BAD LANGUAGE MAKES GOOD POLITICS

Adam F. Gibbons
Department of Philosophy, Lingnan University
Hong Kong Catastrophic Risk Centre

Abstract. Politics abounds with bad language: lying and bullshitting, grandstanding and virtue signaling, code words and dogwhistles, and more. But why is there so much bad language in politics? And what, if anything, can we do about it? In this paper I show how these two questions are connected. Politics is full of bad language because existing social and political institutions are structured in such a way that the production of bad language becomes rational. In principle, by modifying these institutions we can reduce the prevalence of bad language. However, as I show, such practical efforts are fraught with difficulties. After first outlining an account of bad language (Section 1), I examine the rationality of three different types of bad language: inaccurate language, insincere language, and unclear language (Section 2). Next, I discuss the possibility of implementing institutional reforms to improve the quality of political discourse (Section 3). However, I then outline and discuss two serious complications for institutional reforms—namely, they create risk of abuse, and they could preclude instances of seemingly bad language that, in fact, are socially beneficial (Section 4). I conclude with some thoughts about how to pursue institutional reform in an appropriately circumspect manner (Section 5).

Keywords: misinformation; virtue signaling; dogwhistles; fact-checking; speech regulation; epistocracy; incentives; institutions

Introduction

In his famous essay 'Politics and the English Language', George Orwell wrote that "political language has to consist largely of euphemism, question-begging and sheer cloudy vagueness". Political leaders, he notes, often find themselves needing to defend seemingly indefensible acts. When such a need arises, it is expedient to obscure the nature of the acts in question by describing them euphemistically. State-sponsored abduction of foreign nationals to circumvent domestic laws becomes *extraordinary rendition*, the unintended loss of civilian life during military operations becomes *collateral damage*, and so on. Obfuscatory language is rational when faced with a choice between the safety of dishonest obscurantism or the dangers of honest clarity.

Orwell was writing nearly 80 years ago, but politics makes us communicate in peculiar ways as much as ever. However, such peculiarities go far beyond the occasional need to defend the indefensible. Consider the evasive speech of a politician faced with questioning journalists, the trite slogans of the campaign trail, the brash and bombastic headlines of our newspapers, the insincere virtue-signaling so prevalent on social media, or the massive volume of political misinformation shared online. Politics is full of liars and bullshitters, demagogues and dog-whistlers, and more. In short, politics is full of *bad language*.[1]

The prevalence of bad language in politics raises two separate concerns. First, one might wonder what it is about politics that makes it so hospitable to bad language. Liars and bullshitters and

---

[1] This expression is borrowed from Hermann Cappelen and Josh Dever's 2019 book *Bad Language*.

other assorted miscreants can be found outside of politics, to be sure, but their seeming pervasiveness in politics calls out for explanation. Second, one might view the noxious state of political discourse as a matter of practical concern. What, if anything, can we do about bad language in politics?

In this paper I tackle each concern. As we shall see, they are importantly related. I argue that widespread incentives, shaped partly by background institutions with which political agents interact, systemically render the production of bad language rational. In effect, bad language is ubiquitous in politics because social and political institutions are structured in such a way that bad language is often rewarded. This has important practical consequences for those interested in improving the state of political discourse, for it suggests that we should proceed, wherever possible, by adjusting the background structures which furnish the relevant incentives. In so doing we can modify the payoffs and penalties associated with the production of bad language and, in principle, more effectively alter the linguistic behavior of political agents in beneficial ways than if we were to ignore the underlying incentives.

The outline of the paper is as follows. In Section 1 I describe in more detail what philosophers have in mind when discussing bad language, suggesting that much bad language can be characterized as inaccurate, insincere, and unclear. In Section 2 I show how different forms of inaccurate, insincere, and unclear linguistic behavior can be rational in politics, outlining the various benefits political actors hope to attain by such behavior. Turning next to practical matters, in Section 3 I discuss the possibility of institutional reforms that seek to combat the harmful effects of bad language. Focusing on efforts to mitigate harm caused by inaccurate language, I outline three increasingly far-reaching proposals: fact-checking organizations; more extensive speech regulation; and epistocratic institutions. However, in Section 4 I note some serious complications for such reforms. On the one hand, such institutions come with a significant risk of abuse. On the other hand, seemingly bad language may have social benefits that should not be overlooked. We minimize the opportunities to produce bad language at

our own peril. Section 5 concludes with some schematic remarks about how to pursue institutional reforms focused on bad language in an appropriately circumspect manner.

1. What is Bad Language?

One of the central claims in this paper is that bad language is often a rational response to incentives that are common in political settings. But what, exactly, is bad language? And what is it for bad language to be rational?

To say of some language that it is bad is to make an inherently evaluative claim. There is something about bad language that makes it (or those who produce it) morally or epistemically suspect in some manner. While evaluative matters concerning moral or epistemic assessments can be controversial, some linguistic behaviors, in virtue of possessing certain characteristics, elicit negative evaluations more frequently than others. These behaviors thus feature quite prominently in the nascent literature on social and political philosophy of language. The same is true of adjacent fields occasionally concerned with political language such as linguistics, social and political epistemology, political science, and the like.[2] These behaviors are what I intend to pick out when using "bad language".

What are the characteristics in virtue of which certain forms of linguistic behavior are evaluated negatively? In general, much bad language in politics possesses *at least one* of the following characteristics: (i) inaccuracy, (ii) insincerity, or (iii) unclarity. Consider first *inaccurate* bad language. Such forms of behavior elicit negative evaluations due to their propensity to degrade the quality of available political information. *Lying* politicians, for instance, are routinely criticized both by academics and the wider public.[3] Likewise, the preponderance of *bullshit* in politics has been noted by many

---

[2] For some examples, see Lakoff (1990), Joseph (2006), Maitra and McGowan (2012), Beaver and Stanley (2018), Cappelen (2019), Lockhart (2019), Mooney and Evans (2019), McIntosh and Mendoza-Denton (2020), Edenberg and Hannon (2021), Hannon and de Ridder (2021), and Khoo and Sterken (2021).

[3] For some academic discussion, see Bok (1978: Ch. 12) and Oborne (2005). For some popular discussion, see Harris (2019).

philosophers, most notably by Harry Frankfurt in his famous book *On Bullshit* (Frankfurt, 2005).[4] For another, various forms of *misinformation* and *disinformation* have recently received much attention from philosophers and political scientists alike.[5]

What unites these behaviors is the fact that they habitually result in the dissemination of inaccurate information, whether this dissemination is intentional (as in the case of lying and disinformation) or indicates a lack of regard for accuracy (as in the case of bullshit). As the quality of available political information degrades, the ability of political actors to make well-informed decisions is increasingly compromised. And since the ability to acquire and make subsequent use of accurate information is central to the proper functioning of various important social and political institutions, the degradation of information in politics constitutes a serious problem. Linguistic behaviors contributing to this problem are thus paradigmatic instances of bad language.

Next consider forms of linguistic behavior evaluated negatively because of their *insincerity*. Where inaccurate bad language involves agents misrepresenting the world, insincere bad language involves agents misrepresenting themselves. *Bullshit*, which we noted above for its tendency to contribute to the degradation of available political information, is also evaluated negatively in part because of its insincerity. As Frankfurt writes, the only "indispensably distinctive characteristic" of the bullshitter is that "he misrepresents what he is up to" (Frankfurt, 2005: 54). In short, on Frankfurt's account, bullshitters misrepresent themselves as having regard for the facts, something they in fact lack. *Virtue signaling* and *grandstanding* are also criticized partly because they often involve people contributing to public discussions with the intent to persuade others of their moral righteousness,

---

[4] See also Cappelen and Dever (2019: 52-72) and Gibbons (2023).
[5] The recent philosophical literature on fake news is particularly notable in this regard. See e.g., Levy (2017), Rini (2017), Gelfert (2018), Mukerji (2018), Blake-Turner (2020), the collection of articles in Bernecker, Flowerree, and Grundmann (2021), and Fritts and Cabrera (2022a). On misinformation more generally, see Benkler, Farris, and Roberts (2018), Rini (2019), Brown (2018; 2021), and Fritts and Cabrera (2022b).

while presenting themselves as not primarily concerned with boosting their social status (Tosi and Warmke, 2016; Grubbs et al., 2019; Tosi and Warmke, 2020; Levy, 2021: 9546).[6]

Insincere behavior of this kind garners criticism for two main reasons. First, insincerity is often seen as disrespectful. For example, Tosi and Warmke note the fact that much grandstanding involves agents using others to display their virtue, or more generally deceiving others about their character (Tosi and Warmke, 2020: 98). This, of course, is widely viewed as disrespectful. Second, there are important social costs to insincere behavior. In their discussion of grandstanding, for instance, Tosi and Warmke note the fact that grandstanding contributes to increasing levels of political polarization (Ibid, 70). For another, widespread grandstanding in politics might cause us to become increasingly cynical about the intentions of others, thus leading to a "devaluation of the social currency of moral talk" (Tosi and Warmke, 2016: 210; Levy, 2021: 9547).

