Gaslighting Citizens

Eric Beerbohm and Ryan Davis[1]

> [L]eaders...have argued that if their followers or subjects are not strong enough to stick to the resolution themselves, they—the leaders—ought to help them avoid contact with the misleading evidence. For this reason, they have urged or compelled people not to read certain books, writings, and the like. But many people need no compulsion. They avoid reading things, and so on.
>
> — Saul Kripke, "On Two Paradoxes of Knowledge"

Politics invariably involves disagreement—some of it, unreasonable. If deep enough and fundamental enough, disagreement might be taken as a sign not only that one of the opposing disputants must be incorrect, but that someone may be somehow failing to respond to the available evidence in a minimally rational way. So begins a much sharper allegation: that one's opponent is not just mistaken, but crazy. In a partisan world, the rhetorical force of this accusation is easily weaponized. If one's opponents lack basic epistemic capacities, one does them no wrong by ignoring them, and encouraging others to ignore them as well.

"Gaslighting"—or attempting to cause people to doubt their own attitudes or capacities—has quickly gained popularity as an explicitly political charge.[2] Antagonists on the right and left both mutually accuse each other of gaslighting. They define the term similarly, so the disagreement looks substantive.[3] But the opposing outlooks may share little besides the concept.

This essay aims to understand gaslighting as a political phenomenon. It proceeds in six parts. First, we will sketch the concept of gaslighting as it has been developed in the philosophical literature. The next two sections apply the concept of gaslighting 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. Complicating these suggestions, section four describes how effective political agency can sometimes require distinctively shared beliefs that go beyond the evidence. We explain the role that what we call audacious believing plays in persevering in our individual and collective projects. The final sections respond to a dilemma for citizenship as an ideal that governs our evidential

---

[1] This early draft is the first piece of a longer project on the varieties of manipulation to which democratic citizenship is especially susceptible. For useful discussion, we are grateful to Leslie Duhaylongsod, Dan Moller, Eric Nelson, Avishay Ben Sasson-Gordis, Wendy Salkin, Don Tontiplaphol. Since its completion will require us to persist in the face of (strong?) evidence of failure, we would be grateful for your comments and suggestions: beerbohm@fas.harvard.edu; rwdavis@byu.edu

[2] Emma Brockes, "From gaslighting to gammon, 2018's buzzwords reflect our toxic times," *The Guardian*, 18 November 2018; Sara Spary, "What You Need to Know about the Words of the Year 2018," *Huffington Post* 11 July 2018.

[3] For example, Corey Lewandowski, "Trump Haters Are Gaslighting America on the Economy," RealClear Politics, August 2, 2018. Lewondowski: "Gaslighting is when a person manipulates other people into questioning their own perception of reality." Ariel Leve: "The term gaslighting refers to when someone manipulates you into questioning or second-guessing your reality." See Leve, "How to Survive Gaslighting: When Manipulation Erases Your Reality," *The Guardian*, March 16, 2017.

policies. How can we protect ourselves from the gaslighter without rendering ourselves insusceptible to the mobilizing efforts central to democratic politics?

1. Gaslighting

Our concept derives from the play and subsequent movie, *Gaslight*, in which 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, according to which gaslighting aims "to induce in someone the sense that her reactions, perceptions, memories and/or beliefs are not just mistaken, but utterly without grounds."[4] This process is diachronic. Gaslighting involves not merely ignoring or dismissing another person, but a hostility to even the possibility of challenge.[5] Its most extreme success condition is the destruction of the "independent, separate, deliberative perspective from which the disagreement arises."[6]

This proposal is morally and psychologically complex. It will help to look a little more closely at its component parts. On Abramson's account, gaslighting is one form of manipulation. Manipulation is often understood as a way of influencing another person that is pro tanto wrong, but at least may be somehow distinct from coercion or deception—the paradigmatic types of interpersonal wrongdoing. Abramson suggests that the gaslighter's primary strategy is not—at least in the first instance—epistemic. Instead, she argues, the gaslighter poses as a source of normative authority. The gaslighter assumes the pretense of sincere testimony, drawing on the standing to issue demands that others "see things his way."[7] In this kind of case, the gaslighter forgoes efforts to change how the victim sees the evidence, and tries to change her normative situation through the direct application of putative authority.

It follows that gaslighting is an especially pernicious form of manipulation, for two reasons.[8] First, the gaslighter attempts not only to make unwarranted demands on what the recipient believes, but also tries bring it about that the other agent "not occupy a standpoint from which challenges might be issued" in the first place.[9] If the gaslit agent can be made to doubt the deliverances of their epistemic capacities, they can be subject to the gaslighter's demands well beyond any given one-off case. Second, the gaslighter demonstrates a particularly objectionable form of contempt. If contempt is wrong because it treats a member of the moral community as though they were outside of that community, then gaslighting attacks precisely this standing. Membership in the moral

---

[4] Kate Abramson, "Turning Up the Lights on Gaslighting," *Philosophical Perspectives* 28 (2014), p. 2.

[5] Ibid., p. 9.

[6] Abramson, ibd., p. 10

[7] Ibid., p. 15.

[8] In both of the following reasons, we follow Abramson (2014). We will explain how this manipulation impinges on the epistemic and practical authority of the gaslit agent slightly differently, emphasizing gaslighting as a kind of agent relative wrong. We will agree, however, that the wrong of gaslighting can be understood in terms of an attempt to exclude its victim from the moral community.

[9] Abramson, ibid., p. 16.

community depends in part on having an independent standing from which to issue moral complaints, and this standing depends on the presence of various capacities for responding to reasons. If the gaslighter succeeds at undermining these capacities or an agent's confidence in them, then he threatens the agent's status as a member of the moral community.

To shift points of view, suppose we began with the post-treatment effects distinctive to gaslighting. What, exactly, has happened to the practical and epistemic standing of gaslit victims? For now, let's postpone inquiry into the motives and aims of the gaslighter, and see how much progress we can make from a patient-centered perspective. The victim has been supplied evidence of their own epistemic malfunction. Their evidential relations are not what they took them to be; they have been led down the garden path to think that the wiring between their evidence and their beliefs is off.

The scene of the crime, on our view, is the victim's higher-order evidence.[10] Our *evidence about our evidence* tends to have special vulnerabilities. Gaslighting works by supplying higher-order defeaters. It tries to show that the victim's beliefs have been produced through a faulty process, inducing the victim to conclude that her attitudes were never justified in the first place.[11] The gaslit agent thereby comes to occupy an uncommonly perverse epistemic situation.

