

Compromising with the Uncompromising: Political Disagreement Under Asymmetric Compliance*

Alex Worsnip

Forthcoming in the *Journal of Political Philosophy*

Please cite published version if possible

Imagine a fictional society in which there exist two groups of people, the Twisters and the Stickers. Often, but not always, when some new political issue arises, the Twisters have one reaction about what is to be done, while the Stickers have some other, conflicting reaction. But there is also a difference in how the Twisters and the Stickers react when they *learn* about their disagreements. Even though the Twisters do not find the Stickers' arguments particularly persuasive, they take the fact that so many Stickers disagree with them to be a sign that they might be mistaken in their reactions. As such, when they encounter disagreements with the Stickers, they become significantly less confident in their original views. The Stickers, on the other hand, remain intransigent in the face of disagreement, never reducing their confidence. Both groups participate in the political process – protesting, campaigning, discussing, and voting – based on the views (and levels of confidence) that they have ultimately reached, and the process gives equal weight to each participant's view in determining outcomes. As one might expect, the view that the Twisters began with almost never carries the day. By contrast, the view that the Stickers began with often does carry the day. And over time, the political center of gravity systematically shifts toward the (original) views of the Stickers.

What has gone wrong here? One thing that has plausibly gone wrong, at least given many ways of developing the case, and given certain assumptions about the epistemology of disagreement that will be discussed later, is that the Stickers are being unreasonable in their total intransigence. They ought to, like the Twisters, recognize that there are many thoughtful and intelligent people who disagree with them, and moderate their confidence in the truth of their views in response, at least somewhat. But it seems to me – and I hope it seems to you – that more than that has gone wrong. There is something *unfair* about the outcomes of the political process in our imagined society. Specifically, there seems to be some sense – I'll try to precisify what sense later – in which the situation unfairly favors the Stickers (and their views), and fails to adequately represent the Twisters.

Given this, what ought the Twisters to do? Three answers suggest themselves:

- (1) The Twisters should change their epistemic practices, and refuse to moderate *their* confidence, just as the Stickers do.
- (2) The Twisters should keep on doing what they're doing.

* For helpful discussion of the arguments of this paper, I'm grateful to Chris Blake-Turner, Tom Dougherty, Georgi Gardiner, Hanna Gunn, Kathryn Lindeman, Jessica Marsden, Doug MacKay, Daniel Muñoz, Geoff Sayre-McCord, Julia Staffel, three anonymous referees, and audiences at the University of Colorado Boulder (and in Alison Jaggar's graduate political philosophy seminar shortly afterward), the 2019 PPE Society Annual Meeting, and the Political Epistemology Network meeting at the 2019 Pacific APA.

- (3) The Twisters should continue to moderate their confidence in response to disagreement, but they should *act* (vote, participate in discussions, etc.) in a way that doesn't reflect that moderation in confidence.

In this paper, I will argue in favor of answer (3). My discussion has at least three interesting broader upshots. First, it has the upshot that a procedure where everyone votes based on their sincerely-held beliefs, and everyone's vote counts equally, is not necessarily representatively fair. Second – and more boldly – it has the upshot that there are cases where there is a principled rationale for actively voting *against* one's sincerely held beliefs – but one that is distinct, I'll argue, from cynical “strategic” attempts to manipulate the political process to yield one's desired outcome. But third, I'll argue, we can posit a mental state – I will (stipulatively) call it one's “personal take” – that one *should* vote on the basis of in these cases.

Of course, in the real world, we are not in a situation as neat as that of the Stickers and Twisters. We have overlapping political allegiances, and we don't split into two discrete, perfectly internally homogenous political groups. And even when we can classify people into political groupings on the first-order normative issues, it certainly isn't true that everyone in one group responds in a conciliatory way to disagreement, and that everyone in the other responds in an intransigent way. Nevertheless, the same sort of problem can arise, albeit in a less extreme way, in the real world. We can be faced with the intransigence of others, and wonder what we should do about our own views and behavior, in the face of this intransigence. So the more general question that I am exploring here is this: how should we react to disagreement under conditions of asymmetric compliance with the ideal principles about how we should react to disagreement – especially when collective choice is at stake?¹

1. Confidence levels and political behavior

Before I get to the three options mentioned above, let me make our central case, which is thus far somewhat abstract, a bit more concrete by considering some ways in which the Twisters' reactions to disagreement might change their political behavior – thus, in combination with the Stickers' intransigence, moving the center of gravity toward the Stickers. I will briefly survey three ways in which this can happen, at least when people act on the basis of their beliefs.

First, changes in confidence can rationalize support for policies that hedge against risks. For example, suppose you are very confident that there will not be a collapse in the value of the dollar this year, but you then discover that many others disagree. If you reduce your confidence that there will not be a collapse in the value of the dollar – even if you don't go as far as believing that it *will* collapse – this may rationalize supporting economic policies designed to mitigate the risks that such a collapse would pose. A similar phenomenon *may* occur when one reduces one's confidence in one's moral

¹ For another discussion of cases where one party reduces their confidence and the other remains steadfast, see Aikin *et al.* (2010). Aikin *et al.* think such cases create trouble for conciliationism. I do not agree with them about this, but can't argue this here.

views in response to disagreement – if it’s sometimes rational to “hedge one’s bets” under moral uncertainty to avoid the possibility of moral catastrophe.²

Second, there can be cases where one’s loss of confidence in disputed claims leads to a kind of “washing out” of the more controversial considerations in determining my stance toward a given political issue, leaving the uncontroversial considerations to determine my view. Suppose that we are discussing a trade treaty. I believe that the trade treaty will lead to some short-term domestic job losses, but that this is outweighed by (i) the jobs it will create for those giving in greater need in the developing world, and (ii) the long-run benefits it will have for the domestic economy. You, on the other hand, agree that the trade treaty will lead to short-term domestic job losses, but you also think it won’t really help workers in the developing world (and/or that their well-being should matter less for policymaking than that of domestic workers) and that it will be bad for the domestic economy in the long-run. As I reduce my confidence in the various propositions we disagree about, these considerations progressively “wash out” in my decision-calculus more and more: the less confident I am that the treaty really will benefit foreign workers, and the less confident I am that it really will be good for the domestic economy in the long-run, the less weight these factors will have in my overall decision: it’s rational for me to discount for my uncertainty. But since we agree that the policy will lead to short-run domestic job losses, my confidence in this proposition won’t change. And so it seems that if I reduce my confidence in the disputed propositions sufficiently, eventually the (relative) certainty of short-run domestic job losses will become the dispositive factor in my choice, and I will come to oppose the treaty, or at least become broadly neutral about it such that it ceases to become a politically important issue for me.

Third, disagreement may sometimes not just call for a reduction in confidence in one’s views but also for compromising with one’s disputants in the *content* of one’s beliefs.³ For example, if I believe that a passing stranger is 6’4”, and you believe that a passing stranger is 6’0”, there might be at least some pressure on each of us to come to believe that the passing stranger is some more intermediate height. Some political cases may be like this too. Perhaps, for example, if I believe that some tax plan will create rampant inflation, and you believe that the tax plan will create a drastic reduction in inflation, this creates some pressure on each of us to think that the effect of the tax plan on inflation will be less pronounced than we each first thought. And if I moderate my belief in this way, this will have the capacity to change my political behavior.