Thirdly, consider language that is evaluated negatively because it is deliberately *unclear*.[7] Such behaviors involve agents using language in intentionally misleading ways in order to achieve some goal. The sort of *euphemisms* which animated George Orwell are emblematic of this class of behaviors. Euphemistic language—such as labelling accidental loss of civilian life in war as collateral damage— is not false *per se*, and so it isn't aptly characterized as inaccurate. And although there is something insincere about political euphemisms, it is unlike standard cases of merely insincere language insofar as it involves deliberate linguistic distortion. The hallmark of euphemistic language is the obscurantist intent with which it is produced; those who produce it seek to mask some underlying state of affairs.

Sloganeering of a certain sort presents another example of unclear bad language. Campaigning politicians will often resort to bland, inoffensive slogans that don't seem to convey very much

---

[6] See Táíwò (2021) for related discussion of vice signaling.
[7] Language that is *unintentionally* unclear may also elicit negative evaluations. But due to its seeming prevalence in politics, I restrict my focus to intentionally unclear language.

information.[8] Nondescript slogans that provide little information while eliciting the appropriate emotional reaction from the relevant target audience are often more useful than more informative, less catchy slogans—hence slogans such as 'Change You Can Believe In' and 'Make America Great Again' (slogans used by Barack Obama and Donald Trump respectively).

Finally, consider *code words* and *dogwhistles*.[9] Such linguistic tools involve speakers using an expression that conventionally communicates a certain content while intending to tacitly communicate some other content. Purported examples discussed in the relevant literature include the use of "inner-cities" to refer to poor African American neighborhoods and the use of "illegal immigrants" to refer specifically to undocumented Latin American immigrants (Khoo, 2021: 148). Agents in politics might find it useful to use code words whenever they want to tacitly communicate something that may prove controversial, whether with voters who could impose costs for controversial statements, rival politicians, or some other group.

Some caveats about this way of categorizing different forms of bad language are in order. First, these categories are not comprehensive. Some forms of bad language are not inaccurate, insincere, or unclear.[10] When rival politicians trade immature insults with one another, their insults may be accurate, sincere, and painfully clear. If there is something bad about such behavior, it must be because of some other characteristic (its incivility, say).[11] I restrict myself to inaccuracy, insincerity, and unclarity partly for convenience, with more comprehensive categorizations becoming overly burdensome. Still, the three categories used above capture much of the purportedly bad language discussed in the literature.

---

[8] One might think that we shouldn't expect slogans of this sort to be precise. Their aim, rather than to convey information in a precise fashion, is to marshal support, mobilize voters, and the like. But while this may be correct, it does not thereby follow that slogans are not a form of bad language, for on the account offered in this paper, something qualifies as bad language if it is linguistic behavior that is routinely evaluated negatively by others. Thanks to an anonymous for pushing me to clarify this point.

[9] For more on code words, see Khoo (2017). On dogwhistles, see Saul (2018) and Santana (2021).

[10] With that said, nothing precludes the extension of the analysis offered here to other categories of bad language. Indeed, the underlying incentives-based analysis can be applied to much linguistic behavior in general, whether good or bad.

[11] See Frimer et al. (2022) for a recent discussion of increasing incivility among American politicians.

More importantly, grouping the seemingly disparate forms of bad language in this way allows us to highlight commonalities among them. Many different linguistic behaviors are driven by the same sort of incentives. These categorizations thus allow us to generate unified explanations for the prevalence of an otherwise disunified group of behaviors. This latter feature is important because by highlighting the relevant incentives we can, in principle, better tailor appropriate institutional responses. If we know what sort of incentives are driving such behavior, perhaps we'll be better placed to engineer counterincentives that render the production of bad language less rational.

Second, one might wonder whether this categorization implies that good language in politics is accurate, sincere, or clear. If inaccurate language is bad because it degrades the quality of available political information, is accurate language good because it enhances this quality? If insincere language is bad in part because it conveys disrespect, is sincere language good by conveying respect? And if unclear language is bad because it obscures, is clear language good because it reveals?

Although this suggestion might initially seem plausible, it does not stand up to scrutiny. As mentioned above, some bad language is accurate, sincere, and clear. Hence, accuracy, sincerity, and clarity are *insufficient* for language to be good. Are these properties instead *necessary* for language to be good? This question will be addressed in more detail in Section 4. To preempt the discussion, some seemingly bad language may in fact serve a useful function. Since much of what seems bad about bad language is its purported social costs, any beneficial consequences of *prima facie* bad language heavily suggests that some bad language is, in fact, good (at least consequentially). Accordingly, prima facie good properties such as accuracy, sincerity, and clarity are neither necessary nor sufficient for good language. At most, we can say that *many* instances of good language possess such properties, much like *many* instances of bad language possesses the contrary properties.

Before moving on, more needs to be said regarding what it means for bad language to be *rational*. Rational agents choose courses of action which, given their beliefs, are suitable means for

achieving their ends (Kolodny and Brunero, 2018).[12] When adopting this sort of means-end framework, the ends are treated as given, allowing one to focus on ascertaining the suitability of chosen means (though one can of course independently scrutinize the ends). We may not like it when powerful political figures use their power to spread harmful misinformation, for example, but such behavior may nonetheless be rational for them given their ends. It is in this sense that I claim that bad language is so often rational in politics. As we shall see in the following section, extant social and political institutions are structured in such a way that bad language provides a suitable means for agents to pursue their chosen ends. Thus, the production of bad language becomes heavily incentivized.

2. The Rationality of Inaccuracy, Insincerity, and Unclarity

When confronted with some bad language, it is helpful to first ask ourselves what those who produce it might be hoping to achieve by behaving in such a manner. What benefits do they seek by using bad language? What costs do they hope to avoid or diminish? What ends are they ultimately pursuing? By doing this consistently, one quickly realizes that many distinct linguistic phenomena are produced for similar reasons.

The underlying rationality of deliberately inaccurate bad language, for example, is quite straightforward. Political actors sometimes value the truth, but the truth is not the only thing they value. When their commitment to the truth clashes with their other goals, political actors often opt for the latter over the former. Consider the position of lying politicians. They are frequently engaged in electoral competition with other politicians for a limited number of positions. Though they may value the truth, electoral success (and the attendant benefits such success brings) is typically foremost among their ends. Whenever the truth conflicts with their electoral ambitions, politicians are

---

[12] There is some controversy regarding whether rationality requires beliefs that are in some sense well-grounded, as well as controversy regarding how well-grounded beliefs must be in order for an agent to be rational (Gaus, 2008: 9-12). I set this controversy aside in this paper.

incentivized to hide it. Sometimes this might just involve *lying by omission*, as when politicians fail to disclose information that may harm their electoral prospects. Think of a politician wishing to keep the electorate in the dark about their past ties to unpopular extremist groups, for instance. But naturally there are cases where politicians need to actively lie about such matters. If the electorate entertains worries about the politician's ties to unpopular groups, the politician can either come clean or attempt to lie. If the former is infeasible, the politician may decide that lying is the best course of action. In general, politicians will lie when the benefits of doing so exceed the costs, whether this involves lying to preserve their reputation, lying to damage an opponent's reputation, or lying for some other reason.[13]

Other agents in politics will lie for different reasons. For instance, producers of disinformation and fake news may be motivated by financial gain (Rini, 2017; McBrayer, 2021: 24-39). If inaccurate distortions of the truth sell better than the truth itself, media outlets driven by profit have a clear incentive to lie.[14] The financial incentives of such outlets are analogous to the electoral incentives driving the behavior of politicians. In both cases, personal and professional ends take primacy, and any clashes with the truth motivate behaviors that variously conceal, obscure, or fail to disclose the truth.

One might at this point question the rationality of lying, whether from politicians or members of the media, because lying can bring steep reputational costs (Altay, Hacquin, and Mercier, 2022). These reputational costs can translate into electoral or financial costs respectively, *disincentivizing* lying when they are high enough. However, such reputational costs will often not be high enough to counterbalance the incentives to lie.

