Gaslighting Citizens
Eric Beerbohm and Ryan Davis

Gaslighting, as an interpersonal wrong, brings its victims to doubt the sources of their evidence. This paper holds that political gaslighting, by leading citizens to hold beliefs disconnected from the available evidence, poses a distinctive threat to democratic politics. But holding “audacious beliefs” — beliefs that are ahead of the evidence — can serve as a core ingredient for democratic movements. This creates a dilemma for citizens, who must choose between two kinds of evidential policies. How can they protect themselves from the gaslighting without rendering themselves insusceptible to the mobilizing efforts central to democratic politics? Citizens, then, face a standing challenge: to remain open to the bully pulpit while vigilant against the gaslighter’s epistemic bullying.

Members of political movements face daunting odds. Even minor changes to policy are hard won. Most attempts fail. If the movement succeeds, it’s hard to connect the success of the group to any individual member’s contribution. If you want to change the world, the objective chance that you’ll prevail is probably bleak. So it’s unsurprising that citizens collectively engaged in efforts to put a dent in the world have to adopt and maintain beliefs that — in some ways — extend beyond the evidence available to them. We’ll describe these beliefs as audacious. The empirical literature suggests that they play an important role in successful democratic efforts. But this distance between beliefs and evidence is also symptomatic of victims of gaslighting, whose perceptions and attitudes are at odds with reality. The charge of gaslighting — bringing people to doubt their own attitudes or capacities — has quickly gained popularity as an explicitly political charge (Hoberman 2019).

This essay explores gaslighting as a political phenomenon. We argue that gaslighting operates as a higher-order attack, threatening its victims’ identities. This sets it apart from ordinary deception in politics, and explains the distinctive threat that it poses to democratic citizenship. Our argument proceeds in six parts. First, we will sketch the concept as it has been developed in the philosophical literature. The second section applies the concept to cases of manipulation by political elites. It’s tempting to think that the problems associated with gaslighting might be
alleviated either by a policy of believing only on the evidence, or else by avoiding relations of epistemic dependence on others. But effective political agency can depend on shared beliefs — about ourselves, our opponents and our institutions — that go beyond the evidence. We explain the role that audacious beliefs play in persevering in our individual and collective projects. The final sections respond to a dilemma for citizenship as an ideal that governs our evidential policies. How can citizens protect themselves from the gaslighter without rendering themselves insusceptible to the mobilizing efforts central to democratic politics?

1. Gaslighting: The Personal

In the play and subsequent movie, Gaslight, Gregory intentionally attempts to cause Paula to doubt her sensory inputs. In the titular act of manipulation, Gregory changes the brightness of the gaslights in their home but insistently denies that there is any difference when Paula repeatedly notices that the lights have dimmed. This moment serves as a paradigm for Kate Abramson’s account. For her, gaslighting “induce[s] in someone the sense that her reactions, perceptions, memories and/or beliefs are not just mistaken, but utterly without grounds” (Abramson 2014, 2). The gaslighter’s final goal is not to lead the victim to doubt their own perceptions and beliefs, but to protect the gaslighter’s own conception of the world by eliminating even the possibility of challenge or disagreement (2014, 9-10). What’s crucial about this relationship is that the victim comes to endorse the gaslighter’s own beliefs (Abramson 2014; Spear 2018, 230). To secure this kind of epistemic capture, gaslighting involves a diachronic process in which the perpetrator seeks to undermine and ultimately destroy the “independent, separate, deliberative perspective from which the disagreement arises” (Abramson 2014, 10). Only when the victim’s own perspective has been thoroughly undermined can the gaslighter be confident that their worldview has been secured against the possibility of challenge.

The loss of the gaslit agent’s point of view has practical as well as epistemic consequences. The gaslighter aims to displace the victim’s self-trust with trust in — or at least, reliance on — their
beliefs, so in the first instance they seek control over the victim's belief-forming processes (Cf. Spear 2018, 2321). In so doing, they undermine the victim's ability to make assertions about the world, or their own experiences, apart from the gaslighter's signals about what may be said or believed. As epistemologists have recently pointed out, the practice of assertion carries in tow a claim to a kind of authority (Wanderer 2012; Goldberg 2015). In asserting a proposition as true, an agent implicitly claims to be in position to provide reasons in support of the assertion, or to be held accountable by other agents who accept and act on it (Tanesini 2016). To play these roles is to claim authority within a community of speakers (Goldberg 2015, 187-188). This authority is ethically significant, since holding and being held accountable in turn are morally serious. So the status as an asserter partly constitutes the agent's moral status. By compromising an agent's epistemic authority, gaslighting reduces one to the role of object rather than subject in the practice of sharing knowledge (cf. Fricker 2007, pp. 132-133). In this way, gaslighting also threatens the agent's moral status (Abramson 2014, 16).

We can draw together the shared properties from this conceptual work. At its core is the idea that its victims' rational faculties have been captured in a distinctive way:

*Gaslighting:* Agent [A] wrongly induces another agent [B] to doubt B's ability to respond rationally to evidence, in order to make B epistemically reliant on A.

This account marks out an end and the characteristic means. Gaslighters create a relationship of objectionable dependence by playing with their victim's grip on the evidence before them. A frequent attack vector is the victim's higher-order evidence (Feldman 2005; Kelly 2010). Our evidence about our evidence tends to have special vulnerabilities. Gaslighting can work by supplying higher-order defeaters, exposing the victim's beliefs as the upshot of a faulty process and leading victims to conclude that their attitudes were never justified in the first place.