In all of these kinds of cases, if the Twisters *do* moderate their beliefs and the Stickers *don’t*, and so the Twisters change their political behavior accordingly and the Stickers don’t, we can expect this to produce a systematic political shift toward the Stickers’ original views.

2. First option: stop conciliating

² Lockhart (2000), MacAskill (2014), and Sepielli (2009, 2017) argue that this can happen. Weatherson (2014) and Harman (2015) argue that it can’t. Putative examples of such cases are controversial, but examples that have been said to have this structure include abortion (see Lockhart 2000: ch. 3) and meat-eating.

³ This possibility is occasionally mentioned – see e.g. Gardiner (2014: 85-6); Wiland (2017: 495-6); Jackson (2019: 2483) – but rarely (if ever) explored in detail.

The first option in response to the case of the Twisters and the Stickers is to say that the Twisters should simply stop moderating their confidence in response to the Stickers' disagreement. It might be thought that the first sub-view can be motivated by appeal to a "steadfast" view of the epistemology of peer disagreement. However, this is not as clear as it might first seem. Consider the following view:

Minimal Conciliationism. Suppose you discover that a large group of people disagree with your belief in p . Unless you are justified in thinking that this group is *no better than (or worse than) chance* with respect to issues like whether p , learning of their disagreement calls for at least *some* moderation of your confidence in your belief in p .

Notice that minimal conciliationism is, in keeping with its name, a fairly weak claim. More robust forms of conciliationism say that regardless of the circumstances and one's other evidence, peer disagreement always calls for a *significant* adjustment in confidence. The most famous such view is the equal weight view, which says that in a two-person disagreement, you are required to give equal weight to your own view and that of your disputant, thus arriving at a credence equidistant between the two, regardless of your other evidence. While minimal conciliationism is compatible with these kinds of views, it incurs no commitment to them: it just says that disagreement calls for *some* moderation of your confidence in your belief. This allows that the amount of moderation might be quite low in some cases, depending on the circumstances and your other evidence.

The rationale for minimal conciliationism is simple. When you learn that a person (or group) who is better than chance believes not- p , this is some evidence for not- p . And, holding all else fixed, gaining new evidence for not- p calls for a reduction in your confidence in p . So, when you learn that group of people who are better than chance believes not- p , this calls for a reduction in your confidence in p . Notice that this argument requires no assumption to the effect that the disagreeing group are your "epistemic peers"; all it requires is that they be better than chance.⁴ Moreover, unlike arguments for more robust forms of conciliationism, it requires no special assumptions about complex epistemological matters such as whether permissivism is true or the extent to which higher-order evidence defeats first-order evidence.⁵ The argument looks rock-solid.

Do advocates of the "steadfast view" about peer disagreement reject *minimal* conciliationism, and thus, this argument for it? In fact, they generally do not. For example, consider Thomas Kelly's (2010) well-known "total evidence" view, usually classified as a "steadfast" or "anti-conciliationist" view. On Kelly's view, the disagreement of others is just one more piece of evidence to be weighed alongside whatever other evidence you have. This is manifestly compatible with minimal conciliationism, which – to repeat – merely says that disagreement calls for *some* adjustment in confidence, and says nothing about how it weighs against other evidence in order to determine the

⁴ As King (2012) and others have argued, exact peerhood is rare. However, it is a mistake to think that as long as you are somewhat more reliable than some disputant of yours, their disagreement provides no relevant evidence calling for a reduction in confidence – or that ignoring their disagreement and relying on your own judgment will constitute a more reliable procedure than giving their verdict some weight. Rather, as long as they are better than chance, their disagreement does provide some evidence, and giving their verdict some weight will constitute a more reliable procedure than ignoring it. This is easy to show mathematically; see White (2009), as well as the Sachin and Rahul example below.

⁵ See also Chen & Worsnip (forthcoming).

overall rational level of confidence. What Kelly rejects is the aforementioned more *robust* form of conciliationism, which he objects to on the grounds that it washes out the significance of one's non-disagreement evidence. If one's non-disagreement evidence for p is very strong, he thinks, then the reduction of confidence that disagreement warrants might be quite small. But minimal conciliationism, once again, is compatible with this, since it maintains only that disagreement is *some* pro tanto evidence calling for *some* reduction in confidence. Similar points hold for the vast majority of other "steadfast" views.⁶ So despite the real controversies in the epistemology of disagreement literature, minimal conciliationism is something that almost everyone can sign up to.

Now, of course, the qualification in Minimal Conciliationism does allow that if the Stickers are justified in thinking that the Twisters are no better than (or worse than) chance when it comes to the kind of issue that their disagreement concerns, they might not need to conciliate with them. I accept this. However, since I am interested in the most interesting version of the case at hand, my response is to simply stipulate that the Stickers are *not* justified in thinking this. Lest this make the case seem too remote from real-world political disagreements, however, I will say something about how hard it is to satisfy the condition of being justified in thinking that one's political opponents are no better than chance.

Suppose that you think that your political opponents are *worse* than chance, that is, that they are *anti-reliable*. This commits you to thinking that, all else held fixed, their believing p is evidence *against* p . Thus, if you already believe p , and you discover that your opponent also believes p , you would have to then become *less* confident in p . If you attribute anti-reliability to a whole group – say 25% of the population – you're committed to thinking that p is *less* probable conditional on 100% of the population believing p than it is conditional on 75% of the population believing p and the group in question believing not- p . This is a *really* extreme commitment. Even in our highly polarized political climate, I suggest, we would be much more inclined, faced with some issue upon which *even our opponents* agree with us about p , to think that this is a sign of p 's obviousness or undeniability, rather than of its dubitability. This suggests that we rarely think (and that we rarely should think) that our opponents are actually worse than chance across a suitably representative range of political issues.⁷ Nor is it plausible that we think (or should think) that they are *precisely* as good as chance – which, again, would not justify taking their *agreement* as a sign of truth.

What can be confusing here is that even someone who is somewhat reliable in general might be wrong more (much) often than not *in cases of disagreement*. However, such a person isn't anti-reliable in the strong sense needed to establish that their disagreement does not call for some reduction in

⁶ Other examples of views that resist at least some classical elements of the conciliationist picture, but allow for disagreement to be *some* evidence warranting *some* reduction in credence, include Lackey (2010a, 2010b), Enoch (2010), and Worsnip (2014). The only papers I know of that deny that disagreement is *any evidence at all* are Kelly (2005) and Titelbaum (2015). As I've already made clear, Kelly abandons this view in his (2010). And similarly, Titelbaum takes back this very strong claim in his (2019).