---

[13] See Tullock (1972: Ch. 9) for a classic treatment of the rationality of lying in politics. See also Jay (2010) and Mearsheimer (2011).

[14] More generally, media outlets may present information in a distorted manner because of partisan bias. See Larcinese et al. (2011), Puglisi and Snyder Jr. (2011), and Gentzkow et al. (2015) for relevant discussion.

First, many agents in politics will remain ignorant about the relevant lies. The acquisition of political information is costly, taking significant time and effort, and these costs frequently outweigh any benefits attained from acquiring political information. This rational ignorance means that fewer lies will be discovered than if, say, the electorate were heavily motivated to acquire information about the conduct and character of electoral candidates.[15]

Second, partisan agents will under-penalize lies from those with politically congruent views, while over-penalizing lies from those with incongruent views, in much the same way that partisan affiliation biases judgment in other ways (Miller and Conover, 2015; Michael and Breaux, 2021; Oyserman and Dawson, 2021; Jenke, 2023). Indeed, partisan assessments of the behavior of others may themselves involve a certain degree of lying (as well as bullshitting). For example, even honest politicians may be unfairly accused of lying, thus bearing the reputational costs of a known liar without having lied. In short, then, a certain degree of lying is safe—that is, free from costs—when partisan supporters or allies are concerned, while reputational costs from partisan opponents will arise regardless of one's behavior. So long as the benefits of lying arising from the former group outweigh the penalties associated with the latter group, lying remains rational.

Before moving on to insincere language, it is worth noting that the rationality of *unintentionally* inaccurate language is quite different to the rationality of intentional lies and disinformation. Unintentionally inaccurate language such as bullshit can arise when conflict between a regard for the facts and one's other ends emerges. But bullshitting agents do not decide to deliberately obscure (or fail to disclose) the facts they are acquainted with because it conflicts with their ends. Instead, the costs of engaging in serious truth-directed inquiry outweigh any benefits, and so they fail to become acquainted with the facts they represent themselves as knowing (Gibbons, 2023). Some of this bullshit

---

[15] Downs (1957) is largely recognized as introducing models of rational ignorance into political economy. See Somin (2021) for a helpful overview of rational ignorance.

is deliberate in the sense that it involves agents weighing up prospective costs and benefits of either engaging in truth-directed inquiry or directing their efforts elsewhere. For example, a media outlet might produce fake news merely because it is cheap to produce and sells well, but not because they want to spread falsehoods. But some of it arises because of *epistemic insouciance*, a habitual indifference to the facts (Cassam, 2018). This latter form of bullshit is not a deliberate choice. Still, much unintentionally inaccurate language is rational in the sense outlined above.

Next, consider the rationality of insincere language. Where inaccurate language becomes rational when it is advantageous to misrepresent the world, insincere language becomes rational when it is advantageous to misrepresent oneself. It can be advantageous to misrepresent oneself when placed in situations where sincerely expressing oneself either yields penalties or fails to bring benefits that are otherwise desirable. As before, politicians engaged in electoral competition provide a useful illustration. In conditions where the electorate rewards public avowals of commitment to some cause or ideal, self-interested politicians motivated by electoral success will act accordingly. Politicians who do not support the relevant causes are incentivized to insincerely act *as if* they support them. Even politicians who genuinely support the relevant causes may be incentivized to behave somewhat insincerely. Consider a situation where every competing politician sincerely expresses a commitment to some cause or ideal. To stand out from the pack, politicians will exaggerate the *intensity* of their commitment to secure the electoral rewards at the expense of their competitors. Their commitment *per se* is not insincere, but they are insincere about the degree to which they are committed.

This latter cases involves a phenomenon known as *ramping up*—the use of moral discourse to appear more morally impressive than one's peers (Tosi and Warmke, 2020: 51). Ramping up is one manifestation of grandstanding. This suggests, then, that much grandstanding that takes place in political contexts is driven by a desire to secure benefits of various kinds (Savejnarong, Pornsukjantra, and Manley, 2022). Politicians naturally seek electoral benefits. Some members of the media are

perhaps driven to insincerely express commitment to whatever ideal maximizes their revenue, while others pursue status. Citizens may seek social benefits such as the praise and esteem of their peers. For example, agents may feel pressure to publicly express certain views to maintain their position within their social group (Kuran, 1995; Williams, 2021; Gibbons, 2022b:11-3). Importantly, much contemporary political discussion takes place on social media platforms that are structured in such a way that the rewards for one's communicative acts are made salient (McDonald, 2021; Nguyen, 2021). The prospect of getting likes and shares drives people to communicate in ways that secure these rewards. Among others, the insincere expression of one's commitment to certain ideals is one method to secure these rewards.

Like inaccurate language, overly insincere language can bring costs rather than benefits. Politicians looking to secure electoral benefits by insincerely expressing commitment to some ideals run the risk of bearing costs from an electorate that also dislikes disingenuous behavior. On the one hand, skillful insincerity can deceive others into thinking that one shares their values. On the other hand, easily detectable insincerity can leave one looking manipulative. There are no easy ways to avoid such costs if one cannot feign sincerity in an effective manner. But as before, partisan irrationality provides politicians some room to maneuver. Accusations of grandstanding and virtue signaling are typically levelled against one's opponents, not one's allies. Thus, the subset of voters likely to penalize one for insincerity are already likely to be one's opponents, while one's supporters will scrutinize one's actions less stringently. When costs related to the former group are outweighed by benefits related to the latter group, insincere language can be rational.

Lastly, consider the rationality of deliberately unclear language. First, unclear language is rational when an agent seeks to obscure something—whether some underlying state of affairs or their own intentions—that would be costly if revealed. As we have seen, such is the motivation behind euphemisms. A political leader who supports the extrajudicial kidnap and transfer of foreign

belligerents from one territory to another might characterize their actions as *extraordinary rendition*. Their actions will still draw criticism from opponents, as well as citizens who are paying sufficient attention (assuming they evaluate the actions negatively). Euphemisms cannot offer complete protection. But if the leader is lucky, the euphemisms they use to describe their actions will be the language that spreads through countless testimonial chains among the general public. Inattentive, rationally ignorant voters may come across headlines discussing the administration's practice of extraordinary rendition. Many of these voters who would otherwise oppose such actions will not even know what this euphemistic expression purports to pick out, shorn as it is of clear language such as "illegal", "kidnap", "torture" and so on. Euphemisms, while not insulating one from costs, can thus be used to mitigate costs, and their use is to that extent rational.

Similar reasoning lies behind the strategic use of dogwhistles and code words.[16] These linguistic tactics allow one to communicate potentially costly contents to some target audience while bypassing the awareness of others who would strongly penalize such contents. If accused of attempting to communicate the relevant content, agents deploying code-words and dogwhistles can maintain a certain degree of plausible deniability, repudiating their intent to spread the costly message.[17] In such cases, agents rationally obscure their communicative intent. Moreover, successfully discovering that some agent has been using code words is substantially more difficult than discovering that they have used euphemistic language to obscure something. In the latter case, there may be concrete evidence of the relevant misdeeds (say, the kidnap of some foreign agent). In the former case, unless one can find some record of a prior agreement or commitment to use code words, one must rely on

---

[16] See also Stanley and Beaver (2018: 508).
[17] For more on plausible deniability, see Lee and Pinker (2010), Peet (2015), Camp (2018), Mazzarella (2021), and Dinges and Zakkou (forthcoming). For closely related discussion, see Neufeld and Woodard (forthcoming).

circumstantial evidence. Accordingly, agents are less likely to bear costs for using code words than using euphemisms.[18] Using code words to secure some benefits is thus relatively safe.

A final example of rationally unclear language comes in the form of sloganeering. Here one attempts to use vague but appealing language to either earn the support of as many as voters as possible or avoid alienating too many voters.[19] Precise and informative slogans that render one's commitments clear are risky. While sizable portions of the electorate are likely to oppose these commitments, others are unlikely to have a firm opinion on them. The risk of the former case is obvious. If voters will penalize these commitments and one wants to avoid penalties, one will render one's commitments obscure. The risk of the latter case arises because making one's commitments precise is unlikely to yield benefits whenever the electorate is largely indifferent towards them (assuming that the other available candidates are not disfavored). Vague language can mitigate both risks. Only when assured of the support of sufficiently many voters does precision in one's public-focused statements become rational for those pursuing electoral success.