This explains how gaslighting targets the victim's own point of view (Abramson 2014; Spear 2018; Stark 2019, 224). Whether you enjoy authority within a practice of assertion depends on your judgements about your own capacities. Although you might doubt your beliefs, you will not be
driven to question your authoritative standpoint as long as your confidence in your belief forming capacities remains intact. If the wrongness of gaslighting is grounded in compromising this standing, then it makes sense to understand it as undermining an agent’s confidence in these capacities (Cf. Sliwa and Horowitz 2015).

Everyday deception and manipulation, when successful, undermine our first-order evidence. Gaslighting, in contrast, functions as a higher-order attack, threatening its victim’s particular identity. In ordinary cases, “the import of a particular bit of evidence may depend on the thinker’s background beliefs, but it does not depend on who the thinker is” (Christensen 2010, 190). However, if you learn that your particular belief-forming process was defective, then you have a reason to treat the evidence differently than you otherwise would in virtue of facts about you, rather than facts about the first-order evidence alone. The gaslighter “claims that the target’s judgments lack credibility” as a result of “a defect in her” (Stark 2019, 224).

Contrast this with ordinary challenges to our first-order evidence. Suppose you become convinced that your beliefs in some domain are systematically incorrect, due to some clearly identifiable error. Perhaps you have a cluster of false beliefs about the city you’re visiting because you’ve been reading the map upside down, or you have a series of false moral beliefs because you have been persuaded of an extensionally incorrect theory. In cases like these, discovering that one’s beliefs are systematically mistaken — or even utterly without grounds — wouldn’t induce any skepticism about your authoritative standing within a community of moral agents. Your false beliefs might only say something about the tools you were using or the evidence you had. They don’t say anything about you. You may take comfort in the thought that this error could have happened to anyone.

The wrong of gaslighting is always relative to a particular agent or group. Put differently, whether a gaslighting speech act would challenge my epistemic agency depends on indexical facts — who I am affects how the utterance affects me. Part of what makes gaslighting insidious is that it
exploits the vulnerability of some agents while leaving others (characteristically outside the target group) perplexed about how anything seriously wrong could have been perpetrated at all.

The significance of the identity of members of the target group finds support in sociological work on gaslighting. Sweet (2019) explains that “gaslighting is effective when it is rooted in social inequalities, especially gender and sexuality, and executed in power-laden intimate relationships” (852). Intersecting inequalities may amplify these effects, suggesting that “women of color, poor women, immigrant women” and others may be especially vulnerable to gaslighting (Sweet 2019, 856). While not strictly necessary for the concept to apply, we will suggest in the next sections that this sociological finding can help inform understanding of gaslighting in political contexts as well.

2. Political Gaslighting

If deep enough, political disagreement can signal not only that our opponents are incorrect, but that they are failing to respond to the evidence in a remedial way. So begins a much sharper allegation: that our opponent is not just mistaken, but “crazy.” In a partisan world, the rhetorical force of this accusation is easily weaponized. If our opponents lack basic epistemic capacities, we do them no wrong by ignoring them, and encouraging others to ignore them as well. A polarized politics can seem like a target rich environmental for prospective gaslighters.

But our paradigm cases have taken place in intimate settings, involving close personal relationships, or at least their perverse facsimiles. It is not obvious that all of these ingredients could be in place in politics. And yet, politics is now the bellwether of the concept’s prominence in popular discourse. Does the charge of gaslighting travel from the personal to the political? Consider three examples:

(a) During the Republican Primary process, Trump repeatedly insinuated (and then denied he was suggesting) that George W. Bush was responsible in some way for 9/11, and floated conspiracist narratives conjecturing that the public did not understand the attacks (cf. Glueck 2016; Bort 2019).
(b) After his inauguration, Trump and his surrogates famously maintained that the inaugural crowds “went all the way back to the Washington monument” and “was the largest audience to ever witness an inauguration,” publicly disputing media and DC Metro accounts and photographs disconfirming these claims. (Robertson and Farley 2017).

(c) In July, 2019, Trump tweeted that four progressive congresswomen should ‘go back’ to where they came from, falsely implying that they were not US citizens. He then claimed, contrary to video evidence, that he did not originate the phrase, and that he tried to prevent a crowd at a rally from chanting it. (Rupar 2019; Crowley 2019).

These episodes have been the basis for accusations of gaslighting in popular media (Rupar 2019; Carpenter 2018). We’ll now consider how our account can explain the political wrongs at stake in these cases.

First, each of these cases involves denying facts which were either easily confirmed through secondary sources, or else could be perceived directly by onlookers. In order for the claims to be believed, citizens would have to discount the deliverances of their own capacities for observation and evidence gathering, or else refuse to exercise those capacities entirely. Second, Trump and his surrogates insisted that opposing evidence could not be trusted, including evidence from supporters’ own observations. “Just remember, what you’re seeing and what you’re reading is not what’s happening,” Trump advised one rally crowd (Rupar 2019). Third, the claims could be interpreted as part of a campaign to rule out counterargument and make members of the target audience epistemically reliant on the speaker (cf. Hahl, Kim, and Sivan 2018). Fourth, the claims invoked power asymmetries, including those created by intersectional inequality, in order to rule out certain deliberative viewpoints. In the case of the “send them back” tweet, the target group was comprised of minority women, including one immigrant – echoing the identity categories prominent in first-personal cases of gaslighting (Sweet 2019, 856).
There is evidence that these tactics do help facilitate the gaslighter’s central objective: insulating the speaker’s assertions against the possibility of challenge by occluding even the possibility of disagreement. Remarkably, when shown pictures revealing the crowd to be visibly smaller at the 2017 inauguration than in 2009, 15 percent of Trump voters denied the apparently obvious evidence and incorrectly identified the smaller crowd as larger. Politically-engaged supporters were more than twice as likely to choose incorrectly as their unengaged counterparts (Schaffner and Luks 2018). These results cohere with broader findings in political psychology that partisan identity strongly influences attitudes, and that partisans will simply follow elite signals when reporting their political views – including about seemingly factual matters (Bartels 2002; Flynn, Nyhan, and Reifler 2017; Bisgaard and Slothuus 2018). Partisan loyalties make politics an especially congenial domain for the gaslighter’s tactics.

