Polarization is Epistemically Innocuous

Mason Westfall*

Johns Hopkins University

Abstract

People are manifestly polarized. On many topics, extreme perspectives are much easier to find than 'reasonable', 'moderate' perspectives. A natural reaction to this situation is that something epistemically irrational is afoot. Here, I question this natural reaction. I argue that often polarization is epistemically innocuous. In particular, I argue that certain mechanisms that underlie polarization are rational, and polarized beliefs are often fully justified. Additionally, even reflective subjects, who recognize themselves as in a polarized or polarizing situation shouldn't necessarily reduce confidence in the relevant beliefs. Finally, I draw attention to some often overlooked epistemic *benefits* associated with polarization. A fuller understanding of the epistemology of polarization requires incorporating both the potential costs and the potential benefits, and being more precise about exactly what is—and is not—epistemically objectionable in these situations.

Word Count: 9252

This is a pre-print of a paper to be published in *Synthese*. Please cite the published version when possible.

People are concerned about polarization. Some have argued that we're more polarized than ever. Many mechanisms that drive polarization have been discussed—virtue signaling, the group polarization effect, social media, epistemic bubbles, echo chambers, social sorting. . .—the list goes on. Often in discussions of polarization, it is taken for granted that polarization is bad. I'd like to question this assumption. In my view, anxiety about

polarization is misplaced. More specifically, I'll suggest that polarization as such is epistemically innocuous. Polarization may be practically or politically unfortunate, but polarizing is often the epistemically rational response to situations in which subjects find themselves, and polarization holds potential epistemic benefits that have not been sufficiently appreciated.

1 Preliminaries

Let's begin with a paradigm of polarization. A group of people starts out with some distribution of views on a political question—say, what regime of gun control laws is preferable. Over time, these people are exposed to more evidence, e.g. about the efficacy of stricter gun control laws, and, as a result, their beliefs change. We can imagine that these changes exhibit an interesting pattern: people who started out modestly supporting stricter gun control laws tend to wind up more strongly in favor of such laws and the those who started out modestly opposed wind up more strongly opposed. Here are two ways in which people can be understood to be 'more strongly in favor(/opposed)'. First, they might become more confident in the belief they started out with—initially somewhat confident that we should have stricter laws, eventually quite confident. Second, they might adopt qualitatively more extreme beliefs—initially supporting a handgun ban, eventually supporting a total ban. So understood, polarization is a group-level process. We can derivatively understand a state of polarization as the output of this process: a group is in a state of polarization when it has undergone the process of polarization. To a first approximation, this paper is about the epistemology of situations that are relevantly similar to this paradigm. I expect some indeterminacy and uncertainty about the boundaries of the category, but I trust I have pointed to a recognizable phenomenon worth theorizing about.

My characterization of polarization is agnostic about mechanisms. Presumably, many mechanisms subserve the group pattern of belief change characteristic of polarization. Which mechanisms explain particular instances of polarization, or a more general tendency towards polarization is an interesting empirical question I won't say much about. That's not to say I am uninterested in mechanisms. Quite to the contrary, my argument will focus on two mechanisms that can be expected to contribute to polarization.[1] In particular, I'll argue that *these* polarizing mechanisms are epistemically innocuous. How relevant my argument is to polarization in the real world depends on the empirical questions I am setting aside. If the mechanisms I discuss don't explain much real world polarization, then the fact that these mechanisms are epistemically innocuous isn't particularly important to the epistemic evaluation of real world polarization. If, instead, these mechanisms explain a lot of real world polarization, then perhaps quite a bit more real world polarization is epistemically innocuous than many suppose. Even if the mechanisms I

---

[1]The mechanisms are Biased Assimilation and Selective Exposure, to be described below.

discuss explain relatively little real world polarization, my argument is still of interest to epistemologists. It reveals interesting things about how agents like us *should* behave in circumstances like those we are in, even if we don't often behave that way *in fact*.[2]

My aims can be put into sharper relief in reference to the adjacent and vigorous empirical controversy over the psychology of reasoning and belief fixation. Here is a (broadly) empirical question: how much of human belief fixation is driven by (approximately) rational processes, and how much is driven by arational (or irrational) processes? To hear some people tell it, the answer is that belief fixation is largely the result of processes that have little to do with rationality.[3] If so, then it may seem that the scope of my argument is harshly delimited. Sure, the thought goes, in some abstract way there might be rational routes to polarized belief, but we have *learned from psychology* that those are not operative in practice, so who cares about the abstract rational routes?

Now, the rationality of belief fixation is hotly debated and I myself suspect (though will not argue) that it involves a motley of both rational and arational (or irrational) processes. However the empirical controversy is resolved though, it bears only an indirect relation to the questions of interest here. I am interested in how agents *should* behave, under realistic conditions. That question retains its interest even if it is demonstrated that, as an empirical matter, humans rarely (or even never) in fact behave in that way. What would follow, if that is how the empirical controversy is resolved, is that actual polarized subjects by and large (or even universally) did not form their beliefs rationally. Analogously, applied ethicists query how people should behave in various realistic situations, and should not (as theorists) be dispirited if it turns out empirically that people rarely (or never) behave in the ways they should.[4] So the normative assessments that form the core of this paper are independent of the empirical questions about belief fixation, though whether actual polarized subjects are rational *does* turn on empirical questions about belief fixation.[5]

Polarization can be considered at both the individual and the group level. We can epistemically evaluate individuals who have been polarized, and we can evaluate the group as a whole. There's reason to doubt that an epistemic evaluation of the group can

---

[2]Thanks especially to Jake Quilty-Dunn, Allan Hazlett, and an anonymous reviewer for encouraging me to clarify my thinking on this point.

[3]For variously pessimistic perspectives, see (Kundra and Sinclair, 1999; Haidt, 2001; Mandelbaum and Quilty-Dunn, 2015; Mandelbaum, 2019).

[4]One might worry that a maximalist arationalist conclusion would threaten the distinction drawn in the text. If it were shown empirically that, as a psychological matter, humans *could not* form beliefs rationally, then, assuming ought implies can, it is also false that we ought to do so, given our situation. It would be interesting to consider the merits of this argument, but for present purposes, it is sufficient to note that the argument actually supports global epistemic nihilism. *Qua* work of applied epistemology, I make the controversial but widely endorsed assumption that epistemic nihilism is false. So either the maximalist arationalist view of the psychology of belief fixation is false, or the argument presented in this footnote is unsound.

[5]Though, if belief fixation in general is a largely arational affair, then non-polarized subjects are plausibly also not rational in their beliefs.

be read off of our evaluation of individuals (Mayo-Wilson, K. J. Zollman, and Danks, 2011). Perhaps the individuals are doing the best they can in their circumstances, but this has bad epistemic consequences for the group (K. J. Zollman, 2007). It may also be that, under certain conditions, a group composed of multiple polarized groups enjoys some epistemic benefits over an unpolarized group (K. J. S. Zollman, 2010). I'll set aside these interesting questions. Instead, my focus will be on the individuals: how should we evaluate individuals who have been polarized by specific mechanisms?

I need to say what I mean by 'epistemically innocuous'. Two sorts of evaluations will concern us. First, we can evaluate a suite of *epistemic actions*, second, we can evaluate *beliefs*. With respect to epistemic actions, I'll focus on gathering and scrutinizing evidence. When subjects gather and scrutinize evidence well, I'll say that they are *reasonable*. I'll argue that some ways of gathering and scrutinizing evidence are both reasonable and apt to lead to polarization. With respect to beliefs, I'll say (roughly) that a belief is rational when it is supported by and properly based on the subject's evidence.[6] When a polarized subject's beliefs are the result of the mechanisms we'll discuss, I'll argue that they are often rational.[7] So, in the cases we'll consider, subjects behaved reasonably—they gathered and scrutinized evidence well—and they believed rationally—their beliefs are (roughly) supported by and properly based on their evidence. I'll also argue that reflective subjects—those who apprehend their situation—are still justified, and should not reduce their confidence in virtue of recognizing that they have undergone polarization. Finally, I'll argue that polarization has distinctive epistemic *benefits* that have been overlooked.