A couple of illustrative examples will help to underscore this point. First, imagine a case in which an agent becomes convinced that her beliefs in some domain are systematically incorrect, albeit due to some clearly identifiable error. Perhaps she has many false beliefs about the city she's in because she is reading the map upside down, or she might have many false moral beliefs because she has been persuaded of an extensionally incorrect theory. It's plausible to think that in cases like these, discovering that one's beliefs are systematically mistaken--or even utterly without grounds--would not induce any skepticism about one's authoritative standing within a community of moral agents. The agent might become convinced that some subset of her beliefs are formed incorrectly without experiencing the further thought that she does not have or cannot exercise the capacity to form correct beliefs in response to the available evidence. Her false beliefs might only say something about the tools she was using or the evidence she had, but not say anything about her.

On the other hand, imagine an agent who comes to learn that although some set of her beliefs are both true and justified by suitable grounds, she very easily could have had false beliefs. Standard Gettier cases furnish relevant examples. She believes the correct time by observing the hands of a stopped clock. Or she correctly judges that there is a sheep in the field, notwithstanding that what she observes is a dog wearing a sheep costume. In the epistemologist's terms, her beliefs are true

[10] Thomas Kelly, "Peer-Disagreement and Higher-Order Evidence " in Goldman and Whitcomb (eds.) *Social Epistemology: Essential Readings* (Oxford: Oxford University Press, 2010); Richard Feldman, "Respecting the Evidence" *Philosophical Perspectives* (2005) uses "second-order evidence," p.: 100; David Christensen, "Higher-Order Evidence," *Philosophy and Phenomenological Research* 18:1 (2010): 185-215.

[11] David Christensen underscores this "retrospective" aspect, although not with respect to gaslighting in particular. In the epistemologists' cases, agents acquire higher order evidence about how their particular justifications have a defective history, but not now their capacities themselves are faulty.

and justified, but lack safety.[12] In cases like these, it seems plausible that an agent might sometimes be led to doubt whether her epistemic capacities could reliably and correctly respond to the evidence. For example, if I became convinced that I had been duped into holding my current political views by unscrupulous ideologues, I might worry about my ability to independently assess the evidence in forming political beliefs--even if it so happened that my ideological brainwashers happened to have true beliefs. In short, it is possible that agents might suffer the bads of gaslighting even without thinking that their beliefs were either false or without grounds.

Although understanding gaslighting in terms of causing another to think that their attitudes are "not just mistaken, but utterly without grounds," offers a helpful insight, cases like these recommend an additional clarification of this standard. It is possible, that is, to be a victim of gaslighting without being led to believe that one's beliefs are either mistaken or without grounds. At the same time, one can be caused to think that one's beliefs are both mistaken and utterly without grounds, without thereby becoming a victim of gaslighting. Abramson's (2014) account helps explain the disconnect. Gaslighting is a moralized concept, and the reasons against gaslighting have to do with its connections to manipulation and contempt suffered by those agents who are denied a certain kind of normatively authoritative standpoint. But as the above cases indicate, whether an individual's standpoint is compromised in a manipulative or contemptuous way is a separate matter from how they assess their beliefs.

We are now in a position to say something about why this is so. Whether I enjoy a standpoint of normative authority (a notion we will say something more about, below) within some community of agents, or even whether I perceive myself as enjoying such a status, depends on my judgements about my own capacities. Although I might doubt--to varying degrees--my beliefs, I will not be driven to question my authoritative standpoint as long as my confidence in my belief forming capacities remains intact. If the wrongness of gaslighting is grounded in compromising this standing, then it makes sense--in the first instance--to understand gaslighting as undermining an agent's confidence in these capacities. It's for this reason that our proposal focuses on an agent's higher order evidence, which includes the evidence that bears on agents' abilities to evaluate their first-order evidence (Cf. Sliwa and Horowitz 2015, p. 12). We might then revise the account as follows:

> Gaslighting occurs when one agent [A] induces another agent [B] to doubt B's ability to respond rationally to evidence, characteristically by providing higher-order defeating considerations about B's reasoning capacities.

This is still a rough idea, but we suggest that it can help illuminate the why gaslighting is typically morally wrong. It's commonplace to think wrongdoing is conceptually connected to a bipolar normative structure (Dawall 2006). When I wrong someone, I don't merely bring about a worse state of affairs, but I act toward another agent so as to warrant the recipient's agent-relative complaint. Understanding gaslighting in terms of an agent's higher order evidence can reveal how

[12] Duncan Pritchard, "Epistemic Luck, Safety, and Assertion," in Littlejohn and Turri (eds.) *Epistemic Norms: New Essays on Action, Belief, and Assertion* (Oxord: Oxford University Press, 2014)

gaslighting is an agent-relative phenomenon, and so also a candidate for wrongdoing within a biopolar normative relation. To see why, it will help to contrast how evidence works in a standard case with a case of higher order evidence. 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."[13] However, if an agent learns that her particular belief-forming process was defective, then she has an agent-relative reason to treat the evidence differently than she otherwise would. Gaslighting not only third-personally affects the evidence available to agents, it also second-personally affects how agents think about themselves. In this way, it is a type of action that connects two agents in a morally serious way. Gaslighting, then, is an agent-relative wrong.

Compare two ways of spreading misinformation. In one case, imagine a speaker who incorrectly asserts that some body of evidence is not credible. For example, a politician might assert that the body of climate science is not good evidence about how the Earth's temperature will change. Second, imagine a speaker who asserts that the testimony of members of a certain group (say, migrants, or the news media) cannot be trusted as good evidence. Both kinds of assertion are false, and certainly objectionable in various ways. However, the second more clearly raises the specter of gaslighting. It does so, we propose, because its uptake would challenge the reasoning capacities of members of specific groups. Put differently, whether the speech act would challenge my epistemic agency would depend 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 perplexed about how anything seriously wrong could have been perpetrated at all. This combination of perlocutionary consequences is puzzling only until we realize how the agent-relativity of higher-order evidence can account for the same utterance's differential effects on different audience members.

The final idea Abramson's account emphasizes is that of authority or standing. The gaslighter tries to undermine another's standing or authority to issue challenges from a separate point of view. In gaslighting, the abuse of authority is janus-faced. Gaslighters inflate their own normative status while undermining the epistemic status of another agent. One might object that surely these strong claims must be in some way metaphorical. Certainly as matter of fact, the gaslit agent still does occupy a perspective from which moral reasons might be issued, and still is a member of the moral community. What's wrong with gaslighting, after all, is precisely that it fails to treat another as having the status that she actually does have. In what sense, one might wonder, does gaslighting make it the case that a person lack moral standing, or is outside of the moral community?