We think that there aren’t two concepts here, but a common one. To extend the concept from the personal to the political domain, we propose this analogue:

**Political Gaslighting**: A political actor wrongly induces a group of citizens to limit the exercise of their belief-forming and revising capacities in ways that serve the political ends of the gaslighter, in order to bring about that the group becomes epistemically reliant on the gaslighter.

The exploit lies in the gap between the victim’s beliefs and their available evidence about the political world. When successful, gaslighting changes how they hold and handle those beliefs. On this account, victims needn’t come to doubt their own epistemic capacities. Far from questioning their capacities, the audiences of the tactics described in (a) - (c) reproduced the speaker’s attitudes with a high level of confidence. The manipulator expressly went out of his way to praise the audience’s capacities. This can seem at odds with the earlier aim of undermining confidence. In the case that launched the concept, the victim loses confidence in her beliefs. By stark contrast, strong partisans tend to be especially confident in their own experiences and perspective. Indeed, it’s this very confidence that inspires them to turnout in higher numbers at the polls (Ortoleva and
Perhaps paradoxically, those partisans who are most susceptible to manipulation also display the most confidence in their political beliefs (Anson 2018).

Our account allows for this variation by fixing on efforts to limit citizens’ exercise of their rational capacities. In the interpersonal case, the gaslighter sought a kind of control, making the target dependent on them. Political cases make clear that this can involve both undermining and inflating a target’s self-confidence. Both tactics may appear – though perhaps less conspicuously – in interpersonal cases as well. It’s not uncommon for victims of interpersonal gaslighting to confabulate false perceptions or evidence in support of the gaslighter’s claims, and thereby to become overconfident as well (Spear 2019). Likewise, some targets of political gaslighting report anecdotally the experience of having their reality destabilized, as in personal cases (cf. Carpenter 2018, 67). Either way, gaslighter’s central aim is to protect their own worldview against the possibility of challenge. Sometimes, this aim might be better achieved by removing, rather than creating, the sensation of self-doubt. The gaslighter’s aims cannot be fulfilled without the target being “in some way invested in what the manipulator believes” (Stark 2019, 223). That is, the target audience must somehow share an identity with the gaslighter.

In politics, partisanship furnishes the shared identity that motivates such an investment in avowing the party line. Tribal allegiances risk erasing the individual citizen’s independent point of view. Barber and Pope take advantage of Trump’s tendency to assert both conservative and liberal opinions in a variety of policy areas (e.g. minimum wage, taxes, abortion, etc.). They find that if primed with a statement from the President, Republicans mostly followed Trump’s lead, regardless of whether he was saying something liberal or conservative (2018, 42). They conclude that “partisan loyalty is more relevant to a large group of Republicans than is any kind of conservative issue preference” (43). Because of their investment in what their respective elites say, partisans are especially susceptible to manipulation. If beliefs are constituted in part by a kind of modal stability — or resistance to reconsideration (Friedman 2017), then the partisan follower might lack any
beliefs at all (cf. Mason 2018; Bisgaard 2015). Even if they don’t exhibit any actual doubts about either their beliefs or their belief forming processes, they display a functionally similar epistemic profile. They speak and act like people who are dependent on signals from others for their own beliefs, or attitudes that don’t display the stability of belief states. Just as in the paradigm case, successful gaslighting produces followers who do not just defer to elites on a single issue, but will follow them across a wide variety of issues without much regard for accuracy (Swire et al. 2017). In the edge case, the gaslit citizen’s point of view is ushered off the deliberative stage of democratic politics altogether.

This points to the core continuity between gaslighting in the personal and political orbits. The charge issues an accusation: the gaslighter wants to rule out challenges to their own view, and so must utilize strategies to prevent the target from threatening their favored story. In all cases, the destination was not the self-doubt of the target in its own right, but the elimination of the target’s distinctive standpoint. By “standpoint,” we have in mind the perspective from into which evidence and observations are collected, and from which judgments about the world and intentions about how to act are issued. In other words, the gaslighter is not seeking to eliminate another’s standpoint by – say – drugging or killing them, even if such outcomes could be achieved without consequence. The gaslighter wants to leave the target’s basic agency intact, albeit to be deployed according to the gaslighter’s discretion.

The standpoint of the gaslighter’s target is not the only perspective at risk. For in the domain of politics, when things go well, there isn’t just a collection of individual points of view, but also a point of view that can be shared. This idea doesn’t turn on an extra-strong view about collective agency. If we assume there’s another agent in the room, “the people,” an argument revealing a collective victim of gaslighting comes relatively easily. But we think that there is a more ecumenical premise that will allow the argument to go through. All that’s needed is the attractive idea that fellow citizens, when they reason with each other, are engaged in something like shared
action (Stilz 2016). It suffices that my reasoning about what we should do together be sincerely open to considerations provided by other agents, and that my judgments about shared plans be responsive to them. Without an Archimedean point like joint deliberation, the very concept of a “common good” looks like a category mistake (Beerbohm and Davis 2017).

If this premise about joint deliberation is plausible, then political gaslighting does, after all, threaten a morally valuable point of view. Recall that the gaslighter seeks to cut off the target from other sources of information, instead becoming an exclusive source of inputs for beliefs and planning. Sidelining deliberative inputs from other sources thwarts citizens from gathering normative and empirical data from each other, impairing their ability to forge a shared “common good” (Cf. Westlund 2009). It shouldn’t surprise us that when people reason only with co-partisans who already hold a common outlook, they are more likely to show directional rather than accuracy motivation (Klar 2014). We cannot get at the truth of common values when gaslighters have their way, and competing sources of information are eliminated. Gaslighting compromises the shared perspective of citizens together.

