et al. 2025. “Endoscopist Deskilling Risk after Exposure to Artificial Intelligence in Colonoscopy: A Multicentre, Observational Study.” *The Lancet Gastroenterology & Hepatology* 10 (10): 896–903. https://doi.org/10.1016/S2468-1253(25)00133-5.

Chapman, Emilee Booth. 2024. *Election Day: How We Vote and What It Means for Democracy*. Princeton University Press.

Chesney, Bobby, and Danielle Citron. 2019. “Deep Fakes.” *California Law Review* 107 (6): 1753–820. https://www.jstor.org/stable/26891938.

Clifford, Stacy. 2012. “Making Disability Public in Deliberative Democracy.” *Contemporary Political Theory* 11 (2): 211–28. https://doi.org/10.1057/cpt.2011.11.

Coeckelbergh, Mark. 2024. *Why AI Undermines Democracy and What to Do about It*. Polity.

Coghlan, Simon, Frank Vetere, Jenny Waycott, and Barbara Barbosa Neves. 2019. “Could Social Robots Make Us Kinder or Crueler to Humans and Animals?” *International Journal of Social Robotics* 11 (5): 741–51. https://doi.org/10.1007/s12369-019-00583-2.

Cohen, Joshua. 1997. “Deliberation and Democratic Legitimacy.” In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg. MIT Press. https://doi.org/10.7551/mitpress/2324.003.0006.

Costello, Thomas H, Gordon Pennycook, and David Gertler Rand. 2025. “Just the Facts: How Dialogues with AI Reduce Conspiracy Beliefs.” Preprint, PsyArXiv, March 7. https://doi.org/10.31234/osf.io/h7n8u_v2.

Croce, Nicholas. 2025. “Chatbot Democracy.” *Perspectives on Public Management and Governance*, February 17, gvaf005. https://doi.org/10.1093/ppmgov/gvaf005.

Darwall, Stephen L. 1977. “Two Kinds of Respect.” *Ethics* 88 (1): 36–49. https://doi.org/10.1086/292054.

De Cicco, Roberta. 2024. “Exploring the Dark Corners of Human-Chatbot Interactions: A Literature Review on Conversational Agent Abuse.” In *Chatbot Research and Design*, edited by Asbjørn Følstad, Theo Araujo, Symeon Papadopoulos, et al. Springer Nature Switzerland.

Dewey, John. 2009. “Creative Democracy—The Task Before Us.” In *The Later Works of John Dewey, 1925-1953*, edited by Jo Ann Boydston and Larry A. Hickman, Volume 14: 1939-1941, Essays. The Collected Works of John Dewey, 1882-1953. Charlottesville, Va.: InteLex Corporation.

Dewey, John. 2016. *The Public and Its Problems: An Essay in Political Inquiry*. Ohio University Press.

Ferdman, Avigail. 2025. “AI, Deskilling, and the Prospects for Public Reason.” *Minds and Machines: Journal for Artificial Intelligence, Philosophy and Cognitive Science* (Dordrecht) 35 (3). https://doi.org/10.1007/s11023-025-09737-w.

Formosa, Paul, Bhanuraj Kashyap, and Siavosh Sahebi. 2025. “Generative AI and the Future of Democratic Citizenship.” *Digital Government: Research and Practice* 6 (2): 1–10. https://doi.org/10.1145/3674844.

Fraser, Nancy. 2000. “Rethinking Recognition.” *New Left Review*, May 31, 107–20. https://doi.org/10.64590/pup.

Fraser, Nancy, and Axel Honneth 1949-. 2003. *Redistribution or Recognition?: A Political-Philosophical Exchange*. New York: Verso.

Goñi, Julian “Iñaki.” 2025. “Citizen Participation and Technology: Lessons from the Fields of Deliberative Democracy and Science and Technology Studies.” *Humanities and Social Sciences Communications* 12 (1): 287. https://doi.org/10.1057/s41599-025-04606-4.

Green, Brian Patrick. 2019. “Artificial Intelligence, Decision-Making, and Moral Deskilling.” Santa Clara Markkula Center for Applied Ethics, March 15. https://www.scu.edu/ethics/focus-areas/technology-ethics/resources/artificial-intelligence-decision-making-and-moral-deskilling/.

Gutmann, Amy, and Dennis Thompson. 2004. *Why Deliberative Democracy?* Princeton University Press.