However, vague language carries its own risks. Some voters, seeing through the superficial veneer of respectability, will be frustrated with empty and vacuous political slogans, especially if not accompanied by clarity in other contexts.[20] But even here the use of slogans can still be rational, so long as even more voters will impose costs on precise and clear language that lays bare the relevant politician or party's commitments (or, more generally, when voters will penalize precision more than vagueness). For the careerist politician, there is no risk-free strategy. The goal, though, is not to

---

[18] It is also important to point that, given the epistemic difficulties of reliably identifying dogwhistles or code words, it is highly likely that agents will mistakenly accuse others of dogwhistling or using code words. Presumably, they will also sometimes manufacture accusations of dogwhistling while knowing them to be spurious.

[19] On the former, see Lakoff (1990: 257-82).

[20] In fact, there is some evidence indicating that failing to pick sides in contentious disputes can cause agents to appear deceptive or untrustworthy to others (Silver and Shaw, 2022).

eliminate risk entirely, but to minimize risk where possible. Superficial campaign slogans can achieve this end.[21]

3. What Can We Do About Bad Language in Politics?

Bad language is everywhere in politics. Politicians obscure and distort the truth when it is useful for them to do so, media outlets share fake news when it is profitable, users of social media engage in shallow and insincere virtue signaling to maintain or increase their social status, and so on. Such behaviors are widely viewed in a negative light, and so some may naturally want to do something about them. But what can we do about bad language in politics?

As we have just seen, part of the problem is that, given the structure of extant social and political institutions, agents can often secure payoffs—the ends they pursue—by means of bad language. Accordingly, one might think that by modifying these institutions in such a way that bad language is instead penalized rather than rewarded, we might be able to reduce the prevalence of bad language in politics. To examine this strategy further, consider three different proposals to mitigate the harmful effects of inaccurate language: (1) fact-checking organizations, (2) more extensive speech regulation, and (3) epistocratic institutions.

Fact-checking organizations could flag inaccurate statements so that people are made aware of their inaccuracy (Amazeen, 2013; Amazeen; 2015; Rini, 2017). In principle, if made aware of the dubious nature of the material in question, people would be less likely to assume the information is reliable, less likely to share the information, and the like.[22] If people are less likely to accept and share inaccurate information, agents can anticipate fewer benefits from attempting to spread such

---

[21] Additionally, slogans can bring benefits if they mobilize and rally a sufficient number of supporters. Such benefits should be factored into any overall assessment of the utility of sloganeering. Thanks to an anonymous referee for discuss on this point.

[22] This of course makes certain empirical assumptions about belief formation that are ultimately questionable (Mandelbaum, 2014). However, I wish to set these aside to focus on describing how defenders of such institutions might motivate them.

information in the first place. For political lies to be successful, agents need to manipulate others into believing some falsehood. But if these falsehoods are consistently being flagged by fact-checkers, they are less likely to be believed. Thus, the lies are less likely to achieve their intended effect, and the expected benefit of lying correspondingly diminishes. Something similar will apply to the creation and transmission of fake news, misleading propaganda, and so on.

Unintentionally inaccurate bad language (such as some forms of misinformation and political bullshit) operates differently. By hypothesis, agents are not sharing such information with the intent to deliberately deceive others. Instead, as we have already seen, such agents merely lack regard for the facts. Since the expected benefits of intentionally manipulating others into believing falsehoods do not factor into their cost-benefit calculus, any diminution of those benefits will not affect their decision-making qua bullshitters. Still, fact-checking has a potential role to play in these cases. The costliness of caring seriously about the facts, together with the benefits of departing from truth-directed inquiry, incentivize the production of political bullshit. But fact-checkers might bring reputational costs worth avoiding. Politicians who gain a reputation as consistent bullshitters might find themselves taken less seriously, media platforms with bad reputations may lose out to competing media platforms with better reputations, and so on. When these costs are steep enough, they will shift some agents into taking greater care to get the facts right. Indirectly, then, fact-checking organizations can play a role in reducing the benefits of unintentionally inaccurate bad language.

Extensive speech regulation offers a more heavy-handed approach to reducing the prevalence of bad language than the somewhat oblique approach offered by fact-checking organizations.[23] Legislators could directly penalize those who produce inaccurate bad language (Sim, 2019; Brown, 2021; Fritts and Cabrera, 2022a; Millar, 2022). Alternatively, legislators could impose costs on social

---

[23] I assume that the proposed regulation seeks to regulate a wider range of linguistic behavior than is currently subject to regulation in the relevant jurisdiction.

media platforms who fail to play a greater role in moderating the content that appears on their platforms (Rini, 2019).[24] These platforms, looking to avoid such costs, are in turn incentivized to impose costs on those who use the platforms to spread misinformation (whether by suspending their accounts, assigning them negative reputation scores, or by some other mechanism). In the former case, legislation directly increases the expected costs of producing bad language. In the latter case, legislation directly increases the expected costs of failing to act against producers of bad language. In either case, the costs of producing or enabling bad language may increase to such an extent that its production or enablement no longer remain rational.

Lastly, epistocracy is a proposed form of government where the political power of citizens who possess a sufficient amount of political knowledge is amplified relative to their less knowledgeable peers. While there are many potential forms of epistocracy, common to each is the intent to mitigate the harmful effects of pervasive voter ignorance, a phenomenon which, as we noted earlier, decreases the likelihood that agents are penalized for lying (or, more generally, producing false statements).[25] It is important to acknowledge that part of the rationale for epistocracy does *not* appeal to increasing the expected costs of producing bad language. Instead, by decreasing the proportion of political power allocated to politically ill-informed citizens, epistocratic institutions would prevent such citizens from "polluting the polls", thus in principle combatting the negative effects of voter ignorance by limiting the influence of the relevant citizens (Brennan, 2009). However, at least *some* of the rationale for epistocracy does in fact appeal to putative beneficial consequences of increasing the expected costs that campaigning politicians can expect to face for producing inaccurate bad language. By placing greater amounts of political power into the hands of more knowledgeable citizens, politicians can no

---

[24] For an argument that social media companies have a duty to regulate themselves, see Smith and Niker (2021).
[25] For an outline of several different forms of epistocracy, see Brennan (2016). See also López-Guerra (2014), Jeffrey (2018), Mulligan (2018), Manor (2022), and Gibbons (2022a). For an overview of the empirical literature documenting voter ignorance, see Caplan (2007).

longer expect to gain as much by producing falsehoods or other misleading statements. If such politicians want to gain the support of an electorate constituted by disproportionately many well-informed citizens, they cannot afford to lie so readily. And as above, the negative reputational costs of unintentionally inaccurate language could incentivize politicians to have greater regard for the facts, taking greater care to avoid unnecessary mistakes, thereby reducing the extent to which politicians bullshit. If so, epistocratic institutions would disincentivize the production of some common forms of inaccurate bad language.

Naturally, reformers could attempt to implement more than one of the above reforms (or others), hoping that the joint effect of each is to reduce the prevalence of inaccurate language without also creating additional costs that detract from their overall value. Perhaps, for example, fact-checking organizations could operate alongside more extensive speech regulation to more comprehensively tackle harmful misinformation and disinformation in politics. Implementing multiple reforms is especially feasible when each reform is relatively modest—making only slight changes to existing institutional structures—since the downstream effects of more radical institutional reform are harder to predict. For instance, the formation of an independent fact-checking organization does not require any substantial alteration to central political institutions, more extensive regulation of speech could proceed via incremental adjustments to pre-existing law, and so on. Neither case requires wholesale revision of longstanding institutions in the way that many proposed epistocratic institutions do (though this of course does not *entail* that epistocratic institutions ought not be implemented).

Likewise, though this section has examined reforms focused on inaccurate language, it may be worth further exploring institutional proposals focusing on other forms of bad language. As a simple example, consider fact-checking organizations that also flag attempts by political agents to deliberately use language in obfuscatory ways. Much like these organizations could increase the expected costs of producing inaccurate language (thus disincentivizing its production), they could increase the expected

costs of deliberate obscurantism in such a way that elected officials are incentivized to use clear and accurate language.

In sum, fact-checking organizations, extensive speech regulation, and epistocratic institutions (or some combination thereof) would each, in principle, combat the harmful effects of inaccurate language by modifying the costs and benefits of its production in such a way that it becomes disincentivized. As we'll see in the following section, though, the foregoing cases for these institutional reforms overlook some serious complications. And even setting aside their efficacy, one might have serious reservations about the propriety of laws further restricting freedom of expression, or about institutional proposals to exclude ignorant citizens from the franchise (or to otherwise diminish their power).[26] However, I discuss these proposals not because they are philosophically uncontroversial, but to present clear and widely discussed examples of practical efforts to combat bad language that proceed by way of changing the background structures which incentivize the production of bad language in the first place. Whether one finds these reforms plausible or not, they each purport to directly tackle the problem at hand rather than simply ignoring the underlying incentives to produce bad language.