Nor is this the last result. We exercise our own reasoning capacities best when we have to sort through reasons that interact in complex ways. The messiness of politics requires that we reflect on differing inputs. So we should worry that ruling out listening to political opponents – as the political gaslighter would hope – will compound difficulties with thinking on our own. And there is reason to think this fear is born out in real-world politics. Citizens who are more reflective, or more willing to scrutinize their own political beliefs, are more able to see the virtues of those they disagree with (Arceneaux and Vander Wielen 2017, 135-151). Hating one’s opponents predicts diminished reflectiveness. If we follow the old Kantian idea that part of being an agent is being able to exercise a capacity to reflect on one’s own attitudes, then inducing incendiary views toward opponents might compromise an individual’s own agency, after all. In short, political gaslighting can damage a target’s point of view, and also undermine citizens’ collective capacity to think from a
shared point of view. And that, in turn, compounds the problems for each individual’s own agential perspective. Once begun, the gaslighting’s contagion can spread throughout the civic landscape.

The mechanisms of political gaslighting may be more variable than in pairwise cases, and may rely on shared in-group identity rather than personal intimacy. In effect, the political gaslighter seeks to install themselves in the role of trusted friendship, but without any of the features of an ordinary personal relationship making that trust appropriate. Then the gaslighter seeks to exploit the relationship to rule out voices other than their own. Is such manipulation unavoidable in democratic politics? In the next section, we will consider whether gaslighting can be circumscribed in such a way as to avoid covering modes of citizenship that are morally important for democratic practice.

3. Audacious Believing

Gaslighting targets an agent’s confidence in their rational capacities. Deployed in politics, it attempts to undermine the confidence of groups of citizens in their rational capacities, or at least induce them to act as if their confidence were undermined. The effect, again, is to eliminate a perspective – either of an individual or of a group – from deliberative consideration. In personal and political cases alike it renders the target agent dependent on the gaslighter as a source of beliefs and other commitments. When it succeeds, the victim’s beliefs are responsive to the gaslighter, and so become disconnected from the evidence. Yet severing the connection between belief and evidence “raises the specter of an inflexible dogmatism” among citizens who take their cues from party elites, making such dependence “problematic for many conceptions of good citizenship” (Druckman, Fein, and Leeper 2012). In general, political scientists have been sympathetic to the idea that citizens’ beliefs ought to aim at accuracy or the truth – that is, that they should be apportioned only according to epistemically relevant considerations (Hochschild and Einstein 2015). What complicates this picture is that there are plenty of practical cases in which a
little inflexibility — even what looks like selective dogmatism — can look like a part of practically rational agency. This section will introduce the role of motivated believing for democratic politics. It will bring into focus a dilemma for the citizen who wants to both avoid gaslighting and participate in political movements.

*Individual Audacity*

Suppose you commit to climb Mount Everest, or to run a marathon. Each day of training, let’s presume that you believe that you will complete this project. But you are also aware of the considerable evidence that suggests that you will fail to make it to the summit, or the finish line. Perhaps you’ve seen the hard numbers on the success rates of people with your similar training and background. If your belief that you will succeed persists, it does so in spite of the relevant evidence. Your hope that you will complete these difficult projects rest on the expectation that your training will pay off (Duckworth 2016). We will describe rationally permissible beliefs that place greater credence in the prospects for one's success than the evidence requires as audacious beliefs.

This definition prescinds from a variety of epistemological possibilities. Audacious beliefs may be formed in part in response to pragmatic rather than evidential considerations, may discount evidence of failure for actions that are somehow “up to us” (Marušić 2015), or may land on the optimistic side of the rationally permissible range of possible beliefs (Preston-Roedder 2013) — among other possibilities.

An agent with audacious beliefs will differ with what a rational third-party believes about them (McCormick 2014; Marušić 2015). Consider the exchange:

C-3PO: "Sir, the possibility of successfully navigating an asteroid field is approximately three thousand seven hundred and twenty to one."
Han Solo: “Never tell me the odds” (Brackett et al. 2006).
Here C-3PO is inviting Solo to take up a predictive stance — perhaps as a result of his mechanical make-up, he may not be capable of taking P.F. Strawson’s participant stance (Strawson 1962). But we can bracket that. Solo is seeking to avoid evidence here, preventing him from acquiring a belief. On one interpretation, Solo has already committed to fly through the field, and has no other choices, so learning the odds would have no value. But perhaps Solo appreciates the threat this testimony poses. A second take suggests that he resists the evidence because it may lead him to abandon his project. Does Solo really need to believe that he will navigate the asteroid field? It may be enough for him to avoid holding C-3PO’s probabilistic report. If he came to believe that the odds were vanishingly small, he couldn’t sincerely retain his commitment to the mission.

Even if we suppose that beliefs should be responsive to evidence of what is true and only such evidence, there is a further question about policies towards evidence. How strong must opposing evidence be for it to rationally demand that a belief be abandoned or revised. Morton and Paul call this the “evidential threshold” for belief change. Even an evidentialist about belief could allow that equally rational agents might yet have different thresholds (Morton and Paul 2019, 191; Lawlor 2014). If evidence alone does not settle where to set the threshold for belief revision, then an agent’s implicit policy about when to revise a belief could depend in part on practical considerations.¹ Individuals could hold beliefs sensitive to the evidence, but also have high thresholds for evidence contrary to the future success of their aims. Such agents would hold epistemic policies exhibiting a kind of “grit” (Morton and Paul 2019).

