Joseph Heath*

Post-deliberative Democracy

Abstract: Within any adversarial rule-governed system, it often takes time for strategically motivated agents to discover effective exploits. Once discovered, these strategies will soon be copied by all other participants. Unless it is possible to adjust the rules to preclude them, the result will be a degradation of the performance of the system. This is essentially what has happened to public political discourse in democratic states. Political actors have discovered, not just that the norm of truth can be violated in specific ways, but that many of the norms governing rational deliberation can also be violated, not just without penalty, but often for significant political gain. As a result, the level of noise (false or misleading communications) has come to drown out the signal (earnest attempts at deliberation). The post-truth political condition is the cumulative result of innovations developed by actors who adopt an essentially strategic orientation toward political communications.

Keywords: deliberative democracy, post-truth, communications strategy, social media

During the heady days of the early internet, the U.S. Defense Advanced Research Projects Agency (DARPA) became interested in the power of crowdsourcing as a way to resolve previously intractable intelligence problems. In 2011 the agency launched the 'Shredder Challenge,' in which they provided images of five hand-written documents that had been sliced, both vertically and horizontally, by high-quality paper shredders, offering a large cash prize to any group that could reconstruct all five documents. One of the late entrants to the contest was a team of researchers at the University of California San Diego (UCSD), who instead of trying to solve the problem algorithmically, created an online platform where members of the public could log in and work on the puzzle at their leisure, in return for a small share of the expected reward for each piece successfully placed. The platform quickly attracted thousands of users, who proceeded to reassemble

_____

*Corresponding author: Joseph Heath, Department of Philosophy, University of Toronto, Toronto, Canada, E-mail: joseph.heath@utoronto.ca

the first few documents, rapidly advancing the group to second place in the overall contest.

And yet, just as it looked as though the UCSD team was on track to win the prize, the project suddenly stalled. The platform they built had been fully open, based on the assumption that users would make a good-faith contribution to solving the puzzle. And yet when the team began to investigate the reasons for the stall, they found evidence of intentionally disruptive behavior—users not just undoing portions of the puzzle that had been solved, but also stacking pieces on top of one another, or moving them to the far edges of the board where they were difficult to find. An alert was put out to the community, who rallied to undo the damage, and began to monitor the project in real time. Their efforts proved fruitless. Hundreds of users quit in frustration, contributions declined, recruitment stalled, and the UCSD group wound up losing the challenge.

Years later, computer scientists studying the project data logs made a surprising discovery (Stefanovitch et al. 2014). Although the attackers claimed to have recruited a troll army from 4chan, in reality almost all of the damage had been inflicted by just two individuals. With a bit more detective work, it was discovered that one had been a member of the team that won the challenge. In a later interview with a journalist, the perpetrator expressed no regrets, pointing out that he had merely exposed a weakness in the UCSD model (Harris 2015). In a competitive interaction, it would be foolish to assume that everyone will be equally committed to the success of each project. This was the conclusion that DARPA drew as well. The radical asymmetry that was revealed—the fact that it took only two people to undermine the work of thousands—was so disconcerting that it led the agency to abandon the entire crowdsourcing research program.

This story is one that should be of considerable concern to proponents of deliberative democracy, particularly those who characterize the broader public sphere as an important deliberative forum (e.g. Chambers 2012). In standard discussions of the wisdom of crowds, it is commonly assumed that participants will make a good-faith contribution to the deliberative enterprise (e.g. Landemore 2012). And yet this is plainly not realistic, especially given that democracy, like the Shredder Challenge, is a competitively structured interaction. The issue is not just that individuals may lie or misrepresent their preferences. In many cases, there will be groups who can advance their interests by disrupting deliberation (or by blocking 'democratic will-formation'), so that no collective action can be mobilized. The experience of the Shredder Challenge illustrates a general principle, which is that it takes only a few disruptive actors (or 'trolls'), acting strategically, to prevent any orderly process of collective problem-solving from occurring. Furthermore, the development of social media

has clearly potentiated this capacity for disruptive intervention in the public sphere.

Since the election of Donald Trump to the U.S. Presidency in 2016, it has become common to describe the current era as one of 'post-truth' politics (Davis 2017). I would like to propose the following analysis of the phenomenon. Within any adversarial rule-governed system, it often takes time for strategically motivated agents to discover effective exploits. Once discovered, these strategies will soon be copied by all other participants. Unless it is possible to adjust the rules to preclude them, the result will be a general degradation of the performance of the system (Heath 2014, 177–185). This is essentially what has happened to public political discourse in democratic states over the past century. Political actors have discovered, not just that the norm of truth can be violated in specific ways, but that many of the norms governing rational deliberation and debate can also be violated, not just without penalty, but often for significant political gain. And because it is impossible to rule out these strategies, they have become universally adopted. As a result, the level of noise (in the form of false or misleading communications) has come to drown out the signal (earnest attempts at deliberation).[1] As I will attempt to show, the post-truth political condition is the cumulative result of innovations developed by actors who adopt an essentially strategic orientation toward political communications. Since the prospect of a return to the earlier set of norms is slim to non-existent, we must confront the fact that open public deliberation on matters of general concern has become impossible. The interesting questions of democratic governance now involve the design of institutions for a post-deliberative age.

1 The Dark Arts

Proponents of deliberative democracy have never been as concerned as they should be about the problem of strategic behavior. If one glances at the rival public choice literature, or the work done on aggregative models of democracy, one of the first questions that is asked, with respect to any voting procedure, is whether it is 'strategy-proof' (Barberà 2001). Models of preference aggregation are typically presented and analyzed first under ideal conditions, in which individuals vote their preferences honestly (and myopically). This is immediately followed by

---

1 The claim that political discourse has become 'post-truth' should not be taken to imply that citizens have ceased to care about truth, or that there has been a decline in their level of concern. The claim is that public discourse and deliberation has been disrupted, so that truth is no longer a prevailing norm. As the Shredder Challenge shows, even if thousands of people still want to solve the puzzle, it is still possible to disrupt their efforts.

a strategic analysis, in which the possibility of individuals *misrepresenting* their preferences is introduced. In certain multi-stage voting models, for example, it can make sense to vote for an amendment that one does not genuinely support, if it will foreseeably result in the defeat of the final motion, which one would like to see fail. With a strategy-proof procedure, this sort of advantage can only be obtained if some participants vote honestly while others vote strategically, but if everyone votes strategically, the effects are neutralized, and so the outcome is the same as if everyone voted honestly. Strategy-proofness is considered an important desideratum in a voting procedure.