⁷ There is a very tricky issue here for all accounts of disagreement about the level of generality at which our reliability should be typed. On the one hand, we want to allow that people can be reliable on *some* issues but not others, and not to simply measure reliability on *all* issues. On the other hand, if we individuate the issues too finely, it arguably becomes too easy to write off one's disputants on unreliable *only* on the precise matter at hand, and thus to avoid any conciliatory pressure. I have used the term "suitably representative" to stay vague on the best way to solve this issue.

confidence.⁸ To illustrate this, suppose, for simplicity, that there are just two individuals: Rahul, who is 90% reliable, and Sachin, who is 60% reliable – still better than chance, but a lot less reliable than Rahul. Sachin is right 60% of the time across all cases, many of which Rahul and Sachin will agree about. But *when Rahul and Sachin disagree*, Rahul is going to be right (and thus, Sachin is going to be wrong) the vast majority of the time (about 86%). That means that if you learn, all at once, that Rahul believes p and Sachin believes not- p , your confidence in p should go up. But for all that – and this is the crucial point – if you *already know* that Rahul believes p , and then you learn that Sachin believes not- p , your confidence in p should still go down. After all, the probability of p conditional just on Rahul believing p is 0.9, whereas the probability of p conditional on Rahul believing p and Sachin believing not- p is 0.86. Thus, learning of Sachin’s disagreement with Rahul still calls for a reduction in credence that Rahul is correct, *even though* when Rahul and Sachin disagree, Rahul is right the vast majority of the time.

Encountering disagreement with your own view is like this. You already know what your own view is, we are assuming, and then you discover that someone else disputes it. Now, you might think – or even be justified in thinking – that *when* you disagree with this person, they tend to be wrong and you tend to be right – that you are like Rahul and your disputant is like Sachin. But as the above mathematical illustration shows, this isn’t enough to show that you shouldn’t become at least somewhat less confident of your own view when you learn that your disputant disagrees. As long as your disputant is better than chance *over the full range of cases* – including not just those in which you disagree, but those in which you agree – you should become somewhat less confident in your view.

I suggest that many of our political opponents are such that, although we think they tend to be wrong about issues *about which we disagree*, we do not think they are actually anti-reliable (even about political matters) *overall*. This would explain why we don’t treat their agreement with us about political matters as reason to become less confident in our own views. But it’s also enough for *disagreement* with them to be some reason to become less confident in our own views. Thus, I don’t think that the assumption that the Stickers are better than chance – and thus, that their disagreement calls for some conciliation – is a far-fetched one.

I have said that minimal conciliationism shows that the Twisters should not stop conciliating with the Stickers. But it might be objected that this changes when the Twisters realize that the Stickers aren’t conciliating with *them*. There are two ways to interpret this suggestion. On the first, the idea is that this gives the Twisters an *epistemic* reason not to conciliate with the Stickers. On the second, the idea is that this gives the Twisters a *practical* reason not to conciliate with the Stickers. Let’s take those in turn.

Why would the discovery that the Stickers aren’t conciliating with the Twisters give the Twisters an epistemic reason not to conciliate? The most obvious suggestion is that it shows that the Stickers are not very reasonable people, and hence aren’t worth taking seriously.⁹ There are two problems with this suggestion. First, whether someone is good at appropriately responding to disagreement is at least somewhat independent of whether they are reliable in assessing what their

⁸ Contra, e.g., Hallsson (2019: 2197), who suggests that if it is “less likely that [my disputant’s] belief is correct [than incorrect] when we disagree,” there is “no motivation to adjust my view” on the basis of my disputant’s ability.

⁹ Georgi Gardiner, among others, suggested this to me.

(non-disagreement-constituted) evidence supports. Someone could be very good at the latter, and yet insufficiently humble to do the former well. Different virtues of mind seem to be involved in these two different abilities. Thus, finding out that someone is too intransigent in the face of disagreement needn't give us an extremely weighty reason to think them less reliable than we originally thought. Now, admittedly, the fact that the Stickers don't conciliate in response to disagreement may give us *some* reason to think that their assessments of the (non-disagreement-constituted) evidence are off: perhaps failure to conciliate correlates with a biased treatment of first-order evidence against one's prior views.¹⁰ And this would justify downgrading one's estimation of the Stickers' reliability *somewhat*. But this leads into the second, and more important, problem – namely, that even if the discovery that the Stickers are intransigent in the face of disagreement gives us reason to think them *somewhat* less reliable than we originally thought, there is no guarantee that it justifies us in going as far as thinking that they are no more reliable than chance. But unless this is so, minimal conciliationism will require us to conciliate with them at least somewhat.

What about the suggestion that learning of the Stickers' intransigence gives the Twisters a *practical* reason to do the same? Here the idea, presumably, is that by remaining steadfast, the Twisters can avoid the unfair outcome described in the setup of the case, where the political center of gravity systematically shifts toward the Stickers. Of course, this doesn't show that minimal conciliationism is false *qua epistemic norm*: at most, what it shows is that the Twisters have practical reasons to violate it. It's of course a matter of controversy how (if at all) these sorts of practical considerations interact with epistemic norms like minimal conciliationism to produce all-things-considered verdicts about what one ought to believe or do. Let's grant for the sake of argument, though, that a practical consideration such as this one could in principle override the Twisters' epistemic reasons to conciliate. Still, this rationale for remaining steadfast is not compelling. For, as I've already indicated and will go on to explore later, the Twisters don't need to remain steadfast (and break the epistemic norm of minimal conciliationism) in order to avoid the unfair outcome. Instead, they can conciliate in their beliefs, yet still *act* as if they had their old beliefs. Of course, this is ultimately a satisfactory response only if this alternative option can be defended. Later, I will be doing exactly that.

3. Second option: keep on doing what they're doing

The second option in response to the case of the Twisters and the Stickers, recall, is to say that the Twisters should simply keep on doing what they're doing. This view can be motivated by appeal to the conjunction of minimal conciliationism (which I've already argued for) with the following principle:

The Sincerity Norm. In the political arena, you ought to act (vote, participate, etc.) in accordance with your sincerely-held beliefs (or credences).¹¹

¹⁰ Thanks to an anonymous referee for this suggestion.

¹¹ For support for something like the sincerity norm, see, e.g., Schwartzman (2011: esp. 385); Landemore (2013: 154-6); Maskivker (2019: 80); van Wietmarschen (2019: 27). Cf. also Meehl (1977), Cohen (1989: 24); Rawls (1993: 241-2); Brennan & Lomasky (1993: esp. ch. 10), Wolgast (1994), Murphy (1998), Wodak (2019), and Ridge (2021: esp. 405).

Let's stipulate that the sincerity norm should be understood so that it says, for example, that in a referendum you should vote for the option you sincerely believe to be best (or that maximizes utility given your credences). Similarly, in an election, you should vote for the candidate to sincerely believe to be best (or whose election would, given your credences, maximize expected utility). In political discussion, you ought not to misrepresent your own views about what ought to be done. And so on.

Given this way of understanding the sincerity norm, it's easy to see how it combines with minimal conciliationism to suggest that the Twisters should keep on doing what they're doing. Minimal conciliationism tells them to moderate their confidence in their views. If they do this, they will end up with new, sincerely-held credences (and in some cases, new outright beliefs, or absences thereof). And the sincerity norm tells them to act (vote, participate, etc.) in accordance with these sincerely-held doxastic states. That is exactly what they are doing at present, as we told the story.

