


Against Bot Democracy: The Dangers of Epistemic Double-Counting

Ana Tanasoca

The article focuses on the question of how each of us should deliberate internally when forming judgements. That is a matter of political consequence, insofar as those judgements stand behind our votes. I argue that some violations of epistemic independence like message repetition can, if the receivers are not aware of the repetition, lead them to double-count information they have already taken into account, thus distorting their judgments. One upshot is that each of us should ignore or heavily discount certain sorts of inputs (e.g., bot messages or retweets) that are likely just to be repetition of what we have already taken into account in our internal deliberations. I propose various deliberative norms that may protect our internal deliberations from epistemic double-counting, and argue that opinion leaders have special epistemic duties of care to shield their audiences from clone claims.

People learn from one another. Generally, the more people there are reporting the same thing, the more credible it is likely to be. That makes good sense if they are all providing genuinely independent evidence.¹ It makes no sense if, as in the 2016 U.S. presidential election, many of the apparently “confirming reports” are really just replicas of the original report, reposted automatically by some computerized web robot, or bot.² Treating a replica as if it were a new, independent piece of evidence amounts to double-counting the original report—and double-counting evidence is as bad from an epistemic point of view as double-counting votes would be from a democratic perspective.

Research on this article was supported by the Australian Research Council, grant FL140100154. I thereby thank the ARC and the grant holder, John Dryzek, for supporting my work. Many thanks to the three journal reviewers for their questions and comments that strengthened the article. For their helpful comments, suggestions, and conversations, my gratitude also goes to Tim Ainstrope, Simon Cotton, Ned Dobos, Toni Erskine, Bob Goodin, and Jensen Sass.

Ana Tanasoca  is a postdoctoral research fellow at the Centre for Deliberative Democracy, University of Canberra (ana.tanasoca@mq.edu.au). Recent works include *The Ethics of Multiple Citizenship* (Cambridge University Press, 2018) and “Ritual Deliberation” (with Jensen Sass) in the *Journal of Political Philosophy*. She is currently working on a book titled *Deliberation Naturalized*.

Bots are just the extreme case. Sometimes real people (politicians among them) engage in message repetition deliberately to epistemically mislead their audience. Other times, people just repeat what they have been told, without pausing to assess the truth of those claims. From an epistemic point of view, it would be just as wrong in those cases, as in that of bots, to double-count the parroted evidence if we have already taken it into account.

In this article I focus on how each of us should deliberate internally when forming judgments. I propose various deliberative norms to help prevent beliefs from being compromised by double-counting. One is that each of us should ignore or heavily discount certain sorts of inputs when making up our own minds.³ Even though such double-counting affects our votes, in contrast to advocates of “epistocracy” (Brennan 2011; 2016), I emphatically do not propose that anyone be denied a vote.

This article is concerned with individual epistemic performance and how double-counting the same information can degrade it, rather than with the collective epistemic performance of the electorate as a whole.⁴ Any violation of independence among voters compromises the latter, whereas only certain, very special types of violations of independence compromise the former.⁵ If X influences Y and both influence Z , all three of those influence relationships undermine the collective epistemic performance of the electorate, but only the relation between X and Y undermines the individual epistemic performance of Z in the ways I am interested in here.⁶

There are reasons to worry less about the impact of non-independence on collectives than on its impact on individuals and on how non-independent informants

compromise their recipients' competence in forming accurate judgements to guide their votes. Some of the factors that arguably undermine independence among voters (such as public discussion and party competition) may nonetheless improve collective epistemic performance by helping impose structure on the collective choice situation (Goodin and Spiekermann 2018, ch. 9; List et al. 2013; White and Ypi 2016).⁷ Furthermore, democratic aggregation procedures are robust against failures of independence in ways that individuals are not when deciding what to believe on the basis of evidence contained in one another's assertions.⁸

My approach is different from the collective "wisdom of the multitude" model (Condorcet [1785] 1976; Goodin and Spiekermann 2018). It does share, however, one important feature with it, which I must declare and defend at the outset. Like that other model, mine too assumes that at least some components of political judgment are fact-based, even though political judgments are also infused with values, which are arguably importantly different from facts (Hume [1739] 1896, 335).⁹ Insofar as they are based on facts, it is important that our beliefs track the truth about those facts. Furthermore, whereas when voting facts and values might typically be bundled together (Goodin and Spiekermann 2018, 40–41), there is no such forced bundling when it comes to our beliefs. We can, and typically do, form separate beliefs about matters of fact and matters of value.

The other key assumption in my approach is that other people might have information about those facts that I do not, which I should take into account in revising my beliefs. That information may take many forms, such as evidence, experience, and argument. Where it is possible to engage fully with the other, those types of information should be assessed and incorporated into my own beliefs in different ways. But even in circumstances where that is not possible, the sheer fact that others have independent, considered beliefs different from mine should give me some reason to reconsider how confident I should be in my own.

Familiar Concerns, Differently Approached

There are many long-standing arguments against an extensive democratic franchise on the ground that voter ignorance compromises the epistemic quality of social decisions. Such critiques of democratic inclusion, philosophically familiar from the ancient Greeks and politically prominent among opponents of nineteenth-century extensions of the suffrage, are heard in some quarters still today (Brennan 2011, 2016; Caplan 2007; Mill [1861] 1977, ch. 8; Plato 2006).

Worries about the independence of voters are familiar from history as well. Denying the franchise to people on grounds that they lack the requisite independence is an

unfortunately familiar political trope. Nineteenth-century antidemocrats opposed extending the franchise to wage laborers or married women on the grounds that their wills were not properly independent of those of their employers or their husbands.¹⁰ In modern times, we have become increasingly worried over the effect of opinion leaders more generally.¹¹

My argument, however, does not concern whether people should be included in or excluded from the electorate, nor does it rest on any claims about citizens' ignorance: I accept that everyone should be given an equal vote in all decisions affecting them. Yet even if citizens have (each or on average) better-than-random competence at making decisions, they can still fail to be independent of one another in the way required for good individual epistemic outcomes. My focus is on the internal realm of deliberation (Goodin 2000; Goodin and Niemeyer 2003; Landmore and Mercier 2012). What sort of challenges might grave violations of epistemic independence pose for people's internal processes of belief formation?

Weighing Truth Claims

As citizens we often need to reflect on the truth of claims made in the public sphere. I use the umbrella term "truth claims" to cover a wide range of inputs that have communicative and epistemic content, be they arguments (per Habermas), reasons (per public reason theorists), or testimonies (per Young). Each of us must assess how these truth claims fare in comparison to one another, and hence what epistemic weight each claim should exert on our individual judgment (and through that on our vote).

Weighing versus Counting?

Deliberative democrats often insist that, whereas voting is about counting heads, deliberation is about weighing arguments. Deliberation, they say, requires us to consider arguments or truth claims on their own merits—and we should do so regardless of the number of people voicing them. We should recognize the force of the better argument when we encounter it.¹²

The classic deliberative model rests on protracted, dynamic discursive exchanges among a small(ish) set of people, however. The forceless force of the better argument is at its best in such settings, which allow back-and-forth exchanges, reciprocal prodding, and questioning among deliberators. Yet most of the information we receive everyday does not come through such structured deliberations or ideal speech situations. Instead, we have to assess what weight we should give it solitarily. In the absence of protracted, discursive engagement with our informants, it becomes harder to assess the "betterness" of the information and arguments we come across. And that may be especially so if what is at stake is some matter of fact rather than of logic or moral conviction.

Perhaps the epistemic merits of some truth claims are plain on their face, such that we immediately recognize their correctness as soon as they are uttered.¹³ The merits of others are not so clear, however. Whether those claims are true or some arguments are better than others is far from self-evident to us. In such cases, the force of the better argument is too feeble to provide clear guidance.