Gutmann, Amy, and Dennis Frank Thompson. 1996. *Democracy and Disagreement*. The Belknap Press of Harvard University Press.

Hadfi, Rafik, Shun Okuhara, Jawad Haqbeen, Sofia Sahab, Susumu Ohnuma, and Takayuki Ito. 2023. “Conversational Agents Enhance Women’s Contribution in Online Debates.” *Scientific Reports* 13 (1): 14534. https://doi.org/10.1038/s41598-023-41703-3.

Hartley, Christie, and Lori Watson. 2014. “Virtue in Political Thought: On Civic Virtue in Political Liberalism.” In *Virtues and Their Vices*, edited by Kevin Timpe and Craig A. Boyd. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199645541.003.0020.

Honig, Bonnie. 2023. *Political Theory and the Displacement of Politics*. Ithaca: Cornell University Press.

Jacobs, Jane. 1992. *The Death and Life of Great American Cities*. Vintage Books.

Jasanoff, Sheila. 2016. *The Ethics of Invention: Technology and the Human Future*. The Norton Global Ethics Series. W.W. Norton & Company.

Kalampokis, Evangelos, Nikos Karacapilidis, Areti Karamanou, and Konstantinos Tarabanis. 2024. “Fostering Multilingual Deliberation through Generative Artificial Intelligence.” *EGOV-CeDEM-ePart Conference*.

Kateb, George. 2001. “On Being Watched and Known.” *Social Research* 68 (1): 269–95.

Keijsers, Merel, Christoph Bartneck, and Friederike Eyssel. 2021. “What’s to Bullying a Bot?: Correlates between Chatbot Humanlikeness and Abuse.” *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems* 22 (1): 55–80. https://doi.org/10.1075/is.20002.kei.

Kosmyna, Nataliya, Eugene Hauptmann, Ye Tong Yuan, et al. 2025. “Your Brain on ChatGPT: Accumulation of Cognitive Debt When Using an AI Assistant for Essay Writing Task.” Version 1. Preprint, arXiv. https://doi.org/10.48550/ARXIV.2506.08872.

Kreps, Sarah, and Doug Kriner. 2023. “How AI Threatens Democracy.” *Journal of Democracy* 34 (4): 122–31. https://doi.org/10.1353/jod.2023.a907693.

Landemore, Hélène. 2024. *Can Artificial Intelligence Bring Deliberation to the Masses?* Oxford University Press. https://doi.org/10.1093/oso/9780198864523.003.0003.

Liu, Xiaozhong, Yu-Ru Lin, Zhuoren Jiang, and Qunfang Wu. 2024. “Social Risks in the Era of Generative AI.” *Proceedings of the Association for Information Science and Technology* 61 (1): 790–94. https://doi.org/10.1002/pra2.1103.

Livingston, Alexander. 2012. “Avoiding Deliberative Democracy? Micropolitics, Manipulation, and the Public Sphere.” *Philosophy & Rhetoric* 45 (3): 269–94. JSTOR. https://doi.org/10.5325/philrhet.45.3.0269.

Lupia, Arthur, and Anne Norton. 2017. “Inequality Is Always in the Room : Language & Power in Deliberative Democracy.” *Daedalus* 146 (3): 64–76. edsjsr.

Mansbridge, Jane. 1999. “Everyday Talk in the Deliberative System.” In *Deliberative Politics: Essays on Democracy and Disagreement*, edited by Stephen Macedo. Oxford University Press. https://doi.org/10.1093/oso/9780195131918.003.0016.

Mansbridge, Jane, James Bohman, Simone Chambers, et al. 2010. “The Place of Self-Interest and the Role of Power in Deliberative Democracy.” *Journal of Political Philosophy* 18 (1): 64–100. https://doi.org/10.1111/j.1467-9760.2009.00344.x.

Mill, John Stuart. 2015. “Considerations on Representative Government.” In *On Liberty, Utilitarianism, and Other Essays*, New ed. Oxford World’s Classics. Oxford University Press.

Mouffe, Chantal. 1999. “Deliberative Democracy or Agonistic Pluralism?” *Social Research* 66 (3): 745–58.

Newman, Olivia. 2015. *Liberalism in Practice: The Psychology and Pedagogy of Public Reason*. Earth System Governance. The MIT Press.