Some – call them political cynics – would reject the whole spirit behind the sincerity norm. They would suggest that, at least when their cause is righteous, citizens are justified in misrepresenting their views in order to secure their desired outcome. For example, they would say that quite generally (and not only when I'm facing intransigent opponents like the Stickers), I might be justified in representing my own views as much more extreme than they are, in attempt to engineer a “compromise” with others that, in reality, effects exactly the policy that I wanted from the start, and doesn't really grant anything to my opponents. Such cynics will be antecedently sympathetic to the view I'm ultimately going to argue for here – that the Twisters should vote in a way that doesn't reflect their moderation in confidence. However, I don't want to rest the case for this on a cynical view (which, for what it's worth, I don't share). Rather, I want to suggest that even if one is attracted to the spirit behind the sincerity norm (as I am), it should be qualified to make an exception in the special case of the Twisters and the Stickers.¹²

To begin to make the case for this, let's consider what might be *objectionable* about the cynical approach. One compelling thought is the kind of ploy mentioned in the previous paragraph constitutes a kind of *manipulation* of the political process, which is supposed to ensure equal representation of each person's view, in order to achieve an outcome where, in effect, one's opponent's view is neutralized in its effect, and one simply gets one's own way. This is *unfair*, because it violates the democratic principle that each person's view should have equal weight in determining the outcome of whatever decision-making we have entrusted to democratic control.

Now, as we'll see a bit later, I don't think that this principle is *quite* right as it stands. However, I do think it's plausible that the problem with the cynical approach is that it licenses behavior that violates constraints of fairness of broadly this kind. Now, the sincerity norm is plausible in part because it looks like typically, violations of it will also violate such constraints of fairness. If I act (vote,

¹² Let's also ignore objections to the sincerity norm that take issue with how it instructs even people with *highly irresponsible beliefs* to vote in accordance with them. This is a general objection to any norm that makes reference to actual attitudes that can be unjustified, and there are a range of well-known strategies for solving it: presenting the norm as a “subjective” one, modifying it to talk about *justified* attitudes, restricting it to with agents with justified attitudes, going “wide-scope” about the norm, etc. I won't try to arbitrate between these here. I take it that the core insight of the sincerity norm is that there is (at least typically) *something* definitively criticizable about political insincerity, which can be isolated in a way that brackets whether one's beliefs themselves are justified or not. Of course, acting on the basis of highly unjustified beliefs is *also* criticizable, but in a different way.

participate, etc.) out of accordance with my sincerely-held beliefs, I am subverting the process in which everyone's views count equally; and in general, I can be expected to do so in a way that tries to make my own influence *outsized* that of others.

But recall, now, the problem in the case of the Twisters and the Stickers. At the outset, I tried to bring out the intuitive sense in which the outcome in this case is itself *unfair*. It is unfair that the political center of gravity shifts toward the Stickers and away from the Twisters just because the former are intransigent in response to political disagreement and the latter are not. Crucially, note that here, it is *following* the sincerity norm that allows this unfair outcome to obtain. By violating the sincerity norm, then, the Twisters may be able to *restore* the fairness of the process.

What was objectionable about the across-the-board cynical approach was the way that it licensed breaking the sincerity norm *in order to get one's own way*. But this need not be what the Twister does when she violates the sincerity norm. Rather, she may violate it *in order to restore the fairness of the process*. This gives us a principled basis for explaining how and why the sincerity norm should be qualified to exclude the case of the Twisters.¹³ The suggestion is this: we should respect the sincerity norm *except when violating it is necessary to restore the fairness of the process*. And the claim is that this is so in the case of the Twisters and the Stickers.¹⁴

Some advocates of the sincerity norm may maintain that it remains wrong to misrepresent one's beliefs to one's fellow citizens, regardless of whether one does so in order to restore the fairness of the process. I will be better-placed to answer this objection when I've said more about my positive case for the third option, so I defer it until then.

4. Third option: acting out of accord with one's beliefs

I just suggested that the Twisters might be justified in violating the sincerity norm – thus, acting out of accord with their beliefs – if doing so is necessary to restore the fairness of the process. However, what I haven't yet explained is how this could be so. There are two questions to address here. First: how can the process in the case of the Twisters and the Stickers be unfair? After all, by hypothesis, it is one where everyone votes (and otherwise participates) on the basis of their beliefs, and each vote is counted equally. It is *prima facie* puzzling how this can constitute an unfair process. Second: how can the Twisters' voting and participating out of accordance with their beliefs – more particularly, voting in a way that doesn't reflect their moderation in confidence in response to the Stickers' disagreement – constitute a correction of the unfairness of this process? These are the questions I will aim to answer in this section.

¹³ Cf. Lister (2017) and Leland (2019), who both argue that Rawlsian requirements of public reason should be qualified to apply only when one can expect reciprocity from others.

¹⁴ This general approach may vindicate other exceptions to the sincerity norm too. For example, the sincerity norm seems to forbid "strategic voting" where one votes for a candidate that one thinks not to be best to avoid throwing away one's vote. But there are much stronger incentives to engage in such strategic voting in electoral systems that are *unfair* in the first instance, in that they make lots of peoples' votes *de facto* worthless. It may be that strategic voting (in violation of the sincerity norm) can be justified when it is needed to correct such unfair outcomes. Interestingly, Wodak (2019) argues, roughly, that what is objectionable about "plurality rule" (a.k.a. first-past-the-post) electoral systems is that they effectively force people to choose between violating the sincerity norm and having their vote (unfairly) carry no weight.

As a preliminary, note that the focus here is on the *fairness* of the democratic process, with respect to how it represents citizens' views.¹⁵ As many democratic theorists have urged, there are plausibly other constraints on the legitimacy of democratic system, beyond its representing citizens' viewpoints fairly.¹⁶ I do not mean to deny this; I simply assume that the fair representation of citizens' viewpoints is one crucial desideratum on a democratic process. For short, I will talk about whether a democratic process is 'representatively fair' to refer to whether the way in which it represents citizens' viewpoints is fair.

Consider, then, a precisification of the putative principle of democratic representation that I mentioned in the last section:

Equal Weight for Beliefs. A democratic process is representatively fair just if each person's beliefs [about what should be done] have equal weight in determining the outcome of whatever decision-making we have entrusted to democratic control.¹⁷

If this principle is correct, then it simply follows that the process in the case of the Twisters and the Stickers is representatively fair. It is a stipulation of the case that the Twisters' and the Stickers' beliefs (at the time of participation; that is, after the Twisters have conciliated) have equal weight in determining the outcome of the political process.

Yet, I submit, this is intuitively off. As I tried to bring out at the outset of this piece, there seems to be something unfair about a process that allows the political center of gravity to systematically shift toward the Stickers' views, simply because the Twisters (justifiably!) conciliate and the Stickers (unjustifiably!) don't.¹⁸ The case thus constitutes a problem for the *Equal Weight for Beliefs* principle. So it is worth seeing whether there are other, nearby, principles that can explain how the process is unfair in the case of the Twisters and the Stickers.