Suppose I am confronted with some such proposition, and suppose I have already assessed its substance and used whatever relevant knowledge I have to form a judgment about it. Suppose I then learn that one hundred people, each of whom I regard as my epistemic peer, came to a judgment different from mine. Shouldn't *that* fact figure in my internal deliberations as well? For a (probably quite large) subset of claims based on facts, counting heads can also aid us in weighing the truth of these claims correctly. If multiple people are making the same claim independently from one another, then we ought to revise our beliefs in light of that fact. The more independent assertions there are of the same claim, the more epistemic weight we should ascribe to it, updating our beliefs in light of each independent assertion.

This does not mean that we should “follow the herd” mindlessly, of course. The point is just that if many independent thinkers report that they believe *X*, we should count each of those reports as evidence for *X*'s credibility, and we should take that evidence into account when revising our credence in *X*. How much weight we attach to that evidence is an open question. Yet each new piece of independent evidence should count for something in our internal deliberations, and if many pieces of independent evidence point toward the same conclusion, the result of that updating will be that we will ourselves be more inclined toward that conclusion.¹⁴

Hence if informants are epistemically independent, then something akin to counting has its place in internal deliberation, just as it does externally in voting. But that is true only if those informants are speaking *independently* of one another.

What Weighing Truth Claims Ideally Entails

Something like Bayesian updating ought ideally to be at the heart of internal deliberation.¹⁵ In Bayesian terms, each speaker's claim can be seen as a new piece of evidence in light of which we revise our beliefs, according to its epistemic weight in proportion to its credibility. In simple terms, the epistemic weight that one accords to a claim depends on how reliable one believes the information it conveys to be. In the first instance, we assess the epistemic merits of a proposition on the basis of our prior knowledge: we form an initial judgment of a claim on its substance, as it appears to us. Then, as additional evidence comes to light, we revise our belief in the claim in light of how credible we believe that new evidence to be.

Take the claim, “Russia interfered in the 2016 American elections.” In my internal deliberations, I would attribute some credence—a prior probability—to this claim on the basis of what I already know of Russian-American relations, of Russia's technological capacity to leverage such attacks, of how susceptible social media and voting technology are to such interference, and so on. If I then hear the same claim again from a reputable security analyst, I would adjust upward my credence in this claim, given the relatively high conditional probability that I attach to that claim being true if that expert says it is. My confidence in that claim increases further with each additional credible agent making the same claim. More generally, the more people who endorse a claim, the more likely I will judge that claim to be true (assuming I regard those people's reports as credible at all).¹⁶

Of course, the number of sources may be used as reliable evidence in our internal deliberations only insofar as those sources act independently of one another, exercising critical independent judgment rather than merely slavishly repeating or reporting what others have said. And just as we would not want to count twice the same person's vote in an election, so too should we steer clear of “deliberative double-counting.” Directly or indirectly taking into account the same person's claim twice in our internal deliberations would be to give it undue epistemic weight. And that, of course, is one way the Russians attempted to manipulate the American election—by having their many bots and trolls say the same thing.

Boundedly Rational Agents Being Misled by Message Repetition

The preceding section described what our processes of internal deliberation should ideally entail. Well-known cognitive limitations may hamper our judgements, however, precluding us from weighing others' claims as accurately as we should. Our limited memory capacity and bounded rationality, in particular, can undermine our assessments. The way we react, as listeners, to message repetition is one such cognitive limitation of particular relevance here.

Psychological studies show we are very vulnerable to message repetition, quite generally. Hearing the same message multiple times creates an impression of both familiarity and commonality. The feeling of familiarity with a proposition increases with the number of exposures, regardless of their source: that familiarity (or “exposure”) effect is much the same whether someone hears the same thing three times from the same person or from three different people. But hearing the same report from multiple different people also has a commonality effect. It leaves listeners with the impression that what they hear is the general public opinion across their community. Crucially, both familiarity and commonality effects serve to make them more likely to believe that the proposition is true (Rothbart et al. 1978; Weaver et al. 2007).

Theories of public opinion formation based on models of bounded rationality account for citizens' susceptibility to repetitions of the same information via what is called "persuasion bias." Such studies have shown that individuals' influence on group opinions is not uniquely determined by the accuracy of the information they provide, but also by how well-connected they are; that is, by their position in the social network. Being well positioned in the network allows their information to be repeated many times across the network and have an echo effect that is not epistemically warranted (DeMarzo, Vayanos and Zwiebel 2003, 914).

These findings are extremely relevant for developing a normative account of internal deliberation and belief formation. The deeper patterns of interconnectedness among individuals today are likely to make information repetition ubiquitous in the public sphere. Taking others' claims into account without knowing if they are epistemically independent, or counting claims without being certain of their origin,¹⁷ could thus undermine our internal deliberations and lead to flawed individual (and indirectly collective) epistemic judgments.

This effect of information repetition is well-known and put to good use by some political and economic agents such as political parties and corporations. Both political propaganda and advertisements intentionally use message repetition to exploit their audiences' cognitive vulnerabilities and mislead their beliefs. Yet not all agents engage in message repetition with the malevolent intent of misleading the others' beliefs. Message repetition may occur accidentally in discursive exchanges with a benign intention: to inform, to justify a belief, and so on. Furthermore, in some cases speakers may not even be aware that they are parroting information from another source; they may have internalized that information and forgotten its source—a case of epistemic appropriation.¹⁸

Regardless of the intention behind it, however, message repetition distorts the judgements of listeners if they are unaware that the message is a mere repetition of something they have already taken into account. In what follows, I discuss why grave violations of epistemic independence are problematic, regardless of the intentions of agents engaged in message repetition.¹⁹

Why Others' Epistemic Independence Matters

In internal deliberation, as in Bayesian updating, it is crucial not to update our beliefs twice in response to the same piece of evidence. We should update our beliefs in response to any given speaker's claim only insofar as that claim is indeed a new piece of evidence, one that we have not already taken into account in our internal deliberation. Thus we need to be able to distinguish new evidence from that which was already taken into account.

Of course, it is not uncommon for multiple people to make claims having the same propositional content; for example, "the sun rises in the east." Insofar as each person makes that claim on the basis of an independent assessment—insofar as each asserts this claim qua an independent epistemic agent—each person's assertion should be treated as a new piece of evidence supporting that proposition. Our confidence in it, and the credibility we attach to it, should then increase.

Yet if others are merely repeating a claim from someone else that we have already taken into account, and they have no independent ground for making that same claim themselves, then we should not update our beliefs in light of their repetition. This is a violation of epistemic independence on the speakers' part that can, if unnoticed, lead us to double-count the same evidence and attribute too much epistemic weight to that claim in our own internal deliberations.

How should we understand independence? Epistemic independence is a scalar property: it comes in degrees. Hence, there are more or less grave violations of epistemic independence. Furthermore, we can never be fully epistemically independent from one another statistically. Insofar as we are both competent observers reporting on the true state of the world, our reports (that the sun rises in the east, for example) will not be independent of one another but rather will be caused by the same thing (the direction of sunrise; Dietrich 2008). Or because we all rely on the same body of scientific knowledge and other evidence about the state of the world, we cannot be fully independent from one another (see the later discussion). Those are violations of independence that cannot be avoided. My concern here is with graver violations of independence that we could avoid: those complete violations of independence that occur when one person merely repeats something from someone else.

Clone Claims

As senders of information, agents sometimes repeat information they have themselves received from other agents, without adding any new information of their own in the process. In so doing, they violate epistemic independence. In communicating that repeated information to others, those agents expose yet other agents as receivers of information to the replicated information, which may mislead the latter's judgments if unaware of this replication, making them less epistemically competent.

Take the following example. A chemistry lab runs a series of experiments. Twenty scientists conduct the same experiment, each one reporting the results of their experiment to the team leader, who compiles the lab's findings on the basis of the evidence received from all of the experimenters. Eighteen of them are working alone, but two (trainee Jim and his mentor Kate) are working together. Instead of observing the experiment Jim plays

poker on his iPhone. In need of a report to send, he then copies Kate's report and sends it as his own. Kate sends her report in as well, unaware of what Jim has done.