Deliberative democrats have, with a few notable exceptions (e.g. List and Dryzek 2003), refrained from asking this question about their own models (Warren 2007, 284). Partly this is because deliberation is not actually a decision procedure, in the vast majority of cases, and so it must be coupled with a voting procedure in order to produce a decision. But mainly it is because the answer is so obviously negative. A situation in which some people lie and dissemble is going to be disruptive in any deliberative discourse, and the situation can hardly be remedied by having everyone lie. Furthermore, the range of options that are available to participants in a debate are much greater than to voters (who are generally constrained by what is on the ballot). For example, Jürgen Habermas includes a prohibition on lying as a rule of practical discourse, but also includes the stipulation that "everyone is allowed to introduce any assertion whatever into the discourse" (Habermas 1990, 89). The latter opens the door to the strategy, commonly used to derail deliberation, of constantly bringing up new issues in order to run down the clock and prevent the group from reaching agreement. This serves as an example of how, even without breaking the rules laid down by deliberative theorists themselves, strategically oriented actors can behave in ways that will undermine the effectiveness of any deliberative procedure.

There are two central ways in which deliberative democrats have sought to overcome this difficulty. The first is by distancing themselves from the claim that the democratic public sphere could serve as a locus of meaningful deliberation, and instead focusing attention on institutionally constrained deliberative fora, such as citizens' assemblies or specially constituted 'minipublics' (Fung 2007; Smith and Setälä 2018). The second is by positing the existence of deliberative 'systems,' in which individual contributions need not respect the usual norms of rationality or truth, but where the net effect of the various components working in concert may be to achieve an outcome that approximates deliberation (Mansbridge et al. 2012). I shall focus my discussion on the latter proposal, because it is the only one that is relevant to the broader public sphere. Unlike earlier work on deliberative democracy (e.g. Gutmann and Thompson 2004), which tended to be idealistic about the quality of public debate, or excessively

normative in demanding that citizens exhibit a high level of civic virtue, the deliberative systems approach at least recognizes the importance of such features of political debate as partisanship and its associated rhetorical devices. At the same time, proponents of this perspective have barely begun to articulate the conditions that must be satisfied in order for the various parts of a deliberative system to work together to produce a beneficent outcome, much less to specify the mechanism through which the desirable properties of such systems are maintained (Warren 2007, 278). While it is true that the effects of partisan distortion may be canceled out by countervailing partisan distortions, it is also possible that the two tendencies will exacerbate one another, producing a race to the bottom that pushes public communication further and further from deliberative norms. Indeed, one might easily think that the latter represents the *default* outcome. Thus the deliberative systems approach more closely resembles a promissory note than a fully specified analysis. In the discussion that follows, I will present various reasons for thinking that this note cannot be redeemed.

My central claim is that the post-truth political environment is the product of precisely the sort of race to the bottom that a deliberative system must prevent. Of course, if one begins with an overly idealized conception of deliberation, it is not difficult to show that any open public process will fall short of the ideal in various ways. Yet even if one starts with extremely minimal formulations of the relevant norms, it is not difficult to see how current trends have led to their systematic violation. With respect to truth, for instance, while it would be too much to demand that everyone's public pronouncements be aimed at no other objective, it is reasonable to expect that, upon being caught saying something untrue, agents feel some obligation to retract their claim, or that the negative consequences are sufficient to deter them from repeating it. Similarly, with respect to norms of rationality, it need not be the case that all public actors be held to forensic standards of justification, but it is reasonable to expect that, when pressed on a claim, speakers feel obliged to offer something by way of support, bearing some intelligible connection to the contested claim. The concern that political discourse has become post-truth is based on the concern that even minimal standards such as these are routinely violated with impunity.

Although some have tried to blame the decline of truth and rationality on various currents of academic thought in the 20th century, such as postmodernism, which emphasized the contingent and socially constructed nature of all specific claims to truth (Keyes 2004; Wight 2018), this places far too much power in the hands of academics, not to mention positing hidden channels of influence. The more plausible explanation is the more obvious one – that the post-truth political environment is a direct and immediate consequence of the influence of strategic communications on political discourse, starting with advertising, and

then extending to include debate performance, public appearances, fundraising, and ultimately lawmaking itself. This is not mere conjecture; communications consultants are perfectly forthright in the claims that they make, and politicians make no secret of the role that these strategists play (Lees 2014). Consider Frank Luntz's influential book, *Words That Work* (Luntz 2007). Luntz takes inspiration from the observation, which he takes to reveal a deep truth, that "people will forget what you say, but they will never forget how you made them feel" (Luntz 2007, 18). The implication is that communications should aim at what comedian Stephen Colbert dubbed 'truthiness'—saying things that *feel* true—rather than truth per se. The insistence on truth, in Luntz's view, reflects a simple-minded literalism, which is deeply inconsistent with the requirements of effective communication.

Thus the key to understanding the post-truth condition is to analyze communications strategy. Communications, however, cannot be understood apart from an understanding of media. Unlike deliberative democrats, for whom the model of face-to-face debate remains paradigmatic, strategic communications takes seriously the fact that the preponderance of political communication occurs through a medium, traditionally print, radio or television, but increasingly through an internet platform (the set of which are often referred to as 'social media'). Each medium is, to a certain degree, unnatural, which enables certain strategies of persuasion to be successful that would not be effective in the naturalistic setting of face-to-face communication. While Marshall McLuhan's claim that 'the medium is the message' is something of an exaggeration, it is most certainly the case that different strategies work better, and therefore different content is favored, in different media.

Perhaps the most important feature of traditional media is that, with the exception of paid advertising, access is controlled through gatekeepers, viz. journalists and editors, and so a great deal of traditional communications strategy involves manipulating, circumventing, or undermining these gatekeepers. The biggest change with the transition to social media is that the gatekeepers have been removed, making it possible for political actors to communicate directly with citizens, and also citizens to communicate directly with one another. This calls for an entirely different set of strategies, the contours of which are only beginning to emerge, but none of which seem to favor the triumph of rationality and truth.

The most important medium of political communication, over the past 60 years at least, has been television. Although the novelty of social media, and the fact that these platforms favor different content, has attracted considerable attention and commentary, they remain less important than television (although social media now easily eclipse print) (Mitchell et al. 2020). Nevertheless, media consumption is strongly stratified by generation (a recent study showed that 48% of Americans aged 18–29 rely primarily upon social media for news, versus

only 3% of those over the age of 65), and so it is not difficult to foresee the day when it will become the dominant modality. Thus I will divide up my analysis of communications strategies into those that arose in the age of television, and the new ones that are emerging in the internet age.