The claims I will argue for do not exhaust the epistemically interesting claims that are true about polarization. For example, shifting to a group-level analysis of polarization can be expected to be revelatory. Within polarization we can also expect heterogeneity that is significant to the analysis. Perhaps some groups engage in epistemically pernicious 'recruitment', which could function as a defeater for otherwise rational beliefs. I don't mean to foreclose these or other possibilities. A full understanding of the epistemology of polarization will involve understanding both the rational and the irrational factors, the individual and the group level assessments. My discussion is partial, but it is a part of the full story.[8]

With these preliminaries out of the way, we can turn to the mechanisms of interest.

---

[6]This characterization is rough because of subtleties involving the distinction between ideal and non-ideal rationality, and the role of scrutinizing the evidence one has. These details will be explained in Section 2.

[7]Exactly how often will depend on other epistemological issues about which I'll remain neutral, not least whether externalism or internalism is true.

[8]Thanks to an anonymous reviewer for encouraging me to clarify this point, and, more generally, everything in this section.

2   Belief Polarization I: Biased Assimilation

Belief polarization may be subserved by many different mechanisms. I'll focus on two: selective exposure and biased assimilation.

Selective Exposure People tend to seek out evidence that confirms their beliefs, rather than evidence that disconfirms their beliefs (Frey, 1986).

Biased Assimilation People tend to give confirmatory evidence they encounter more credence than disconfirmatory evidence they encounter (C. G. Lord, Ross, and Lepper, 1979).

Both of these practices sound bad. Presumably, we should gather evidence neutrally, and treat evidence for and against our views evenhandedly. But, I think, that presumption is mistaken.[9]

In a well-known discussion, Tom Kelly (2008) argues that biased assimilation is rational, when it is subserved by a particular mechanism. Kelly argues that when subjects are confronted with apparently confirming and disconfirming studies—as in (C. G. Lord, Ross, and Lepper, 1979)—they often scrutinize the disconfirming studies more closely. As such, they are more likely to discover flaws with the apparently disconfirming evidence. When they discover flaws, the evidential import of the disconfirming study is undermined. Then, they have (undefeated) confirmatory evidence and defeated discomfirmatory evidence. The rational reaction to such a package of evidence is to increase confidence. So, upon being presented with both confirmatory and disconfirmatory evidence, subjects wind up rationally increasing their confidence in their prior beliefs.[10][11]

---

[9]A different model of rational polarization has been expounded by Kevin Dorst (2023). On Dorst's model, rational polarization can be expected under conditions of asymmetrically ambiguous evidence. In particular, when confirmatory evidence is apt to be unambiguous but disconfirmatory evidence is apt to be ambiguous. The model is ingenious, but I struggle to apply it to the cases under discussion, because if evidence is asymmetrically ambiguous in these cases, it is asymmetrical in the opposite direction from what is required for Dorst's model. In particular, when agents scrutinize evidence, they either find or do not find flaws in it. Finding flaws is naturally understood as an unambiguous defeater, and not finding flaws is naturally understood as ambiguously confirmatory evidence. (Compare Dorst's guiding example of a word completion task.) So I struggle to see how Dorst's model can make sense of the cases I am interested in, and a different explanation of how biased assimilation can be rational *in these cases* is needed.

[10]Note that this explanation is symmetrical, so it applies equally to subjects on each side of the question at issue.

[11]A different model of biased assimilation can be found in (O'Connor and Weatherall, 2018). For O'Connor and Weatherall, we can model epistemic agents (they are focally interested in scientists) as treating evidence from other agents with whom they disagree as more uncertain than evidence from other agents with whom they agree. Increasing uncertainty decreases evidential significance (since, under Jeffrey conditionalization, the appropriate credence is what would be achieved via strict conditionalization multiplied by the agent's credence that the evidence is real, which is lower the greater uncertainty is). O'Connor and Weatherall show in simulations that this procedure can produce polarization in the community. I don't dispute that there are

A natural place to question Kelly's reasoning is the unequal scrutiny subjects deploy to confirmatory and disconfirmatory evidence. Indeed, Emily McWilliams (2019) does just this. On McWilliams' view, Kelly is too restrictive in his conception of evidence. We should treat the flaws in the confirmatory evidence as a part of the subject's evidence, and as such, subjects are not rational to increase their confidence. Increasing confidence on the basis of evidence for which you have a defeater is not rational. On McWilliams' view, the problems with the confirmatory evidence are a part of the subject's evidence whether or not they identify them. Why? McWilliams introduces *motivated defeaters*:

Motivated Defeater A defeater that could be grasped just by reflecting further on one's current evidence, that is not grasped because the subject is not motivated to further reflect, that would act as a defeater if it were grasped. (McWilliams, 2019, p. 9)

I agree with McWilliams that such a motivated defeater exists, but I question whether its existence settles the question at issue. Subjects' motivations could vary arbitrarily so, absent further restrictions, motivated defeaters are any defeaters that could be apprehended by indefinite scrutiny of one's evidence. If we included *all* motivated defeaters in a subject's evidence, then, assuming evidentialism, subjects would never be justified on the basis of evidence that contained overlooked defeaters, however subtle or abstruse. Though such a notion of justification may have some uses, I am not inclined to be so restrictive in thinking about justification.[12]

When we consider the rational response to a body of evidence, we should distinguish between ideal rationality and non-ideal rationality. We can assume that the ideally rational response is to apportion your beliefs to the evidence. Ideally rational agents never overlook subtle flaws in an argument, or alternative explanations of data. We are not ideally rational agents, though, and as such we need non-ideal standards of rationality.[13] Non-ideal agents require time and energy to assess evidence, and can be expected to overlook flaws when they do not spend enough time scrutinizing evidence. As agents consider a body of evidence, what is rational for them to believe may change, despite the body of evidence remaining the same.[14] As they uncover flaws, they are rational to discount the

---

cases well-described by O'Connor and Weatherall's model, but the cases I am interested in are interestingly epistemically different because they involve scrutiny of evidence, which is not a component of their model.

[12]McWilliams draws on empirical evidence to argue that the motivated defeater is not beyond a subject's abilities to understand. I grant that understanding is within a subject's abilities, and instead question whether it is a part of their evidence. So the empirical evidence McWilliams draws on is not relevant to my worry.

[13]For recent discussions of ideal and non-ideal rationality, see (Smithies, 2019; Staffel, 2020).

[14]I am hence making a substantive assumption about the nature of evidence that has been denied by some theorists, e.g. (Feldman, 1988; Williamson, 2002, Chapter 9; Weatherson, 2019, Chapter 7) For my purposes, it is useful to have a less brittle conception of evidence, in order to make it plausible that subjects may share evidence. However, those who opt for a more dynamic conception of evidence are apt to find the spirit of this section even more compelling, since, on a dynamic conception, these subjects are in fact responding *ideally* rationally to their evidence, so conceived, and so assuming evidentialism, are not just non-ideally rational, but ideally rational as well.

relevant evidence, even if the ideally rational thing to do is to notice even *more* subtle flaws. These are standards of non-ideal rationality, and I think that they are the relevant standards in this case. From this perspective, we can grant that motivated defeaters are, in one sense, a part of the subjects' evidence, and yet leave open the question about whether subjects in these cases are behaving rationally according to the relevant non-ideal standards.

McWilliams argues that, were we to exclude motivated defeaters from our evidence, we would not align with the common sense understanding of justification. Excluding motivated defeaters does not align with common sense, because, intuitively, subjects who engage in biased assimilation are not justified, whereas subjects who hear testimony from such subjects are justified. The best explanation for this, McWilliams suggests, is that the testifier has a motivated defeater, but the person receiving testimony does not (McWilliams, 2019, p. 24).