The example involves a flagrant, complete violation of epistemic independence on Jim's part, one that could have been easily avoided. How should the team leader count Jim's replica report? If Jim had independently observed the experiment himself, then it would have made sense for the team leader to weigh both Jim's and Kate's findings in deciding the lab's overall conclusion. But because Jim's report is simply a replica of Kate's, the team leader should take no account whatsoever of Jim's report, after having already taken Kate's into account.

Analogously, while internally deliberating we can erroneously attribute too much epistemic weight to a claim someone has just repeated²⁰ after hearing it from another person, whose own claim on this topic we have already fully taken into account (cf. Goodin 2001, 122, fn. 6, and 123–24; Lehrer 1976; 2001; Lehrer and Wagner 1981). If someone is only reporting others' claims (as Jim was doing with Kate's observations), then there is of course a danger of double-counting the same claim (as the team leader risks doing with Kate's report), which would distort our final judgment.

Epistemic Overcounting Short of Double-Counting

Strict double-counting occurs when we take into account the same information twice in our internal deliberation. That might easily happen, if we cannot distinguish new information from old information that we have already incorporated.

Next let us consider something short of that: cases in which a new assertion provides some genuinely new information while also containing some old information that we have already taken into account. Two reports might come from sources that are subject to some common causes, but nonetheless report independently of one another. Or the second report might repeat—but then add to—information that we have already taken into account. To treat the second report as fully on a par with the first one would be to overcount the first report. But it would not be literally to double-count it, because the second report also adds some genuinely new information.

1. Common Causes, Independent Assessments

We all get much of our information from the same sources, and we are all influenced by what others are thinking: “our cognitive lives are never self-made” (Beerbohm 2012, 154). In this sense, we can never be fully epistemically independent. Yet we can nonetheless exercise independent judgement both in processing the information we receive from others and when passing along that information to others. Hence, an agent can be said to be “relatively epistemically independent” if that

agent “keeps some distance” from and “is not excessively reliant” on what others think (Beerbohm 2012, 154).

Consider a variation of the earlier example. Suppose that Jim and Kate conducted their experiment together, but each wrote a separate, independent report of what they observed. While their write-ups were independent of one another's, however, the experiment was same for both. Both were reporting on what happened in the same test tube.

When compiling the results of what happened in the lab that day, the team leader should give weight to both Jim's and Kate's reports. Each brought some independent judgment to bear on their report, unlike in the first version of the example. But the team leader should accord less weight to each of their reports than to the reports of other 18 scientists in the lab reporting on what happened in their different test tubes.

Such “common causes” are ubiquitous. Indeed, even the 18 scientists' reports observing a different test tube cannot be said to be completely independent from one another. All of the scientists were working with the same batch of chemicals, which might have been contaminated. All were working at the same altitude, under the same atmospheric conditions. Those factors were common causes affecting the outcomes of all the scientists' experiments at the same time (Dietrich and Spiekermann 2013).

Still, because each of those 18 other scientists' reports was “independent conditional on all those common causes,” each report is epistemically worth something (and more than the reports from Jim or Kate, who were observing the same test tube). They are just not worth as much, epistemically, as the reports from scientists working in different labs, using different batches of chemicals and at different altitudes.

Recent contributions to epistemic democracy distinguish various ways of violating epistemic independence that can compromise decision making to a greater or lesser extent (Dietrich and Spiekermann 2013; Estlund 1984; 1989; 2008; Goodin and Spiekermann 2018, ch. 5; Grofman and Feld 1989; Waldron 1989). The general message is that agents' reports are often likely to be subject to common causes of various sorts. But although this reduces the epistemic value of their reports, it does not vitiate it altogether. Just so long as those reports are independent conditional on all common causes to which they are subject, they will add something unique in their assessment of the situation. In other words, their assessments will be epistemically independent from one another, bracketing all the common causes that make them non-independent of one another.

This weaker form of independence is the most we can realistically aspire to in political life. People read the same newspapers, watch the same television programs, and read the same webpages; we cannot easily avoid having our judgments influenced by the same common causes. But even this weaker form of epistemic independence is

undermined if people merely repeat what they heard, without any critical reflection on their part. That is an abnegation of independence conditional on all common causes, and it is at least in principle a wholly avoidable one (a later section suggests ways in which that can be done).

If, in contrast, people critically reflect on the content before passing it along, they are exercising judgment that is independent, conditional on all common causes (Goodin 2003; Waldron 1989). Their judgment is somewhat non-independent, insofar as they are subject to the same common cause (the same newspaper, TV program, webpage, or whatever); but it is somewhat independent, insofar as they reflect on the content before passing it on (Goodin 2003; Waldron 1989). Then their repetition of the original message itself contains some new information.

Partial Repetition

Sometimes we may actually have good epistemic reasons to repeat others' claims, and hence our repetition then conveys some new information about that claim that our audience has not taken into account. Suppose I have good reason to regard you as likely to be correct because of what I know (but my hearers do not) about your training or track record. Or suppose I have reflected on your claim and find that, on the basis of my own knowledge, I concur with it. Then my reiteration of your claim conveys some new evidence about that claim's epistemic credibility, as well as repeating evidence already on the table. Others taking into account my reiteration of your claim would therefore lead to something less than complete double-counting—although counting my reiteration fully on a par with the original claim would still amount to overcounting it, assuming they have taken your original claim fully into account.

Everything depends on whether my repetition is epistemically warranted in some way—whether it conveys some new, additional information that the others do not have, over and above that which is already contained in the original claim being repeated. Suppose that I repeat your claim x simply because I want to be in your good graces²¹ with no epistemic reason for repeating it. Then my repetition would not contain any new, relevant information about x 's credibility, over and beyond the information contained in your original assertion of x .

Consider retweeting or reposting in this light. The way online platforms are set up allows us to instantly observe both that someone is simply repeating someone else's assertions and whose they are. We should thus be able to tell if we have seen that message before, and to discount it or disregard it, if so. In that respect, we should be able to easily distinguish new information from old information.²² What we cannot so readily know is what to make of the fact that someone has *retweeted* or *reposted* that message. Does that fact, that she retweeted or reposted it, contain new, independent information,

over and above the information contained in the original message?

Maybe it does, if the retweeter or reposter knows something that we do not about the reliability of the original tweeter or poster, for example. Then we ought to count the retweet or repost for something, even if we do not weigh it as heavily as a wholly new piece of evidence. At least in the case of retweets or reposts, however, we can reasonably doubt that the repeated message will be based on new information, simply because the architecture of online interaction is set up in a way that rewards unreflective, impulsive behavior.²³ Most retweets, reposts, "shares," or "likes"—that is, the sorts of clone claims we most typically encounter online—are probably not the result of any considered judgment about those claims' credibility and do not contain any new, value-added information. Hence, it would ordinarily be wrong for people to take impulsive retweets or reposts as representing opinions that are based, at least in part, on new information not contained in the original tweet or post and that should be given some epistemic credit in their internal deliberations.

A Real Risk

Some may doubt whether epistemic double-counting or overcounting is a real risk to our judgments. How often do individuals merely repeat or report what others have said and these repetitions are utterly devoid of any new information? To what extent is public discourse dominated by "clone," no-information-value-added claims?

Note that such repetition often occurs in our everyday social interactions. We listen to what our colleagues, friends, and families say, and in our subsequent conversations we often repeat or report what we heard. Sometimes we do so in ways that involve weaker violations of independence—for example, communicating our updated belief in light of what we heard, having passed some critical judgment on the claim before communicating it to someone else. But for the moment, concentrate on those cases where we literally just repeat a claim without having any independent evidence for its truth. Insofar as people often frequent the same circles, our interlocutors may well also get direct input from the original independent source whose claim we merely repeated. If so, they may end up attributing too much epistemic weight to that claim by updating their beliefs twice in light of it; that is, by epistemically double-counting it.