In adopting this evidential policy, Solo may have told C-3PO: “Feel free to tell me the odds, but my evidential threshold for considerations against my success is extremely high.” Solo’s insistence that he will be the exception to the rule — to his droid’s modeling — can be defended as

¹ Although we find this outlook appealing, it does presuppose a commitment to a kind of permissivism about belief, which roughly denies that one’s total body of evidence always picks out a uniquely rational doxastic state. For skeptical views on permissivism, see, for example, Horowitz (2019) and White (2007).
resilience. There is something audacious about this kind of policy — no one would accuse Solo of modesty. But it needn’t be irrational either.

**Collective Audacity**

Social movements are prone to self-doubt. Marked by loose connections, small numbers, and powerful opponents, their members regularly encounter evidence that they might fail. Their shared aims, after all, are difficult, long-form and risky — what James Baldwin called the “dangerous road.” No less than marathons and mountain-climbing, the joint commitment to bring about transformational change seems to require audacious believing of some kind. In his classic work on the 1880s Farmers’ Alliance, Lawrence Goodwyn picks out “collective self-confidence” as a central building block of mass democratic politics (Goodwyn 1978).

To see collective confidence — even over-confidence — at work, consider the logistical backend of the Montgomery Bus Boycott, which lasted 381 days. The night Rosa Parks was arrested, on December 1, 1955, Jo Ann Robinson mimeographed 52,000 leaflets. Within a day, these were distributed to sixty-eight black social organizations across Montgomery. The boycott was planned to last one day, on Monday, December 5. It was far from clear that the boycott could succeed as a one-off protest. To sustain it through December 20, 1956 required a vast carpool network, “325 private cars transported passengers from 43 dispatch stations and 42 pickup stations from five in the morning to ten at night” (Tufekci 2017, 64).

Can audacious belief, seen as a first-personal attitude, scale to collective projects? If so, what role can shared beliefs play in the persistence of joint efforts in the face of evidence about their diminishing chances of success? At the group level, the policies of resilience from the last section are complicated in two ways. First, members of movements needn’t only hold beliefs about
themselves, but in each other. So we have to make sense of what this other-directed attitude — believing in our co-citizens — could look like.²

Consider Martin Luther King Jr.’s confidence in the arc of the moral universe’s direction. This metaphor has the potential to mislead, if we take King to have announced a belief that he has simply read off from the world. His confidence wasn’t merely a function of proportioning his (cosmic) beliefs about humanity’s future to the evidence. It took considerable effort and epistemic resilience. In response to the Black Power movement’s rejection of nonviolence, King conceded that this approach was a function of “despair and disappointment.” The evidential basis for this despairing was as clear to King as his opponents. It was, in Brandon Terry’s words, “a response to several ugly facts, including the frequency with which white brutality continued to go unpunished (even after Selma)” (Terry 2018, 278). King was alive to the sense in which members of the Black Power movement were proportioning their beliefs to the cruel evidence in front of them. If they dug down in their response to that evidence, King insisted that alternatives to nonviolence would backfire, furthering the sense of powerlessness and raising the temptation of sour grapes-style reasoning.

King’s worry about the self-defeat of such tactics wasn’t volitional, but epistemic: by a process of elimination, nonviolence was the only stance compatible with preserving the confidence of movement members about their victory. Put into practice, King thought that alternatives to nonviolence would lead members to believe that there was no hope. King’s Nobel Prize Acceptance Speech helps explain, if not fully defend, his justified optimism: “With an abiding faith in America and an audacious faith in the future of mankind,” King employs language compatible with William

² In “Believing in Others,” Sarah Paul and Jennifer Morton defend the permissibility, and sometimes requirement, to believe our friends’ capacity to stick to their projects, even when the evidence points in the other direction. Their defense of this way of believing in others closely parallels their defense of raising our evidential threshold to protect our beliefs about our success; they think we can raise that threshold for evidence that bears on the success of intimates. Our question is whether this stance can extend to political relationships (Paul and Morton 2018).
James’s pragmatism, “I refuse to accept the idea that the ‘isness’ of man’s present nature makes him morally incapable of reaching up for the eternal ‘oughtness’ that forever confronts him” (King and Washington 1991). The acknowledgement of its audacity is probative evidence that this kind of believing isn’t textbook evidentialism. King adopted an evidential policy that permitted him to act — in this case through speech acts — in the absence of, or against the grain of evidence.

King faced pressure to account for the basis of his optimism. Robert Williams, in his extended debate with King over nonviolent means, asked: “Can a program of nonviolence... realistically expect to deal with such an enormous, entrenched evil?” (King Jr 2010, 52) If audacious believing has rational credentials that wishful or other kinds of motivated reasoning lack, King has resources to defend his epistemic resilience from this challenge. In individual and collective cases of audacious believing, moral and political considerations play a legitimate role in how we handle evidence. They raise the evidential threshold for considerations that suggests that ought doesn’t, in the end, imply can. These considerations support agents’ confidence that they can succeed in doing what they morally ought to do.

For us to form audacious beliefs that are shared about our cause, it is not enough to believe in any given co-member of our movement — in a pairwise way. For it seems that you need to hold commitments whose direct objects are first-personal plural. The object of confidence, then, is doubly collective: not only must I believe that “we” will succeed in our project. It may be important for us to share this attitude. So you will need a belief about my confidence in our project, and vice versa. Efficacy-directed slogans like “yes we can” shouldn’t be passed over as mere cheap talk or overblown rhetoric. The confidence, even overconfidence, of members of social and democratic

---

3 This passage is cited in Terry 2018, 307.
4 Beliefs in group efficacy must be shared, since “individuals are more focused on achieving group goals through the joint effort of collective action.” (Van Zomeren, Leach, and Spears 2010, 1056).
5 In some cases group efficacy beliefs are produced by reasoning together. When group members discuss plans together, they become more confident in the success of joint projects. Such attitudes are fundamentally shared. See, for example, Thomas and Louis 2013.
movements makes possible signals of their strength that are outsized. There is considerable instrumental value in the protestor’s portfolio: rallies, occupations, speechifying marches. All of that talking and gathering is a proof of concept of the depth and breadth of the front-line.