4. Bad Language or Bad Policy?

If the previous section's argument is sound, then institutional responses to bad language may prove to be effective. However, there are serious problems with such proposals that the preceding section left unaddressed. On the one hand, such reforms carry significant risk of abuse. On the other hand, there are social benefits to some instances of putatively bad language which successful institutional

---

[26] On the latter, see Ingham and Wiens (2021), Malcolm (2021; forthcoming), and Kogelmann (forthcoming).

reforms would threaten.[27] These two problems greatly complicate the case for institutional reforms aiming to tackle bad language in politics. Let's consider each problem in turn.

First, risk of abuse. Each of the above cases made certain presuppositions which, upon closer examination, are deeply contentious. Among other things, the case for fact-checking organizations presupposed the willingness and ability of fact-checkers to reliably document inaccurate or misleading statements, flagging them so that people are made aware of their dubious nature.[28] If fact-checkers were instead flagging accurate statements as false or misleading, or if they were disproportionately focusing on some groups but not others, they would be presenting a distorted view of the informational landscape in a way that runs directly counter to the purpose of fact-checking organizations.[29]

Similarly, the case for more extensive speech regulation presupposed the willingness and ability of legislators to impose costs on either those who produce bad language or those who enable its production and transmission. If, instead, costs were imposed on those not producing bad language, or disproportionate costs were imposed on some groups but not others, this would inappropriately alter the incentives to produce bad language in a way that runs counter to the legislation's purpose.

Lastly, the case for epistocratic institutions presupposed the willingness and ability of suitably positioned agents to alter the relevant institutions in such a way that the political power of well-informed citizens is amplified relative to their less informed peers. But if instead those tasked with

---

[27] For related discussion of possible benefits of epistemic vice, see Morton (2014), Smart (2018) and Bland (2022).

[28] Another controversial presupposition of the case for fact-checking organizations is the claim that people will take the verdicts of fact-checkers seriously. This, of course, might simply be false for very many people (Garrett and Weeks, 2013; Shin and Thorson, 2017). It is plausible to assume that this is especially likely to be the case if fact-checking organizations are perceived as biased. Additionally, one might have reservations about the value of fact-checking in its application to political disputes that very often involve substantive normative (and not just empirical) disagreements (cf. Lepoutre, 2020).

[29] The empirical literature on fact-checking is quite mixed, with some researchers concluding that it can reduce misinformation (Fridkin, Kenney, and Wintersieck, 2015) and others more negatively appraising its efficacy (Nyhan and Reifler, 2010; Garrett and Weeks, 2013; Jarman, 2015). Notably, in a recent meta-analysis, Walter et al. (2020) stress that there is "no simple main effect of fact-checking". However, in the very same article, they also find that "the effects of fact-checking on beliefs are quite weak and gradually become negligible the more the study design resembles a real-world scenario of exposure to fact-checking". Further empirical research to ascertain to what extent this is due to (perceptions of) bias on the part of fact-checkers would be desirable. At the very least, this possibility should not be ruled out.

designing such institutions focused their efforts on amplifying the political power of citizens more likely to support them (or their party) rather than the political power of politically knowledgeable citizens, this would alter the composition of the electorate in a way that runs counter to extant epistocratic proposals. Instead of focusing on the harmful effects of inaccurate language and political ignorance, the relevant agents would be focusing on buttressing their own power.[30]

Why might such institutional reforms fail in this manner? The agents who staff these institutions, like other agents in politics, are not omniscient angels. They are not always willing and able to impartially discharge their duties. They have their own biases and preconceptions which distort their view of the facts. They sometimes pursue power and position, seeing themselves as locked into zero-sum competition with political opponents. Given their position, fact-checkers will be able to harm the reputation of their political opponents while ignoring the bad language of their allies. Those with the power to apply legislation aimed at combating the spread of bad language will be able to use their position to impose disproportionate costs on their political opponents. And those with the ability to change the distribution of fundamental political power may do so in a way that harms their political opponents while benefiting them. In each case, partisan agents will be able to impose costs on their opponents while securing benefits for themselves. The feasibility of such institutional reforms is thus compromised by the existence of powerful incentives to abuse the power they grant.

Similar problems would arise for institutional reforms purporting to tackle other forms of bad language. The sincere language of political opponents would be smeared as insincere grandstanding, the unclear speech of one's political allies will be overlooked, and so on. Some agents in politics are less prone to abusing their power, of course. Even if they *could* harm their political opponents by abusing their position, these agents will refrain from doing so. The point, though, is not that all agents

---

[30] Several critics of epistocracy have suggested that the process by which the political power of politically knowledgeable citizens would be amplified could be manipulated or abused by self-interested actors seeking to further their own ends. See for example Bagg (2018), Klocksiem (2019), Vandamme (2020), Somin (2022), and Gibbons (2022c).

in politics are willing to abuse their position, but that there is a risk that *at least some* will be willing to do so; and it is the existence of this latter group of agents which complicates the case for institutionalizing mechanisms to combat bad language in politics. The case for institutional reforms should not rest on hypothetical versions thereof where everybody involved is maximally fair and competent. Instead, we should take people as they are—imperfect and prone to rivalrous behavior—before considering whether to implement them. Perhaps, on balance, they will be worth pursuing nonetheless. But the risk of abuse should not be ignored.

Suppose, though, that these institutional reforms perform exactly as their proponents intend. Everybody involved in fact-checking, for example, is even-handed and competent, not disproportionately fixating on their political opponents; legislators attempting to curb the spread of misinformation do not simply exploit their position to harm their opponents; and agents tasked with designing epistocratic mechanisms do not do so in a way that intentionally tries to harm their opponents or otherwise abuse their position. Would we then be justified in implementing such institutional reforms? Even here, I argue, there are serious complications worth bearing in mind.

Academic treatments of various forms of bad language accentuate the negative aspects of bad language. This is understandable. After all, we're dealing with things like lying politicians, propagandizing, disinformation and fake news, virtue-signaling, and the like. These are phenomena we single out for investigation *because* they are evaluated so negatively. Still, to focus only on the social costs of such phenomena is to offer a lopsided analysis. We should instead consider both costs *and* benefits (if there are any), before deciding whether to implement institutional mechanisms to reduce their prevalence.[31] An analysis of the potential social benefits of all forms of putatively bad language goes beyond the scope of this paper. But let's briefly consider the potential social benefits of the

---

[31] Cf. Levy (2021: 9547-8).

following three examples of inaccurate, insincere, and unclear bad language respectively: misinformation, virtue signaling, and deliberate ambiguity or vagueness.

The expected costs of *failed* interventions focusing on inaccurate bad language are clear. Not only do they risk abuse, they also risk depriving the public of potentially helpful truths when well-meaning but error-prone agents get things wrong. But what costs could *successful* interventions bring? In *On Liberty*, John Stuart Mill famously argued that the censorship of falsehoods renders our beliefs "dead dogma, not living truth" (Mill, 2008: 40).[32] Insulating ourselves from falsehoods that run contrary to prevailing opinion prevents us from attaining a better understanding of the truth. By engaging with those who disagree with us, we can sometimes come to learn the shortcomings of our own views, even if they are ultimately correct to some extent; and by doing this we can come to refine our partly true views in ways that would be difficult if we never encountered competing falsehoods. On other occasions, competing views on some issue may mix truth with falsehood. Each view gets us some way towards the truth, but each is also mistaken in some way. It is only by assessing these competing views, weighing their respective merits and demerits, that we can come to discharge the relevant falsehoods while retaining the aspects of those views that are correct. Institutions that would preclude the public transmission of such views threaten to eliminate any benefits that could be derived from engaging with them.