A first attempt might go as follows. In the case of the Twisters and the Stickers, the political center of gravity systematically shifts away from the Twisters' *original* beliefs – that is, the beliefs that they held prior to conciliating with the Stickers. So, one way of explaining the unfairness of the process in this case would be to posit the following principle:

¹⁵ I am assuming here that democracy aims to represent citizens' *views*, in a broad sense of 'views' where they are some sort of cognitive state in the neighborhood of – but, as we'll see later, not necessarily identical to – beliefs. As van Wietmarschen (2019: 26) notes, this "cognitivist" approach has a long history, including e.g. Wollheim (1962), Cohen (1986), Coleman & Ferejohn (1986), Estlund (1990), Beerbohm (2012), and Landmore (2013: ch. 8).

¹⁶ This is a common refrain in deliberative conceptions of democracy: see, among many others, Cohen (1989) and Gutmann & Thompson (1998). It's also a refrain among those who hold that the legitimacy of democracy is not purely procedural but also at least in part a function of its capacity to deliver substantively justified outcomes: see, again among many others, Anderson (2006), Estlund (2008), and Landmore (2013).

¹⁷ The principle that individuals ought to have equal influence in determining the outcome of democratic decision-making finds one canonical source in Rawls (1971: 221-8; for developments of this idea, see Brighouse 1996; Brighouse & Fleurbaey 2010; Krishnamurthy 2012, 2013; Kolodny 2014; cf. also May 1952: 681). Given a cognitivist interpretation of voting (which Rawls seems to share: see Rawls 1993: 251), and the point that we care about democratic representation beyond merely ensuring the opportunity to vote, it's plausible to develop this into a principle along the lines of *Equal Weight for Beliefs*. See also van Wietmarschen (2019: 32-35).

¹⁸ You might say it's not the process that's unfair, but rather the Stickers' behavior. But why is the Stickers' behavior unfair? If Equal Weight for Beliefs is the correct principle, the Stickers' behavior doesn't skew the outcome away from the representatively fair one, so what's the problem?

Equal Weight for Original Beliefs. A democratic process is representatively fair just if each person's *original* beliefs [about what should be done] – that is, the beliefs that they had prior to responding to any disagreement – have equal weight in determining the outcome of whatever decision-making we have entrusted to democratic control.

I think there is something right about the *Equal Weight for Original Beliefs* principle. However, at it stands, it has at least two problems. First, it looks hard to give a principled theoretical rationale for. Why would the beliefs that we care about representing fairly be ones that were held in the *past* but have now been abandoned, rather than those held at the time of participation? It seems hard to say why this would be so. Second, while the *Equal Weight for Original Beliefs* principle may apply neatly to the idealized case of the Twisters and the Stickers, it seems harder to apply to messier, real world cases. The case of the Twisters and the Stickers involves both groups forming their initial beliefs in isolation, then encountering each others' disagreement, and then revising (in the case of the Twisters) or not revising (in the case of the Stickers) their beliefs in response. But real world cases are rarely this neat. We form our initial beliefs about many political issues already aware of significant disagreement about the relevant subject matter. As such, for many real world issues, there simply *is* no “original” belief that one had prior to encountering disagreement, and as such the *Equal Weight for Original Beliefs* principle seems not to apply.

4.1 *Detour: Introducing the notion of a personal take*

To work up to something more adequate, I want to make some points – that I think are insufficiently stressed in the peer disagreement literature – about what exactly is involved in the kind of conciliation that disagreement calls for. Normally, when ordinary folks talk of changing one's mind in response to disagreement, they have in mind someone's being persuaded by, or at least coming to see the force of, the first-order *arguments* that the other side offers. However, this is *not* what's involved in the kind of conciliation that the peer disagreement literature focuses on. This literature addresses itself to cases where the two sides have exchanged their arguments, and each is unpersuaded by the arguments of the other side. The claim that the conciliationist (even what I earlier called the “minimal conciliationist”) makes is that when this happens to you, you are nevertheless required to reduce your confidence in your belief, in response to *the fact of disagreement as such*. That is: the fact that a well-informed and intelligent person (or group of people) disagree with you is *itself* a reason for you to (at least somewhat) reduce your confidence, even if their arguments do not seem persuasive to you. After all, as we saw in examining the case for minimal conciliationism, the fact that a well-informed and intelligent person (or group of people) believe not-*p* is at least some evidence that not-*p*.

The claim I now want to make is that there is a real psychological difference between the kind of reduction in confidence that is prompted by being persuaded (or at least moved) by someone's arguments and the kind of reduction in confidence that is (or can be) prompted by the fact of disagreement as such. Call the former ‘moderation by persuasion’ and the latter ‘moderation by disagreement-as-such’. The psychological difference, I think, is this. In cases of (pure) moderation by

disagreement-as-such, although your credence in your original view goes down, there is a sense in which your view still *seems* right to you, just as much as it ever did. Specifically, when you reflect just on the arguments and (first-order) evidence pertaining to the matter in question, they still strike you as supporting your original view. For example, suppose that you believe that euthanasia is morally wrong, but you then encounter disagreement with intelligent, well-informed people who believe that it is morally permissible. Suppose you find their arguments utterly unpersuasive. Still, as a good (minimal) conciliationist, you reduce your confidence that euthanasia is morally wrong at least somewhat. This reduction in confidence does not change the fact that, when you reflect on the arguments for the wrongness of euthanasia, they still strike you as very powerful; and when you reflect on the arguments for the permissibility of euthanasia, they still strike you as very weak. By contrast, this is not so in cases of moderation by persuasion.

In cases of moderation by disagreement-as-such, then, it seems that there is a kind of state of finding-plausible that you bear to the proposition ‘euthanasia is morally wrong’ that *persists* even after you have conciliated in response to disagreement. I will call this your “personal take”.¹⁹ We can understand your personal take on whether p as the belief (if any) that seems correct to you when you reflect on the first-order arguments and evidence for p , evaluating them for yourself and bracketing the disagreement of your disputants. (We can also define a graded notion of a personal take, corresponding to credence, in terms of *how* plausible p seems to you when you do this.) In cases of moderation by persuasion, your personal take changes after encountering disagreement; in cases of moderation by disagreement-as-such, it persists – even as your beliefs and credences may change.²⁰

4.2 *Back to representative fairness*

Like the notion of one’s “original” belief or credence (i.e. that which one had before encountering disagreement), one’s personal take can loosely be understood as one’s “pre-disagreement” belief or credence. The crucial difference, however, is that we no longer interpret this *temporally* – in terms of the belief or credence that one historically had before encountering disagreement – but rather in terms of the belief or credence that seems right to one when one *brackets* the fact of disagreement-as-such, and focuses solely on the arguments and (first-order) evidence at hand. Thus, the notion of a personal

¹⁹ Compare Barnett (2019) on the notion of “disagreement-independent inclination.”