4. Gaslighting vs. Mobilizing

When faced with evidence that their political aims may fail, citizens who persevere demonstrate a kind of epistemic resilience. They come to hold beliefs that are less than tightly connected to the evidence. Or at the very least, they hold evidential policies that are informed, perhaps implicitly, by moral considerations that sustain their ability to keep their commitments. To secure steadfastness, mobilizers attempt to ensure that members’ shared beliefs in their eventual success will resist defeat by the inevitable presence of countervailing evidence. The mobilizer, in other words, seeks to influence members’ epistemic policies in a way that insulates them against the diversity of opposing evidence. While here we will focus on evidence relating to self-efficacy, collective audacious beliefs make up a much larger set of views, relating to the evidence about the good will of one’s political opponents, the strength of prevailing norms and their susceptibility to violation, the capacity of one’s home institutions to change.

This slack between beliefs and evidence is diagnostic of gaslit citizens. To hive off followers from beliefs that may undercut their authority, gaslighters tell their audiences that core sensory inputs of the citizen — the news media — is nothing more than a counterfeiter of evidence. Any beliefs that citizens have acquired through this testimonial mechanism should be, by extension, abandoned. Gaslighters will protect their followers’ higher-order evidence from certain classes of counter-evidence about their evidence. Or they will present defeaters designed to challenge their target’s higher-order evidence.

One might counter that there is no deep dilemma between gaslighting and audacious believing. After all, it’s relatively straightforward to think that one should set a high evidential threshold when it comes to confidence in one’s individual or collective projects, but still regard it as
an obvious error to ignore evidence that a leader is lying or exploitive. Why, the objection goes, should we worry that an audacious believer is at some special risk of gaslighting?

To get at this issue, it will help to work through specific cases. For simplicity, we’ll work with an individual case of audacious belief, offered by Paul and Morton:

When the coach cheers her trainee on from the sidelines, the swimmer need not be outright deceived about what the coach is up to (though he might well refrain from reflecting explicitly on her motives). And if he believes that being confident is going to help him succeed, he might reasonably expect his coach to spin the evidence positively to encourage him (Paul and Morton 2018, 92).

The story features perfectly familiar motivational techniques. The coach is trying to get the swimmer to adopt audacious beliefs, but certainly is not gaslighting. For one thing, the swimmer retains an independent standpoint on the evidence. One way of seeing this is that we are told the swimmer and coach share an end, so the swimmer can “reasonably expect” the coach’s spin on the facts. When listeners and speakers share an end that is advanced by a given assertion, and there is a kind of implicit agreement that speech will be used in this way, then it seems morally innocuous.

However, suppose the coach decided it might help to put just a little more spin on the facts. Say she starts telling him things that are not – strictly speaking – true, but which might be made true if he became sufficiently convinced of them. Or, in a step further, suppose she tells him things which she doubts could be made true by any level of belief, or which she actively disbelieves. Still, she might reason, such false beliefs would be instrumental to inspiring still faster times. Perhaps it would also help the swimmer to develop a more positive self-image generally – not just about his capacities as a swimmer. Little by little, the coach increases the range of her advice – extending to swimmer’s intelligence, popularity, relationships, etc.

Eventually the coach may exceed any implicit agreement on how speech should be used within the relationship. Flagging this fact, the swimmer might ask, “Do you really believe any of these words, or are you just trying to get me to believe it?” Now imagine that the coach replies, “Don’t worry about it. Just keep your mind on the prize. Focus on what I tell you till you get yourself
to believe it.” Now it looks like the coach has entered the territory of the gaslighter. She is not just encouraging the swimmer to set a high evidential threshold, but is attempting to get him to replace his own judgements with hers. She is actively talking him out of holding his own, independent deliberative standpoint on the evidence.

What’s ambiguous is the space in between the case as described by Paul and Morton, and this final case. Where, we might wonder, does the implicit agreement (signaling that swimmer’s independent sharing of the end) run out? Perhaps there is no final fact of the matter about precisely how far their agreement extends. The coach, motivated merely to maintain the swimmer’s audacious beliefs, might slip into gaslighting without realizing it.

Let’s turn to a political case. Consider this exchange between a Latino canvasser and persuadable voters:

[I canvassed] 18 individuals that were, Mexicanos, Hispanos, even an Asian couple, they all opened their doors and actually hear what I had to say... But more importantly, were enjoying it and were like, “okay, we’re going to vote for that.” It wasn’t out of the jaded perspective that no matter what we say, they’re going to still do what they want. It was more “I feel what you’re saying and I feel your passion. And I believe you now.” And I was like, “Yes!” It’s important that we believe that we have the power to make changes instead of feeling like we're going to continuously be inferior to everyone else.” (Carson, Abrajano, and Bedolla 2020, p. 126)

We take this to be a paradigm case of an audacious belief. Our mobilizer is attempting to change the target audience’s sense of their social position and power. There is clearly an affective and group-oriented component. The canvasser is enthusiastic about spreading a shared sense of collective empowerment. There is also a doxastic component. It’s not just that the mobilizer is highly motivated. This motivation is connected to — and partly maintained by — a belief that the group does have the power to make a difference. Third, this belief in group-efficacy is regarded by the speaker as “important” not merely because it is justified, but because of what it helps to do. The significance of the belief is that it helps to counteract feelings of inequality and subservience to other groups — perhaps those with greater extant political influence. Believing in one’s power to
make a difference is part of defeating the political temptation to feel “continuously inferior to everyone else.”