As we have seen, though, absent some restriction on which motivated defeaters to include, McWilliams' view itself will produce the counterintuitive result that subjects are *never* justified when they overlook flaws, no matter how subtle. So I am not sure that McWilliams' view enjoys the support of intuition. Moreover, I wonder how much independent motivation this intuitive difference can provide. The intuition McWilliams aims to vindicate seems quite close to what was at issue in the first place. Indeed, part of what makes Kelly's argument interesting is its apparent vindication of a counterintuitive result. Biased assimilation *does* seem intuitively problematic, and yet Kelly has produced an argument that this intuition is misleading. Adjudicating this dispute seemingly requires independent reason to think the intuition at issue should be relied upon.

McWilliams endeavors to provide just such a motivation. She notes that by iterating the process of biased assimilation, subjects might become very confident 'based on shallow consideration of bad evidence' (McWilliams, 2019, p. 6). We need not admit that 'shallow consideration of bad evidence' secures justification in order to defend the rationality of polarization. Of course many people with polarized beliefs are unjustified. If a person never subjects confirmatory evidence to more than 'shallow consideration', perhaps they won't be justified. However, shallow consideration is inessential to Kelly's explanation. Polarization will be rational, on Kelly's view, so long as there is *unequal scrutiny*—that is, apparent counter-evidence is *more* scrutinized than apparent confirmatory evidence.[15] We can satisfy McWilliams' intuitive objection by setting some minimal level of scrutiny, below which (bad) evidence provides no justification. Insofar as subjects are engaged in unequal scrutiny above the bar though, rational polarization is predicted.

I am not sure how to set the minimum level of scrutiny, but I think we can reasonably expect such a distinction to exist. Consider two examples from Errol Lord. First, Lord fails to notice his wife telling him that she can't pick up their son that day because he's reading

---

[15]Unequal scrutiny, rather than shallow consideration, is what is essential to Kelly's explanation, because unequal scrutiny is what explains why subjects are more likely to find flaws in discomfirmatory evidence. The absolute level of scrutiny is not relevant to the explanation.

a magazine, and second, a slip of paper that reads 'I ate pizza on January 3rd, 2004' is wedged between two pages of a book on his shelf (E. Lord, 2018, p. 91). In each case, knowledge is, in some sense or other, available to Lord. Nevertheless, Lord takes it that in the first case, he is in a position to know, whereas in the second case he is not. How exactly to render this distinction precise is not obvious, but I suggest that we should recognize the same distinction in the *a priori* domain. We are in a position to know some features of our evidence, and insofar as we overlook them, we are not justified. Other features of our evidence, though, are only available through the *a priori* analogue of walking over to our bookshelf and riffling through *American Psycho*. Of course, I haven't analysed the notion of being in a position to know—I don't know how to draw this distinction in a principled way—but I think we should expect a distinction between oversights that defeat justification and those that do not to be available.

So we can dismiss shallow consideration of bad evidence as a red herring. The question is whether *unequal scrutiny* is reasonable. This question is an instance of a more general one: how should non-ideally rational agents like ourselves scrutinize evidence? This question has practical and epistemic aspects. Scrutinizing evidence is an action; it's something we do instead of other things. But it is also a distinctively epistemic action. Scrutinizing evidence affects what it is (non-ideally) rational for us to believe. As such, scrutinizing evidence is subject to general norms of practical rationality, and proprietary norms of practical-cum-epistemic rationality. The distinctively epistemic aspects are most relevant to our discussion, so I will focus on them, but more general practical norms also affect the reasonableness of scrutiny. For example, if one has pressing business, it may be unreasonable to scrutinize one's evidence rather than attend to their pressing business.[16] Turning to the distinctively epistemic norms, one position is *egalitarianism*:

Egalitarianism Agents should equally scrutinize all evidence.

Egalitarianism is not plausible. An egalitarian subject would apply the same scrutiny to ordinary perceptual knowledge and abstruse mathematical proofs. Whatever absolute level of scrutiny they applied would be often unreasonable. Someone following a demanding version of Egalitarianism would waste time and energy on straightforward issues. When a friend they know full well to be a nice person did something nice, they might spend hours weighing alternative explanations before concluding they behaved as they did because they are nice. Doing so would be a neurotic waste of time. Someone following a more permissive Egalitarianism would be objectionably uncritical. Treating all evidence uncritically could fairly be criticized along the lines McWilliams suggests—subjects becoming very confident on the basis of shallow consideration of bad evidence is problematic, even by non-ideal standards of rationality.

Once we reject egalitarianism, how should subjects apportion unequal scrutiny? As I see it, at least two considerations seem relevant. First, subjects should inversely propor-

---

[16]For related discussion, see (Friedman, 2020).

tion scrutiny to how 'straightforward' the evidence is. Second, subjects should inversely proportion scrutiny to how plausible the evidence is.[17] A subject who apportions scrutiny according exclusively to 'straightforwardness' can be thought of as a 'banded' egalitarian. They divide evidence into bands based on how complex it is and then apply equal scrutiny within each band. I don't think banded egalitarianism is advisable. Recall our friend who we know to be a nice person. However we cash out 'straightforwardness', I assume that apparently nice behavior and apparently not nice behavior will be equally straightforward evidence. As such, the banded egalitarian will apportion equal scrutiny to both sorts of behavior. But I take it that doing so is unreasonable. Subjects who know full-well that their friend is nice should scrutinize apparently not nice behavior more closely.

Why should a subject scrutinize apparently not nice behavior more closely? I suggest they should do so because such evidence supports an antecedently less plausible proposition. That is, from the subject's perspective, a less straightforward interpretation of the evidence is more likely to be true. Their prior evidence and justified beliefs suggest that evidence of being not nice is more likely to be misleading than evidence of being nice. So, from the subject's perspective, scrutinizing this evidence more closely is more likely to reveal that the evidence is misleading than applying scrutiny to evidence to the contrary. As such, scrutinizing evidence of being not nice is a better use of time than scrutinizing evidence of niceness, and subjects should prioritize it in deciding where to allocate time and attention to evidential scrutiny. To do so is to allow antecedent assessments of plausibility to guide allocation of scrutiny. It's hard to see how subjects could assess the plausibility of evidence without relying on their prior beliefs. So, for non-ideally rational subjects like us, unequal scrutiny is reasonable. It's not that they owe a debt of loyalty to their friend; rather, they have good reason to expect such evidence to be misleading.[18]

Our overall assessment of an agent who apportions scrutiny on the basis of antecedant plausibility assessments rests on multiple factors, including whether their plausibility assessments are themselves justified. In realistic situations, agents already have some evidence that speaks to how plausible a proposition is when it is time to apportion scrutiny. So, like other beliefs, plausibility assessments can be responsive to an agent's evidence, or not. When the plausibility assessment is justified, the agent is doing the best they can given their limitations. When the plausibility assessment is not justified, though, we enter normatively murky waters. On the one hand, *given* their plausibility assessment, their apportionment of scrutiny seems reasonable. On the other, they *ought* not to assess plausibility as they have, so there's a sense in which they ought not to apportion scrutiny as they have either. The situation is analogous to a subject inferring an obvious entailment of an unjustified belief. Should a subject who unjustifiedly believes they are Superman infer

[17]These proposals are not in competition. I'll suggest that reasonable scrutiny is affected both by 'straightforwardness' and plausibility.

[18]I take it that what the reasoning here supports is that prior plausibility assessments are one factor that guides rational allocation of scrutiny. It may be that complexity is also a factor. As long as plausibility is one factor, unequal scrutiny is licensed.

that they can fly? On the one hand, *given* that they believe they are Superman, it seems that they should believe they can fly. On the other, they *ought* not believe they are Superman in the first place, so they ought not believe they can fly either.[19] I won't attempt to settle how we should ultimately think about cases like this, but it seems clear that when an agent's plausibility assessments are unjustified, our overall assessment should be somewhat less than fully positive.