Increasingly, public deliberation takes place online, which magnifies these dangers many times. Many clone claims are produced and amplified by bots or web robots, algorithms used in social media networks that look like real users. More than half of web traffic is attributed to bots (Lafrance 2017). There are many types of bots, from bots that can "like" your posts and vote in online polls, to bots that can look like your "followers," to ones that can post comments or even hold conversations on Twitter.²⁴

To make matters worse, we are not good at distinguishing bots from real users (Gorwa and Guilbeault 2018).

Bots are used not only for commercial purposes but for political ones as well. Between the first and second presidential debate of the 2016 U.S. presidential campaign, one third of pro-Trump tweets and one fifth of pro-Clinton tweets came from bots (Guilbeault and Woolley 2016). Considering their magnitude, our internal deliberations could be easily manipulated with the help of clone claims coming from completely non-independent sources. There is a real risk that the outpouring of non-independent bot “clone” claims will drown out or marginalize epistemically independent claims coming from real people with real stakes in the matter, by making these claims seem relatively less common and hence less important than they actually are, which poses not only an epistemic problem but also a democratic one.

Just how vulnerable are we, epistemically, to clone claims? Take this small but telling anecdote illustrating the political pervasiveness of clone claims and our vulnerability to them. The U.S. president himself recently tweeted his gratitude to a social media fan, Nicola Mincey, only for this user’s account to be suspended the very next day on the grounds of it being fake and probably part of Russia’s disinformation campaign (Phillip 2017). The fact that even the U.S. president was fooled by a bot, that he was unable to distinguish an independent citizen from a non-independent robot, shows that we are not well equipped to spot such sources, and hence that our judgments risk being derailed by the large number of bots participating in public deliberation.

Bots are just the latest manifestation of this phenomenon, however. There have been earlier notorious attempts to influence both domestic and international politics through non-independent, “clone claims” of various sorts. In domestic politics, legislators used to rely on “counting the mail” as a way of determining just how important any given issue was for their constituents. Getting a large volume of letters from constituents for or against some piece of legislation was an important signal for representatives. Yet with the advent of copiers and even more of the computer, legislators’ mailbags and email inboxes were filled with innumerable copies of the same “form letter” to which constituents had in fact merely affixed their signatures (Schribman 1982).

Similar things happen in the international realm. Take the case of the UN’s Universal Periodic Review (UPR) of member states’ human rights records. Under that procedure, NGOs are invited to report on each state’s human rights record. In the initial UPR review cycle, some states tried to manipulate the process by creating a plethora of state-funded NGOs that then sent numerous reports, many of them identical, in an effort to drown out and make less visible the reports sent by genuinely independent organizations (Chauville 2014,

105–6).²⁵ Had this not been noticed, disproportionate weight would have been given to multiple biased perspectives having one single origin: the government of the state that was under review.

Guarding against Double-Counting and Overweighting

How can we protect our internal deliberations from double-counting or undue attribution of epistemic weight? There are a range of approaches, varying both in their epistemic strengths and in the costs they would impose on other democratic principles.

New Discursive Norms

The least intrusive approach would try to solve the problem through the adoption of a new set of deliberative norms by citizens and particularly by influentials. Those norms generically entail doing whatever one realistically can to avoid epistemically misleading other people. Perhaps ordinary citizens should be subject to a weaker form of that requirement, trusting that hearers, once adequately informed of the existence of repeated claims, would discount them. But influentials, by reason of their power to command large audiences, should be held to a higher epistemic duty of care not to help promulgate clone claims.

Citizens, Reveal Your Sources. The problem of epistemic double-counting arises from the fact that, all too often, we do not know whether someone is making a claim as a relatively independent epistemic agent (as a firsthand source) or as a relatively non-independent epistemic agent (as a secondhand source). If that is the problem, the most obvious and most modest solution is the voluntary adoption of deliberative norms requiring speakers to reveal (1) when they are simply passing along others’ messages, (2) from whom, and (3) for what reasons. This would give listeners the information they need to increase the accuracy of their internal deliberations. It would allow them to distinguish new from old information that they have already taken into account, thereby helping them to avoid double-counting or misattributions of epistemic weight to repeated claims.

Such deliberative norms should be a preferred solution insofar as they do not pose the same sort of threat to free speech that alternative solutions might do. These norms do not prohibit speakers from saying any particular things. They actually require deliberators to say more, not less.²⁶ Speakers would have to elaborate on the things they have already said, providing some background for the claims they are advancing—details they would not have otherwise provided.

On the downside, the speakers’ commitment to these norms is essential to the success of this approach, especially if the norms are only informal and are not supported by any enforcement mechanism. Doubtless some people will fail, at least from time to time, to internalize and act on these norms. Even those who, in principle, consider the

norms useful and legitimate might find, in practice, that living by them proves a nuisance. Providing a full reference for every claim one makes in a deliberation may be burdensome and disturb the natural flow of conversation in an obnoxious way.

In their study of social networks and persuasion bias, DeMarzo, Vayanos, and Zwiebel (2003, 919, fn. 16) briefly consider this solution but promptly dismiss it.²⁷ They do so precisely on the grounds that such communication would be “extremely complicated”; indeed, “the information the agents would need to recall and communicate would increase exponentially with the number of rounds” of communication. To fully eliminate repetition, agents would have to know and report the full structure of their discursive network; that is, who/what was the source of their info, who/what was the source of their source, and so on. Furthermore, to abide by full-disclosure norms, the information that agents would need to communicate would increase with each discursive interaction, imposing an unreasonable burden on them.²⁸

Fully countering clone claims in this way might thus be infeasible. But we might at least reduce their prevalence in public discourse by inculcating a norm of revealing one’s sources at least two steps back; that is, a norm of revealing one’s own source and the source’s source. Such a norm of *limited* disclosure would be much less of a burden on speakers’ memories and natural conversational flows.

Some might object that even such norms of limited disclosure might be too demanding and obtrusive to be adopted voluntarily as an informal social norm. The alternative would be for them to be formally adopted and socially enforced. Yet, in some environments the norms may be virtually unenforceable. The online sphere is one example, because of the sheer volume of communications that need monitoring. The social media platforms themselves would have to step in to monitor whether the norms are being respected and impose some kind of sanctions when they are breached. To be sure, they are already occasionally penalizing users for some offenses, but usually they can do so only by relying on other users to report the offense. And if the norms are seen as too demanding to begin with, and few users comply with them, then the rest of the user community may not bother reporting norm violations insofar as noncompliance is pervasive. One may well ask, then, whether it is naive to consider such deliberative norms as a solution.²⁹

This objection, however, overlooks the interests and motivations that all speakers and listeners naturally have in conversation. Both listeners’ and speakers’ incentives are mutually compatible with the deliberative norms that I propose. Speakers have an interest in being listened to, and listeners have an interest in listening to new information. This means listeners will have an incentive to

listen to information whose “newness” they can assess, in preference to information whose newness they cannot assess. If so, speakers who have an interest in being listened to will have an incentive to abide by the norms and reveal their sources, so that others will listen to what they have to say.

Thus, we should not be unduly skeptical about the uptake these deliberative norms would enjoy. Provided speakers care about being listened to, and listeners care about listening particularly to new information, there are incentives for all parties to adopt and abide by this modest version of the deliberative norms.

Influentials, Know Your Sources. It would suffice to hold ordinary citizens to a limited epistemic standard (insofar as it is practical to do so), because any given one of them has a limited capacity to influence public debate and distort collective rationality. But there may be reasons to hold influential public officials and prominent opinion leaders to higher epistemic standards, because they can influence a great many people at once: those whom we know to be boundedly rational agents who may not necessarily be able to keep track of the source (and the source’s sources) of everything they hear.