Our mobilizer is presenting an audacious belief. It is a belief in the prospective efficacy of the group of which the speaker is a part, based partly on evidence (in fact, the speaker and the group really are not the political inferiors of others in society) but also partly on the desirability of the belief (it is important to think that one has the power to make a difference in order to realize this power). The latter component shows that the belief is not required by the evidence; its grounds include evidence and pragmatic usefulness.

Now consider two ways the speaker could proceed:

(a) The speaker could continue to (i) talk about how the speaker and the listener can make a difference through exercising their capacities (ii) in ways that would help them realize ends they selected themselves and (iii) that demonstrate their equality to other members of the political community.

(b) The speaker could talk about how (i) the listener should go along with the speaker’s assertions without further reflection (ii) in ways that would help them realize ends that may or may not be shared by both parties, while (iii) insisting on or praising the listener’s loyalty to the speaker or other leaders, as opposed to other sources of information.

We think that in the first case (which maps onto the actual passage), there are indicators the canvasser and listener alike are exercising their belief-forming and revising capacities. The canvasser quotes a listener to show uptake of the fact that they actively shared the same ends. Their exchange indicates they also share an affective motivation and sense of shared group-efficacy. The speaker also indicates that both of their beliefs are sincere, expressed by the tone of the exchange as well as the observation that interlocutors did not have a “jaded perspective.” The canvasser has a political project, but there is a sense of empowerment creating conditions of equality between agents. In the second case, the speaker does not want the listener to exercise their own capacities, either to form beliefs or select political ends. The assurance of equal standing has disappeared, and an effort is afoot to exclude other sources of information.
Suppose we further learned that the speaker and listener would continue to support their current political causes, even if the leaders of their cause promptly switched sides. Or suppose that if the speaker acquired disconfirming evidence, they would update their beliefs (or at least treat this evidence in other ways as deliberatively relevant). Such actions would suggest the speaker’s audacious beliefs were maintained through capacities whose exercise did not depend wholesale on someone else. Alternatively, suppose that the speaker had the audacious beliefs as described, but was robustly disposed to switch directions entirely upon receiving differing signals from the relevant elites. That would suggest that while the speaker had avowed their own “power”, it was more a proxy for an external power.

In both the swimmer and canvasser cases, we can connect gaslighting to a belief’s pedigree, as well as to its maintenance. Two swimmers might have the same sets of audacious beliefs, but if one had them out of unconstrained deference to a coach’s insistence, in the absence of clearly shared ends, then that pedigree is probative evidence of gaslighting. Two canvassers might try to inspire others in the audacious agenda, but if one would easily shift beliefs at the whim of a party elite, then we could worry their attitudes had been maintained through gaslighting. Given the ambiguities of either relationship, the line between mobilizing audacious beliefs and gaslighting will not be self-explanatory.

5. A Dilemma for Citizenship

Two aspects of our ideal of active citizenship stand in tension. The first is an ideal of shared democratic agency. Without the ability to pursue long-term projects with others, we can feel hopeless — without a sense of co-agency. The second is an ideal of independence. We want to maintain our own point of view within a community of self-reflective, reasoning agents. But the
same evidential policies that can, under some circumstances, empower us as gritty participants supporting inspired leaders may, in less fortuitous times, leave us duped by gaslighters.  

Can citizens find policies to maintain active citizenship without falling prey to gaslighting? One possibility is to make ample space for audacious belief, giving the benefit of the doubt to ambitious mobilizers. But the pedigree of such beliefs, as well as the policies for maintaining them, may reveal one’s agency has been compromised. Gaslighting can be morally costly. When you believe self-serving falsehoods about your political opponents, you risk wronging them with your beliefs (Schroeder 2018; Preston-Roedder 2013). Failing to see evidence from the other side — even the inability to "code" their assertions as evidentially relevant — puts you in a position to inflict testimonial injustice on them (Fricker 2007).

Alternatively, one could keep one's focus on the evidence, not allowing oneself to be tied too closely to any political group. While this would avert the risks associated with gaslighting and concomitant failures of justice, it would impose costs on mobilizing and maintaining the solidarity of groups of citizens.

These two strategies both simplify to a fault. They do this by reducing the conceptual space between the evidentialist citizen and the activist citizen to a single dimension. Citizens might opt instead for a middle ground, division-of-labor strategy. Some members of a political group could serve as thoroughgoing, hard-nosed evidentialists, while others are visionary believers. On this proposal, what we need is not a single “correct” norm to follow, but a variety of citizens with different epistemic and practical virtues. We saw King’s self-reports about his epistemic confidence, which included the disclaimer that they were indeed audacious. But King didn’t expect that the activists’ epistemic policy came in one size. There was room for non-theists buy-in so long as they continued to believe that “something in the universe unfolds for justice” (King and Washington

---

6 William James mused that evidentialism was based on the "preponderant horror of becoming a dupe." He countered that there are “worse things than being duped may happen to a man in this world” (McDermott 2013, 727).
This belief is surely weaker than King’s own “infinite hope.” It’s even less rosy than its variant, the arc of the moral universe standard of patient optimism. For King, the Civil Rights Movement can hold up even with, or perhaps benefit from, a division of evidential policies among its members.