Here it is important to recall how the gaslighter reacts to a speaker's assertions. The gaslighter doesn't just disagree with his interlocutor's assertions. He acts as though she could not have any reasons for what she asserted, and even tries to get her to believe that she didn't mean to assert what she did. The gaslighter deliberately refuses to treat the other agent's statements as inputs into a common deliberative process, instead rejecting them out of hand. In so doing, the gaslighter

[13] Christensen 2010, p. 190.

violates what epistemologists describe as the "ethics of assertion."[14] Because assertion is norm governed, making an assertion involves an implicit claim to being in a position to know that the assertion is true (or at least to reasonably believe that one knows). As Goldberg writes, "to ignore or dismiss an assertion on insufficient grounds, then, is to disrespect the speaker as an epistemic subject by inappropriately repudiating her claim to relevant authoritativeness."[15] The gaslighter goes further, in trying to get the speaker to believe that she is not in a position to know what she asserts, to offer reasons in its defense, or to be held accountable if others act on her assertion.[16] If the gaslighting works, the gaslit agent abandons her conception of herself as a participant in the practice of assertion, and instead adopts the gaslighter's disrespectful attitude toward her own agency. In this sense, the gaslit agent "has been turned against herself."[17] So while it is true that such agents still have moral standing, they have come to inhabit a role of objects rather than subjects in the practice of sharing knowledge.[18] Gaslit agents becomes a purely passive recipient of assertions, and so third parties will feel pressure to stop treating them as subjects or participants. The gaslighter thereby conscripts other agents in eroding her standing within a deliberative community.

We have tried in this section to *narrow* the concept of gaslighting from a kind of epistemic manipulation in general to the case in which one agent demands in particular that another accept higher-order evidence undermining the justificatory force of the latter's own reasoning. This species of agent-relative wronging can shed some additional light on gaslighting. We will also suggest that it can help to *extend* the concept's application to a wider array of cases, including its recent invocations in American politics.

2. Political Gaslighting

Politicians and their allies in the media routinely charge their political adversaries with gaslighting. But gaslighting, the last section argued, is an especially serious way of undermining another person's membership in the moral community. It often involves repeated attempts at disrupting a person's search for evidence, and further efforts to distort the ways in which they make assertions based on the evidence they've collected. Paradigm cases often involve close personal relationships, or at least perverse facsimiles of such relationships. One might reasonably wonder if all of these ingredients could be in place in politics.

Suppose we begin with the original account of gaslighting, according to which agents are made to feel that their attitudes are not just false, but "utterly without grounds." With this conception in mind, there is reason to doubt that there is any phenomenon like 'political gaslighting.' If political

---

[14] Stanford Goldberg, *Assertion* (Oxford: Oxford University Press, 2015); Wanderer, J. "Addressing Testimonial Injustice: Being Ignored and Being Rejected," *The Philosophical Quarterly* 62:246 (2012): 148-169.

[15] Gorldberg, ibid., pp. 187-188.

[16] Tanesini, A, "'Calm Down, Dear': Intellectual Arrogance, Silencing and Ignorance," *Aristotelian Society Supplementary Volume* XC (2016): 71-92.

[17] Abramson, ibid., p. 16.

[18] Fricker, *Epistemic Injustice* (Oxford: Oxford University Press, 2007), pp. 132-133.

figures are engaged in gaslighting, there must be some agents who are being gaslit. The most likely candidates for these agents are either the political figure's supporters, or else the figure's opponents. As we'll underscore below, it makes little sense to say that political leaders are gaslighting their opponents, since those who oppose them do not accept their assertions, and so do not appear epistemically affected. It's more plausible to say that leaders are gaslighting their followers, whose beliefs often are influenced by elite signals.

However, one important aspect of gaslighting is importantly absent from these political cases. In the paradigm case of interpersonal gaslighting, the gaslit agent 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.[19] Freed from attention to the complexities of evidence, partisans have little occasion to worry that they might be wrong. Rather than coming to doubt themselves, those partisans who are most susceptible to manipulation also display the most confidence in their political beliefs.[20]

If political gaslighting must involve reducing people's confidence in their beliefs, then perhaps there isn't any such thing. We might conclude that the recent popularity of accusations of gaslighting in the political sphere belies a kind of systematic conceptual error. Before making that inference, however, we will consider whether we can make sense of the phenomenon on our revised conception--emphasizing agent's higher order evidence about their epistemic capacities. There is evidence, for example, that partisans frequently follow their favored political party into espousing clearly mistaken attitudes. One recent study presented a random group of citizens with pictures from the inaugurations of Barack Obama in 2009 and Donald Trump in 2017.[21] The contrast between the pictures obviously show that the 2009 event had a much larger crowd. Playing on the well-known controversy surrounding the Trump administration's insistent denial of this evidence, the researchers asked respondents to identify which of the pictured crowds was larger. Respondents were not given any information about the pictured gatherings (i.e., they were not told which inauguration was which). They were simply shown two pictures and asked which crowd was larger. While negligible numbers of non-voters and Clinton voters reported that the Trump inaugural picture contained the larger crowd, 15 percent of Trump voters denied the apparently obvious evidence and incorrectly identified the smaller crowd as larger.[22] Politically engaged Trump supporters were more than twice as likely to choose incorrectly as unengaged Trump voters.

Does this mean that Trump was, in fact, gaslighting his supporters? There's a case to be made, at least. Just as Paula is led by Gregory's manipulation to disbelieve her own obvious perceptions of the gaslights, it seems that partisans can be led to misreport their perceptions of similarly obvious sensory data. Social scientists have known for some time that partisan identity strongly influences

---

[19] Cf. Ortoleva and Snowberg 2015.

[20] Druckman 2004; Anson, I.G., "Partisanship, Political Knowledge, and the Dunning-Kruger Effect," *Political Psychology*, forthcoming.

[21] Shaffner and Luks, "Misinformation or Expressive Responding?" *Public Opinion Quarterly* 81:1 (2018).

[22] Shaffner and Luks, p. 140

political attitudes.[23] But it's not just partisanship. In the case above, Trump's insistence against the evidence appears to influence his strongest supporters to adopt his own representation of reality, or anyway to express that as if it were their perception. This is not an isolated case. It's a long-standing observation that partisans will disagree about how the economy is doing based on which party is in power. Recent work has shown that the reason for this is that partisan citizens are simply taking on the opinions of elites. When leaders in power acknowledge that the economy is not doing well, their followers accept this view as well. When leaders deny it, their partisans follow the denial.[24]

Does this amount to gaslighting? The difficulty is that we don't know exactly what is happening in the minds of citizens described by studies like these. For clarity, it will help to distinguish between three general possibilities:

(1) Citizens-as-Victims. Perhaps partisan citizens update their beliefs according to signals sent by elected leaders and other party elites. On this account, citizens unknowingly adopt the false beliefs sold to them by their parties. They mistakenly think their beliefs are supported by the evidence.
(2) Citizens-as-Enablers. Citizens are aware that their partisan assertions are not accurate, but they take cues from political elites as a way of expressing support of their own political tribe. They are aware that their statements are false, but they don't regard political statements as expressing beliefs.
(3) Citizens-as-Self-Manipulators. Citizens are neither unwitting victims of deception by elites, nor are they knowing accomplices in spreading misinformation. Instead, they occupy a kind of middle ground, in which they are complicit in their own epistemic manipulation. Citizens' partisan assertions may express genuine beliefs, albeit ones lacking the evidential support citizens would require in other domains.