In the case of epistocratic institutions, epistemic democrats worry that diminishing the political power of politically ill-informed citizens may lead to a collectively *less* competent electorate (Landemore, 2013; Goodin and Spiekermann, 2018). They argue that, under appropriate conditions, individually ignorant citizens can together constitute a collectively competent electorate—indeed, a more competent electorate than one constituted by individually better informed but less numerous

---

[32] For related discussion, see Joshi (2022).

and diverse citizens.[33] If they are correct, then epistocratic institutions would threaten potential social benefits of the seemingly bad language produced by systematically ignorant citizens. Likewise, by reducing the extent to which elected officials could engage with the views of politically ill-informed citizens, epistocratic institutions would yield epistemic costs rather than epistemic benefits.[34]

Critical discussions of virtue signaling and grandstanding emphasize both the morally vicious character of those who engage in such behavior and the related social costs such as increased cynicism about moral discourse, increasing polarization, widespread public shaming, and more (Tosi and Warmke, 2016; 2020). But as with inaccurate bad language, such analyses overlook potential benefits of what at first can seem like uncontroversially negative behavior. For instance, Neil Levy argues that virtue signaling can serve a useful *epistemic* function, providing higher-order evidence regarding moral matters in the form of information about the confidence and numbers of agents who share some moral judgement (Levy, 2021: 9548).[35] Additionally, Westra argues that virtue signaling and grandstanding can play a positive role in moral change, motivating the adoption of new moral norms (Westra, 2021). Whether the benefits of such behaviors outweigh the costs is, of course, an empirical question. But the important point for our purposes is that we neglect potential benefits of these behaviors to our own detriment. If the benefits outweigh the costs, any attempt to reduce the prevalence of virtue signaling would be misguided.

Lastly, consider intentionally vague or ambiguous language. Campaign slogans, for instance, are notoriously often light on concrete detail but heavy on provocative or emotional language. Or

---

[33] See Hannon (2022) for a criticism of epistocracy according to which politically knowledgeable citizens are also more partisan and less rational about politics than other citizens. In effect, if his arguments are sound, the net harms caused by the bad language of well-informed but irrational partisans would outweigh the net harms caused by the bad language of politically ill-informed citizens. However, see Gibbons (2022a) for a response.

[34] For a criticisms of the various formal models to which epistemic democrats appeal, see Brennan (2016: 172-203), Houlou-Garcia (2017), and Hédoin (2021: 7-8). However, see van Bouwel (2022) for a more empirical argument for the epistemic superiority of democratic institutions compared to epistocratic institutions.

[35] However, see Hill and Garner (2021) for a skeptical take on the epistemic merits of widespread virtue signaling. For related critical discussion, see Hill and Fanciullo (2023).

consider the frequent unwillingness of politicians to make their views on some issue sufficiently clear. When pressed by the media, they respond with deflection and dissimulation, skirting around the topic without ever committing to anything specific. One might initially view such language as yet another manifestation of political careerism, with the intentional ambiguity reflecting a strategy to avoid needlessly alienating some potential supporters, or to appeal to as wide a base as possible, and so on.

This, of course, is true of *some* intentionally unclear language. But there are also cases where unclarity is not motivated primarily by a regard for one's own electoral prospects. Israel, for instance, maintains a stance of deliberate ambiguity regarding whether it possesses nuclear weapons.[36] If it made clear that it in fact possesses nuclear weapons, it may incriminate itself as violating anti-proliferation laws. If it made clear that it does not possess nuclear weapons, however, it might lose some of its ability to deter external aggression. In principle, ambiguity allows it to create enough uncertainty to potentially deter external aggression while avoiding the explicit violation of anti-proliferation law. Any Israeli politicians who remain intentionally unclear on such issues, then, may be carefully adhering to longstanding geostrategy rather than engaging in self-serving behavior. Of course, one might view the strategy negatively for a variety of reasons. Perhaps one thinks that nuclear weapons should not be used as deterrents. Perhaps one thinks they are ineffective deterrents.[37] Regardless, this example is merely meant to illustrate that intentional unclarity can be motivated for reasons unrelated to selfish electoral pursuits. It can serve an important purpose, and the potential upshots of intentional unclarity need to be factored into attempts to force political actors to communicate more clearly.

Together, the risk of abuse of institutional mechanisms targeting bad language and the prospect of even successful reforms doing more harm than good should caution us against institutionalizing the power to target bad language. The impulse to do something about the

---

[36] For discussion of Israeli nuclear policy, see Cochran (1996)
[37] See Wilson (2013) for a sustained defense of this claim.

degradation of political discourse is understandable, and many reformers likely possess noble intentions. But a realistic appraisal of proposed institutional reforms suggests that the cures they claim to offer might very well be worse than the underlying illnesses.

5. Conclusion

Bad language is pervasive in politics because it is rational, and it is rational because agents in politics are embedded in institutional settings where bad language yields rewards. And while institutional reforms could in principle reduce the prevalence of bad language in politics by diminishing the extent to which it is incentivized, reforms carry risks that ought not be overlooked. They may be abused by self-interested agents who benefit from doing so, or they may preclude genuinely beneficial instances of bad language (or both). But it does not follow from this that we should not do anything about the poor quality of contemporary political discourse. In this section I conclude with some schematic remarks about how to pursue institutional reform in an appropriately circumspect manner.

First, we must think more carefully about the overall costs and benefits of different forms of linguistic behavior. There is a pronounced tendency within the existing philosophical literature to emphasize the downsides of so-called bad language. But this overlooks potential benefits from the very same behaviors. Accurately determining the overall distributions of costs and benefits of different forms of linguistic behaviors is crucial to avoid the implementation of reforms that would decrease the prevalence of net-beneficial behaviors.

Second, the possibility that institutional reforms designed to curb the spread of bad language could be abused needs to be factored into an assessment of the overall plausibility of such reforms. Even if the social costs of some form of bad language clearly outweigh any compensating benefits, reforms may still not be feasible if the expected costs of abuse are sufficiently high. Different reforms will carry different levels of risk, much as different behaviors will possess different distributions of

costs and benefits. And there may be ways to design the relevant reforms such that they are more difficult to abuse. But in all cases an accurate estimation of the overall costs and benefits of various behaviors, together with an analysis of the potential risks of candidate reforms, should be sought before we implement any institutional reforms.

Ascertaining the relevant costs and benefits will be a complicated empirical undertaking. Given the complexities involved, an accurate *ex ante* assessment will be difficult to secure. Perhaps we will in some cases overstate the expected costs of abuse, thus arriving at unduly pessimistic conclusions. By the same token, however, perhaps we will understate the relevant costs, rushing headlong unwisely into potentially damaging reforms. Accordingly, it is prudent to adopt a cautious, incremental approach to the implementation of any institutional reform. Reforms should first be implemented on a local level where, if they fail, the damage is relatively contained. Implementing risky reforms on a national level increases the prospect of damage from institutional failure ramifying throughout the entire system.

Relatedly, some sort of mechanism ought to be in place where the relevant reform can be reversed or eliminated.[38] Without such a mechanism, we could find ourselves saddled with a reform that merely exacerbates the costs it was designed to address, or a reform that is overly susceptible to being abused, and so on. If reforms prove successful on a local level, we may then cautiously proceed to extend the range of such reforms in an incremental fashion, all while keeping an eye on their overall success. By such means, we could discover effective, net-beneficial reforms that tackle bad language in politics.

Still, one last, frequently overlooked complication remains. In much the same way that institutionalizing the means to combat bad language is susceptible to abuse, the process by which political leaders experiment with novel institutional reforms is also susceptible to abuse. If failed

---

[38] On the importance of such mechanisms, see Barrett (2020: 122).

reforms that yield greater aggregate costs than benefits deliver concentrated benefits to those in power, there is no guarantee that they will be reversed. The mechanisms by which failed reforms are scaled back are not automatic fail-safes; they instead require the good faith and knowledge of suitably positioned political actors to make the appropriate changes. But as we have seen, such actors often find themselves facing powerful incentives to abuse their position. This is an inescapable feature of collective decision-making, something we must learn to live with. And like the possibility of abuse of institutional reforms more generally, it does not entail that we should do nothing about various social ills, whether bad language or otherwise. Again, though, it should not be ignored entirely. For while bad language can be very bad, other things can be worse.[39]

References

Altay, S., Hacquin, A.-S, and Mercier, H. (2022) 'Why Do So Few People Share Fake News? It Hurts Their Reputation', *New Media and Society* 24 (6): 1303-24

Amazeen, M.A. (2013) 'Making a Difference: A Critical Assessment of Fact-Checking in 2012', *New America Foundation Media Policy Initiative*, research paper

Amazeen, M.A. (2015) 'Revisiting the Epistemology of Fact-Checking', *Critical Review: A Journal of Politics and Society.* 27 (1): 1-22

Bagg, S. (2018) 'The Power of the Multitude: Answering Epistemic Challenges to Epistocracy', *American Political Science Review* 4 (112): 891-904

Barrett, J. (2020 'Social Reform in a Complex World', *Journal of Ethics and Social Philosophy.* 17 (2): 103-32

Beaver, D. & Stanley, J. (2018) 'Toward a Non-Ideal Philosophy of Language', *Graduate Faculty Philosophy Journal.* 39 (2): 503-47

Benkler, Y., Faris, R., and Roberts, H. (2018) *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics.* Oxford, UK: Oxford University Press

Bernecker, S., Flowerree, A.K., and Grundmann, T. (2021) *The Epistemology of Fake News.* Oxford, UK: Oxford University Press

Blake-Turner, C. (2020) 'Fake News, Relevant Alternatives, and the Degradation of our Epistemic Environment', *Inquiry,* https://doi.org/10.1080/0020174X.2020.1725623.