One cost to a division-of-labor strategy is that it backs away from an ideal of genuinely shared deliberation. Citizens reason together, but they may have trouble finding ways of reaching agreement. If the evidential norms they follow diverge, then they may permanently differ on what considerations count as reasons, and on how to combine those considerations. This solution risks leaving out the shared perspective, described earlier, from which “deciding together” is possible (Westlund 2009).

If a mixed strategy could be forged that ensured genuine deliberative cooperation while also resisting gaslighting, it might shed light on how to retain the virtues of civic mobilization while avoiding the perils of gaslighting. One thing to bear in mind in this process is that audacious believers can – and characteristically do – remain sensitive to new contrary evidence, even when such evidence does not reach their threshold for belief revision. King did not encourage members of the civil rights movement to ignore evidence of obstacles to their success, though he also didn’t want them to lose confidence either. There is no inconsistency in this middle ground. Evidence might affect an agent’s attitudes in other ways – such as modifying her plans and intentions – even if it does not change her beliefs (Morton and Paul 2019, 196; Friedman 2017). Remaining open to new evidence also ensures that when the threshold for belief revision is reached, agents will be able to recognize it and respond accordingly.

Here, then, is another contrast between the mobilizer and the gaslighter. Unlike audacious believers, gaslighters do not characteristically want their followers to appreciate opposing evidence at all. Their message is not that their political group can overcome obstacles, but that there are no obstacles, or that the obstacles have already been defeated, or are mere illusion propagated by their
enemies to discourage them. Recall the examples of prominent political gaslighting. The first relied on insinuation to avoid being held responsible for one’s claim, thereby avoiding any demands to provide evidence for it (cf. Camp 2018). In the second case, gaslighting took the form of a bold or transparent lie. If we treat transparent lies as supplying faux evidence at the first-order level, they can appear downright puzzling. Why attempt to deceive in such a bald-faced way? But seen as a second-order challenge — as an attack on our evidence about our evidence, the transparent lie can be a core part of the gaslighter’s toolkit. For they can curry favor with members of one’s political in-group, showing how one’s dedication to their cause defies any opposing evidence (cf. Hahl, Kim, and Sivan 2018). In the third case, the “send them back” chant tried to exclude opponents from the set of agents with standing to provide evidence at all.

In each of these cases, the gaslighter deploys tactics to refuse or avoid evidence, rather than acknowledge it while remaining steadfast in one’s beliefs. To avoid the risks of gaslighting, the audacious believer can conscientiously attend to evidence opposing her group’s success, including evidence provided by political opponents. The error in gaslighting is not just that it raises the evidential threshold for belief revision, but that it denies or avoids opposing evidence altogether.

This difference, in turn, sheds light on what makes the role of the mobilizer morally important. The mobilizer will find ways to direct followers to appreciate all of the evidence while remaining confident in their own views. For example, given the chance to passively ignore false disparaging information about a political opponent, the mobilizer will correct followers’ misperceptions – even if those misperceptions might have aided their political cause. Or, given the option to accept favorable conspiracy theories which, if true, would make success more likely, the mobilizer will take steps to eschew such beliefs. In this way, mobilizers begin to make space for epistemic contributions from epistemic authorities outside of their group (e.g. scientific or journalistic sources), as well as acknowledging valid contributions from their political opponents.
This last approach tries to marry resistance to gaslighting together with active citizenship by way of an ideal of shared deliberation. Thinking with others resists gaslighting and improves political activity. Of course, this is the most familiar of philosophical proposals for democratic citizenship. It also is only the beginning of a path forward negotiating the tension between being open to active citizenship but wary of potential gaslighters. Exactly how and when to be open to political enemies and allies are matters of recent ethical and epistemic controversy (cf. Fantl 2018; Rini 2018). Like the joke about offering directions by suggesting, “I wouldn’t start from here,” there may be times when our prior beliefs create epistemically adverse conditions. It can happen that there’s some doxastic response that your epistemic position makes appropriate, but that there’s no epistemically good dynamic route from your current attitudes to that response. Why? Because you may also have (unjustified) beliefs in U that get in the way of having a doxastically justified belief in Q. And it might not be permissible for you to refrain from believing Q, either, since after all believing Q is what your evidence does support.

**Conclusion**

No friend of democracy, James Fitzjames Stephen held that in democratic politics, “the ruling men will be the wire-pullers and their friends” (Stephen 1967, 239). It’s tempting for citizens to see themselves as either victims of manipulation or one of its agents. Here we’ve switched back and forth from these two points of view. Each supplies us with a way of “seeing as” a citizen: victim or agent. Gaslighting and mobilizing both target how citizens handle evidence about their political world. We’ve argued that gaslighting manipulates citizens in a way that is as morally distinctive as it is disturbing, posing a threat to our agency that can perversely make us complicit in the process.

---

7 Ludwig Wittgenstein uses the rabbit-duck illusion to distinguish between “seeing as” and “seeing that.” Here we’ve attempted to take seriously the former, first-personal perspective on citizens as they manage the evidence before them (Wittgenstein 2009, 194-99).
Yet insisting on exercising our own epistemic capacities, and fortifying against gaslighting, can threaten our ability to sustain meaningful political action with others. What’s at stake is the felt need of citizenship — to see our own agency in our home institutions. In ordinary life, we experience our agency most dramatically when we manage to stick to commitments over time, allowing us to complete difficult and risky projects. That need is no less significant in democratic politics, where confidence that we will be jointly efficacious is part of the price of admission. Mobilizers, when successful, manage to change the way their audience handles the damning evidence in front of them. They get them to hold beliefs, or evidential policies, that protect them from a sense of futility and hopelessness. We are left with a standing challenge: to remain open to the bully pulpit while vigilant against this form of epistemic bullying that targets citizenship.

---

8 This theme runs through Waldron 1999, 101-118 and Tuck 2016.
References