2 The Age of Television

The 1960 televised debate between U.S. presidential candidates John F. Kennedy and Richard Nixon is usually identified as the turning point, at which the key medium in American politics switched from radio to television. The most salient feature of this transition was, of course, the addition of visuals to the audio stream, with the major surprise being the ease with which the power of these images overwhelmed the content of what was said. (Among the more immediate lessons, male politicians learned the importance of a fresh shave, along with the need to wear both makeup and powder.) In the long term, the tendency of television was therefore to crowd out linguistic discourse, as strategists discovered ways to communicate through the construction and presentation of effective images. One can see this clearly in the history of product advertising (Heath 2014, 201–205). And yet apart from squeezing out discourse, certain visual effects or modes of presentation also have an effect on the content of discourse. Diana Mutz (2015), for example, has presented fascinating research on how the development of the 'close up' shot has affected political communications.[2] In the discussion that follows, however, I will set aside the construction of visuals and focus entirely on the effects of television on the structure of linguistic communication.

The most common form of discourse on television is that of the 'interview' between a journalist and a political actor (or a 'scrum' involving multiple journalists). The major effects of this interaction format on the content of discourse are a consequence of *time compression* (Rosensteil et al. 2007, 35). Interactions with print journalists are reasonably leisurely – subjects have some time to think about what they want to say, and journalists have hours, and sometimes days, to think about how they want to present the material. Radio is significantly less leisurely. It allows individuals to speak at length, and permits considerable back-and-forth between speakers, but it does require rapid response (because of the abhorrence of 'dead air'). Television, by contrast, is extremely rushed. Interview

---

2 Her key observation is that the close up shot of a person's face is highly unnatural perspective on a stranger, and therefore triggers emotional responses more appropriate to intimate settings, requiring a consequent adjustment in the way that *public* communications must be carried out.

subjects have no time to think when formulating a response, must give extremely short answers, with extremely limited opportunity for back-and-forth between speakers. It is therefore (as anyone who has been interviewed on television can attest) an extremely unnatural form of discourse. Although it often has the formal structure of dialogue, it is much closer to performance than conversation. Most of the communications strategies that have been developed for use on television take advantage, in one way or another, of opportunities that are made available by these features of the medium.

2.1 Message Discipline

The first and most important principle of modern communications—whether it be private advertising or political speech—is repetition. Most members of the public exposed to a particular communication will not be paying careful attention. Furthermore, the majority of the population does not have a well-developed political ideology, or even what could be described as political beliefs (Converse 1964). What they have are a set of associations, between their feelings, particular recognizable individuals, and set of political phrases. The primary focus of political communications is to build up, through repetition, these mental associations. Most people do not grasp the propositional content of political claims, much less the inferential and logical relations between multiple propositions. If, however, every time they hear the phrase 'tax cut' they also hear the word 'Conservative party,' they will over time come to think of tax cuts whenever they hear about the Conservative party. If the phrase 'tax cut' has positive affective resonance, this will increase their propensity to support the Conservative party.

It is, of course, well understood that this is not how the entire population processes political information. In a political system with a relatively small number of ideologically distinct parties, however, it is an accurate representation of how the majority of persuadable voters process information, and in the standard run of cases this is the group that politicians are most interested in addressing. Furthermore, repetition has powerful effects even on individuals who engage in more sophisticated cognition, in part due to the phenomenon of 'source amnesia,' i.e. the fact that we often forget where we learned what we 'know'. As a result, merely hearing something again and again can cause people to believe it, regardless of whether any evidence for it has been put forward.

Politicians have therefore adopted a more-or-less rigorous adherence to the practice of deciding upon a message (often based on focus-group or public opinion research), and then repeating it in all communications, even when doing so violates conversational norms of relevance, and of course, truth. Indeed, it can be advantageous to make a false claim against one's opponent, because in order to

deny it that person will often feel compelled to repeat the key phrase, which serves to solidify the association in the minds of many (hence the mantra for politicians seeking to defend themselves: 'never repeat a negative'). Politicians will also tone down or cancel entirely the expression of any idea or opinion that is 'off message,' in order to ensure that the media has nothing to report other than the designated message. By making their speech boring, they can further reduce the attention level of the audience, increasing the effectiveness of message repetition.

2.2 Talking Points

Perhaps the most woeful development of the television age was the invention of the 'talking point' as a strategy for dealing with interviews. The traditional notion of the interview is that a journalist identifies a person of interest, asks that person questions, which the person then tries to answer. The interaction is often adversarial, in the sense that the questions are often not really attempts to discover new information, but rather an exercise in holding the person to account, by requiring them to defend their actions or decisions. The central limitation of television, however, as a medium for transmitting this interaction is that it is very difficult for the interviewer to follow up on a given question, because of time pressure. Thus the standard interview is one in which a journalist asks a question, listens to the response, and then moves on to the next question. (This is an even more pronounced feature of a scrum, in which individual journalists seldom have the opportunity to ask more than one question in sequence.) This is what led to the fateful discovery, on the part of interview subjects, that it is not actually necessary to answer the questions posed; it is possible to use the questions as a pretext for saying what one likes.

The development of talking points solves several problems at once for the interview subject. The idea is relatively simple. One begins by preparing, in advance, a list of points that one would like to make, formulated in simple, short phrases. These are the talking points. When asked a question during the interview, one tries to find some verbal bridge between a word used in the question and one of these talking points. One can then repeat the key word, in order to sound like one is answering the question, before going on to present the prepared phrase ('you mentioned x; I agree, and think it is also important to emphasize that. . . '). If the connection is less direct, one can instead employ a 'pivot,' which involves redirecting the response toward the talking point ('I know that there has been a lot of talk about x, but I think it's important not to lose sight of the issues that matter most to hard-working families. . . '). This strategy has numerous advantages. Because television offers no time to think up a response, it allows one to answer quickly and confidently, by drawing on previously prepared lines.

And because of the format, it is extremely unlikely that the journalist will draw attention to the non-responsiveness of the answer.

The effect of this strategy, when universalized, is a significant degradation of public discourse.[3] Most obviously, it generates a speaking style that the French refer to as *langue de bois* (or 'speaking with a wooden tongue'), which many people find highly alienating. (Perhaps the most important shared trait of populist politicians is that they avoid speaking this way. The public fails to respond negatively to their 'gaffes' or vulgarity, because these constitute breaches of the standard political communication style, and therefore feel more honest.[4]) More generally, the use of the pivot leads to a breakdown in the rational structure of conversational exchange and debate, since verbal contributions become increasingly unrelated to one another. Apart from directly violating norms of rationality, this contributes to erosion of the norm of truth, not just because it becomes impossible to press people on false claims, but because it becomes impossible to discern which claims are well-supported and which are not.