One worry about the preceding line of reasoning is that it violates a plausible principle—the commutativity of evidence:

Commutativity of Evidence What it is rational to believe depends on one's total evidence, and does not depend on the order in which one encountered the evidence.[20]

If subjects allocate scrutiny on the basis of plausibility assessments, then the significance of new evidence for them will depend on their prior beliefs. As such, their scrutiny of evidence will depend on what evidence they have already been exposed to. Kelly holds that his view can respect the Commutativity of Evidence, because while the order in which subjects gain evidence may *causally* affect the total evidence they come to possess, it does not *epistemically* affect what is rational to believe *given the total evidence they actually end up with* (Kelly, 2008, p. 23). As McWilliams notes, Kelly requires a specific notion of evidence, namely, one on which subjects *do not* possess defeaters that would be revealed were they to further scrutinize the evidence they have. Such a notion of evidence makes it extremely hard for subjects to 'have the same evidence'. Not only must subjects 'have the same evidence' in the sense that they have been presented with e.g. the same studies or behavior, but they also must scrutinize the evidence in the same way. It's fair to assume such a correspondence will be quite rare. Kelly's view also renders a strange verdict concerning what happens when subjects consider a body of evidence. Intuitively, we can consider *the same body of evidence*, and uncover surprising things.[21] On Kelly's view, considering a body of evidence inevitably changes the evidence. So I don't think that we should vindicate the commutativity of evidence by appealing to such a fine-grained notion of evidence.

Instead, I suggest that we appeal again to the distinction between ideal and non-ideal rationality. The commutativity of evidence is a principle that holds for ideally rational agents. They never overlook problems with their evidence, nor are they time and energy limited as we are. As such, they apportion their beliefs to the total evidence. We are not so lucky, though. We are forced to do our best given our limitations. My view is that the best we can do violates the commutativity of evidence. In endeavoring to respect the commutativity of evidence, we will allocate our scrutiny unreasonably—systematically either over-scrutinizing or under-scrutinizing evidence. To respect the commutativity of

---

[19]For discussion of these delicate issues, see (Broome, 1999; Kolodny, 2005; E. Lord, 2014; Worsnip, 2015; Pryor, 2018; Barnett, Forthcoming).

[20]In addition to (Kelly, 2008), see (Lange, 2000; Weisberg, 2009).

[21]These considerations were raised by Julia Staffel in a talk on material from (Staffel, 2020).

evidence would be to make ourselves predictably worse-off epistemically speaking—either missing defeaters that we could have found had we deployed our scrutiny more judiciously, or wasting time and energy scrutinizing highly plausible evidence. To do so would be a form of 'rule worship' that we should reject. We do better epistemically by taking our limitations into account, and doing our best *given* our limitations, rather than acting as if we didn't have them. The best we can do given our limitations, I suggest, involves allowing prior plausibility guide our scrutiny of evidence, even though doing so leads to violations of the commutativity of evidence.[22]

To sum up, non-ideally rational subjects like us require time and energy to scrutinize our evidence, and doing so affects what it is (non-ideally) rational for us to believe. The best we can do in this situation is to apportion our scrutiny in ways that are most likely to be epistemically useful. Doing so involves being guided by our antecedent plausibility judgments, and this in turn results in unequal scrutiny being reasonable. So biased assimilation, when it is the result of this kind of unequal scrutiny, is the output of reasonable epistemic agency.

3 Belief Polarization II: Selective Exposure

The second mechanism is selective exposure. Recall:

Selective Exposure People tend to seek out evidence that confirms their beliefs, rather than evidence that disconfirms their beliefs.

In a recent discussion, Thi Nguyen distinguishes between two aspects of selective exposure—epistemic bubbles and echo chambers. An epistemic bubble is an information sharing network that tends not to include all relevant information. As Nguyen puts it, epistemic bubbles have poor 'coverage'—subjects are apt to miss relevant information or arguments. By contrast, an echo chambers is 'an epistemic community which creates a significant disparity in trust between members and non-members' (Nguyen, 2020, p. 10). Echo chambers need not operate by systematically excluding counter-evidence, as in epistemic bubbles. Rather, counter-evidence is often freely shared, but the trust disparity renders the counter-evidence epistemically inert. Rather than telling against the question, it serves as further evidence of the untrustworthiness of outsiders. As the process unfolds, members' trust of one another is further amplified, and non-members are further epistemically discredited.

Both of these features of networks can drive polarization. Suppose, in particular, that members of an epistemic bubble tend to be organized around an answer to particular question or questions. That is, members of the bubble share a view on that question, and

---

[22]I take my view to be vindicated by more formal work from Justin Dallmann (2017), who argues that refraining from updating on new evidence on questions for which we have already gathered substantial evidence is often the superior cognitive policy from the perspective of expected credal accuracy.

moreover the bubble tends to exclude counter-evidence on that question.[23] If people are systematically not exposed to counter-evidence, then the available evidence will support the predominant view on the question more so than a more holistic set of evidence, leading members of the bubble responsive to available evidence to become more confident than counterfactual versions of themselves who had access to the more holistic set of evidence. In some cases, feedback effects may occur, such that as subjects become more confident, the bubble becomes less permeable, over time rendering subjects within the bubble less and less likely to encounter counter-evidence.

Subjects in echo chambers, by contrast, may enjoy plenty of 'coverage', but the epistemic import of counter-evidence has been distorted. Evidence from out-group members is systematically discounted. Rather than being taken as evidence against their view on the question, counter-evidence is taken as evidence that out-group members are even less credible than previously thought. After all, they have just (knowingly?) provided false or misleading evidence! Feedback effects are prevalent with echo chambers as well. The more subjects discount out-group members' credibility, and boost in-group members' credibility, the more asymmetrically they'll treat apparent evidence from these sources, exacerbating the credibility gap.

Often, polarization driven by epistemic bubbles and echo chambers is epistemically worrisome, but is that an *essential* feature of echo chambers and epistemic bubbles? Let's start with epistemic bubbles. Failure to be exposed to counter-evidence is worrisome, but this worry has to be weighed against the potential benefits. At first, it may not seem obvious that epistemic bubbles have epistemic benefits, but it turns out to be true.

Before proceeding, let me address a verbal issue. We can either use 'epistemic bubbles' and 'echo chambers' to pick out a particular structure of epistemic communities, or we can use these terms to pick out only the epistemically vicious instances of those structures. I am adopting the first usage, because it is the structural features that explain polarization, not whether those structural features are epistemically virtuous or vicious. As such, the first usage eases exposition. On this usage, whether these structures are sometimes, usually, or always epistemically vicious is an open question. The same points can be made adopting the second usage as well. On that usage, the question is whether epistemic communities meeting the relevant structural descriptions are *really* 'epistemic bubbles' or 'echo chambers'. My argument—on this alternative usage—is that often they are not.

In brief, I'll suggest that credibility assessments play a similar role with respect to epistemic bubbles and echo chambers as plausibility played with respect to biased assimilation. The idea that credibility assessments play an important role in social epistemology is of course not new. Saliently, Regina Rini (2017) argues that co-partisanship can be an individually epistemically virtuous basis for giving more weight to testimony, and Endre Begby (2022) argues that sharing partisan beliefs can serve as rational grounds for iden-

---

[23]I take this supposition to be a plausible one for many epistemic bubbles, especially with respect to politically relevant questions.

tifying (Begby prefers to think of 'bestowing' the status of) epistemic peers. My point is complimentary to these authors. In addition to others being sources of testimonial knowledge, and potential epistemic peers, they are also means of managing our access to evidence and arguments that we evaluate directly.

Just as reasonable scrutiny of evidence turns on antecedant judgments about plausibility, reasonable responses to social-epistemic structures turn on antecedant assessments of credibility. The reasonable approach to evidence and arguments from different sources cannot be assessed absent a view about the credibility of the sources. When subjects respond to different sources in light of their justified views about source credibility, they are acting in an epistemically appropriate manner. The epistemically appropriate response, then, may align with echo chambers and epistemic bubbles, or not, depending on what subjects are justified in believing about the relevant sources.

Let's begin with a 'good case'. Reflecting on such a case serves as a 'proof of concept' that subjects sometimes reasonably construct epistemic bubbles and echo chambers. I'll consider a social-epistemic structure that aligns with my view on the merits.[24] In such a case, I'll invite the reader to agree with me that, despite being a paradigm of an epistemic bubble/echo chamber structurally speaking, the subject is behaving reasonably. Such a verdict suggests that what is ultimately epistemically worrisome about some bubbles or echo chambers doesn't turn on the structural features relevant for polarization as such.