²⁰ There are independent reasons to want the notion of a personal take for the epistemology of disagreement. For example, as Tal (2021) points out, if we understand conciliatory principles as requiring you to reach a credence intermediate between your pre-disagreement credence and your disputant’s, this yields the wrong prediction in cases where you’re (rationally) persuaded by your disputant’s arguments so as to come all the way over to their credence. The naïve way of interpreting the conciliatory view seems to suggest that this is impermissible, but surely it is permissible. But armed with the notion of a personal take, we can fix conciliationism to avoid Tal’s objection. The thought would be that what conciliationism really requires is a credence intermediate between your (graded) *personal take* and that of your disputant. It’s not your temporally pre-disagreement credence that matters, but your credence *bracketing* the fact of disagreement itself, i.e. your (graded) personal take. Those will be the same only when your disputant hasn’t actually persuaded you with her arguments. By contrast, in the case where you *have* been persuaded, your personal take has changed – indeed, if you’ve been *fully* persuaded, it is now identical to your disputant’s personal take. So, when this happens, conciliationism permits a credence that matches the personal take that you and your disputant (now) share – rather than forbidding it. I think there will be still other ways in which the notion of a personal take comes in handy beyond this one, too.

take looks like it could make for a good substitute for the notion of an original belief in our principle of representative fairness. This gives us:

Equal Weight for Personal Takes. A democratic process is representatively fair just if each person's (current) personal takes [about what should be done] have equal weight in determining the outcome of whatever decision-making we have entrusted to democratic control.

Like *Equal Weight for Original Beliefs*, *Equal Weight for Personal Takes* can explain why the process is unfair in the case of the Twisters and the Stickers. For, in this case, the personal takes of the Stickers get more weight than those of the Twisters. This is because the personal takes of the Stickers influence the political process *both* via influencing their own beliefs (and thus their own political participation) *and* via influencing the Twisters' beliefs (and thus their political participation). By contrast, the personal takes of the Twisters are not at all reflected in the Stickers' beliefs (and participation), and only partially reflected in their own beliefs (and participation). Consequently, the political center of gravity shifts away from the (enduring) personal takes of the Twisters and toward those of the Stickers.

At the same time, *Equal Weight for Personal Takes* makes progress in addressing both of the problems that I raised for *Equal Weight for Original Beliefs*. To take the second problem first, *Equal Weight for Personal Takes* does not apply only when one has an identifiable historical belief that one held about the matter at hand before ever encountering disagreement. All it requires is that there is something that strikes one at right when one reflects only the arguments and evidence bearing on the matter at hand, and brackets the fact of disagreement as such. So it will not be limited to being applied in idealized cases like that of the Twisters and the Stickers.

Turning now to the first problem, unlike one's original belief, one's personal take is not a state that one has *given up* by the point that one participates. Rather, it is a state that one still has at the time of participation. Thus, I think, there is less mystery about how it could matter (from the point of view of fairness) to represent people's personal takes equally than there is about how it could matter to represent people's original beliefs equally. Notice that the unfairness here can be appreciated *from the Twisters' point of view*, even after they have conciliated.²¹ Precisely because their personal take persists even after conciliation, a Twister who accepts *Equal Weight for Personal Takes* can diagnose the unfairness of the situation when their personal take isn't weighted equally.²² By contrast, it's unclear why a Twister should care about their *original belief* (per se) not being represented equally if they've now given up this belief.

Moreover, when personal takes and original beliefs come apart, it seems that *Equal Weight for Personal Takes* gets the right result and *Equal Weight for Original Beliefs* gets the wrong one. The cases where one's (current) personal take does not match one's original belief are those of moderation by *persuasion*. But when one has been genuinely *persuaded* by one's political interlocutors' arguments, it seems perverse to insist on representing one's *original*, pre-persuasion beliefs, rather than one's current

²¹ Thanks to an anonymous referee for prompting me to clarify this.

²² It's also worth noting that the unfairness can be appreciated by a *Sticker* who accepts *Equal Weight for Personal Takes*. Even if they consider their own intransigence reasonable, it remains true that the two sides' personal takes aren't getting equal weight, which violates *Equal Weight for Personal Takes*.

ones. *Equal Weight for Personal Takes* can accommodate this, since one's personal take changes when one is persuaded.

Thus, *Equal Weight for Personal Takes* seems to get the best of both worlds. Unlike *Equal Weight for Beliefs*, it can account for the fact that something unfair has happened in the case of the Twisters and the Stickers. Yet it overcomes the problems that attended *Equal Weight for Original Beliefs*. Still, someone might ask for a deeper explanation – beyond appeal to intuition about the case of the Twisters and the Stickers – of why it is personal takes, rather than all-things-considered beliefs, that we ought to be concerned with representing equally. A notion of a personal take may sound alien compared with that of a belief. But I think this is misleading. In a pluralistic, democratic society, we find ourselves facing complex normative issues on which there are a multiplicity of ways of evaluating and interpreting the same arguments and evidence – and we must reach a collective decision about what to do. *Equal Weight for Personal Takes* incorporates the idea that when making such a collective decision, everyone's judgment about how to evaluate the arguments and evidence should count equally.²³ When we conciliate in response to disagreement, we incorporate the judgments of others into our own (all-things-considered) beliefs. But when they don't reciprocate – as they may well not – representing everyone's beliefs equally will lead to some people's judgments being given more weight than others'. This problem is what motivates the switch from *Equal Weight for Beliefs* to *Equal Weight for Personal Takes*.

4.3 *How this justifies acting out of accord with one's beliefs (but in accord with one's personal takes)*

We can now pull this together to address how it can be justified for citizens to vote (and otherwise participate) out of accord with their beliefs under conditions of asymmetric compliance with conciliatory norms. As we have seen, when there is asymmetric compliance with conciliatory norms, a process whereby everyone votes (and otherwise participates) on the basis of their beliefs, and this participation is afforded equal weight in determining political outcomes, will result in a representatively unfair outcome (given *Equal Weight for Personal Takes*). In such a situation, one can correct (or contribute toward correcting) this unfair outcome by participating not on the basis of one's beliefs, but rather on the basis of one's personal takes. So we have vindicated the idea that doing this can be justified on fairness-based grounds.

As noted earlier, this requires violating the sincerity norm as stated. But we can now also explain why this kind of violation is very different from a violation of the sincerity norm that would be licensed by what I earlier called the "cynical approach".

First, such approaches license breaking the sincerity norm in order to engineer outcomes that are representatively *unfair*, in order to get one's own way. By contrast, the view I've argued for licenses breaking the sincerity norm in order to *prevent* such unfair outcomes, and engineer fair ones.

²³ Of course, there are probably some issues where we would not want this – where the judgment of experts ought to carry more weight. But these are issues that we would not want to subject to (direct) democratic control in the first place, and so that fall outside the scope of any of the principles we are considering. My own view is that this will be confined to descriptive questions (albeit perhaps normatively relevant ones). Essentially normative questions, by contrast, are such that for the purposes of public decision-making, no-one's judgment should count more than anyone else's. Of course, much more would need to be said to make out the full case for this.

Second, armed with the notion of a personal take, we can understand how someone who participates out of accord with their beliefs – but in accord with their personal takes – is not being as insincere or misrepresentative to their fellow citizens about their views as the cynic is. After all, the views that they are expressing in discussion, voting on the basis of, and so on *are* their authentic personal takes – a state that I’ve argued is a psychologically real one, reflecting *what strikes one as right* when considering only the arguments and (first-order) evidence bearing on the relevant question, and bracketing the fact of disagreement-as-such. In fact, you may well identify *more strongly* with your personal takes – the things that strike *you* as right when you look at the arguments – than with beliefs or credences that are tempered to reflect the judgments of others who disagree. For this reason, participating on the basis of one’s personal take, rather than one’s beliefs, also need not be experienced from within as requiring a kind of fighting against oneself or mental gymnastics, in the way that trying to act on the basis of views one does not believe might be thought to do.