Non-independent agents like bots can come to dominate public debate when their claims are amplified by persons who have many followers, such as the U.S. president. As an avid retweeter, President Trump has routinely amplified tweets from dubious sources (Timberg, Dwoskin, and Entous 2017).³⁰ He is not the only person to undermine public deliberation in this way, to be sure. Various other political figures, celebrities, and media personalities have amplified clone claims coming from the fake accounts run by Russian operatives (Timberg, Dwoskin, and Entous 2017), making it more likely for citizens to come across what in fact are clone claims they can easily double-count, not being aware of their *sub rosa* common origin.

High public officials such as the U.S. president, media outlets, and other high-profile opinion leaders are in a privileged position to influence a great many people at the same time. Given their prominent epistemic position and the (apparent) epistemic authority they enjoy, they should be under a particularly stringent *epistemic duty of care* to exercise independent judgment and to pay careful consideration when passing along information from other unchecked sources, like bots—which makes their audience more vulnerable to double-counting. Some political theorists have already discussed at length the nature and extent of citizens’ epistemic obligations (Beerbohm 2012, ch. 6). Perhaps it goes without saying that these should extend, in an even more demanding way, to influential political decision makers whom we should expect not to be easily swayed by non-independent agents deployed in public deliberation at someone else’s command.

Minimally, the norms of public deliberation ought to impose a special epistemic duty of care on influentials to “know their source”—and, as in the banking analogy, not to accept things from dubious sources when passing along information to their many followers.³¹ With power comes responsibility. There is much to be said in favor of imposing such duties, particularly on prominent public persons, as safeguards against the falsification of public opinion, the corruption of public deliberation, the distortion of collective reasoning, and ultimately the undermining of popular sovereignty.

Not only are there good epistemic reasons to do that but there are also good democratic reasons. Allowing claims made by non-independent agents to infiltrate our deliberations is likely to increase existing power imbalances and exacerbate structural inequalities. Powerful and resourceful agents will be the first to try to amplify their voices by using other non-independent (human or non-human) agents as mouthpieces. There are myriad examples of powerful lobby groups and individuals (from tobacco companies to Monsanto to the Koch brothers to President Putin) resorting to such tactics—whether astroturfing,³² deploying bots and trolls, or financing scientifically dubious research—to increase their power or protect their advantaged status (Oreskes and Conway 2010). They all try in those ways to exaggerate the weight and support of their claims.³³ The deployment of epistemically completely non-independent agents may distort our internal deliberations to the point where our individual judgments and ensuing collective decisions are not only epistemically flawed but also can no longer genuinely claim the citizens’ meaningful consent. At least in principle then, epistemic hazards can have serious political consequences, especially when we are dealing with problems (complete epistemic non-independence) and processes (internal deliberations) that are invisible to the naked eye.

How exactly can such norms requiring influentials to know their sources and not pass on dubious information stick? One solution would be for online platforms to enforce them as part of their “conditions of use.”³⁴ Or, perhaps, they could be merely enforced by social pressure as in the previous section. Even in the absence of external enforcement, however, assuming once again that listeners have an interest in distinguishing good quality information from bad, and that influentials have an interest in being listened to, then influentials will naturally have an incentive to reassure their listeners as to the quality of the information they communicate. More specifically, influentials have an interest in (1) checking carefully the origin of information they intend to promulgate and the credibility of its source, before passing along that information to their listeners, and (2) signaling clearly to their listeners that they have done epistemic due diligence in this respect. The norms requiring influentials to exercise a special

epistemic duty of care would thus be fully compatible with listeners’ and speakers’ interests.

Excluding or Removing

In addition to these suggested discursive norms, two other approaches may counteract epistemic overinclusion. Both promise to be equally epistemically effective in preventing clone claims from impinging on our internal deliberations, although they come with differential costs in other respects. The first, more extreme version works by banning anyone voicing any such claims from public sphere discussions altogether. The second, less extreme approach removes clone claims from the deliberative space. The first approach bans the speaker, whereas the second bans only the specific offending clone claim.

Both come with democratic costs. The former approach amounts to an act of social exclusion, banning a person from the public forum altogether,³⁵ which would be democratically unacceptable (at least in the case of real people, if not bots). The latter approach amounts, at most, to a violation or a limitation of free speech. That too would generally be regarded as democratically problematic, although perhaps less so once we realize how these limitations are similar to others we already allow and consider legitimate.

Excluding the Offenders. The first exclusionary approach is one that social media networks such as Twitter and Facebook have begun implementing. They have started, albeit hesitantly and belatedly, weeding out fake user accounts associated with bots that mimic real, independent citizens participating in public deliberation (Wu 2017). Considering the magnitude and political implications of clone claims coming from these accounts recently and their potential for undermining internal deliberation and collective rationality, that is surely good news.

Although their exact number is unknown, existing estimates indicate that a large number of non-independent agents are currently influencing public debate. On some calculations, as many as 48 million Twitter active users or “nearly 15 percent—are automated accounts designed to simulate real people.” Facebook has disclosed that it may have many automated accounts; its initial estimate in 2017 was 60 million,³⁶ but in November 2018, Facebook announced that it removed more than one billion fake accounts, many of them bots. That shows that online platforms have the capacity to identify and remove bot accounts (Romm and Dwoskin 2018).

There are at least two different problems with managed bots: they are spreading deliberately false information, and they are mimicking one another in spreading the same false information. Both obviously pose epistemic problems, but it is the latter that gives rise to risks of epistemic double-counting and overcounting.

While some are individual accounts (e.g., the account of citizen John Doe), a number of them pretend to represent collective organizations or groups, hence larger numbers of individuals. One example is @Ten_GOP, an account claiming to speak for Tennessee Republicans that was recently shown to be set up by Russian operatives (Timberg, Dwoskin, and Enous 2017). This type of fake accounts claiming to represent interest groups have a greater potential of undermining internal deliberation by giving the impression that a view is endorsed by a myriad of independent epistemic agents—when in fact it can be traced to a sole agent. The phenomenon of “followers” is similar. Many real users take pride in having large numbers of followers, who are supposedly citizens independently associating themselves with the user’s opinions. Yet in fact many are bots, deployed to make it look like the user’s opinions are epistemically weightier than they really are (Confessore et al. 2018). Buying bots and fake user accounts has unsurprisingly become a profitable business.

Removing fake automated user accounts is minimally problematic democratically. Excluding real human agents from public deliberations (online or otherwise) is much more so.³⁷ Banning citizens³⁸ from public debate on the grounds that they engage in message repetition would constitute a serious violation of their civic rights³⁹—rights that robots do not have.⁴⁰ Hence we should look instead for other ways to counteract the negative effects of message repetition, options that do not infringe on citizens’ freedoms or gravely stifle public debate.

One such solution is suggested by my earlier argument that the online architecture structuring public debate rewards impulsive rather than reflective behavior, which makes it unlikely that humans’ online message repetition is the result of much independent judgment. Furthermore, even if users can be aware of every single instance of message repetition occurring via retweeting or reposting, they may well be unable to keep track of all instances of message repetition they come across and to remember all claims and their firsthand sources as they have been disclosed to them. Both factors make people vulnerable to the exposure effects of message repetition, causing them to overestimate the epistemic weight of particular claims, simply because their frequency creates the illusion that they are widely endorsed, when in fact they are not.

To remedy this, we could simply disable retweeting or reposting, thereby removing one all-too-easy way for citizens to clone others’ claims. Or we could flag to the audience those users whose content contains a large proportion of repeated claims: just as we display the number of followers, for example, we could display the proportion of retweets/reposts of the user’s total number of tweets/posts. At the very least, this would put their readers on notice that their content might well be unreliable. This might also disincentivize citizens from behaving non-

independently, insofar as a large number of reposts or retweets would signal to the audience that a user does not post original content and therefore might not be worth “following.” Displaying the count of repeated messages might thus undermine the profiles of agents who are incorrectly perceived by the masses as “opinion leaders”—whose, by definition, independent judgment deserves particular attention—but who in fact are “followers” themselves, because most of their content comprises retweets or reposts.