The case of Citizens-as-Victims offers the clearest example of gaslighting. While they may not lose confidence in their beliefs, they are being led into adopting systematically false attitudes, and this seems to be at least a part of the concern for those who invoke worries of 'gaslighting' within the democratic context. The most difficult case, on the other hand, is posed by the second category: Citizens-as-Enablers. If citizens are not being deceived, and remain confident about their beliefs, is there any meaningful sense in which we can say they are being gaslit?

Certainly there is a sense in which this scenario carries a democratic cost. If citizens parrot the attitudes of political elites--however knowingly--then it becomes unclear how there is an independent set of attitudes which those elites could represent more or less well. When elite cues include instructions to discount other sources of information (the "fake news"), partisans may become more reliant on others for their views than on the deliverances of their own capacities.

---

[23] Bartels 2002; Flynn, D.J., Nyhan, B. and Reifler, J. "The Nature and Origins of Misperceptions: Understanding Flase and Unsupported Beliefs about Politics," *Advances in Political Psychology* 38:1 (2017): 127-150.

[24] Bisgaard, M., and Slothuus, R., "Partisan Elites as Culprits? How Party Cues Shape Partisan Perceptual Gaps," *American Journal of Political Science*, forthcoming.

Even if they do not consciously judge that they should not rely on their own capacities, enablers are acting as if that were the case. And this, itself, shares an important family resemblance with gaslighting. Elites may not be turning partisans 'against themselves,' but they are contributing to the erasure of some citizens' perspectives from the democratic landscape. If—as third parties—all we need to know is what elites are saying to successfully predict what ordinary political participants will say, then it seems that those politically active citizens have at least compromised their independent point of view. They cannot fully participate in holding others responsible for what they assert about political realities. As Bisgaard and Slothuus observe, results like these paint "a grim outlook for the idea of accountability."[25]

With this in mind, we suggest that we can extend our concept of gaslighting to capture the concern about gaslighting in democratic politics. Consider:

> Political Gaslighting occurs when one political actor induces some citizen or group of citizens to doubt, or to act as if they doubt, their ability respond rationally to evidence independently of the political actor's guidance. The political actor may characteristically provide higher-order defeating considerations about some citizens' reasoning capacities.

This modification acknowledges that gaslit citizens may not exhibit any actual doubts about either their beliefs, or their belief forming processes. However, citizens may still display a functionally similar epistemic profile. That is, they may engage in the kind of unreflective modification of beliefs and epistemic deference to a figure who has acted to undermine their perspectives as an independent location of deliberative engagement. As in Abramson's original account, this compromises the perspective of partisan citizens as a place from which challenges could be issued. It thereby raises the familiar worries of manipulation and contempt.

Of course, as the third category of citizens above suggests, citizens may be actively complicit in their own manipulation. It's philosophically familiar to think that individuals can engage in self-deception, either by diachronic compartmentaling of their beliefs, or by psychologically cordoning off some beliefs from others. Self-manipulation might operate in a similar way. For example, partisans may be led by elites to affirm beliefs they had previously rejected. Likewise, activating partisan identity might cause citizens to accept patterns of reasoning they would eschew when acting under some other identity. Certainly partisanship seems to correspond to directional motivation, rather than accuracy motivation.[26] Citizens may not be deceived, but their reasoning may be resistant to the evidence. Even in cases when opposing partisans do come to agree on facts about the world (say, regarding economic conditions), they maintain they will alter their beliefs about who is responsible for those conditions so as to sustain their partisan allegiance.[27]

---

[25] Bisgaard and Slothuus, p. 12.

[26] Flynn, Nyhan, and Reifler 2017; Druckman 2004; Arceneaux, K., and Vander Wielen, R.J., *Taming Intuition: How Reflection Minimizes Partisan Reasoning and Promotes Democratic Accountability*. Cambridge: Cambridge University Press, 2017).

[27] Bisgaard, "Bias Will Find a Way," *Journal of Politics* (2015).

Our suggestion, then, is that we can at least make sense of the popular invocation of "gaslighting" by identifying those cases in which citizens form or express beliefs in a way that defers to the unreliable guidance of political elites. Again, the idea of higher order evidence has played an important role, in two respects. First, citizens in the grip of a partisan identity may functionally discount deliverances of their own epistemic capacities. Second, elites may amplify the effects of partisan identity by trying to undermine some sources of evidence--particularly so as to insulate partisan directional reasoning against the prospect of challenge.

3. Problems with Political Gaslighting

Gaslighting aims to instill relationships of troubling epistemic dependence. The gaslighter often attempts to make their own testimony a kind of epistemic monopoly by ruling out all other channels of evidence gathering. In the epigraph to this paper, Kripke considers an individual who attempts to protect a pre-existing belief by avoiding evidence that runs counter to it. This could entail "reading the wrong books (for they contain nothing but sophistry and illusion), associating with the wrong people, and so on."[28] Displacing other sources of evidence can be a start to gaslighting — at least for those followers who already treat the gaslighter as an epistemic authority. While not a success condition, it's now clearer how gaslighters convince an audience to treat their testimony as the only reliable source to ground first-order beliefs in town. Here we think that our account's sensitivity to how defeaters operate at the two-level structure has significant explanatory force.

The agent-relativity of gaslighting with higher-order defeaters also usefully illustrates how political gaslighters can have different effects on members of the same audience. Imagine, for example, a politician who insinuates that members of rural communities systematically lacked the level of educational attainment necessary for civic participation.[29] If this statement were issued to a mixed group of citizens, rural listeners might uniquely experience higher-order defeaters to their political inferences. We can see how the same language that mobilizes some listeners could gaslight others. Consider, for example, the following exchange:

> Lucy: Do you see this tree? it is a fir tree. it's called a fir tree because it gives us fur for coats. it also gives us wool in the winter time.
>
> Linus: I never knew that before, Lucy. That's very interesting.
>
> Charlie Brown: Now wait a minute Lucy! I don't mean to interfere, but...
>
> Lucy: And way up there, the little stars and planets make the rain that often showers and when it's cold and winter is upon us the snow comes up! just like the flowers.
>
> Charlie Brown: Now Lucy, I know that's wrong! Snow doesn't come up, it comes down!!
>
> Lucy: After it comes up, the wind blows it around so it looks like it's coming down, but actually it comes up out of the ground, just like grass. It comes up, Charlie Brown, snow comes up!
>
> Charlie Brown: Oh, good grief! (he exits, from off stage there is a hollow thumping sound.)