Often, epistemic bubbles exclude low quality and misleading evidence.[25] Suppose, for example, a student comes to be enmeshed in an epistemic bubble of professors who work on the social determinants of crime, in virtue of taking their classes. Such a student's epistemic bubble has limited coverage; they do not have access to lots of counter-evidence and counter-arguments published on racist forums, for example. We can suppose the evidence that the student is not exposed to is 1. low quality and misleading, and 2. would have the psychological effect of reducing the student's confidence in the views supported by the evidence presented in class. What should we think about such a case? My own view is that, in this case, the student's epistemic bubble is *epistemically virtuous*. In virtue of being isolated from counter-arguments, the student enjoys a superior epistemic position. They would be worse off, epistemically speaking, were they to begin to spend time on racist forums.

The same lesson recurs for echo chambers. We can imagine a second student who is in an 'echo chamber', in that they treat their professors as enjoying substantially higher credibility than racist forum posters. Although they frequently encounter the counter-evidence (perhaps they are doing sociological research), its effect is not to induce a reduction of confidence about the causes of crime, but rather to make them more confident that

---

[24]People who disagree on the merits are invited to substitute a case that aligns with their own views. My point is structural, rather than substantive.

[25]I am here stacking the deck in my favor, in assuming that the subject's justified beliefs about credibility are also objectively correct. I think this is reasonable given that I am offering a 'proof of concept'. I'll explore what happens when we relax this assumption below.

their in-group members are credible and the forum posters are not credible. Again, my reaction to such a case is to take the student to be responding rationally.

None of this is to object to Nguyen's characterization of epistemic bubbles or echo chambers. As Nguyen says, echo chambers 'are perversions of natural, useful, and necessary attitudes of individual and institutional trust' (Nguyen, 2020, p. 14). The question is: what makes an echo chamber a 'perversion' rather than a straightforward instance of these natural, useful and necessary attitudes? As I see it, the most promising answer for echo chambers turns on the epistemic status of the credibility assessments that explain the echo chamber. If that answer is correct, then we cannot epistemically assess (positively or negatively) echo chambers absent a view about the underlying credibility assessments.

How the basic idea that we evaluate echo chambers and epistemic bubbles relative to an epistemic evaluation of the underlying credibility assessments is worked out will depend on one's epistemological scruples—not least whether one is an internalist or an externalist. For 'radical' externalists like Amia Srinivasan (2020), only the objectively correct credibility assessments are justified, and so, conjoined with the basic idea, only the echo chambers and epistemic bubbles that result from objectively correct credibility assessments will meet our standard. But for internalists, potentially many more social-epistemic structures meet this standard. The most straightforward way to spell out an internalist version of the proposal is to replace objectively correct credibility assessments with justified beliefs about credibility. On this internalist version, instances of symmetric polarization may involve two camps, each of which is constructing a reasonable social-epistemic structure by their own lights.

In either case, the thing that is objectionable about polarization driven by echo chambers is not really about polarization. What's objectionable does not turn on the structural facts driving polarization—since those occur in epistemically virtuous as well as epistemically vicious instances—but on the fact that people are misallocating credibility. My proposal vindicates the intuitive verdicts about our students. The 'epistemic bubble' and 'echo chamber' that they inhabit meets our epistemic standard, because it conforms to the appropriate allocation of credibility of in-group and out-group members.[26]

One worry about the argument of this section is that subjects will not in general be able to tell whether they are in a virtuous or a vicious bubble(/echo chamber). Because the structural features are symmetrical, subjects trapped in epistemic bubbles failing our standard will find the very same reasoning available as those in the virtuous analogue. Wouldn't a more 'open' epistemic structure better facilitate subjects being self-aware with respect to their epistemic situation?[27]

We can distinguish between two aspects of this worry. First is a concern about symmetry. To the extent that subjects in bubbles(/echo chambers) that meet our standard are able to justify their situation, subjects in bubbles(/echo chambers) that don't will appar-

---

[26]Because I've chosen a good case, this is correct on both internalist and externalist elaborations of the basic idea.

[27]Thanks to Brett Karlan for encouraging me to consider this possibility.

ently be able to do so as well. We might think the apparent availability of this reasoning in the bad case undermines the soundness of the reasoning in general. Second is a concern more directly related to the structure of a bubble(/echo chamber). If it were the case that, in general, subjects situated in more 'open' epistemic structures were better able to assess their own epistemic situation, that might give us positive reason to prefer those structures independent of the arguments of this section.

I'll defer a discussion of symmetry to the next section, for now, let's consider the structural argument. Are more open structures—those that share evidence for and against questions freely, and allocate credibility more evenly—epistemically better than more closed epistemic structures? I am not sure whether, in general, the answer to this question is 'yes'. Broadly, we might expect subjects with more 'open' epistemic structures to more commonly suspend judgment on questions they shouldn't, and subjects with more 'closed' epistemic structures to not suspend on questions they should. One would then need a view about how to trade off between type 1 and type 2 errors, which is unlikely to be domain general. Additionally, the putative benefits of an 'open' epistemic structure must be weighed against the potential costs. When substantial credibility disparities exist between in-group and out-group, a subject who fails to construct an echo chamber will be substantially worse off epistemically *ceteris paribus*. Rather than living in an echo chamber, they will wind up living in an epistemic junkyard. So I think the answer to the objector's question is not straightforward, even on the objector's terms.

However, I think we should reject this objection for a deeper reason. The objector evinces a 'pragmatist' sensibility that I wish to resist. Perhaps it's true that more 'open' epistemic structures are in general better for the reason the objector suggests, but subjects *should not* apportion credibility with an eye to constructing open epistemic structures. They should apportion credibility in accordance with the evidence, and then they should construct social-epistemic structures that accord with their (justified) credibility assessments. If that at times results in social-epistemic structures that are suboptimal, then, although that may be in some sense unfortunate, that is what epistemic normativity demands. Put another way, a subject would be *irrational* to misalign their credibility assessments with the evidence, or fail to be guided by their justified credibility assessments in constructing social-epistemic structures, even if doing so would have some other benefits.

4 Reflective Subjects

Supposing that what I've said so far is correct, polarization can occur via reasonable processes. So, for unreflective subjects, polarized beliefs are often justified. Some subjects, though, reflect on their situation. For example, readers of this paper are in a position to understand the processes that resulted in their polarized beliefs. In doing so, perhaps their beliefs are undermined to some degree. Indeed, Tom Kelly endorses this view:

One might then think that one ought to correct for the operation of the rel-

evant psychological mechanisms, by being less confident of those beliefs that are likely to have been the past beneficiaries of the mechanisms. In short, to the extent that the invisible hand becomes visible, one ought to correct for its operation. I believe that this last thought is correct. Those few of us who are aware of the phenomenon of belief polarization—a group which includes, presumably, readers of the present paper—ought to be less confident of beliefs that are likely to have benefited from the underlying psychological mechanisms. The psychological mechanisms in question constitute biasing factors inasmuch as they influence the evidence which one ends up with in a systematic, directed way. (Kelly, 2008, p. 629)

Kelly's line of thought has an intuitive appeal. After all, we could have wound up in a different situation, brought different prior beliefs to bear on the evidence, enmeshed ourselves in different epistemic communities. Had we done so, we would have believed differently, and taken ourselves to be rational, just as we do in the actual circumstance. Given this observation, shouldn't we reduce confidence once we reflect on the situation we find ourselves in?

I think not.