2.3 Wedge Issues

At the same time that communications strategy has evolved, political actors have also become a great deal more sophisticated about electoral strategy. In particular, improved understanding of voter behavior has had a significant impact on the way that politicians approach all aspects of communications, and even legislation. A crude approximation of the major discovery is that 'issues don't matter,' when it comes to determining who wins elections. A more accurate statement would be that 'most issues don't matter.' Although voters claim to be motivated by a range of concerns, their stance on most issues seldom drives their voting decisions. There are, however, a small number of issues that, for certain segments of the electorate, are important enough to generate vote-switching. Thus politicians have increasingly learned to focus their attention on these specific issues (Dumouchel, Ouellet, and Giasson 2021). To take a specific example, management of the health care system is by far the most important responsibility of the Canadian state, and voters routinely identify health care as the most important issue of concern to

---

3 To see how much has changed, it is interesting to watch interviews with athletes before the development of modern communications. Professional athletes are now given a half-dozen stock phrases and instructed not to deviate from the script. The result is that most 'interviews' have become a bizarre sort of performance, in which answers are completely devoid of content.

4 Many commentators were baffled by the fact that so many Trump supporters considered 'honesty' to be one of his great virtues. This is much less difficult to understand if one pays attention to his speaking *style*. In contrast to many other politicians, there can be little doubt that what Trump says is precisely what is going through his mind.

Canadians. More careful investigation, however, reveals that most voters do not see any effective difference between the major political parties, when it comes to their management or commitment to the health care system. As a result, it is not necessary for any political party to have general policy positions with respect to health care, because there are no votes to be won in this area. Despite its general importance, health care does not actually drive voter behavior, in the sense that no one actually votes for one party, rather than some other, because of their stance on health care. One issue, however, on which political parties are able to differentiate themselves in the domain of health policy is with respect to safe injection sites for IV drug users, an issue that is both controversial and polarizing. Because of this, the debate over safe injection sites—an issue that is relatively unimportant to the health of most Canadians—has easily eclipsed public discussion of wait times, cost containment or quality management in the system.

So while political philosophers tend to think of public policy as a broad domain, encompassing all activities of the state, politicians increasingly think in terms of a much smaller set of wedge issues, on which they are able to clearly differentiate themselves from the other parties (and thus, on which they are able to win votes). The major tactic in public communications is therefore to accentuate wedge issues that advantage one's own party (i.e. where the majority of voters, or persuadable voters, are closest to one's position), while deflecting or obfuscating the difference between oneself and one's opponent on wedge issues that advantage others. This is known as the 'swords and shields' strategy. The goal is retain wedge issues that are in one's interest (swords), while denying wedges to one's opponents (shields) (Wells 2015). It is the latter component of the strategy that is the most corrosive of public discourse, since it typically involves deliberately sowing confusion, so that low-information voters cannot tell, or cannot remember, which party supports which positions on a given issue. (For example, in Canada there actually are major differences between the federal political parties on environmental questions. And yet these would be difficult to discern from an examination of their communications, because all the parties sound alike. This is the result of a shield strategy aimed a neutralizing the power of environmentalism as a potential wedge issue.)

The shield strategy is the one that generates the most 'troll-like' behavior. Unlike lying, which involves trying to persuade others to believe a falsehood, the strategy here involves trying to generate confusion. For example, it has become common for right-wing parties to attack left-wing parties from the left, and vice versa, not because doing so expresses their actual stance, but because the net effect is to confuse inattentive or unintelligent voters about which party stands for what. Thus many political parties no longer make an effort to defend an ideologically consistent line or even to develop a comprehensive electoral platform.

Perhaps the most pernicious consequence of the overall strategy is that it truncates public discourse, so that the most important issues facing voters are seldom discussed, because politicians are constantly redirecting conversation toward a small set of wedge issues that they can derive electoral benefit from.

2.4 Overall Effect

Each of the enumerated strategies is, in its own way, corrosive of political discourse. But beyond that, they are also mutually complementary, and so can be combined in ways that are even more destructive. Once a set of wedge issues that favor one's party has been identified, for example, these can be used to derive a set of talking points, which can be used to redirect any interview toward those topics. These can, in turn, be repeated *ad nauseum*. And so, for example, political parties may issue a set of 'talking points for the week' to all elected representatives, political staff, and party workers, who will then be expected to bring them up in all public appearances (speeches, interviews, etc.).

These communications techniques, combined with an increasingly ritualistic partisanship, results in the development of a growing disconnect between the style of political discourse and that of ordinary conversation and debate. Intellectuals may become adapted to this communication style, losing track of how stilted and unnatural it has become. Furthermore, they may become skilled at 'reading the tea leaves,' in order to extract information from what they hear (e.g. detecting slight shifts in strategy, etc.) For the average citizen, however, the entire spectacle becomes almost completely incomprehensible. The most obvious consequence is that they lose interest in politics, and refuse to make the investments in attention and knowledge-acquisition that are necessary in order to become competent participants. Furthermore, as we have seen in recent years, many people accumulate resentment toward both purveyors of the discourse and those who are able to understand it, based on the sense of being excluded and condescended to. This in turn feeds the populist backlash that has become a prominent feature of electoral politics in several Western democracies.

3 Social Media

The most important feature of the transition to social media is the loss of gatekeepers. Most importantly, it has practically eliminated the role that *editors* used to play in determining what entered public communication, and in what form it entered. As Jill Lepore has observed, prior to the development of the internet, practically everything that was published had also been edited ("in the sense

that it passed through the hands of at least one person whose entire job was to consider the judiciousness and reasonableness of the argument and the quality of the evidence" [Lepore 2017]). By now the opposite situation prevails, where very little that the average person reads or views will have been edited. This is often heralded as a more 'democratic' mode of communication, and in certain respects it is. At the same time, the results are somewhat appalling from the standpoint of both rationality and civility. Indeed, it is difficult to avoid the impression that a great deal of the claims made on behalf of deliberative democracy were based on a failure to appreciate just how much was being filtered out of public discourse in the pre-internet age. For example, if one compares the 'letters to the editor' published in a traditional newspaper to the comment section in a modern, online version of the same paper, it is not difficult to see how editorial selection radically elevated both the tone and quality of commentary that was purported to be coming from the general public.

The second major consequence of the internet is that it destroyed the business model of many traditional media firms, especially those based on print. The old-fashioned print newspaper engaged in considerable cross-subsidization between sections, with the 'serious' content, such as news, features and editorial, being extremely expensive to produce, and yet generating the lowest levels of reader engagement, while the 'fluff' sections, such as sports and automotive, being inexpensive to produce but having the highest readership. Both readers and advertisers were given no choice but to purchase the bundled product. The internet disrupted this, not only by making specialty publications available for the majority of readers, who had no interest in news, but also by making reader engagement transparent, so that advertisers could see exactly how few people were actually reading the news stories. This made it increasingly difficult to subsidize news production, which in turn led to a decrease in the quantity produced (particularly in local newspapers, many of which have disappeared) as well as overall erosion of quality.