Removing Claims. The second exclusionary approach to counteracting clone claims would be to simply expunge such claims from the deliberative environment. In the online space, this would require removing certain contributions.⁴¹ Just like the previous approach, this approach would also impinge on the democratic value of freedom of expression, although less so, because what is being banned is the post or claim rather than the person making it. We might find even that practice problematic at first, on grounds of free speech. But perhaps it might appear less problematic if we came to see the contamination of public discourse through the promulgation of clone claims as being epistemically analogous to “falsely shouting fire in a crowded theater.”

Some restrictions on freedom of speech are considered legitimate by courts and political theorists alike. In the judgment just alluded to, Supreme Court Justice Oliver Wendell Holmes argued that limitations on free speech are legitimate, where speech is bound to have dangerous consequences.⁴² J. S. Mill goes further, arguing that if one’s statements (true or not) will cause harm to others, then it is acceptable for the state to limit one’s liberty to express them at least in certain circumstances.⁴³ Clone claims could be viewed in an analogous light.

Even where their content is true, clone claims are nonetheless deceptive. They mislead the public into attaching more confidence to those claims than is epistemically warranted, given that they are copycat claims. Corrupting other people’s judgment—or even worse, a nation’s collective judgment—constitutes a grave moral and epistemic harm. Depending on how debilitating its effects are for citizens and their interests, we might consider the making and propagation of clone claims epistemically analogous to causing bodily harm, and we might, on that basis, justify limitations on the freedom of such speech in special cases as well.

Conclusion

Completely non-independent epistemic agents like bots can have an important impact on public opinion. First, half of web traffic comes from bots, and up to dozens of millions of Facebook and Twitter accounts are bot accounts (Confessore et al. 2018; Lafrance 2017). Second, individuals are vulnerable to message repetition and likely to overestimate the epistemic weight of repeated messages;

they are also poorly equipped to distinguish a clone claim from an independent one. Third, a substantial proportion of the U.S. population gets their news from social media platforms: as of August 2017, 67% of Americans reported that they get “at least some of their news on social media,” with Twitter in particular increasing its usership by 15 percentage points (Shearer and Gottfried 2017). Combining these findings there is a plausible reason for thinking that completely non-independent epistemic agents like bots could undermine internal deliberations to an extent that might actually flip an election. They would do it by artificially “boosting” the credibility of some claims, leading citizens to overestimate their epistemic weight.

There is clearly a case for banning bots from political discussions, but the reason for doing so—the epistemic damage that can be done to people’s judgment when they are unwittingly exposed to message repetition—extends well beyond bots. Plenty of message repetition occurs in human interactions as well. To know what to make of what others tell them, people need to know where they got their information and whether it is something they have already taken duly into account. New norms and policies governing discursive practices are required, particularly given the way in which online interaction makes it increasingly difficult for people to ascertain what they need to know to accurately assess the information they are receiving.

Notes

- 1 By “epistemic independence” I mean “independence conditional on all the common causes” (Dietrich and Spiekermann 2013). See my extended discussion in the subsection “Common Causes, Independent Assessments.”
- 2 Both recent U.S. and French presidential elections, as well as the Brexit referendum, have been plagued by the deployment of robot propaganda, as the Computational Propaganda Research Project at Oxford has revealed. See <https://www.oii.ox.ac.uk/research/projects/computational-propaganda/>.
- 3 “Discounting” may mean not listening at all to those propagating clone claims or simply not taking these claims seriously into account when making up our minds. Such “mental exclusions”—that are importantly different from excluding these people from the democratic conversation altogether—are justified epistemically. I elaborate on these distinctions later.
- 4 Directly, anyway, although lower individual competence ultimately leads to lower collective competence, of course.
- 5 Focusing on as it does on the aggregation of votes, the Condorcet Jury Theorem (CJT) is unsuited to exploring preelection interactions among voters who must decide what to believe and hence how to vote. Some voters’ non-independence may drive

down the competence levels of the rest, as the latter update their beliefs in light of the views of the former, not realizing those views are not independent of one another. In that scenario, the higher the proportion of non-independent voters in the electorate, the larger the drop in competence levels for the rest.

- 6 Notice also that in the CJT setup, relationships of epistemic non-independence are problematic because they effectively reduce the number of competent voters—the law of large numbers being the powerhouse driving the CJT’s results. Here I am instead focusing on how, in Bayesian updating, epistemically non-independent message repetition artificially inflates the number of ostensibly independent sources, which can undermine the judgments and thereby competence of each voter.
- 7 There is thus a trade-off between independence and competence, and it is an open empirical question whether losses in the former dimension are compensated fully by gains in the latter.
- 8 Goodin and Spiekermann (2018, 54–62, 164–77) show, for example, that the CJT is robust against everyone following the same opinion leader to a certain degree or different people following different opinion leaders to a greater degree.
- 9 Black (1958, 163) and Miller (1992, 56) argue that the CJT is inapplicable to democratic elections for this reason. Maybe it is, insofar as every vote is based on a combination of fact and value propositions. But among the multitude of beliefs that stand behind each vote, at least some are purely factual. It is those latter beliefs that are of interest here.
- 10 See Goodin (1993) for an argument about volitional (independence of will) rather than epistemic independence.
- 11 If there are a million voters, but each follows one and the same opinion leader, then there are not a million independent sources of judgment but merely one—and the probability of the majority being correct is just the probability of that one opinion leader being correct (Goodin and Spiekermann 2018, 54).
- 12 In Nozick’s (1981, 4) words, a good argument “forces someone to a belief”; for example, by pointing out that if one endorses premises *a*, *b*, and *c*, then one must also endorse conclusion *d*. But while the force of an argument comes from its use of universally accepted rules of logic, an argument’s premises may be based on facts and hence be nonetheless open to contestation. This is why the “betterness” of an argument may in fact be less self-evident than the Habermasian accounts make it look.
- 13 Maybe if we were a group of mathematicians deliberating over the correctness of a proof, all would recognize an error once one person pointed it out.