Admittedly, how we respond to taking ourselves to be in such a situation depends in part on what we think about ongoing debates about disagreement, higher-order evidence and irrelevant influences (I will collectively refer to these debates as 'the epistemological puzzles'). The scope of this discussion does not include settling those debates. Instead, I'll suggest that the mechanisms under discussion do not pose any special problem, over and above the epistemological puzzles. As such, polarization does not pose a distinctive epistemic challenge to our knowledge. Those who are inclined towards a conciliatory response to the epistemological puzzles[28] should take the same attitude towards polarization, whereas those opposed to such a view,[29] should not be any more worried about polarization. In a sense, I am arguing that some mechanisms that subserve polarization are not epistemically objectionable, and in another sense I am not. I argue that there is nothing *epistemically distinctive* about these mechanisms as such. To the extent that reflective subjects should reduce confidence in light of understanding their situation, it is because they appreciate their situation as being one that gives rise to the epistemological puzzles (together with auxiliary commitments to a particular view on how to resolve the epistemological puzzles). So while many will conclude, given their views, that appreciating polarization *does* rationally require us to reduce our confidence, that is not *because* the

---

[28]e.g. (Christensen, 2007; Elga, 2007; Feldman, 2006).

[29]e.g. (Kelly, 2005; Srinivasan, 2015; Srinivasan, 2019; Schoenfield, 2018; Smithies, 2019; Barnett, n.d.). This way of categorizing views substantially simplifies several complex debates. More carefully, my view is that people should apply their general views about the epistemological puzzles to the case of polarization. So those who endorse more nuanced views (e.g. (J. Lackey, 2010; Kelly, 2010; Sliwa and Horowitz, 2015)) should also apply them to polarization.

situations involve polarization. It is instead for familiar reasons to do with the epistemological puzzles.

Let me emphasize that epistemic questions about mechanisms that subserve polarization do not simply *reduce* to the epistemological puzzles. Though an all things considered verdict on what reflective subjects ought to do requires a resolution of the epistemological puzzles, we can ask distinct questions about subjects understanding themselves as engaged in e.g. biased assimilation. Cases of disagreement need not be cases of biased assimilation, and biased assimilation can occur in the absence of an epistemic peer with whom one disagrees. So although often biased assimilation and disagreement cooccur, we can consider them separately.

However we resolve the epistemological puzzles, we should not ascribe any special epistemic significance to polarization as reflective subjects. If reflective subjects should reduce confidence on the basis of recognizing themselves as polarized, then obviously counterproductive epistemic practices, both subjectively and objectively, would be mandated. Structurally, we find ourselves paradigmatically polarized with respect to questions about the causes of crime, climate change and vaccines, to name a few.[30] But to the extent that we are polarized as a result of biased assimilation and selective exposure, we are behaving as responsible epistemic agents, doing our best given our limitations. We are reasonable to avoid their arguments—reviewing them is a waste of time (selective exposure)—and, when we encounter them, perhaps at Thanksgiving dinner, we should either summarily reject them and downgrade our estimate of speaker credibility (echo chamber) or scrutinize them with an eye to pointing out their errors (biased assimilation). That is to say, in cases like this, the mechanisms that underwrite polarization are *epistemically virtuous*, and we should deploy them even once we understand that this is what we're doing. The fundamental point is that *the content matters*. But if the content matters, then polarization doesn't, because polarization is a structural phenomenon and as such is content neutral.

I don't claim that this attitude is always justified, nor that most people who have this attitude are justified. In my view, a number of factors affect whether this attitude is appropriate. But to the extent that the counter-evidence is low quality, the excluded parties are not credible, and other members of one's epistemic community have engaged with excluded parties and their evidence, a more exclusionary attitude strikes me as appropriate. The aforementioned factors are group level features, but individual level features are relevant as well. For example, if a particular subject is not especially good at assessing evidence, or relatedly, will likely have their views changed by exposure to counter-evidence, that subject may more appropriately take on an exclusionary attitude. I have phrased these factors externalistically, but their significance can be accommodated by an internalist as well, in terms of what a subject rationally believes about these factors, though of course the internalist and externalist factors may come apart.

Moreover, selective exposure may enable us to avoid forms of irrationality that we

---

[30]See (J. A. Lackey, 2013) for a discussion of cases like this.

might expect would befall us were we more ecumenical about the evidence to which we are exposed. For example, suppose a subject was persuaded by Mandelbaum and Quilty-Dunn (2015) that subjects by default accept propositions that are presented to them.[31] According to Mandelbaum and Quilty-Dunn, attention and executive function are required to reject presented propositions. If this 'Spinozan' model of belief formation is correct, then we should expect that mere exposure to propositions contrary to our views will eventuate in false beliefs. How should the reflective subject react to being convinced this is true about cognition? If they care about avoiding false beliefs, endeavoring to avoid exposure to them is a reasonable strategy. As the subtitle of Mandelbaum and Quilty-Dunn's paper suggests, 'liberals shouldn't watch fox news'. Viewed this way, given purported limits of human rationality, perhaps the best we can do epistemically is to engage in 'managerial control' over our beliefs in some cases.[32] Our cognitive architecture does not always facilitate directly rational responses to our environment, but we can deploy our agency to make it the case that we are in relatively more congenial environments, epistemically speaking. Selective exposure accomplishes this, by evading the most epistemically pernicious environments. Or so, I suggest, the reflective subject should reason.[33]

But now, let's return to the symmetry worries. The reasoning I've been suggesting to justify not reducing confidence upon recognizing polarizing mechanisms is apparently available to those on the other side of the question. But surely, *they* should reduce confidence—they're wrong! So there must be something suspect about the reasoning that seems to support remaining steadfast. So we should reject that reasoning, and reduce confidence.

We should ask two questions about the symmetry worry: is the same reasoning really available, and is it true that that reasoning does not support a steadfast response for people on the other side of the question? I don't think it's obvious that either question should be answered as the objector assumes.

In some circumstances, I don't think that symmetrical reasoning is available. Though of course subjects on the other side may *take* themselves to be reasoning in the same way, they might simply be wrong, not just on the merits, but also in how they take themselves to be reasoning. In many cases, the evidence they have does not actually support their position as strongly as they believe, and their attributions of credibility come apart from the attributions that would be warranted. In such cases, though they may take the same attitude, their attitude is not justified. As such, they really should reduce confidence. The reason they should reduce confidence, though, is not because they have recognized them-

---

[31]See (Mercier, 2020) for a skeptical take on this model of belief formation.

[32]See (Hieronymi, 2006).

[33]Recall, I am remaining neutral on the empirical disputes about belief fixation. My point is that an agent's view of their own mind—and the epistemic pitfalls they identify themselves as being susceptible to—can affect how it is reasonable to construct their epistemic environments. The Spinozan model of belief fixation offers a useful example, but the point generalizes to whatever epistemic pitfalls agents are aware they are vulnerable to.

selves as subject to mechanisms that subserve polarization, but because they have not responded appropriately to the evidence.

But to the extent that the symmetrical reasoning is really available, I don't think it's obvious that subjects should reduce confidence. After all, they are appropriately apportioning their beliefs to the evidence, and attributing credibility as they should. Rather than thinking that they should reduce confidence, it strikes me as more plausible to think of them as doing their best in an epistemically unfortunate circumstance.[34] These two replies counter the symmetry objection synergistically: to the extent that the same reasoning is really available, the intuition that subjects should reduce confidence is attenuated, and to the extent that the same reasoning is not available, treating the cases symmetrically becomes less plausible.

So while I don't mean to dismiss or to endeavor to resolve the epistemological puzzles—disagreement, higher-order evidence and irrelevant influences—that intersect with polarization, I don't think that polarization as such is epistemically worrisome for the reflective subject. In many cases, the reflective subject ought to feel secure in their beliefs, even after appreciating their origins in polarizing processes.

5   The Objection from Q

Overall, this paper paints a much more optimistic picture of polarization than is often assumed. Perhaps what is most viscerally implausible about this picture are real world examples of polarization that are manifestly epistemically objectionable. For example, some Americans believe a loose cluster of conspiracy theories that go under the head 'QAnon', or just 'Q'. Among the many beliefs characteristic of Q are that John F. Kennedy Jr. is still alive, and perhaps identical to a person who does not look very much like JFK Jr. at all, and that a cabal of cannibalistic, satanic pedophiles conspired against Donald Trump during his presidency.[35] I'll call beliefs like these 'characteristic Q beliefs'. Characteristic Q beliefs are taken to be both a paradigm of polarization, and manifestly epistemically objectionable. If so, my more optimistic picture strains credulity.