Thus at the same time that social media communications have been increasing in importance, traditional media have been in sharp decline. This has been more apparent in English-speaking countries, simply because English is the dominant language of internet communication outside the 'great firewall' of China. It should be noted that the decline of traditional media has been offset to some degree in countries where state-funded media play a larger role than in the U.S. It is important to observe, however, that the traditional strategies employed by public broadcasters, aimed at limiting the damage from the communications strategies outlined in the previous section, are of limited applicability when it comes to controlling internet-based communications. The fact that public discourse in many European countries has historically been protected by much more

powerful gatekeepers may result in the loss of power of these gatekeepers being felt even more acutely.

On social media, the central mechanism through which content is reproduced is one or another type of recommendation system. Content is published directly by users, and disseminated to a (typically) small group. Members of that group can then decide whether pass it along to other users on the network, either directly, or else indirectly by liking, endorsing, or recommending it. Other users are then shown the content, based on the endorsement of other users, or through an algorithm that displays content based on what those with similar interests and tastes liked. Each platform is slightly different, but they all generate rather similar bandwagon effects, where some content 'goes viral,' acquiring a vast readership/viewership that easily eclipses the reach of traditional media. The system is highly decentralized, and so the only way to achieve wide dissemination is by generating content that a large numbers of users will be motivated to pass along to others. This in turn favors a set of distinctive communications strategies, which have been employed by political actors with increasing success.

3.1 A/B Testing

With traditional communications, a great deal of time and energy was spent devising a message that would make it past the gatekeepers. This was possible in part because the gatekeepers were relatively few and fairly predictable in their judgment. Social media, by contrast, provides direct access to the public. This is challenging, because the public is a great deal more unpredictable, and the dynamics of propagation through network platforms is chaotic. If one considers, for example, the two most-followed Tiktok users at the time of writing—Charli D'Amelio (110 million followers) and Addison Rae (78 million)—it would be safe to say that no one could have predicted, ex ante, that these two individuals would ascend to such a position of fame and influence. It is difficult even to offer a convincing explanation ex post, other than to say that it is a bandwagon effect.

The internet, however, also provides a solution to this problem of unpredictability, in the form of A/B testing (Siroker and Koomen 2013). Rather than investing a great deal of time crafting a single message (or a single set of talking points), one can instead produce multiple messages (i.e. A and B), each one slightly different from the other, then put them all out in order to see which is the most effective. User behavior can be monitored in real time, with the ineffective content being progressively culled. In effect, the communications strategist, rather than trying to predict which message will be most successful in a particular media environment, constructs an evolutionary system that will discover the most

successful message through selection. Examples of the effects can be found in the 'clickbait' links that appear on many pages. Websites that repackage content will initially give an article a dozen or more different headlines and thumbnail images, which are randomly displayed in the links. Those that fail to generate user engagement (i.e. 'clicks') are gradually removed by the site's algorithm, until eventually only one headline and image is retained – the one that maximize engagement, and thus will be the most virulent (Marantz 2019, 94). The algorithm is obviously unconstrained by the norm of truth, but more importantly, it usually selects content that violates it, simply because 'what is true' seldom coincides precisely with 'what maximizes engagement.' One can easily verify this by clicking on some clickbait, in order to see that the headline usually misrepresents the story (or the thumbnail misrepresents the video).

Political actors have not been slow to adopt these techniques. They are commonly used in fundraising appeals, where every email sent or advertisement posted is essentially a randomized control trial, aimed at determining which message will elicit the greatest sum of donations. By far the most corrosive effect on public discourse, however, has been achieved not by political actors seeking electoral advantage, but rather by private actors motivated by pecuniary interests. The term 'fake news' was originally coined in order to refer to news stories, mainly political in focus, that were fabricated entirely, with the objective of generating advertising revenue by driving high volumes of traffic to websites set up for that purpose (Allcott and Gentzkow 2017). For example, during the 2016 American electoral campaign, a small town in Macedonia became the epicenter of fake news dissemination, with over 100 websites serving up fabricated stories about American politics (Hughes and Waismel-Manor 2021). Again, the procedure involved putting out hundreds of stories, then gradually culling those that were less successful. Numerous purveyors of fake news claim to have produced as much left-wing as right-wing content, but the right-wing content came to predominate simply because it generated greater engagement (Hughes and Waismel-Manor 2021, 22). The results were dramatic. The fact that the stories were false is an important part of the business model, since "if all stories were true, then—in comparison with regular media—the website or Facebook page would offer no added value and generate no traffic" (Hughes and Waismel–Manor 2021, 22). According to one purveyor, the optimal balance is about 80% truth and 20% falsity, with the former providing believability, but the latter providing the product differentiation required to generate engagement (Hughes and Waismel-Manor 2021, 22).

Much of this content, it should be noted, was produced without any overt political intent (indeed, the decision to focus on politics seems to have been purely a business decision, a consequence of the greater engagement generated by these

stories). Since that time, however, it has become clear that intelligence agencies are also deeply involved in spreading fabricated content on social media. The Russian-based Internet Research Agency, for example, has been extremely active in attempting to stoke racial conflict in the U.S. (Glaser 2018). Groups like Black Lives Matter, which eschew traditional centralized leadership, are particularly vulnerable to this strategy, because it is extremely difficult to verify the authenticity of any group claiming to speak on its behalf. Thus Russian-backed trolling and misinformation, aimed at fomenting civil strife in the U.S., has become a significant feature of the discursive landscape (Hao 2021). Many of these foreign intelligence organizations control botnets as well, which allows them to generate not just fake news, but also thousands of seemingly-authentic likes and shares of these stories, which then encourages ordinary readers to treat the stories as credible (Sanovich 2019, 27–29). The fact that so much of the content produced by East European troll farms is indistinguishable, in both style and content, from the earnest political commentary of actual Americans, is itself a telling reflection of the degradation of political discourse in that country.

3.2 Conspiracy Theories

One of the most surprising effects of the internet has been a resurgence of conspiracy theories, along with a growing debate over what is driving the phenomenon. A certain amount of attention has been focused on the recommendation algorithms employed by sites such as YouTube and Facebook, which have been major conduits for the propagation of conspiracy theories (O'Donovan et al. 2019). There is, however, another aspect to it, which is a social effect that is merely enabled by the technology. Beliefs tend to be sustained within communities of like-minded individuals, and can be difficult to sustain in contexts in which one is surrounded by those who reject them. Part of this involves rational processes, as non-believers will raise doubts and press objections that believers may not have considered or taken seriously. Part of it is purely a social effect, as individuals feel pressure to align their beliefs with those around them. As a result, irrational beliefs can often be sustained only in deviant subcultures, in which others refrain from challenging them.