Bland, S. (2022) 'In Defense of Epistemic Vices', *Synthese* 200 (1): 1-22

Bok, S. (1978) *Lying: Moral Choice in Public and Private Life.* New York, NY: Vintage Books

Brennan, J. (2009) 'Polluting the Polls: When Citizens Should not Vote', *Australasian Journal of Philosophy* 87 (4): 535-49

Brennan, J. (2016) *Against Democracy.* Princeton, NJ: Princeton University Press

Brown, E. (2018) 'Propaganda, Misinformation, and the Epistemic Value of Democracy', *Critical Review* 30 (3-4): 194-218

Brown, E. (2021) 'Regulating the Spread of Online Misinformation', In M. Hannon and J. De Ridder (eds.) *The Routledge Handbook of Political Epistemology.* New York, NY: Routledge

Camp, E. (2018) 'Insinuation, Common Ground, and the Conversational Record', in D. Fogal, D.W. Harris, and M. Moss (eds.) *New Work on Speech Acts.* Oxford, UK: Oxford University Press

Caplan, B. (2007) *The Myth of the Rational Voter.* Princeton, NJ: Princeton University Press

Cappelen, H. & Dever, J. (2019) *Bad Language.* Oxford, UK: Oxford University Press

Cassam, Q. (2018) 'Epistemic Insouciance', *Journal of Philosophical Research,* 43: 1-20

Cochran, E.S. (1996) 'Deliberate Ambiguity: An Analysis of Israel's Nuclear Strategy', *Journal of Strategic Studies.* 19 (3): 321-42

Dinges, A. and Zakkou, J. (forthcoming) 'On Deniability', *Mind*

Downs, A. (1957) *An Economic Theory of Democracy*. New York, NY: Harper and Row

Edenberg, E. and Hannon, M. (2021) *Political Epistemology*. Oxford, UK: Oxford University Press

Frankfurt, H. (2005) *On Bullshit*. Princeton, NJ: Princeton University Press

Frimer, J.A., Aujla, H., Feinberg, M., Skitka, L.J., Aquino, K., Eichstaedt, J.C., and Willer R. (2022) 'Incivility is Rising Among American Politicians on Twitter', *Social Psychological and Personality Science*. https://doi.org/10.1177/19485506221083811

Fritts, M., and Cabrera, F. (2022a) 'Fake News and Epistemic Vice: Combating a Uniquely Noxious Market', *Journal of the American Philosophical Association*, 1-22. Doi: https://doi.org/10.1017/apa.2021.11

Fritts, M. and Cabrera, F. (2022b) 'Online Misinformation and 'Phantom Patterns': Epistemic Exploitation in the Era of Big Data', *Southern Journal of Philosophy* 60 (1): 57-87

Garrett, R.K. and Weeks, B.E. (2013) 'The Promise and Perils of Real-Time Corrections to Political Misperceptions', in *Proceedings of the 23rd Conference on Computer Supported Cooperative Work*. San Antonio, TX: Association for Computing Machinery

Gaus, G.F. (2008) *On Philosophy, Politics, and Economics*. Belmont, CA: Wadsworth Cengage Learning

Gelfert, A. (2018) 'Fake News: A Definition', *Informal Logic*. 38 (2): 84-117

Gentzkow, M., Shapiro, J.M., and Stone D.F. (2015) 'Media Bias in the Marketplace: Theory', in S.P. Anderson, J. Waldfogel, and D. Strömberg (eds.) *Handbook of Media Economics Volume 1*. New York, NY: North-Holland, 623-45

Gibbons, A.F. (2022a) 'Is Epistocracy Irrational?', *Journal of Ethics and Social Philosophy* 21 (2): 251-81

Gibbons, A.F. (2022b) 'Rational Conceptual Conflict and the Implementation Problem', *Inquiry: An Interdisciplinary Journal of Philosophy*. https://doi.org/10.1080/002174X.2022.2037455

Gibbons, A.F. (2022c) 'On Epistocracy's Epistemic Problem: Reply to Méndez', *Social Epistemology Review and Reply Collective* 11 (8): 1-7. https://wp.me/p1Bfg0-722.

Gibbons, A.F. (2023) 'Bullshit in Politics Pays', *Episteme*, 1-21. https://doi.org/10.1017/epi.2023.3

Goodin, R.E. and Spiekermann, K. (2018) *An Epistemic Theory of Democracy*. Oxford: Oxford University Press

Grubbs, J.B., Warmke, B, Tosi, J., James, A.S, and Campbell, W.K. (2019) 'Moral Grandstanding in Public Discourse: Status-Seeking Motives as a Potential Explanatory Mechanism in Predicting Conflict', *PLoS ONE*, 14 (10), e0223749. https://doi.org/10.1371/journal.pone.0223749.

Hannon, M. (2022) 'Are Knowledgeable Voters Better Voters?', *Politics, Philosophy, and Economics* 21 (1): 29-54

Hannon, M. and de Ridder, J. (2021) *The Routledge Handbook of Political Epistemology*. New York, NY: Routledge

Harris, M. (2019) 'Lying Politicians Are Killing Our Democracies', *The Tyee*, URL = https://thetyee.ca/Opinion/2019/11/28/Lying-Politicians-Killing-Democracies/

Hédoin, C. (2021) 'The 'Epistemic Critique of Epistocracy' and its Inadequacies', *Social Epistemology* https://doi.org/10.1080/02691728.2021.1882609

Hill, J., and Fanciullo, J. (2023) 'What's Wrong with Virtue Signaling?', *Synthese* 201, 117. https://doi.org/10.1007/s11229-023-04131-4

Hill, S. and Garner, R. (2021) 'Virtue Signalling and the Condorcet Jury Theorem', *Synthese*. 199 (5-6): 14821-41

Houlou-Garcia, A. (2017) 'Collective Wisdom, Diversity, and Misuse of Mathematics', *French Journal of Political Science* 67 (5): 899-917

Ingham, S. and Wiens, D. (2021) 'Demographic Objections to Epistocracy: A Generalization', *Philosophy and Public Affairs* 49 (4): 323-49

Jarman, J.W. (2015) 'The Failure of Fact-Checking', in B. Garssen, D. Gooden, G. Mitchell, and A. Snoeck Henkemans (eds.) *Proceedings of the 8th International Conference of the International Society for the Study of Argumentation.* Amsterdam: Sic Sat

Jay, M. (2010) *The Virtues of Mendacity: On Lying in Politics.* Charlottesville, VA: University of Virginia Press

Jeffrey, A. (2018) 'Limited Epistocracy and Political Inclusion', *Episteme* 15 (4): 412-32

Jenke, L. (2023) 'Affective Polarisation and Misinformation Belief', *Political Behavior* https://doi.org/10.1007/s11109-022-09851-w

Joseph., J.E. (2006) *Language and Politics.* Edinburgh, UK: Edinburgh University Press

Joshi, H. (2022) 'The Epistemic Significance of Social Pressure', *Canadian Journal of Philosophy* 52 (4): 396-410

Khoo, J. (2017) 'Code Words in Political Discourse', *Philosophical Topics.* 45 (2): 33-64

Khoo, J. (2021) 'Code Words', in J. Khoo and R. Sterken (eds.) *The Routledge Handbook of Social and Political Philosophy of Language.* New York, NY: Routledge

Khoo, J. and Sterken, R. (2021) *The Routledge Handbook of Social and Political Philosophy of Language.* New York, NY: Routledge

Klocksiem, J. (2019) 'Epistocracy is a Wolf in Wolf's Clothing', *The Journal of Ethics* 23 (1): 19-36

Kogelmann, B. (forthcoming) 'Finding the Epistocrats', *Episteme*

Kolodny, N. and Brunero, B. (2020) 'Instrumental Rationality', *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2020/entries/rationality-instrumental/>.