I'll begin on a concessive note. In principle, I grant that an agent could come to believe characteristic Q beliefs via the mechanisms I have argued we should be epistemically untroubled by. Assuming an internalist framework, such an agent would thereby be justified in their characteristic Q beliefs. Such an agent would be in a profoundly unfortunate epistemic position, in many ways akin to a skeptical scenario, but they are doing the best they can given their evidence. If I am right about the mechanisms under discussion, then a rational Q believer is in principle possible.

---

[34]I am here, concurring with internalist verdicts in new evil demon-style cases (see (Cohen, 1984; Pollock, 1984; Feldman, 1985; Foley, 1985). Many externalists will be inclined to reject this verdict, but note that doing so militates against thinking of the two poles symmetrically, and hence invites the first reply.

[35]For a much more detailed consideration of QAnon, see (Marwick and Partin, 2022).

Nevertheless, my argument does not entail that most (or any) actual Q believers are rational, nor does it deprive us of resources to explain why their beliefs are irrational. My defence of biased assimilation and selective exposure imposed substantive conditions on when these behaviors are reasonable, and we have little reason to believe these conditions are met for Q believers. We would need reason, for example, to think that as people come to hold characteristic Q beliefs, they are judiciously apportioning scrutiny for and against the Q conspiracies on the basis of justified antecedant plausibility judgments, and that assigning high credibility to an anonymous forum poster making extraordinary claims is justified, and that their total evidence, even as scrutinized by them, in fact supports Q beliefs. In order for Q to pose a problem for my argument, these conditions (and others) would need to be met. I see little reason to think that they are, and, absent a very unusual initial epistemic position, substantial reason to think that they are not. So our intuitive judgment that real world characteristic Q beliefs are epistemically objectionable is not in tension with my argument.[36]

Part of the lesson of our exploration is that not all polarization is created equal. We should embrace the consequence that some polarization might be epistemically good, while many other instances of polarization might be wildly irrational. But interrogating particular cases, and weighing the epistemically relevant features of different cases is obscured by the default assumption that polarization as such is epistemically objectionable. As we've seen, it's hard to say what would be objectionable about polarization as such. Indeed, I hold that there are distinctive epistemic benefits to polarization. This lesson is consistent with holding that many particular cases of polarization—including characteristic Q beliefs—are epistemically objectionable, but our explanations of their being objectionable should be more specific, and indeed, not in terms of polarization as such.

6 The Epistemic Upsides of Polarization

In this somewhat more speculative section, I want to thematize some of the distinctive epistemic benefits of polarization. My goal is not to argue that polarization is all things considered good, or that every instance of polarization is epistemically beneficial. Instead, my goal is corrective. The focus on the epistemic problems with polarization has obscured some distinctive epistemic benefits. A fuller understanding of the epistemology of polarization will include both its costs and benefits. I have already argued that often polarization is epistemically innocuous; my more speculative claim is that polarization also often enjoys substantial epistemic virtues.

For questions that align with the spectrum of political views in a society, it's hard to imagine an *a priori* reason to believe that the truth generally lies 'in the middle' of the

---

[36]Though, again, I grant that *in principle* these beliefs could result from the mechanisms I am optimistic about. The most realistic cases I can think of that plausibly meet these conditions are people born and raised in isolated cults, whose evidential environment is extremely restricted *de novo*.

contemporary discourse, wherever that happens to be. To the extent that the truth on any particular question lies 'at the extreme', then the subjects who are polarizing in the right direction with respect to that question are doing better with respect to accuracy. So we need to weigh the costs of one side 'doing worse' against the benefits of the other side 'doing better'.[37] Further, polarization need not be symmetric, so in some cases, we may enjoy the benefits of one community polarizing 'in the right direction', without corresponding negative polarization.

But now, let's suppose, as many do, that society is suffused with *ideology*, in the bad sense.[38] An ideology is a social system that conditions how we think and act. When the ways our thought and actions are conditioned are bad, the social system is an ideology in the bad sense. One way in which a bad ideology conditions our thought is by obscuring or distorting normatively important facts. By doing so, ideology enervates our epistemic position.[39] We would be epistemically better off were we able to extricate ourselves from the grips of ideology. I suggest that often polarization is conducive to doing so.

Ideology often functions by presenting itself (accurately) as the cultural default, the 'reasonable consensus'. Insofar as we start out suffused in ideology, and so near the 'reasonable consensus', it's hard to escape. Ideology exerts a powerful gravitational force on all of us, and even-handedly weighing plausible-sounding arguments for and against positions is not an especially effective method of achieving escape velocity. The mechanisms that subserve polarization, by contrast are just the ticket. As our epistemic position moves away from the 'reasonable consensus'—as we become polarized—ideology's grip on us is attenuated.[40] Views that we accepted uncritically become dubious, speakers we took to be reasonable, increasingly appear to be making questionable assumptions. Our estimates of their credibility decrease. Things that seemed obvious begin to appear apt for critical scrutiny. To the extent that we begin scrutinizing things *that are epistemically good to scrutinize*—and, on the assumption that ideology is pervasive, we can expect there to be many such things—polarization improves our epistemic position.

How persuasive this argument is depends, as always, on its assumption. Those who do not believe that we are suffused with ideology will not be impressed. On the other hand, those who do think that ideology is pervasive—and especially those who think it is *very important* to extricate ourselves from ideology to the degree we can—should take very seriously the epistemic benefits of polarization.

---

[37]Neil Levy (2020, pp. 6–7) makes this observation about gun control today, and race and gender in the antebellum south.

[38]See (Haslanger, 2000; Haslanger, 2007; Shelby, 2014; Haslanger, 2017). A vigorous debate is ongoing about how best to understand ideology. My purpose is not to enter into that debate. Rather, I assume what I take to be common ground in that debate: namely, that ideology in the pejorative sense—however we understand it metaphysically—impoverishes our epistemic position.

[39]As well as our political position and more general normative position. This paper is about epistemology, so I focus on the epistemic effects.

[40]That's not to say that simply being polarized *constitutes* an escape from ideology. I assume that escape is much harder than that, and it will obviously depend on the question under consideration.

Finally, and most speculatively, we might consider another potential benefit to polarization. *Ceteris paribus*, polarized subjects may have an epistemic advantage with respect to understanding alternative perspectives, for two reasons. First, diachronically: almost by definition, polarized subjects *used to inhabit a different perspective*. As such, they have access to an alternative perspective first-personally. Assuming that first-personal acquaintance provides distinctive insight, we can expect polarized subjects (again, *ceteris paribus*) to have distinctive insight into at least one alternative perspective. Second, synchronically: on the assumptions we have made about ideology, it may be reasonable to expect that *everyone* has access to the dominant perspective. Perhaps that's part of what makes it dominant. As such, polarized subjects maintain a kind of access to the dominant perspective. Non-polarized subjects may have to expend quite a bit of effort to achieve corresponding access to alternative perspectives.[41]

7 Conclusion

I feel the impulse to conclude—not by saying what I've done, but—by saying what I haven't done. My aims have not been to vindicate every instance of polarization, nor to minimize the potential downsides of a polarized society. Nor, even within the domain of academic epistemology, have I endeavored to vindicate the epistemic virtue of polarization all things considered and in every case. Rather, my aims have been modest, and corrective. In order to understand polarization more fully, we need to consider it in all its guises—the rational and the irrational, the epistemically beneficial and the counterproductive. Perhaps unspurprisingly, given the political climate, focus has been most squarely on the warts of polarization. I think it behooves us to take a moment to consider the upsides as well. Indeed, on the view I've argued for, even those who are not ultimately persuaded are apt to have their perspectives enriched by seeing things my way, if only for a bit.

References

Barnett, David James (Forthcoming). "Internalism, Stored Beliefs, and Forgotten Evidence".
  In: *Memory and Testimony: New Essays in Epistemology*.
— (n.d.). "Reflection Deflated". In: ().
Begby, Endre (2022). "From belief polarization to echo chambers: A rationalizing account".
  In: *Episteme*, pp. 1–21.
Broome, John (1999). "Normative requirements". In: *Ratio* 12.4, pp. 398–419.