In the past, geographical dispersion has been a serious impediment to the formation of such subcultures. A farmer in Iowa who is disposed to believe that politicians are shape-shifting reptiles (Icke 2001), would traditionally have found it quite difficult to find others sympathetic to these claims. Indeed, such a person might find it uncomfortable even to broach the topic with friends and neighbors for fear of ridicule. And yet on the internet, it is quite easy to search (anonymously) and find a community of believers, who will not only support these convictions

(e.g. https://davidicke.com/), but also encourage the suggestion that the believer is being 'gaslit' by skeptical friends, family and neighbors. Furthermore, there are a significant number of conspiracy enthusiasts, keen to introduce those who are sufficiently open-minded to a range of other theories.

Historically, conspiracy theories have had a relatively marginal impact on politics (with the most significant effect probably being the motivation they provide to violent extremists). Much of this changed, however, with the emergence of the QAnon conspiracy theory, which developed astonishing reach and influence, ensnaring several million Americans and making a significant contribution to the efforts to overturn the results of the 2020 U.S. presidential election. It has since given rise to several offshoots in Europe, becoming particularly influential in German far-right politics (Keady 2021). The QAnon conspiracy has several features that contribute to its virulence. Many such theories burn themselves out over time, because they have a limited evidential base. Speculation about the Kennedy assassination or the World Trade Center attacks eventually declines because there is not much more to be said, and no new revelations are forthcoming. QAnon, by contrast, was constantly replenished with new content, as the mysterious 'Q' figure put out more 'breadcrumbs' for his followers. Much of it also has the structure of an A/B test, encouraging followers to conduct their own investigations. The other important feature of QAnon is that, because of the ongoing provision of content, the conspiracy can be guided to particular ends. While it is unclear what the goals of the actors behind QAnon are, this feature has attracted a number of new entrants to the field, seeking to attract adherents from the QAnon base, or else to develop their own conspiracy *de novo*. And finally, it is important to be clear about how completely QAnon has hijacked the political conversation, and in some cases the legislative agenda, in many regions of the U.S. Even if the force of reason eventually prevails, the mere fact that a set of completely delusional beliefs have been the central focus of political debate remains a win for those who want to undermine the governance capacity of democratic publics.

3.3 Flooding the Zone

The internet has also given rise to new communications strategies, this time based on the ease with which traditional gatekeepers can be circumvented. During the early months of the Trump administration, his chief strategist Stephen Bannon described his approach to dealing with the media as 'flooding the zone with shit.' The centerpiece of this strategy involved generating such a profusion of falsity, misdirection and diversion that it overwhelmed the media's ability to demand a single, coherent account. Pressed to justify any particular false claim, officials respond by making three or four more false claims. Similarly, scandals can be

overcome by generating a new scandal. This strategy forced many American media outlets to abandon the practice of simply reporting what politicians have said (while relying on others to challenge the veracity) and instead to immediately label false claims as false. (The attempt to resist the tidal wave has resulted in sentences such as the following in a news story from the *New York Times*: "'It's not a heavy steroid,' Mr. Trump said of the heavy steroid he's been taking, dexamethasone.")

Observers have noted significant similarities between this approach and what is known as the 'firehose of falsehood' pioneered by the government of Vladimir Putin in Russia (from which the verbs 'to firehose' and 'firehosing' are derived) (Paul and Matthews 2016). The approach differs from traditional 'party line' propaganda in that no attempt is made to establish a consistent position or narrative (Pomerantsev 2019). Completely different claims will often be disseminated in different media channels. If any particular claim is debunked, the government merely shifts to some other. Often different officials will offer conflicting accounts simultaneously. Furthermore, having said both *x* and not-*x*, the government can always claim to have been right all along. More generally, the constant repetition of contradictory accounts gives citizens license to believe whatever they want to believe, or to disbelieve everyone if they prefer.

There are interesting parallels between this approach and the social media strategy of the government of China. Apart from censorship, the Chinese state pays a vast network of private individuals to engage in social media posting under central direction (referred to as the '50 cent army' based on the rumor that they are paid half a *renminbi* per post). In an extraordinary paper, King et al. (2013) published an analysis of leaked emails from the Internet Propaganda Office in Zhanggong District, which both confirmed the existence of the 50 cent army and revealed interesting features of the strategy. Their analysis suggests a substantial effort. They estimate that during a one-year period the government fabricated approximately 448 million social media posts. The effort was highly coordinated, with large bursts of content being generated at the time of politically sensitive events. The most interesting finding, however, was that despite the general perception that the 50 cent army argues in defense of party positions and promotes nationalism, in fact it does neither. Most of the posting seemed to be aimed at misdirection and distraction. For example, after the Shanshan riots in Xinjiang there was a large burst of paid activity, mostly involving posts about President Xi's 'Chinese Dream,' and support for veterans and martyrs. This is quite different from traditional state propaganda strategies. Instead of attempted to promote some specific set of beliefs, it seems to involve shifting the noise-to-signal ratio on social media, in order to prevent online social networks from acting as a conduit for collective will-formation or mobilization.

3.4 Rearguard Strategies

Internet communication has shown a powerful capacity to potentiate social deviance, and thus, to promote political extremism. One can see the effects of this in both the 'alt right' and 'woke left' (Nagle 2017). Excessive attention on the part of analysts to the role of filter bubbles and echo chambers has tended to distract from other mechanisms underlying this phenomenon. Part of it, as we have seen, it simply the ease with which the internet allows individuals with extremist proclivities to find other like-minded individuals, and thus to form deviant subcultures (Pierre 2016). But the internet also makes it easier for extremists to find opponents who exactly fit their preferred conception of the enemy, which in turn confirms their extremist beliefs. Thus one can find social justice warriors engaged in vituperative debate with incels and white supremacists, each group tacitly confirming the other's worst caricature of what its opponents are like, and thus pushing them both toward less moderate stances. This is in many ways the *opposite* of a filter bubble. Twitter in particular often serves to amplify extremism, because users seize upon the most outrageous or ill-considered comment by their adversaries, retweeting it thousands of times, so that the one unfortunate view comes to dominate the perception of what others think.[5] This dynamic erodes the rationality, not just of public debate, but of individual worldviews.