---

[28] Kripke 2011, p. 49.

[29] Kathy Cramer observes the reaction of rural citizens to this perceived form of disparagement. See Cramer, *The Politics of Resentment* (Chicago: University of Chicago Press, 2016).

> Linus: Why is Charlie Brown banging his head against that tree?
>
> Lucy: To loosen the bark to make the tree grow faster! Clouds can make the wind blow, Bugs can make the grass grow so, there you go these are little known facts that now you know![30]

Lucy makes a series of false assertions to Linus. Like a low-information partisan with nothing else to go on, Linus is happy to update his priors accordingly. He is subject to a systematic deception that is veering towards brainwashing. Charlie Brown, on the other hand, is not taken in. He is not misled in his beliefs, although his capacities are strained by her confident imperviousness to the evidence. Brown's sense of helplessness is palpable when he objects — "I *know* that's wrong." Lucy offers a rejoinder that accepts his proffer of evidence. Notice that she doesn't deny his first-order sensory perceptions, but she suggests that Brown's perceptions are misleading him. He shouldn't trust his direct sensory inputs, for they have generated a false belief.

Even citizens who don't experience gaslighting have a basis for objecting to it. For this way of manipulating individuals isn't a pairwise activity. It wrongs third-parties in characteristic ways. Our account can explain the distinctive moral complaints of individuals, who are themselves not gaslit, but who experience distinctive bystander wrongs. Perhaps it's tempting to think that overhearing one party's attempt to gaslight another may be disturbing, but only in the general sense in which witnessing serious injustice of any kind — epistemic or institutional — gives us a complaint. On other occasions, third-parties may allege gaslighting even when they don't experience the phenomenon we've described. When third-parties claim to be gaslit we can explain what they mean: Gaslighting imposes collateral damage to third parties that is distinct from the serial deceiver, who works by challenging the first-order beliefs of an audience. Gaslighting can gradually induce a kind of epistemic fatigue that comes from challenging falsehoods and experiencing its futility. Charlie Brown style, they feel powerless because they assumed that the evidentially-backed charge of falsehood would be "terminal." As responsible practitioners of assertion, they had assumed that falsehood would be a "fatal objection to assertions."[31]

Even if they are not misled, one might think that they could object to such obvious disregard for the truth. Harry Frankfurt concluded "On Bullshit" with a claim that seemed to many to overreach. The liar, he insisted, is morally better than the bullshitter, because at least the liar cares enough about the truth to understand that he asserts its contrary. We don't think that this perverse way of caring about the truth should make us feel any better about the liar, even when compared to the bullshitter. Political gaslighting can vindicate a more modest version of Frankfurt's surprising conclusion. Lying, one might think, is susceptible to exposure, and so ultimately to the prospect of correction. But politicians who jettison all responsiveness to evidence, and target the ways that citizens assess their body of evidence, can look impervious to efforts at exposing abuse of the truth. That, itself, is a cause for frustration. Such speech whittles away any common framework through which citizens can gain a shared outlook on their common political community. Gaslighting targets the higher-order evidence of citizens, but it may simultaneously wear down a common point of

---

[30] Clark Gesner, "Little Known Facts," *You're A Good Man, Charlie Brown* (Play, 1967).
[31] Williams, *Truth and Truthfulness*, 68.

view and make the idea of the "common good" a kind of category error. We will return to this possibility in the final section, once we've gotten a better handle on how gaslighting works.

4. Audacious Believing

So far, we've described how political gaslighting might create relations of epistemic dependence, and offered some reasons to worry about this phenomenon. Social scientists understand these perils well. In one informative study on political communication, Druckman, Fein, and Leeper allow study participants to choose what kinds of information they are exposed to. Strikingly, rather than seeing initial opinions questioned over time as participants acquire conflicting information, the authors find two phenomena. Either subjects flipped opinions based on whatever they've been exposed to most recently, or else insisted on not updating at all. Of the latter group, the authors write:

> [O]nce they had formed initial opinions, [they] clung to those opinions, became certain of them, sought out consistent information, and rejected what otherwise looked like a reasonable counter-argument. Although such stable opinion facilitates responsiveness, it also raises the specter of an inflexible dogmatism, stemming from biased information search, which could be problematic for many conceptions of good citizenship.[32]

Some citizens will easily switch between opinions without much reason, while others will dogmatically insist on what they have accepted. As the authors observe, neither condition portends successful democratic representation. Insofar as their beliefs are being guided by other political agents, citizens in either group are candidate victims of political gaslighting. Although their beliefs might be rational at any given moment with respect to the body of evidence they were considering, their choices over time about which evidence to consider and ignore suggests a dangerous susceptibility to epistemically irrelevant influences. They lack what Sarah K. Paul describes as autonomy in belief.[33] Considered as diachronic agents, they may be gathering and weighing evidence rationally at a given moment, but over time they are either being pushed to revise too easily, or else refusing to revise even when the body of evidence changes. Either way, they depend too much on things outside of the exercise of their own epistemic capacities.

Individual Audacity

A common view among democratically concerned political scientists favors adopting what we can call a kind of evidentialism about belief: the political beliefs of citizens ought to be apportioned only according to epistemically relevant considerations.[34] What complicates this picture is that there are plenty of practical cases in which a little inflexibility--even a little dogmatism--can look like a part of rational agency. Suppose you commit to climb Mount Everest, or to run a marathon. Each day of

---

[32] James N. Druckman, Jordan Fein, and Thomas J. Leeper, "A Source of Bias in Public Opinion Stability," *American Political Science Review* 106:2 (2012): 430-454, p. 447.

[33] Sarah K. Paul, "Doxastic Self-Control," *American Philosophical Quarterly* 52:2 (2015): 145-158.

[34] For example, Hochschild, J.L., and Einstein, K.L. "Do Facts Matter? Information and Misinformation in Politics," *Political Science Quarterly* 130:4 (2015): 585-624, p. 587.

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.[35] We will describe rationally permissible beliefs that place greater credence in the prospects for one's success than the evidence requires as audacious beliefs.

Audacious beliefs represent a departure from the evidentialist standard. In cases like these, a belief that you'll succeed can be partly self-realizing. Accordingly, it can sometimes be rational for agents to resist apportioning their beliefs exactly according to the evidence--again a familiar point from action theory.[36] The point has appeal well beyond the philosopher's armchair. 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."[37]

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 up a Strawsonian participant stance (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-3P0's probabilistic report. If he came to believe that the odds were vanishingly small, he couldn't sincerely retain his commitment to the mission.