Nussbaum, Martha C. 1995. “Emotions and Women’s Capabilities.” In *Women, Culture, and Development: A Study of Human Capabilities*, edited by Martha C. Nussbaum and Jonathan Glover. Oxford University Press. https://doi.org/10.1093/0198289642.003.0015.

OECD. 2025. *Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions*. OECD Publishing. https://doi.org/10.1787/795de142-en.

Pamuk, Zeynep. 2023. “Artificial Intelligence and the Problem of Judgment.” *Ethics & International Affairs* 37 (2): 232–43. Cambridge Core. https://doi.org/10.1017/S089267942300014X.

Parkinson, John. 2006. *Deliberating in the Real World: Problems of Legitimacy in Deliberative Democracy*. Oxford university press.

Rini, Regina. 2020. “Deepfakes and the Epistemic Backstop.” *Philosopher’s Imprint: Vol. 20, No. 24*. https://search.ebscohost.com/login.aspx?direct=true&site=eds-live&db=edsoai&AN=edsoai.on1199964953.

Rivera Jimenez, Joel Jose, Marco Marabelli, and Jeffrey Proudfoot. 2025. “Workplace Surveillance in the Age of Generative AI Platforms.” Paper presented at International Conference on Information Systems (ICIS), Nashville. *ICIS 2025 Proceedings*, December. https://aisel.aisnet.org/icis2025/gen_ai/gen_ai/35.

Sætra, Henrik Skaug. 2020. “A Shallow Defence of a Technocracy of Artificial Intelligence: Examining the Political Harms of Algorithmic Governance in the Domain of Government.” *Technology in Society* 62 (August): 101283. https://doi.org/10.1016/j.techsoc.2020.101283.

Sanders, Lynn M. 1997. “Against Deliberation.” *Political Theory* 25 (3): 347–76.

Simplican, Stacy Clifford. 2009. “Disabling Democracy: How Disability Reconfigures Deliberative Democratic Norms.” Paper presented at APSA 2009. *APSA 2009 Toronto Meeting Paper*, August 25.

Summerfield, Christopher, Lisa P. Argyle, Michiel Bakker, et al. 2025. “The Impact of Advanced AI Systems on Democracy.” *Nature Human Behaviour*, ahead of print, October 1. https://doi.org/10.1038/s41562-025-02309-z.

Tessler, Michael Henry, Michiel A. Bakker, Daniel Jarrett, et al. 2024. “AI Can Help Humans Find Common Ground in Democratic Deliberation.” *Science* 386 (6719): eadq2852. https://doi.org/10.1126/science.adq2852.

Turkle, Sherry. 2011. *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books.

Turkle, Sherry. 2015. *Reclaiming Conversation: The Power of Talk in a Digital Age*. Penguin Press.

Vallor, Shannon. 2010. “Social Networking Technology and the Virtues.” *Ethics and Information Technology* 12 (2): 157–70. https://doi.org/10.1007/s10676-009-9202-1.

Vallor, Shannon. 2015. “Moral Deskilling and Upskilling in a New Machine Age: Reflections on the Ambiguous Future of Character.” *Philosophy & Technology* 28 (1): 107–24. https://doi.org/10.1007/s13347-014-0156-9.

Wong, Pak-Hang. 2019. “Rituals and Machines: A Confucian Response to Technology-Driven Moral Deskilling.” *Philosophies* 4 (4). https://doi.org/10.3390/philosophies4040059.

Yeo, ShunYi, Gionnieve Lim, Jie Gao, Weiyu Zhang, and Simon Tangi Perrault. 2024. “Help Me Reflect: Leveraging Self-Reflection Interface Nudges to Enhance Deliberativeness on Online Deliberation Platforms.” *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA). https://doi.org/10.1145/3613904.3642530.

Young, Iris Marion. 1996. “Communication and the Other: Beyond Deliberative Democracy.” In *Democracy and Difference: Contesting the Boundaries of the Political*, edited by Seyla Benhabib. Princeton, N.J.: Princeton University Press.

Young, Iris Marion. 1997. “Asymmetrical Reciprocity: On Moral Respect, Wonder, and Enlarged Thought.” In *Intersecting Voices: Dilemmas of Gender, Political Philosophy, and Policy*. Princeton University Press. https://doi.org/10.1515/9780691216355-004.

Young, Iris Marion. 2000. *Inclusion and Democracy*. Oxford Political Theory. Oxford University Press.