The traditional position of internet firms with respect to these issues was a principled refusal to be held responsible for the content propagated on their platforms. Their argument, which was not without merit, was that they were not publishers in the traditional sense, since they were merely providing tools that allowed users to communicate with one another. In this respect they are more like the telephone company than a traditional newspaper – and one can hardly hold the telephone company responsible for the fact that political extremists use telephones to communicate with one another. This was the position, for example, maintained by Facebook when ethnic violence broke out in Sri Lanka, fueled by rumors and incitements shared on Facebook and WhatsApp (Taub and Fisher 2018). And yet the company proved unable to adhere to this position in a domestic context, eventually banning President Trump as well as hundreds of QAnon sites

---

5 An example of this, salient at the time of writing, was a tweet by Black Lives Matter activist Bree Newsome, after police in Ohio shot a teenager, Ma'Khia Bryant, who was attempting to stab another girl with a knife: "Teenagers have been having fights including fights involving knives for eons. We do not need police to address these situations." This was retweeted thousands of times and generated a 'debate' that became national news in the U.S., about whether police should be able to use force to prevent attempted murders in progress.

from its platform. The COVID-19 pandemic provided further basis for purging many prominent conspiracy theorists, many of whom could not resist promoting disinformation about public health measures.

And yet it is difficult to find anyone who thinks that the precedents set by these actions could be stable or serve as the basis for coherent policy. The pandemic has in many ways delayed the day of reckoning, by giving social media platforms an overriding reason (i.e. 'public health') to ban the most disruptive set of users. Once a state of normality is restored, and with it the force of traditional free speech norms, the problem will recur with considerable force. And yet given the impressive scale of these platforms (at the time of writing Facebook has an astonishing 2.8 billion active monthly users, and over 1.8 billion daily users), human supervision of content is impossible, especially considering the number of different languages involved. And yet the effectiveness of algorithms at taking down content can be undermined by A/B testing (Ford 2021, 285). Thus the prospects for internet companies to resolve the problems described in this section, or for government regulation forcing internet companies to resolve these problems, are extremely thin.

4 Conclusion

One of the classic problems afflicting aggregative models of democratic decision-making is the 'voter's paradox.' Given that the chances of any one individual's vote actually affecting the outcome of an election is extremely small, it is unclear why a rational individual would be willing to accept the costs involved in casting a ballot. Thus participation in democratic elections seems to be subject to a giant collective active problem. The 'paradox' then arises from the fact that a large number of people nevertheless show up to vote. There are various responses to this paradox, most of which focus on the fact that the cost of voting is very low, many people consider it a duty to vote, and there are expressive reasons for voting other than wanting one's own preferences to prevail. But even if citizens show up to vote, one might still wonder what incentive they have to pay attention to public affairs, in order to ensure that they cast their vote wisely. After all, if any one person's vote has practically no chance of affecting the outcome, there is no feedback mechanism to deter individuals from forming their political opinions in an epistemically invalid or irresponsible manner. Furthermore, the costs involved in following current events and political affairs, for those who derive no intrinsic pleasure from the task, are far greater than the one-off act of voting. And so, keeping in mind that deliberation is typically just a prelude to voting, the fact that each individual's vote makes no difference to outcomes, and there is no feedback

mechanism to punish voters who act irresponsibly, has a tendency to undermine deliberation (Landemore 2012 , 193–195).

This is worth keeping in mind when considering the challenges posed by the post-truth communications environment in modern politics. What the rise of social media has made clear is that the editors of traditional media shouldered a great deal of the cost involved in sustaining the rationality of political discourse. Now that they have been cut out, the responsibility falls on citizens directly. So far citizens have shown limited willingness to take on that responsibility. Even more unfortunate is that all of the proposed solutions to the problem seem to involve more attention, more effort, and more involvement on the part of individual citizens (Chambers 2020; Cohen and Fung 2021). This ignores the underlying collective action problem. Citizens lack the incentive to invest this effort, in part because they suffer no ill consequences from acting irresponsibly (Ananny 2021, 144). Traditional news outlets that published false content had an incentive to publish retractions. Private individuals have no such incentive, when they discover that the meme they just shared was produced by a troll farm. This lack of individual motivation is not a passing phenomenon, but rather a structural feature of political discourse in a democratic society.

In a memorable phrase, Thomas Frank once described the voice of reason in modern politics as like "a kitten's gentle purring while a freight train roars by ten feet away" (Frank 2012, 99). People whose concept of 'public engagement' consists of writing an opinion column for the daily newspaper have grossly underestimated the severity of the communication problems faced by democratic societies. The wild bandwagon effects created by social media, the loss of ordinary inhibitions afforded by anonymity, the prevalence of trolling and attention-seeking behavior, and above all the tsunami of falsehood, have completely overwhelmed the capacity of the average-to-below-average citizen to participate in public discourse in a rational fashion. Even if deliberative democracy at one time provided a plausible reconstruction of the normative significance of the democratic public sphere, it is now the legacy of a bygone age. It is incumbent upon us to focus our attention on the design of institutions and norms for the post-deliberative age. A great deal of this will involve a return to the traditional liberal project of finding ways to maintain the accountability of political institutions to the public, while at the same time insulating decision-making from the excesses of popular sovereignty.

There are numerous avenues along which such an initiative can proceed, many of which are not mutually exclusive. It is always possible to focus on institutional changes aimed at improving deliberative conditions, or insulating specific exercises in deliberation from some of the strategic pressures that prevail in the media space. I myself have made proposals along these lines (Heath 2014, 335–352). But I think it also makes sense to shift away from a preoccupation with

what Bo Rothstein refers to as "the input side of the political system" (Rothstein 2011, 80), involving party politics and electoral democracy, and to pay greater attention to the 'output' side, which involves the actual services that the state provides to citizens, and the quality of the interactions that occur in that context. The preoccupation with democracy, Rothstein argues, has led to a relative neglect of the contribution that state performance on the output side makes to political legitimacy. In the same way that theorists of deliberative democracy expanded our understanding of democratic institutions, beyond those that were directly involved in voting, we must further expand out understanding of democracy to include all aspects of state operations, including administrative agencies and the way that they interact with citizens (Heath 2020; Rosanvallon 2008).

References

Allcott, H., and M. Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31: 211−36.

Ananny, M. 2021. "Presence of Absence: Exploring the Democratic Significance of Silence." In *Digital Technology and Democratic Theory*, edited by L. Bernholz, H. Landemore, and R. Reich, 141−66. Chicago: University of Chicago Press.

Barberà, S. 2001. "An Introduction to Strategy-Proof Social Choice Functions." *Social Choice and Welfare* 18: 619−53.

Chambers, S. 2012. "Deliberation and Mass Democracy." In *Deliberative Systems*, edited by J. Parkinson and J. Mansbridge, 52−72. Cambridge: Cambridge University Press.