If this posture is right, several beliefs may be permissible to hold, and we can rely on non-evidential considerations to choose among the beliefs that bear on our confidence in succeeding in our project. This slack — between the evidence an agent takes herself to have and what it's reasonable to belief, given the evidence — may be considerable, and it makes possible this third view. Jennifer Morton and Sarah Paul defend this epistemic policy as exhibiting a kind of "grit":

> Because this evidence is usually inconclusive, and because grit is generally an aid to achieving one's goals (in resource-moderate contexts, at least), we are rationally permitted to have the disposition to raise the evidential threshold we need in order to give up the

[35] Cf. Duckworth 2016.

[36] For two recent defenders of views in this neighborhood, see Miriam Schleifer McCormick, *Believing Against the Evidence: Agency and the Ethics of Beliefs* (Routledge, 2015) and Berislav Marušić, *Evidence and Agency: Norms of Belief for Promising and Resolving* (Oxford University Press, 2015).

[37] Leigh Brackett and Lawrence Kasdan, The Empire Strikes Back (Original Screenplay).

belief that success is reasonably likely, relative to the threshold we would use to answer the same question before adopting the goal.[38]

At some point your epistemic reasons run out, but a range of permissible beliefs about your chances typically remain. This creates space for you to protect your beliefs about success from evidence that's forward-looking or look back at your own track record. You can then turn to ethical considerations to help you decide among epistemically permissible policies.

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 a necessarily resilience. There is something audacious about this kind of policy — no one will 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.[39]

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."[40]

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

---

[38] Jennifer Morton and Sarah Paul, "Grit," forthcoming in *Ethics.*

[39] The formation of "collective self-confidence, on Goodwyns' view, "permits people to conceive of' the idea of acting in self-generated democratic ways — as distinct from passively participating in various hierarchical modes bequeathed by the received culture.. all significant mass democratic movements in human history have generated this autonomous capacity," Lawrence Goodwyn, The Populist Moment: A Short History of the Agrarian Revolt in America, Oxford University Press, p. 19

[40] Zeynep Tufekci, *Twitter and Tear Gas: The Power and Fragility of Networked Protest* (Yale University Press, p. 64) Tufekci dwells on the logistical backbone of the boycott: ""Despite enormous obstacles, the Montgomery bus boycott persevered for the year long battle—and triumphed, winning much more than its original demand for a bit more decency in the segregated bus system."

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.[41]

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 be 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)." 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 alternative to nonviolence would backfire, furthering the sense of powerfulness 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 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." 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?"[42] 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

[41] 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. *Philosophical Topics*, forthcoming.

[42] Martin Luther King, Trumpet of Conscience, cited by Brandon Terry,

serve to protect beliefs from demoralizing evidence, raising the evidential threshold for considerations that suggests that ought doesn't, in the end, imply can.

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 moment — in a pairwise way. For it seems that you need to hold commitments whose direct objects are first-personal plural.[43] 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.[44] So you will need a belief that about my confidence in *our project*, and vice versa.[45]

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 movements makes possible signals of their strength that are outsized. This is considerable instrumental value in the protester'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.

We can leave it open whether this requires an all-out shared belief, or an elaborate recursive structure in which each of us has a series of beliefs *about our project* and other members beliefs about our projects, but stops short of us holding a co-belief about our chances of success. What is clear is the value of holding beliefs or, at minimum, evidential policies that underwrite our joint commitments, and these will include beliefs about our efficacy, or mere disbelief about our collective failure.

5. Gaslighting vs. Mobilizing

The mobilizer and gaslighter share a common incision point: the higher-order evidence of those individuals they operate on. So far we've taken the point of view of the patient. Both treatments produce distance between professed beliefs and the evidence we take ourselves to have. This gap was key to explaining the success of the citizen at the front lines of a movement. When faced with

---

[43] A belief in group efficacy is one of the "core motivations" for collective action (van Zomeren 2013). Group efficacy beliefs must be shared, with the strongest being those in which "individuals are more focused on achieving group goals through the joint effort of collective action" (van Zomeren, Leach, and Spears 2010, p. 1056).

[44] 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 Thomas and Louis (2013).

[45] James Bryce, The American Commonwealth, vol. II (New York: The Commonwealth Publishing Company, [1888] 1908), 78– 79, cited by Daniel Ziblatt, *Conservative Parties and the Birth of Democracy* (Cambridge Studies in Comparative Politics) Cambridge University Press, p. 35. In The American Commonwealth, James Bryce was envious of the combination of grip and organizational capacity that he saw emerging in the American party system: "The greatest discovery ever made in the art of war was when men began to perceive that organization and discipline count for more than numbers. This gave the Spartan infantry a long career of victory in Greece. The Americans made a similar discovery in politics some fifty or sixty years ago. It was perceived that the victories of the ballot box, no less than of the sword, must be won by the cohesion and disciplined docility of the troops, and that these merits can only be secured by skillful organization and long-continued training."

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, in part, by moral considerations that sustain their ability to keep their commitments. To secure steadfastness, mobilizers attempt to ensure that evidence at this second-order level will not impact their followers' first-order beliefs about success.

It was exactly the distance between beliefs and evidence that was symptomatic 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.

This section shifts from post-treatment effects to the agents that bring them about. How, can we distinguish the democratic mobilizer from the gaslighter? We could, of course, let the substantive merits of their respective aims do all the work. But we think this would overlook the way that the mobilizer manages to respect the epistemic agency of "persuadables" — at least when all goes well in their messaging. So, too, this approach would neglect the diversity of ways gaslighters abuse their authority and the epistemic standing of their victims.

We will proceed inductively, stepping through a series of "treatments" administered on citizens. We start with persuasive techniques that seems permissible, even obligatory. Democratic leaders don't reflect their constituents' preferences, but form joint commitments with citizens and mobilize around them.[46] On this view of joint commitment formation, both parties play an active role. Leaders typically provide testimony through which citizens can gain knowledge "secondhand."[47] Exchanges in which a speaker asserts *p* to a listener are normative transactions. When a leader tells their audience that *p*, the leader assumes a kind of responsibility for the assertion that would not accrue by merely declaring that *p*.[48]

While directed assertions are morally freighted, it still seems permissible for leaders to assert some propositions for which they don't have evidence prior to the asserting. Consider a simple case: leader with a bullhorn tells us she is confident that we will succeed in all driving on the right side of the road, and her very announcement of that belief (nearly) guarantees the belief's truth. In

---

[46] Eric Beerbohm, "Is Democratic Leadership Possible?" *American Political Science Review* 109:4 (2015): 39-652.

[47] Elizabeth Fricker, "Second-Hand Knowledge," *Philosophy and Phenomenological Research,* Vol. 73, No. 3 (Nov., 2006), pp. 592-618.