Indeed, when it comes to at least one important kind of political participation – namely, engaging in public discourse and debate – I think that participating on the basis of one’s personal takes is already – independently of any special considerations about asymmetric compliance – extremely natural. When we engage in political debate with each other with an aim of reaching a consensus, we have a common interest in hearing the strongest case for each point of view being considered.²⁴ As a result, it makes sense for each of us to at least temporarily set aside doubts about our views occasioned by disagreement-as-such (and by other higher-order evidence), and simply put the best case for the view that seems right to us when we consider the arguments and (first-order) evidence. In other words, it makes sense for us to argue for our personal takes.²⁵

A slightly distinct worry to that about insincerity is that acting out of accord with one’s beliefs – and in line with one’s personal takes – amounts to a kind of akrasia.²⁶ Admittedly, the language of “acting out of accord with one’s beliefs” encourages this thought. However, I think that on closer inspection, no akrasia need be involved in the behavior I am recommending. Akrasia in the classical sense occurs when one believes that one ought to Φ , but does not intend to Φ .²⁷ Does the agent who acts on the basis of her personal take have attitudes of this pattern? I think not. Let’s take the example of voting. What we need here is the distinction between the candidate or policy that the agent believes to be best, and the candidate or policy that the agent believes that she ought to vote for. It is true that when the agent votes on the basis of her personal take, she may not vote for the candidate or policy that she believes to be best. However, under the particular conditions of asymmetric compliance, the agent has principled reasons (namely, the ones given in this section) to believe that she *ought* not to vote for the candidate or policy that she believes to be best, and to believe that instead, she ought to vote for the candidate or policy that is best according to her personal take. Thus, when she votes for

²⁴ Cf., e.g., Hallsson & Kappel (2020). They suggest that there is a problem for conciliationism here: by requiring everyone to moderate their views, it stultifies the kind of robust public disagreement and debate that helps us to make progress. But once we note the possibility of participating on the basis of personal takes *rather* than beliefs, this objection to conciliationism is defused: we can still conciliate in our beliefs, while participating on the basis of our personal takes.

²⁵ Cf., again, Barnett (2019), who argues for something like this with respect to *philosophical* disagreement.

²⁶ I thank an anonymous referee for pushing me to address this concern.

²⁷ See, e.g., Broome (2013).

the candidate or policy that is best according to her personal take, she is voting for the candidate or policy that she *believes* she ought to vote for, and hence is not akratic.²⁸

Indeed, there's a *sense* in which she doesn't act out of accord with her beliefs in this case. For while she votes against the candidate or policy she believes to be best (acting out of accord with her beliefs in a natural sense), she also *believes* that she has good reason to do this, given Equal Weight for Personal Takes. Thus, in a different sense, she acts in accordance with her beliefs.²⁹

Finally, it might be worried that if we can just *act* based on our personal takes rather than our beliefs, this takes all the sting out of conciliationism: we may need to revise our *beliefs*, but we needn't change anything about how we act. But here it should be remembered that the rationale for acting on the basis of one's personal take rather than one's beliefs only applies under the specific conditions of asymmetric compliance. So it continues to be the case that conciliationism has ramifications for how we should act across a wide variety of other cases.

5. Conclusion

The case of the Twisters and the Stickers is admittedly artificial. Nevertheless, the purpose of our investigation of it here has been to bring out some philosophically important points that apply even to our messier reality. Let me review the upshots that have emerged.

First, the case provides a particularly stark illustration of the point that that patterns of *asymmetric compliance* with the disagreement norms can – if everyone then votes on the basis of their beliefs – produce unfair outcomes. Though real-world cases are certainly more complicated, it nevertheless can be that different social and political groups react to disagreement in different ways, and that similarly unfair outcomes are produced.

Second, acknowledgment of the first point leads us to see that the principle *Equal Weight for Beliefs* is false. A process that gives equal weight to everyone's beliefs will not necessarily produce a representatively fair outcome.

Third, the first and second points give us a principled rationale for individuals to sometimes vote (or otherwise participate) against their beliefs – thus violating the sincerity norm – but one that is distinct from cynical attempts to manipulate the political process to yield their own desired outcomes.

Fourth, there is nevertheless a mental state that individuals can vote on the basis of in such cases – namely, their personal takes. The idea of a personal take can be independently motivated by appeal to the phenomenology of different ways in which one can change one's mind in response to

²⁸ What if one understands akrasia in terms of failing to do what one believes to be best? Even here, no akrasia need be involved. There is still a distinction between believing that X is the best candidate or policy and believing that voting for X is the best thing to do, and under conditions of asymmetric compliance, these can come apart. The agent does not vote for the candidate or policy that she believes to be the best, but it doesn't follow that in doing so, she fails to perform the action that she believes to be the best.

²⁹ There are obviously broader issues here about what it means in to act in or out of accordance with one's beliefs, which are well worth exploring on a future occasion. Thanks to an anonymous referee for prompting the clarifications about the sense in which the agent doesn't act out of accordance with her beliefs in this case.

disagreement. And personal takes are plausibly independently important for the epistemology of disagreement more broadly.³⁰

Fifth, it is plausible that we should replace *Equal Weight for Beliefs* with *Equal Weight for Personal Takes*. The latter principle can appeal much of the appeal of the former, while also explaining why we get unfair outcomes in cases of asymmetric compliance with the disagreement norms.

Cases of asymmetric compliance with the norms governing disagreement are interesting in part because they dramatize how epistemic and political norms can be tightly interlinked. It might initially seem that the requirement to conciliate is a purely epistemic one. Yet if I'm right that a failure to conciliate (on, for example, the Sticker's part) can produce unfair outcomes, then there may be a sense in which we *politically* owe to our co-citizens to conciliate with them.³¹ As with many other political norms, partial or asymmetric compliance with this norm can also create uniquely perverse outcomes, ones that need to be corrected in further ways. What I have argued here is that the way to make this correction, when faced with the noncompliance of others, is not to ourselves remain steadfast in our beliefs – not to sacrifice epistemic responsibility at the political altar – but rather to politically participate in ways that are in certain respects out of accord with our beliefs. While this may seem initially odd, that appearance can be mitigated once we appreciate the notion of a personal take – or so I have argued. If I am right, there is no need to choose between representative fairness and all the sincerity that political morality requires of us.