- Some models of deliberation presume something like that (e.g., Page's [2007] "rugged landscape" search model).
- 14 Notice that this is very different from saying, after the fashion of the Condorcet Jury Theorem, that I update my beliefs because I think that the majority among a large number of independent thinkers is highly likely to be correct. In my Bayesian framework, we should update our beliefs in light of each piece of new, independent evidence, whether it is in the majority or minority.
 - 15 In epistemology, Bayesianism has become the standard theory of belief change. It focuses less on how we should form our initial beliefs and more on how we should revise them in light of new evidence.
 - 16 Either because we update our credence in a claim every time the same claim is asserted, taking each assertion as a "new" piece of evidence in multiple updating rounds, or because we notice at once that multiple people make the same claim, and hence we attach a higher conditional probability to that claim being true in one updating round.
 - 17 That is, if they come from a bot or from a real person repeating another agent's (a person or a bot) claim.
 - 18 For a different use of "epistemic appropriation," see Davis 2018.
 - 19 The issue of epistemic malevolence is a separate problem that has been addressed by other scholars already. See, for example, Jason Stanley's (2015) discussion of propaganda and virtue ethics accounts of epistemic malevolence, discussed in Cassam (2019).
 - 20 I elaborate on this in the subsection immediately below. But briefly, we are "just" repeating when our repeated assertion does not contain any new information that might be useful to the listener, over and beyond the information already contained in the original assertion the listener has already taken into account.
 - 21 And suppose that is not out of respect for your epistemic credentials.
 - 22 At least in principle: boundedly rational agents who receive a plethora of such messages might not be able to remember the identities of all the sources that they encounter.
 - 23 For a detailed discussion of the ethics of social media platforms and of how social media architectures exploit emotions and impulses rather than rewarding reflective judgment, see Lewis 2017.
 - 24 Some companies like Twitter have started taking legal action against the use of bots and are using machine-learning programs to fight them (Finger 2015). The uncertain geographic origin of bots means also that they can easily be used by foreign powers to intervene in democratic decision-making processes of another state (Persily 2017, 70).
 - 25 The acronym "GNGO"—governmental nongovernmental organization—has been coined for such organizations.
 - 26 While some may say this nonetheless constitutes an interference with the freedom to speak or not to speak, surely requiring (further) speech is far less objectionable than prohibiting speech.
 - 27 As mentioned earlier, persuasion bias refers to people's inability to discount information repetition.
 - 28 That would be so anyway if the content of every communication contained the residue of all previous communications with other agents—which of course it may not.
 - 29 Thank you to a reviewer for pressing me on this issue, which led me to strengthen my argument at this point.
 - 30 On the face of it, the U.S. president's retweets do not pose a danger for double-counting, since such retweets are obvious message repetition and hence easily discounted. The larger problem with these retweets is that they expose huge audiences to the original messages of other agents who may seem epistemically independent but are not. They may, for example, be simply promulgating falsehoods crafted by bot masters, the same falsehoods that their many other bots are also promulgating.
 - 31 The analogy is to the duty imposed on banks to "know their customers" to avoid inadvertently laundering terrorists' funds (see Dhar 2017; Geltzer and Kupchan 2018).
 - 32 The term refers to the use of "fake grassroots campaigns that create the impression that large numbers of people are demanding or opposing particular policies" (Monbiot 2011). See also Monbiot 2010.
 - 33 The use of big data by companies such as Cambridge Analytica only serves to make us even more vulnerable targets of bot attacks, insofar as they can reveal to which type of content we are more likely to be responsive (Monbiot 2017).
 - 34 As others have already noted, Twitter's new conditions would arguably ban President Trump (Meyer 2017; Parkinson 2017).
 - 35 Similarly, others have suggested outlawing using bots altogether (Wu 2017).
 - 36 Confessore et al. 2018 referring to a study by Varol et al. 2017 from the University of Southern California and Indiana University. As of April 13 2017, Facebook discovered more than 30,000 fake accounts in France alone, a number that it expects to grow (Weedon, Nuland, and Stamos 2017).
 - 37 It may also be less clearly warranted epistemically. After all, human agents engaging in message repetition may be exercising some independent judgment when retweeting or reposting, in a way that nonhuman users doing the same thing cannot.

- 38 Perhaps banning foreigners would be permissible. But foreign intervention in our elections is a separate issue, which ought to be treated separately; and I bracket it for purposes of the present article. This article is specifically concerned with the epistemic costs of double-counting clone claims. From an epistemic perspective, the costs of clone claims are the same regardless their origin (i.e., whether they come from fellow citizens or foreigners). I make no distinction between them in this article.
- 39 Perhaps we might think that online platforms, as privately owned rather than public spaces, would be immune to such objections. We might think that just as private shopping centers may remove protesters and picketers from their grounds, online platforms should be able to regulate the speech of their users (or even close down citizens' accounts). Notice however that the courts' decisions on whether private shopping centers owners can legitimately remove protesters and activists on their property are mixed. Although some courts have in the last years asserted private owners' rights to do so, and even the state of California narrowed the applicability of *Pruneyard Shopping Center v. Robins*, 447 U.S. 74 (1980), that decision has yet to be completely overturned. Given this mixed record in the case of private shopping centers, online platforms might well face similar charges of violating constitutional rights of free speech when closing down citizens' accounts. But even if private ownership rights might not be a strong enough ground for flouting freedom of speech, however, other grounds might be—as I argue in the next subsection. I thank a reviewer for raising this point about regulating free speech in private shopping malls.
- 40 Might their masters have any right to speak in this deceptive way, through robots? A U.S. Supreme Court that treats business corporations as persons and money as a form of speech may deem that they do. But that is simply a perversion of U.S. constitutional jurisprudence, widely recognized (even by lawyers within the United States) as such.
- 41 Notice that social media platforms like Facebook already reserve the right to remove posts containing hate speech, graphic violence, nudity, or pornography.
- 42 *Schenck v. United States*, 249 U.S. 47 (1919), <https://www.law.cornell.edu/supremecourt/text/249/47>.
- 43 In the examples he discusses, Mill refers to bodily harm (1879, pp. 100–101).

References

- Beerbohm, Eric. 2012. *In Our Name*. Princeton, NJ: Princeton University Press.
- Black, Duncan. 1958. *The Theory of Committees and Elections*. Cambridge: Cambridge University Press.
- Brennan, Jason. 2011. *The Ethics of Voting*. Princeton, NJ: Princeton University Press.
- . 2016. *Against Democracy*. Princeton, NJ: Princeton University Press.
- Caplan, Bryan. 2007. *The Myth of the Rational Voter*. Princeton, NJ: Princeton University Press.
- Cassam, Quassim. 2019. *Vices of the Mind: From the Intellectual to the Political*. Oxford: Oxford University Press.
- Chauville, Roland. 2014. "The Universal Periodic Review's First Cycle: Successes and Failures." In *Human Rights and the Universal Periodic Review*, eds. H. Charlesworth and E. Larking, pp. 87–108. Cambridge: Cambridge University Press.
- Condorcet Marie Jean, Antoine Nicolas de Caritat, Marquis de. [1785] 1976. "Essay on the Application of Mathematics to the Theory of Decision-making." In *Condorcet: Selected Writings*, trans and ed. Keith Michael Baker, pp. 33–70. Indianapolis: Bobbs-Merrill.
- Confessore, Nicholas, Gabriel J. X. Dance, Richard Harris, and Mark Hansen. 2018. "The Follower Factory." *New York Times*, January 27. <https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html>.
- Davis, Emmalon. 2018. "On Epistemic Appropriation." *Ethics* 128(4): 702–27.
- DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel. 2003. "Persuasion Bias, Social Influence, and Unidimensional Opinions." *Quarterly Journal of Economics* 118(3): 909–68.
- Dhar, Vasant. 2017. "Should We Regulate Digital Platforms?" *Big Data* 5: 277–78.
- Dietrich, Franz. 2008. "The Premises of Condorcet's Jury Theorem Are Not Simultaneously Justified." *Episteme* 5(1): 56–73.
- Dietrich, Franz and Kai Spiekermann. 2013. "Independent Opinions? On the Causal Foundations of Belief Formation and Jury Theorems." *Mind* 122(487): 655–85.
- Estlund, David. 1989. "Democratic Theory and the Public Interest: Condorcet and Rousseau Revisited." *American Political Science Review* 83(4): 1317–22.
- Estlund, David. 1994. "Opinion Leaders, Independence, and Condorcet's Jury Theorem." *Theory and Decision* 36(2): 131–62.
- . 2008. *Democratic Authority*. Princeton, NJ: Princeton University Press.
- Finger, Lutz. 2015. "Do Evil—the Business of Social Media Bots." *Forbes*, February 17. <https://www.forbes.com/sites/lutzfinger/2015/02/17/do-evil-the-business-of-social-media-bots/#33bd1618fb58>.
- Geltzer, Joshua and Charles Kupchan. 2018. "What Counterterrorism Can Teach Us about Thwarting Russian Disinformation." *Washington Post*, February