Chambers, S. 2020. "Truth, Deliberative Democracy, and the Virtues of Accuracy: Is Fake News Destroying the Public Sphere?" *Political Studies* 69: 147−63.

Cohen, J., and A. Fung 2021. "Democracy and the Digital Public Sphere." In *Digital Technology and Democratic Theory*, edited by L. Bernholz, H. Landemore, and R. Reich, 23−61. Chicago.

Converse, P. E. 1964. "The Nature of Belief Systems in Mass Publics." *Critical Review* 18: 1−74.

Davis, E. 2017. *Post-truth*. New York: Little, Brown.

Dumouchel, D., C. Ouellet, and T. Giasson. 2021. Guns for Votes: Wedge Politics in the Canadian Multiparty System. *Parliamentary Affairs*.

Ford, B. 2021. "Technologizing Democracy or Democratizing Technology?: A Layered-Architecture Perspective on Potentials and Challenges." In *Digital Technology and Democratic Theory*, L. Bernholz, H. Landemore, and R. Reich, 274−308, Chicago.

Frank, T. 2012. *Pity the Billionaire*. New York: Henry Holt & Co.

Fung, A. 2007. "Minipublics: Deliberative Designs and their Consequences." In *Deliberation, Participation and Democracy*, edited by S. W. Rosenberg, 159−83. London.

Glaser, A. 2018. *Russian Trolls were Obsessed with Black Lives Matter*. Slate (May 11).

Gutmann, A., and D. Thompson. 2004. *Why Deliberative Democracy?* Princeton: Princeton University Press.

Habermas, J. 1990. "Discourse Ethics: Notes on a Program of Philosophical Justification." In *Moral Consciousness and Communicative Action*, Trans. Shierry Weber-Nicholson. Cambridge, MA: MIT Press.

Hao, K. 2021. *Troll Farms Reached 140 Million Americans a Month Before the 2020 Election*. MIT Technology Review (Sept. 16).

Harris, M. 2015. *How a Lone Hacker Shredded the Myth of Crowdsourcing*. Wired (Feb).

Heath, J. 2014. *Enlightenment 2.0*. Toronto: HarperCollins.

Heath, J. 2020. *The Machinery of Government*. New York: Oxford University Press.

Hughes, H. C., and I. Waismel-Manor. 2021. "The Macedonian Fake News Industry and the 2016 US Election." *PS: Political Science and Politics* 54: 19–23.

Icke, D. 2001. *Children of the Matrix*. Ryde: Bridge of Love Publications.

Keady, J. 2021. "Querdenker, Querfront, and Qanon: On the German Far-Right and its American Occupation." *Europe Now* 42.

Keyes, R. 2004. *The Post-truth Era*. New York: St. Martin's.

King, G., J. Pan, and M. E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107: 1–18.

Lees, J. 2014. *Political Marketing*, 2nd ed. London: Routledge.

Lepore, J. 2017. "Public Thinker: Jill Lepore on the Challenge of Explaining Things." *Public Books*.

Landemore, H. 2012. *Democratic Reason*. Princeton: Princeton University Press.

List, C., and J. Dryzek. 2003. "Social Choice Theory and Deliberative Democracy: A Reconciliation." *British Journal of Political Science* 33: 1–28.

Luntz, F. 2007. *Words that Work*. New York: Hyperion.

Mansbridge, J., J. Bohman, S. Chambers, T. Christiano, A. Fung, J. Parkinson, D. E. Thompson, and M. E. Warren. 2012. "A Systemic Approach to Deliberative Democracy." In *Deliberative Systems*, edited by J. Parkinson, and J. Mansbridge, 1–26. Cambridge: Cambridge University Press.

Marantz, A. 2019. *Anti-Social: Online Extremists, Techno-Utopians, and the Hijacking of the American Conversation*. New York: Viking.

Mitchell, A., M. Jurkowitz, J. B. Oliphant, and E. Shearer. 2020. *Americans Who Mainly Get their News on Social Media are Less Engaged, Less Knowledgeable*. Pew Research Center Report (July).

Mutz, D. C. 2015. *In-Your-Face Politics*. Princeton: Princeton University Press.

Nagle, A. 2017. *Kill All Normies*. Alresford: Zero.

O'Donovan, C., C. Warzel, L. McDonald, B. Clifton, and M. Woolf. 2019. *We Followed YouTube's Recommendation Algorithm Down the Rabbit Hole*. Buzzfeed News (Jan. 24).

Paul, C., and M. Matthews. 2016. *The Russian "Firehose of False-hood" Propaganda Model: Why It Might Work and Options to Counter It*. Santa Monica, CA: RAND Corporation.

Pierre, J. 2016. *Does the Internet Promote Delusional Thinking?* Psychology Today (Jan. 25).

Pomerantsev, P. 2019. *This is Not Propaganda*. New York: Public Affairs.

Rosanvallon, P. 2008. *Counter-Democracy, Trans. Arthur Goldhammer*. New York: Cambridge University Press.

Rosensteil, T., M. R. Just, T. L. Belt, A. Pertilla, W. C. Dean, and D. Chinni. 2007. *We Interrupt this Newscast*. Cambridge: Cambridge University Press.

Rothstein, B. 2011. *The Quality of Government*. Chicago: University of Chicago Press.

Sanovich, S. 2019. "Russia: The Origins of Digital Misinformation." In *Computational Propaganda*, edited by S. C. Woolley, and P. N. Howard, 21–40. New York: Oxford University Press.

Siroker, D., and P. Koomen. 2013. *A/B Testing*. Hoboken, NJ: John Wiley.

Smith, G., and M. Setälä. 2018. "Mini-Publics and Deliberative Democracy." In *Oxford Handbook of Deliberative Democracy*, edited by A. Bachtiger, J. S. Dryzek, J. Mansbridge, and M. Warren, 300—14. Oxford: Oxford University Press.

Stefanovitch, N., A. Alshamsi, M. Cebrian, and I. Rahwan. 2014. "Error and Attack Tolerance of Collective Problem Solving: The DARPA Shredder Challenge." *EPJ Data Science* 3: 1—27.

Taub, A., and M. Fisher. 2018. *Where Countries are Tinderboxes and Facebook is the Match*. New York Times (April 21).

Warren, M. E. 2007. "Institutionalizing Deliberative Democracy." In *Deliberation, Participation and Democracy*, edited by S. W. Rosenberg, 272—88. London: Palgrave MacMillan.

Wells, P. 2015. *Of Shields and Swords and Elections*. Macleans (April 9).

Wight, C. 2018. "Post-truth, Postmodernism and Alternative Facts." *New Perspectives* 26: 17—29.