References

- Aikin, S., Harbour, M., Neufeld, J. & Talisse, R. (2010). "Epistemic Abstainers, Epistemic Martyrs, and Epistemic Converts," *Logos & Episteme*, 1/2: 211-219.
- Anderson, E. (2006). "The Epistemology of Democracy," *Episteme*, 3/1-2: 8-22.
- Barnett, Z. (2019). "Philosophy Without Belief," *Mind*, 128/509: 109-138.
- Beerbohm, E. (2012). *In Our Name: The Ethics of Democracy*. Princeton: Princeton University Press.
- Brennan, G. & Lomasky, L. (1993). *Democracy and Decision: The Pure Theory of Electoral Preference*. Cambridge, UK: Cambridge University Press.
- Brighouse, H. (1996). "Egalitarianism and Equal Availability of Political Influence," *Journal of Political Philosophy*, 4/2: 118-141.
- Brighouse, H. & Fleurbaey, M. (2010). "Democracy and Proportionality," *Journal of Political Philosophy*, 18/2: 137-155.
- Broome, J. (2013). *Rationality Through Reasoning*. Chichester: Wiley-Blackwell.
- Chen, Y. & Worsnip, A. (forthcoming). "Disagreement and Higher-Order Evidence," in Baghramian, Carter & Rowland (eds.), *The Routledge Handbook of the Philosophy of Disagreement*. London: Routledge.
- Cohen, J. (1986). "An Epistemic Conception of Democracy," *Ethics*, 97/1: 26-38.
- (1989). "Deliberation and Democratic Legitimacy," in Hamlin & Pettit (eds.), *The Good Polity: Normative Analysis of the State*, 17-34. Oxford: Blackwell.
- Coleman, J. & Ferejohn, J. (1986). "Democracy and Social Choice," *Ethics*, 97/1: 6-25.

³⁰ See n. 20 above.

³¹ Cf. Ebeling (2017).

- Ebeling, M. (2017). *Conciliatory Democracy: From Deliberation Toward a New Politics of Disagreement*. London: Palgrave Macmillan.
- Enoch, D. (2010). “Not Just a Truthometer: Taking Oneself Seriously (but not Too Seriously) in Cases of Peer Disagreement,” *Mind*, 119/476: 953–997.
- Estlund, D. (1990). “Democracy Without Preference,” *Philosophical Review*, 99/3: 397-423.
- (2008). *Democratic Authority: A Philosophical Framework*. Princeton: Princeton University Press.
- Gardiner, G. (2014). “The Commutativity of Evidence: A Problem for Conciliatory Views of Peer Disagreement,” *Episteme*, 11/1: 83-95.
- Gutmann, A. & Thompson, D. (1996). *Democracy and Disagreement*. Cambridge, MA: Harvard University Press.
- Hallsson, B. (2019). “The Epistemic Significance of Political Disagreement,” *Philosophical Studies*, 176/8: 2187-2202.
- Hallsson, B.G. & Kappel, K. (2020). “Disagreement and the Division of Epistemic Labor,” *Synthese*, 197/7: 2823-2847.
- Harman, E. (2015). “The Irrelevance of Moral Uncertainty,” *Oxford Studies in Metaethics*, 10: 53-79.
- Jackson, E.G. (2019). “Belief and Credence: Why the Attitude-Type Matters,” *Philosophical Studies*, 176/9: 2477-2496.
- Kelly, T. (2005). “The Epistemic Significance of Disagreement,” *Oxford Studies in Epistemology*, 1: 167–196.
- (2010). “Peer Disagreement and Higher-Order Evidence,” in Feldman & Warfield (eds.), *Disagreement*. Oxford: Oxford University Press.
- King, N.L. (2012). “Disagreement: What’s the Problem? Or A Good Peer is Hard to Find,” *Philosophy & Phenomenological Research*, 85/2: 249–272.
- Kolodny, N. (2014). “Rule Over None II: Social Equality and the Justification of Democracy,” *Philosophy & Public Affairs*, 42/4: 287-336.
- Krishnamurthy, M. (2012). “Reconceiving Rawls’s Arguments for Equal Political Liberty and its Fair Value: On Our Higher-Order Interests,” *Social Theory and Practice*, 38/2: 258-278.
- (2013). “Completing Rawls’s Arguments for Equal Political Liberty and its Fair Value: The Argument from Self-Respect,” *Canadian Journal of Philosophy*, 43/2: 179-205.
- Lackey, J. (2010a). “A Justificationist View of Disagreement’s Epistemic Significance,” in Haddock, Millar & Pritchard (eds.), *Social Epistemology*. Oxford: Oxford University Press.
- (2010b). “What Should We Do When We Disagree?,” *Oxford Studies in Epistemology*, 3: 274–293.
- Landmore, H. (2013). *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton: Princeton University Press.
- Leland, R.J. (2019). “Civic Friendship, Public Reason,” *Philosophy & Public Affairs*, 47/1: 72-103.
- Lister, A. (2017). “Public Reason and Reciprocity,” *Journal of Political Philosophy*, 25/2: 155-172.
- Lockhart, T. (2000). *Moral Uncertainty and its Consequences*. Oxford: Oxford University Press.
- MacAskill, W. (2014). *Normative Uncertainty*. DPhil dissertation, University of Oxford.
- Maskivker, J. (2019). *The Duty to Vote*. Oxford: Oxford University Press.

- May, K.O. (1952). "A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision," *Econometrica*, 20/4: 680-684.
- Meehl, P.E. (1977). "The Selfish Voter Paradox and the Thrown-Away Vote Argument," *American Political Science Review*, 71/1: 11-30.
- Murphy, A.R. (1998). "Rawls and a Shrinking Liberty of Conscience," *The Review of Politics*, 60/2: 247-276.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- (1993). *Political Liberalism*. New York: Columbia University Press.
- Ridge, M. (2021). "Voting for Less than the Best," *Journal of Political Philosophy*, 29/3: 404-426.
- Schwartzman, M. (2011). "The Sincerity of Public Reason," *Journal of Political Philosophy*, 19/4: 375-398.
- Sepielli, A. (2009). "What to Do When You Don't Know What To Do," *Oxford Studies in Metaethics*, 4: 5-28.
- (2017). "How Moral Uncertainty Can Be Both True and Interesting," *Oxford Studies in Normative Ethics*, 7: 98-116.
- Tal, E. (2021). "Disagreement and Easy Bootstrapping," *Episteme*, 18/1: 46-65.
- Titelbaum, M. (2015). "Rationality's Fixed Point (Or: In Defense of Right Reason)," *Oxford Studies in Epistemology*, 5: 253-294.
- (2019). "Return to Reason," in Skipper & Steglich-Petersen, *Higher-Order Evidence: New Essays*. Oxford: Oxford University Press.
- Weatherson, B. (2014). "Running Risks Morally," *Philosophical Studies*, 167/1: 141-163.
- White, R. (2009). "On Treating Oneself and Others as Thermometers," *Episteme*, 6/3: 233-250.
- van Wietmarschen, H. (2019). "Political Testimony," *Politics, Philosophy & Economics*, 18/1: 23-45.
- Wiland, E. (2017). "Peer Disagreement and the Dunning-Kruger Effect," *Episteme*, 14/4: 481-498.
- Wodak, D. (2019). "The Expressive Case Against Plurality Rule," *Journal of Political Philosophy*, 27/3: 363-387.
- Wolgast, E. (1994). "The Demands of Public Reason," *Columbia Law Review*, 94/6: 1936-1949.w
- Wollheim, R. (1962). "A Paradox in the Theory of Democracy," in Fishkin & Laslett (eds.), *Philosophy, Politics and Society*, 71-87. Oxford: Blackwell.
- Worsnip, A. (2014). "Disagreement about Disagreement? What Disagreement about Disagreement?" *Philosophers' Imprint*, 14/18.