---

[41]The line of thought in this paragraph is of course indebted to work in feminist standpoint epistemology, although the kind of perspective at issue here bears only a broad similarity to the materialist social positions that are often centrally relevant to that work. See e.g. (Collins, 2002; S. G. Harding, 2004; S. Harding, 2009; Hartsock, 2017; Intemann, 2010).

Christensen, David (2007). "Epistemology of disagreement: The good news". In: *The Philosophical Review* 116.2, pp. 187–217.

Cohen, Stewart (1984). "Justification and Truth". In: *Philosophical Stuidies*, pp. 279–295.

Collins, Patricia Hill (2002). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. routledge.

Dallmann, Justin (2017). "When Obstinacy is a Better (Cognitive) Policy". In: *Philosopher's Imprint* 17.24, pp. 1–17.

Dorst, Kevin (2023). "Rational polarization". In: *Philosophical Review* 132.3, pp. 355–458.

Elga, Adam (2007). "Reflection and disagreement". In: *Noûs* 41.3, pp. 478–502.

Feldman, Richard (1985). "Reliability and justification". In: *the Monist* 68.2, pp. 159–174.

— (1988). "Having evidence". In: *Philosophical analysis*. Springer, pp. 83–104.

— (2006). "Epistemological puzzles about disagreement". In: *Epistemology futures*, pp. 216–236.

Foley, Richard (1985). "What's wrong with reliabilism?" In: *The Monist* 68.2, pp. 188–202.

Frey, Dieter (1986). "Recent research on selective exposure to information". In: *Advances in experimental social psychology* 19, pp. 41–80.

Friedman, Jane (2020). "The epistemic and the zetetic". In: *Philosophical review* 129.4, pp. 501–536.

Haidt, Jonathan (2001). "The emotional dog and its rational tail: a social intuitionist approach to moral judgment." In: *Psychological review* 108.4, p. 814.

Harding, Sandra (2009). "Standpoint theories: Productively controversial". In: *Hypatia* 24.4, pp. 192–200.

Harding, Sandra G (2004). *The feminist standpoint theory reader: Intellectual and political controversies*. Psychology Press.

Hartsock, Nancy CM (2017). *The feminist standpoint: Developing the ground for a specifically feminist historical materialism*. Routledge.

Haslanger, Sally (2000). "Gender and race:(What) are they?(What) do we want them to be?" In: *Noûs* 34.1, pp. 31–55.

— (2007). "" But mom, crop-tops are cute!" Social knowledge, social structure and ideology critique". In: *Philosophical Issues* 17, pp. 70–91.

— (2017). "Racism, ideology, and social movements". In: *Res Philosophica* 94.1, pp. 1–22.

Hieronymi, Pamela (2006). "Controlling attitudes". In: *Pacific Philosophical Quarterly* 87.1, pp. 45–74.

Intemann, Kristen (2010). "25 years of feminist empiricism and standpoint theory: Where are we now?" In: *Hypatia* 25.4, pp. 778–796.

Kelly, Thomas (2005). "The epistemic significance of disagreement". In: *Oxford studies in epistemology* 1.167-196.

— (2008). "Disagreement, dogmatism, and belief polarization". In: *The Journal of Philosophy* 105.10, pp. 611–633.

— (2010). "Peer disagreement and higher order evidence". In: *Disagreement*.

Kolodny, Niko (2005). "Why be rational?" In: *Mind* 114.455, pp. 509–563.

Kundra, Ziva and Lisa Sinclair (1999). "Motivated reasoning with stereotypes: Activation, application, and inhibition". In: *Psychological Inquiry* 10.1, pp. 12–22.

Lackey, Jennifer (2010). "12. What Should We Do When We Disagree?" In: *Oxford Studies in Epistemology Volume 3*, p. 274.

Lackey, Jennifer Amy (2013). "Disagreement and belief dependence: Why numbers matter". In: *The epistemology of disagreement: New essays*. Oxford University Press, pp. 243–268.

Lange, Marc (2000). "Is Jeffrey conditionalization defective by virtue of being non-commutative? Remarks on the sameness of sensory experiences". In: *Synthese*, pp. 393–403.

Levy, Neil (2020). "Virtue signalling is virtuous". In: *Synthese*, pp. 1–18.

Lord, Charles G, Lee Ross, and Mark R Lepper (1979). "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence." In: *Journal of personality and social psychology* 37.11, p. 2098.

Lord, Errol (2014). "The real symmetry problem (s) for wide-scope accounts of rationality". In: *Philosophical Studies* 170, pp. 443–464.

— (2018). *The importance of being rational*. Oxford University Press.

Mandelbaum, Eric (2019). "Troubles with Bayesianism: An introduction to the psychological immune system". In: *Mind & Language* 34.2, pp. 141–157.

Mandelbaum, Eric and Jake Quilty-Dunn (2015). "Believing without Reason: or: Why Liberals Shouldn't Watch Fox News". In: *The Harvard Review of Philosophy* 22, pp. 42–52.

Marwick, Alice E and William Clyde Partin (2022). "Constructing alternative facts: Populist expertise and the QAnon conspiracy". In: *new media & society*, p. 14614448221090201.

Mayo-Wilson, Conor, Kevin JS Zollman, and David Danks (2011). "The independence thesis: When individual and social epistemology diverge". In: *Philosophy of Science* 78.4, pp. 653–677.

McWilliams, Emily C (2019). "Evidentialism and belief polarization". In: *Synthese*, pp. 1–32.

Mercier, Hugo (2020). *Not born yesterday: The science of who we trust and what we believe*.

Nguyen, C Thi (2020). "Echo chambers and epistemic bubbles". In: *Episteme* 17.2, pp. 141–161.

O'Connor, Cailin and James Owen Weatherall (2018). "Scientific polarization". In: *European Journal for Philosophy of Science* 8, pp. 855–875.

Pollock, John L (1984). "Reliability and justified belief". In: *Canadian Journal of Philosophy* 14.1, pp. 103–114.

Pryor, James (2018). "The merits of incoherence". In: *Analytic Philosophy* 59.1, pp. 112–141.

Rini, Regina (2017). "Fake News and Partisan Epistemology". In: *Kennedy Institute of Ethics Journal* 27.S2, pp. 43–64. DOI: `10.1353/ken.2017.0025`.

Schoenfield, Miriam (2018). "An accuracy based approach to higher order evidence". In: *Philosophy and Phenomenological Research* 96.3, pp. 690–715.

Shelby, Tommie (2014). "Racism, moralism, and social criticism". In: *Du Bois Review* 11.1, pp. 57–74.

Sliwa, Paulina and Sophie Horowitz (2015). "Respecting all the evidence". In: *Philosophical Studies* 172.11, pp. 2835–2858.

Smithies, Declan (2019). *The epistemic role of consciousness*. Philosophy of Mind.

Srinivasan, Amia (2015). "The Archimedean Urge". In: *Philosophical Perspectives* 29, pp. 325–362.

— (2019). "VII—Genealogy, Epistemology and Worldmaking". In: *Proceedings of the Aristotelian Society*. Vol. 119. 2. Oxford University Press, pp. 127–156.

— (2020). "Radical externalism". In: *Philosophical Review* 129.3, pp. 395–431.

Staffel, Julia (2020). *Unsettled thoughts: A theory of degrees of rationality*. Oxford University Press, USA.

Weatherson, Brian (2019). *Normative externalism*. Oxford University Press.

Weisberg, Jonathan (2009). "Commutativity or holism? A dilemma for conditionalizers". In: *The British Journal for the Philosophy of Science* 60.4, pp. 793–812.

Williamson, Timothy (2002). *Knowledge and its Limits*. Oxford University Press on Demand.

Worsnip, Alex (2015). "Narrow-scoping for Wide-scopers". In: *Synthese* 192, pp. 2617–2646.

Zollman, Kevin James Spears (2010). "Social structure and the effects of conformity". In: *Synthese* 172, pp. 317–340.

Zollman, Kevin JS (2007). "The communication structure of epistemic communities". In: *Philosophy of science* 74.5, pp. 574–587.