[48] In *Testimony, Trust, and Authority*, Benjamin McMyler notes this distinction and the responsibility shifting it can entail: "Insofar as we think that there is an important distinction between telling a listener that p and merely declaring one's belief that p it seems like this is explained in the way that telling someone that p involves openly offering to take responsibility for the resultant belief, where merely declaring one's belief does not." McMyler, p. 68-9. See also Tanesini 2016.

this case, both her confidence in the belief, and the responsible asserting of the belief, appear justified.

We are now in the position to see how audacious beliefs, in this collective setting, are not the sorts of attitudes that we can typically come to hold on our own. And there are inherent risks in engaging in intentionally biased inquiry. Derek Parfit draws out the danger:

> [W]e might cause ourselves to have some beneficial belief by finding evidence or arguments that gave us strong enough epistemic reasons to have this belief. This method is risky, since we might find evidence or arguments that gave us strong reasons not to have this belief. But we might reduce this risk by trying to avoid becoming aware of such reasons.

We can mitigate this risk by interpersonal persuasion, where a third-party can protect us from false negatives of this kind. Acquiring beliefs through others isn't a passive activity. We can play a strong agential role in coming to believe in a resilient way, and we can take actions that sustain beliefs through policies of epistemic resilience, but typically we need others to help us maintain that overconfidence over time.

Now imagine a knife-edge case that lies at the interstices of mobilization and gaslighting. A leader tries to motivate collective confidence in some belief which she does not affirm, or which she might even presently doubt. There does seem to be something manipulative about encouraging others to have a belief one does not share, particularly if doing so is in one's interest (or, more in the interest of oneself than in other members of the group). To push them to see the evidence differently than you see it might strike us as disrespectful, in a way that is similar to the disrespect sometimes involved in taking the diagnostic stance on them: we are operating on them, even as we are persuaded by appeal to rational considerations.[49]

We can separate two concerns. On the supply-side, the speaker's insincerity may be worrisome whether it manages to alters the first- or second-order evidence of the audience. Is it hypocritical to bolster confidence that one in fact lacks? Or we may look to the demand-side. If the audience's receptivity to the speaker is dependent on them believing that the speaker's self-reporting on their own audacious beliefs, and not merely trying to induce them in other parties, we may bristle that this off-stage assumption is doing the work. If, however, listeners understand and share the manipulative aim, such assertions are also often acceptable. Paul and Morton offer an analogous case:

> 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

[49] For example, Anne Barnhill's "What Is Manipulation?" in *Manipulation: Theory and Practice*, edited by Christian Coons and Michael Weber, 2014) argues that a conceptual feature of manipulation is that it is contrary, or generally contrary, to the interests of the manipulated party.

succeed, he might reasonably expect his coach to spin the evidence positively to encourage him.[50]

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. Members of a movement may recognize that they have certain epistemic needs that can only be attended to by an external party.[51] Or they may not be aware of this, they are objects of a gentle form of epistemic paternalism. In such a case, a listener's knowing that an assertion is aiming to be confidence-bolstering could undermine its intended practical effect.

Compare this with the outsider, say a sociologist of mass movements, who arrives at a rally and grabs the closest microphone: "You really need to raise your evidential policies when you face defeaters that would undermine your belief in your cause. All successful mass movements have held efficacy beliefs that were overblown. Good luck!" To see why this is supplying the wrong kind of reason — not to mention its rhetorical limits — we can distinguish between the first- and third personal of reasons for belief. The social scientist is viewing her beliefs as just another kind of attitude that may be desirable to hold, or suspend, depending on the agent's ends. But first-personally, we don't see our beliefs as items that can be arranged in this way.

If overt awareness of your treatment would likely undermine your capacity to be mobilized, has the would-be mobilizer crossed a bright line? The general worry is that you are now being treated more like a patient than an agent. You are being led to hold attitudes, in this case epistemic policies about how you weigh and act upon evidence, behind your back. So the fact that your awareness would cancel the effects raises the concern that your agency has been bypassed.

6. Citizenship's Two Horns

Two epistemic policies central to our ideal of citizenship stand in tension. Seen together, they present us with something that looks like a dilemma. The first horn is set by the evidentialist constraint. In forming our beliefs in response exclusively to the evidence, we protect ourselves from the gaslighter. This protection closes the door to audacious beliefs, seriously constricting our ability to engage in difficult and risky projects of political reform. Without the ability to pursue long-term projects with others, we can feel hopeless — without a sense of co-agency. The other horn would have us reject evidentialism and become continually vulnerable to gaslighting. This threatens our membership in the community of self-reflective, reasoning agents. If evidentialism is a threat to shared agency, then gaslighting is a threat to one's agency altogether.[52]

Can citizens find policies to negotiate the space in between the horns of this dilemma? You could make ample space for audacious belief, granting that this can veer into wishful thinking. While

[50] Paul and Morton, "Believing in Others."

[51] Hinchman, Edward. (2003). "Trust and Diachronic Agency." Noûs 37, 25–51.

[52] The general form of this tension isn't new. William James mused that Clifford's strict ethics of belief were based on the latter's "preponderant horror of becoming a dupe." While James shared that horror, he countered that "worse things than being duped may happen to a man in this world."

perhaps not an existential threat to one's agency, the moral costs of politically wishful thinking can still be high. When you believe self-serving falsehoods about your political opponents, you risk wronging them with your beliefs.[53] Failing to see evidence on the side of other groups — even the inability to "code" their assertions as evidentially relevant — puts you in a position to inflict testimonial injustice on them.[54]

To guard against these risks, citizens can put in place further safeguards. When responding to defeaters to their higher-order evidence, they could have a policy of responding asymmetrically two different kinds of evidence. Evidence that tends to "downgrade" the status of their beliefs about their chances of success and the good will of others face the highest evidential bar. This would tend to preserve their optimism about projects of reform, while guarding their charity toward the good will of others. This class of defeaters would face the strictest level of scrutiny of all forms of evidence that they encounter. In contrast, second-order evidence that tends to "upgrade" their confidence in executing their plans, or their charity in understanding the aims of opponents, would have to clear a lower evidential bar. This policy of adjusting the bar for what passes muster based on substantive moral claims isn't obviously right. In morally high stakes circumstances, it is plausible that we can adopt policies of double-checking and extra care. But allowing moral considerations to "encroach" on our beliefs is a more radical move that stands in need of greater defense.[55]

Or you can tack closer to the evidentialist, making space for sometimes acting ahead of, if not against, the evidence. There are risks to splitting the difference on this side, as well. Given the primacy of group efficacy beliefs in motivating collective action — especially measured against the relative paucity of evidence supporting this efficacy — it would not take much impersonal attention to the evidence before you could grow politically disillusioned.