Kuran, T. (1995) *Private Truths, Public Lies.* Cambridge, MASS: Harvard University Press

Lakoff, R. (1990) *Talking Power: The Politics of Language.* New York, NY: Basic Books

Landemore H (2013) *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton, NJ: Princeton University Press

Larcinese, V., Puglisi, R., and Snyder Jr., J.M. (2011) 'Partisan Bias in Economic News: Evidence on the Agenda-Setting Behavior of US Newspapers', *Journal of Public Economics* 95: 1178-89

Lee, J.J. and Pinker, S. (2010) 'Rationales for Indirect Speech: The Theory of the Strategic Speaker', *Psychological Review* 117 (3): 785-807

Lepoutre, M. (2020) 'Democratic Group Cognition', *Philosophy and Public Affairs* 48 (1): 40-78

Levy, N. (2017) 'The Bad News About Fake News', *Social Epistemology Review and Reply Collective*. 6 (8): 20-36

Levy, N. (2021) 'Virtue Signalling is Virtuous', *Synthese* 198: 9545-62

Lockhart, M. (2019) *President Donald Trump and His Political Discourse: Ramifications of Rhetoric via Twitter*. New York, NY: Routledge

López-Guerra, C. (2014) *Democracy and Disenfranchisement: The Morality of Electoral Exclusions*. New York, NY: Oxford University Press

Maitra, I. and McGowan, M, (2012) *Speech and Harm: Controversies Over Free Speech*. Oxford, UK: Oxford University Press

Malcolm, F. (2021) 'Epistocracy and Public Interests', *Res Publica* 28 (1): 173-92

Malcolm, F. (forthcoming) 'The Purpose and Limits of Electoral Accountability', *Journal of Ethics and Social Philosophy*

Mandelbaum, E. (2014) 'Thinking is Believing', *Inquiry*. 57 (1): 55-96

Manor, A. (2022) 'Polycentric Limited Epistocracy: Political Expertise and the Wiki-Model' *Episteme* 19 (1): 1-20

Mazzarella, D. (2021) "'I Didn't Mean to Suggest Anything Like That!": Deniability and Context Reconstruction', *Mind & Language*, https://doi.org/10.1111/mila.12377

McBrayer, J.P. (2021) *Beyond Fake News: Finding the Truth in a World of Misinformation*. London, UK: Routledge

McDonald, L. (2021) 'Please Like This Paper', *Philosophy*. DOI: 10.1017/S0031819121000152

McIntosh, J. and Mendoza-Denton, N. (2020) *Language in the Trump Era: Scandals and Emergencies*. Cambridge, UK: Cambridge University Press

Mearsheimer, J.J. (2011) *Why Leaders Lie: The Truth About Lying in International Politics*. Oxford: Oxford University Press

Michael, R.B., and Breaux, B.O. (2021) 'The Relationship Between Political Affiliation and Beliefs About Sources of Fake News', *Cognitive Research: Principles and Implications*. 6 (6) https://doi.org/10.1186/s41235-021-00278-1

Mill, J.S. (2008) *On Liberty and Other Essays. Oxford World's Classics*. Oxford, UK: Oxford University Press

Millar, B. (2022) 'Epistemic Obligations and Free Speech', *Analytic Philosophy* https://doi.org/10.1111/phib.12279

Miller, P.R. and Conover, P.J. (2015) 'Red and Blue States of Mind: Partisan Hostility and Voting in the United States', *Political Research Quarterly* 68 (2): 225-39

Mooney, A. and Evans, B. (2019) *Language, Society, and Power: An Introduction*. New York, NY: Routledge

Morton, A. (2014) 'Shared Knowledge From Individual Vice: The Role of Unworthy Epistemic Emotions', *Philosophical Inquiries* 2: 163-72

Mukerji, N. (2018) 'What is Fake News?', *Ergo: An Open Access Journal of Philosophy*. 5: 923-46

Mulligan, T. (2018) 'Plural Voting for the Twenty-First Century', *Philosophical Quarterly* 68 (271): 286-306

Neufeld, E. and Woodard, E. (forthcoming) 'On Subtweeting', in P. Connolly, S.C. Goldberg, and J. Saul (eds.) *Conversations Online*. Oxford, UK: Oxford University Press

Nguyen, C.T. (2021) 'How Twitter Gamifies Communication', in J. Lackey (ed.) *Applied Epistemology*. Oxford, UK: Oxford University Press

Nyhan, B. and Reifler, J. (2010) 'When Corrections Fail: The Persistence of Political Misperceptions', *Political Behavior* 32: 303-30

Oborne, P. (2005) *The Rise of Political Lying*. London, UK: Free Press

Oyserman, D. and Dawson, A. (2021) 'Your Fake News, Our Facts: Identity-Based Motivation Shapes What We Believe, Share, and Accept', in R. Greifeneder, M.E. Jaffé, E.J. Newman, and N. Schwarz Eds.) *The Psychology of Fake News: Accepting, Sharing, and Correcting Misinformation*. New York, NY: Routledge

Peet, A. (2015) 'Testimony, Pragmatics, and Plausible Deniability', *Episteme* 12 (1): 29-51

Pritchard, D. (2021) 'Good News, Bad News, Fake News', In S. Bernecker, A. Flowerree, and T. Grundmann (eds.) *The Epistemology of Fake News*. Oxford, UK: Oxford University Press

Puglisi, R. and Snyder Jr., J.M. (2011) 'Newspaper Coverage of Political Scandals', *Journal of Politics* 73: 931-50

Rini, R. (2017) 'Fake News and Partisan Epistemology', *Kennedy Institute of Ethics Journal*. 27 (2): 43-64

Rini, R. (2019) 'Social Media Disinformation and the Security Threat to Democratic Legitimacy', *Disinformation and Digital Democracies in the 21st Century*. Publication of the NATO Association of Canada. 10-4

Santana, C. (2021) 'What's Wrong with Dogwhistles', *Journal of Social Philosophy*. https://doi.org/10.1111/josp.12409

Saul, J. (2018) 'Dogwhistles, Political Manipulation, and Philosophy of Language', in D. Fogal, D.W. Harris, and M. Moss (eds.) *New Work on Speech Acts*. Oxford, UK: Oxford University Press

Savejnarong, T., Pornsukjantra, P., and Manley, H. (2022) 'The Interpersonal Consequences of Prestige and Dominance-Based Moral Grandstanding', *Personality and Individual Differences*. https://doi.org/10.1016/j.paid.2022.111656

Shin, J. and Thorson, K. (2017) 'Partisan Selective Sharing: The Biased Diffusion of Fact-Checking Messages on Social Media', *Journal of Communication* 67: 233-55

Silver, I. and Shaw, A. (2022) 'When and Why "Staying Out Of it" Backfires in Moral and Political Disagreements', *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001201

Sim, S. (2019) *Post-Truth, Scepticism, and Power*. Palgrave MacMillan

Smart, P.R. (2018) 'Mandevillian Intelligence', *Synthese* 195: 4169-200

Smith, L. and Niker, F. (2021) 'What Social Media Facilitates, Social Media Should Regulate: Duties in the New Public Sphere', *The Political Quarterly*. 92 (4): 613-20

Somin, I. (2021) 'Is Political Ignorance Rational?', in M. Hannon and J. de Ridder (eds.) *The Routledge Handbook of Political Epistemology*. New York, NY: Routledge

Somin, I. (2022) 'The Promise and Peril of Epistocracy', *Inquiry: An Interdisciplinary Journal of Philosophy* 65 (1): 27-34

Táíwó, O. (2021) 'Vice Signaling', *Journal of Ethics and Social Philosophy* 22 (3): 295-316

Tosi, J. and Warmke, B. (2016) 'Moral Grandstanding', *Philosophy and Public Affairs*. 44, 3: 197-217

Tosi, J. & Warmke, B. (2020) *Grandstanding: The Use and Abuse of Moral Talk*. Oxford, UK: Oxford University Press

Tullock, G. (1972) *Toward a Mathematics of Politics*. Ann Arbor, MI: Ann Arbor Paperbacks: University of Michigan Press

Van Bouwel, J. (2022) 'Strengthening the Epistemic Case Against Epistocracy and for Democracy', *Social Epistemology* https://doi.org/10.1080/02691728.2022.2081823

Vandamme, P. (2020) 'What's Wrong with an Epistocratic Council?', *Politics* 40 (1): 90-105

Walter, N., Cohen, J., Holbert, R.L., and Morag, Y. (2020) 'Fact-Checking: A Meta-Analysis of What Works and For Whom', *Political Communication* 37 (3): 350-75

Westra, E. (2021) 'Virtue Signaling and Moral Progress', *Philosophy and Public Affairs*. 49 (2): 156-78

Williams, D. (2021) 'Signalling, Commitment, and Strategic Absurdities', *Mind & Language*. https://doi.org/10.1111/mila.12392

Wilson, W. (2013) *Five Myths About Nuclear Weapons*. Boston, MA: Mariner Books