22. https://www.washingtonpost.com/news/democracy-post/wp/2018/02/22/what-counterterrorism-can-teach-us-about-thwarting-russian-disinformation/?hpid=hp_no-name_opinion-card-d%3Ahomepage%2Fstory&utm_term=.ab921bb3b52e.
- Goodin, Robert E. 1993. "Independence in Democratic Theory: A Virtue? A Necessity? Both? Neither?" *Journal of Social Philosophy* 24(2): 50–56.
- . 2000. "Democratic Deliberation Within." *Philosophy and Public Affairs* 29(1): 81–109.
- Goodin, Robert E. 2001. "Consensus Interruptus." *Journal of Ethics* 5(2): 121–31.
- Goodin, Robert E. 2003. *Reflective Democracy*. Oxford: Oxford University Press.
- Goodin, Robert E. and Simon Niemeyer. 2003. "When Does Deliberation Begin? Internal Reflection versus Public Discussion in Deliberative Democracy." *Political Studies* 51(4): 627–49.
- Goodin, Robert E. and Kai Spiekermann. 2018. *A Theory of Epistemic Democracy*. Oxford: Oxford University Press.
- Gorwa, Robert and Douglas Guilbeault. 2018. "Understanding Bots for Policy and Research." Paper presented at ICA 2018. <https://arxiv.org/abs/1801.06863>.
- Grofman, Bernard and Scott Feld 1989. "Democratic Theory and the Public Interest: Condorcet and Rousseau Revisited." *American Political Science Review* 83(4): 1328–40.
- Guilbeault, Douglas and Samuel Woolley. 2016. "How Twitter Bots Are Shaping the Election." *The Atlantic*, November 1. <https://www.theatlantic.com/technology/archive/2016/11/election-bots/506072/>.
- Hume, David. [1739] 1896. *A Treatise of Human Nature*. Oxford: Clarendon Press.
- LaFrance, Adrienne. 2017. "The Internet Is Mostly Bots." *Atlantic*, January 31. <https://www.theatlantic.com/technology/archive/2017/01/bots-bots-bots/515043/>.
- Landemore, Hélène and Hugo Mercier. 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." *Analise Social* 47(205): 910–34.
- Lehrer, Keith. 1976. "Rationality in Science and Society: A Consensual Theory." In *Contemporary Aspects of Philosophy*, ed. G. Ryle, pp. 14–30. London: Oriel Press.
- . 2001. "Individualism, Communitarianism and Consensus." *Journal of Ethics* 5(2): 105–20.
- Lehrer, Keith and Carl Wagner. 1981. *Rational Consensus in Science and Society*. Dordrecht: D. Reidel.
- Lewis, Paul. 2017. "'Our Minds Can Be Hijacked': The Tech Insiders Who Fear a Smartphone Dystopia." *Guardian*, October 6. <https://www.theguardian.com/technology/2017/oct/05/smartphone-addiction-silicon-valley-dystopia>.
- List, Christian, Robert C. Luskin, James S. Fishkin, and Iain McLean. 2013. "Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls." *Journal of Politics* 75(1): 80–95.
- Meyer, Robinson. 2017. "Does Twitter's New Hate Speech Policy Cover Trump's North Korea Tweet?" *Atlantic*, December 18. <https://www.theatlantic.com/technology/archive/2017/12/the-trump-exception/548648/>.
- Mill, John Stuart. 1879. *On Liberty and the Subjection of Women*. New York: Henry Holt.
- . [1861] 1977. *Considerations on Representative Government*. Pp. 371–577 in *The Collected Works of John Stuart Mill*. Vol. XIX, *Essays on Politics and Society Part II*, ed. John M. Robson. Toronto: University of Toronto Press.
- Miller, David. 1992. "Deliberative Democracy and Social Choice." *Political Studies* 40(1) 54–67.
- Monbiot, George. 2010. "These Astroturf Libertarians Are the Real Threat to Internet Democracy." *Guardian*, December 13. <https://www.theguardian.com/commentisfree/libertycentral/2010/dec/13/astroturf-libertarians-internet-democracy>.
- . 2011. "The Need to Protect the Internet from 'Astroturfing' Grows Ever More Urgent." *Guardian*, February 23. <https://www.theguardian.com/environment/georgemonbiot/2011/feb/23/need-to-protect-internet-from-astroturfing>.
- . 2017. "Big Data's Power Is Terrifying. That Could Be Good News for Democracy." *Guardian*, March 6. <https://www.theguardian.com/commentisfree/2017/mar/06/big-data-cambridge-analytica-democracy>.
- Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, MA: Harvard University Press.
- Oreskes, Naomi and Erik M. Conway. *Merchants of Doubt*. London: Bloomsbury Press.
- Page, Scott. 2007. *The Difference*. Princeton, NJ: Princeton University Press.
- Parkinson, Hannah Jane. 2017. "Donald Trump Breaks Twitter's Rules, So Why Not Ban Him?" *Guardian*, October 29. <https://www.theguardian.com/commentisfree/2017/oct/28/trump-breaks-twitthers-rules-so-why-not-ban-him>.
- Persily, Nathaniel. 2017. "Can Democracy Survive the Internet?" *Journal of Democracy* 28(2): 63–76.
- Phillip, Abby. 2017. "The Curious Case of 'Nicole Mincey,' the Trump Fan Who May Actually Be a Bot." *Washington Post*, August 7. https://www.washingtonpost.com/politics/the-curious-case-of-nicole-mincey-the-trump-fan-who-may-actually-be-a-russian-bot/2017/08/07/7aa67410-7b96-11e7-9026-4a0a64977c92_story.html.
- Plato. 2006. *The Republic*. New Haven: Yale University Press.
- Romm, Tony and Elizabeth Dwoskin. 2018. "Facebook Says It Removed a Flood of Hate Speech, Terrorist Propaganda and Fake Accounts from Its Site."

- Washington Post*, November 15. https://www.washingtonpost.com/technology/2018/11/15/facebook-says-it-removed-flood-hate-speech-terrorist-propaganda-fake-accounts-its-site/?utm_term=.b06ef9e8d354.
- Rothbart, M., S. Fulero, C. Jensen, and J. Howard, P. Birrell. 1978. "From Individual to Group Impressions: Availability Heuristics in Stereotyping Formation." *Journal of Experimental Social Psychology* 14(3): 237–55.
- Schribman, David. 1982. "Sincerely Yours, Your Congressman's Computer." *New York Times*, August 17. <http://www.nytimes.com/1982/08/17/us/sincerely-yours-your-congressman-s-computer.html>.
- Shearer, Elisa and Jeffrey Gottfried. 2017. "News Use across Social Media Platforms." *Pew Research Center*, September 2017. <http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/>.
- Stanley, Jason. 2015. *How Propaganda Works*. Princeton, NJ: Princeton University Press.
- Timberg, Craig, Elizabeth Dwoskin, and Adam Entous. 2017. "Michael Flynn, Nicki Minaj Shared Content from this Tennessee GOP Account. But It Wasn't Real. It Was Russian." *Washington Post*, October 18. https://www.washingtonpost.com/business/technology/michael-flynn-nicki-minaj-shared-content-from-this-tennessee-gop-account-but-it-wasnt-real-it-was-russian/2017/10/18/8b92fcda-b435-11e7-9e58-e6288544af98_story.html?utm_term=.720d07577ce6.
- Varol, Onur, Emilio Ferrara, Clayton A. Davis, Filippo Menczer, and Alessandro Flammini. 2017. "Online Human-Bot Interactions: Detection, Estimation, And Characterization." International AAI Conference on Web and Social Media (ICWSM), May 2017, Montreal. <https://arxiv.org/abs/1703.03107>.
- Waldron, Jeremy. 1989. "Democratic Theory and the Public Interest: Condorcet and Rousseau Revisited." *American Political Science Review* 83(4): 1322–28.
- Weaver, Kimberlee, S. M. Garcia, N. Schwarz, and D. Miller. 2007. "Inferring the Popularity of an Opinion from Its Familiarity: A Repetitive Voice Can Sound like a Chorus." *Journal of Personality and Social Psychology* 92(5): 821–33.
- Weedon, Jen, William Nuland, and Alex Stamos. 2017. "Information Operation and Facebook." Facebook public release, version 1.0, April 27. <https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf>.
- White, Jonathan and Lea Ypi. 2016. *The Meaning of Partisanship*. Oxford: Oxford University Press.
- Wu, Tim. 2017. "Please Prove You're Not a Robot." *New York Times*, July 15. <https://www.nytimes.com/2017/07/15/opinion/sunday/please-prove-youre-not-a-robot.html?mcubz=0>.