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. For example, you might hold audacious beliefs with respect to the efficacy of your own group, but without discounting the standing of other groups. Perhaps you could maintain confidence in your group by simply ignoring opposing groups. You might even say that they have just as much standing to offer testimonial evidence as members of your group — just with the expectation that both groups will ignore each other. This strategy seeks a separate peace of sorts with political opponents, granting them a reciprocal status — but not one that undermines confidence in one's own political projects.

---

[53] Mark Schroeder, "When Beliefs Wrong," *Philosophical Topics* 46:1 (2018).

[54] Fricker, *Epistemic Injustice*, ibid.

[55] For early attempts at defending "moral encroachment," see Sarah Moss, "Moral Encroachment," *Proceedings of the Aristotelian Society*, Volume 118, Issue 2, 1 July 2018; Rima Basu & Mark Schroeder, "Doxastic Wronging"

In Brian Kim & Matthew McGrath (eds.), *Pragmatic Encroachment in Epistemology*. Routledge (forthcoming). Both papers focus on cases where we recognize that acting on a moral belief (sincerely, but mistaken) would lead to outcomes that were "disastrous," to use Moss's terminology. Here we are shifting up a level, where our evidence about our evidence for these first-order moral beliefs is challenged.

Citizens might also opt for a division-of-labor strategy. Some members of a political group could serve as 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."[56] 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.

A troubling feature of the separate peace and division of labor strategies is that they both back away from an ideal of genuinely shared deliberation — although in rather different ways. In the former case, citizens don't deliberate with their opponents at all, though they do respect them (in a sense). In the latter case, citizens reason together, but even with only the barest of outlines in place for this strategy, it seems likely that they 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 allows us to revisit the worry that gaslighting posed a threat to the agent's own perspective. Despite the common term, political gaslighting appeared not to carry that distinctive risk, since gaslit parties tend toward overconfidence in their own self-reported beliefs. But this disanalogy may have been drawn prematurely. For when leaders gaslight, they undermine the status of their follower's beliefs, *or* the professed beliefs of political opponents. You come to believe that "they" are not people who can be reasoned with. There is a genuine perspective that is lost — the perspective of what is good for us as a group. Gaslighting threatens the existence of a "we" who occupy a point of view from which "deciding together" is possible.[57]

We don't need to suggest a stronger view about collective agency: that fellow citizens are engaged in something like shared action.[58] If we assume there's another agent in the room, "the people," then an argument revealing a collective victim of gaslighting will come relatively easily. But we think that there is a more ecumenical premise that will allow the argument to go through. The idea of joint deliberation is a comparatively weaker notion. To say that you and I are deliberating together, all we need is for my reasons for acting depend *in part* on circumstances constituted by

---

[56] King, "The Power of Nonviolence," http://www.thekingcenter.org/archive/document/power-nonviolence

[57] Andrea Westlund, "Deciding Together," *Philosopher's Imprint* (2009).

[58] Anna Stilz, for instance, holds that citizens engage in shared activity, organizing themselves around the project of the state. On her view, "a shared project to which they are committed... they see themselves as partners in a joint enterprise, acting together to shape the character of their political environment. Nor are they wrong to take this view," "The Value of Self-Determination," *Oxford Studies in Political Philosophy* 2 (2016), 2, 98-127.

our foreseeable plans for action. Without an archimedean point like joint deliberation, the very concept of a "common good" looks like a category mistake.[59]

If this premise about joint deliberation is plausible, then political gaslighting does, after all, threaten a morally valuable point of view. And like the original view of gaslighting, losing this perspective travels in tandem with holding systematically false beliefs. When people reason only with those who share a common partisan outlook, they are more likely to show directional rather than accuracy motivation.[60] Beyond this, rejecting one's political opponents also exposes a kind of limitation in one's own individual point of view. People who ascribe negative traits to opposing partisans are relatively likely to be less reflective—less willing to take their own political beliefs as the objects of their own attention. Citizens who are more reflective are more able to see the virtues of those they disagree with.[61] Gaslighting might be one obstacle to a reflective electorate. "In a polarized polity…political debate among elites has the potential to engender animosity among partisan groups."[62] 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 gaslighting incendiary views toward opponents might compromise individuals' own agency, after all.

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. The reality may be that there is no clear way to navigate between the evidentialist and activist horns. 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.

To put the point in a more empirical way, no amount of reflection is a panacea. To offer one more case: Democratic voters high on racial resentment and reflectiveness moved away from Clinton to Trump. They were able to resolve a tension in their beliefs and identity. Unreflective Democratic voters high on racial resentment stayed with Clinton.[63] In a sense they were less coherent, but coherent reasoning in a way that consolidates racial animosity does not sound like a story about *resisting* gaslighting.

---

[59] Eric Beerbohm and Ryan Davis, "The Common Good: A Buck-Passing Account," *Journal of Political Philosophy* 25:4 (2017): e60-79.

[60] Samara Klar, "Partisanship in a Social Setting," *American Journal of Political Science* 58:3 (2014): 687-704.

[61] Arceneaux and Vander Wielen 2017, ch. 6.

[62] Ibid., p. 151.

[63] Arceneaux and Vander Wielen, ibid.

Conclusion: Citizenship's Duck-Rabbit Illusion

No friend of democracy, James Fitzjames Stephen held that in democratic politics, "the ruling men will be the wire-pullers and their friends." [64] It's tempting for citizens to see themselves as either victims of manipulation *or* one of its agents, among the fellow wire-pullers. In this paper we've switched back and forth, like the rabbit-duck illusion, from these two points of view. Each supplies us with a way of "seeing as" a citizen: victim or agent.[65] Gaslighting and mobilizing both target the lower-level code of citizens: to change how we handle evidence, whether disquieting or hopeful in character. Gaslit citizens have been manipulated in a morally distinctive way, one that targets a particular epistemic mechanism, our evidence about our evidence. Such manipulation poses a real threat to our agency — perversely making us complicit in this process.

Yet insisting on exercising our own epistemic capacities, and fortifying against gaslighting, threatens 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.[66] 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 evidence. They get them to hold beliefs, or evidential policies, that protect them from a sense of futility and hopelessness. The tension motivating this essay lies between these two expressions of agency. Our susceptibility to defeat by higher-order evidence is essential to the co-agency of citizenship, and among the greatest menaces to it.

---

[64] James Stephen, *Liberty, Equality, and Fraternity* (London: Smith, Elder, and Co., 1874), 239.

[65] 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 our reasons and the evidence behind them. See *Philosophical Investigations* (New York: Blackwell, 1958), pp. 194-99.

[66] This need for a decision-procedure that allows us to reasonably see our agency in our political institutions is explored by Jeremy Waldron in *Law and Disagreement* (Oxford University Press, 1999), especially pp. 101– 18. It forms an ongoing thread in Richard Tuck's *The Sleeping Sovereign: The Invention of Modern Democracy* (The Seeley Lectures) Cambridge University Press.