Six Roles for Inclination

Abstract: Initially, you judge that *p*. You then learn that most experts disagree. All things considered, you believe that the experts are probably right. Still, *p* continues to seem right to you, in some sense. You don't yet see what, if anything, is wrong with your original reasoning. In such a case, we'll say that you are 'inclined' toward *p*. This paper explores various roles that this state of inclination can play, both within epistemology and more broadly. Specifically, it will be argued that: (*i*) inclinations can promote the accuracy of inquiring groups; (*ii*) they can support rational participation within philosophy despite pervasive disagreement; (*iii*) they allow us to make sense of an important way in which two people can continue to disagree even after they 'conciliate,' (*iv*) inclinations carry information about individuals' independent judgments and for this reason must be accounted for when updating on the opinions of others; (*v*) inclinations are connected to understanding in a way that belief is not; (*vi*) awareness of the inclination-belief distinction enables us to respond to a provocative challenge purporting to show that critical thinking, or 'thinking for oneself,' typically reduces expected accuracy and hence should be discouraged.

1   Introduction

Ash is a member of a five-player logic team. During a given contest, Ash and her four teammates are presented with various logic problems. The players are shown the same problems, but they are not supposed to collaborate or share answers. Once a player works through a problem, she must vote for either of two possible solutions. The team receives a point whenever three or more team members—a majority—vote correctly.

Imagine that Ash has just arrived at *p* after working through a challenging problem. Before she votes, she steals a glimpse of the answer sheet belonging to her star teammate, Violet. Ash is surprised to discover that Violet arrived at *not-p*. One of them must be wrong, Ash reasons, and given their respective track records, the culprit is more likely to be Ash herself. In light of this observation, how should Ash vote?

One thing to say, of course, is that Ash shouldn't cheat! But suppose we don't care about that. Suppose her team's success is all that counts. If her team's success is all that counts, you might assume that Ash should defer to her more reliable teammate. But this would be hasty. Given a natural way of specifying the example, it can be proven that Ash should actually vote for her original answer, *p.*

How can this be? Ash's vote will affect the outcome only when her vote breaks a tie: two for *p*, and two for *not-p* (including Violet). In this sort of case, since Ash also obtained *p* herself, the team would actually be split three-to-two in favor of *p*. And although Violet is individually more

1

reliable than Ash, it may be that, given a three-to-two split, the side with three is more likely to be correct. So by voting for her original answer, Ash is not voting for the answer likeliest to be correct, but she is voting for the answer likeliest to be correct *conditional on her vote's making a difference*.[1]

A few different lessons can be drawn from this observation. Some might find it surprising that one should not always vote for the proposition one regards as most likely to be correct—or to put it another way, that collective expected accuracy can be promoted at the expense of individual expected accuracy.[2] But the lessons we'll explore here pertain to the attitude Ash takes toward *p*, once she learns that Violet disagrees.

As we will see, Ash's attitude is best characterized as a conditional judgment: a judgment based on only some of her evidence. Conditional judgments of this type, which we'll call *inclinations*, have important roles to play in many different contexts. Some of the functions served by inclinations will be of interest within epistemology; others will have broader application.[3]

To preview these roles: First, as has been alluded to already, reliance on inclinations can promote the accuracy of truth-seeking voting groups.

---

[1] Many authors have discussed this general phenomenon, but this example is adapted from Barnett (2019, p. 116), who outlines the assumptions needed to deliver this result. If we assume, e.g., that Violet's reliability is 90%, that the reliability of all other players is 75%, and that the team members' initial answers are probabilistically independent of each other, the group's reliability will be ~93% when everyone 'votes for herself.' If instead everyone copies Violet, the group's reliability will simply be Violet's individual reliability, 90%. Hazlett (2016, pp. 134–141) discusses a vivid closely related example. The lessons he draws from the case are different, however: In effect, Hazlett suggests that the observation might give Ash reason to *believe* her original answer, *p*, to be correct. In contrast, the claim advanced here is that Ash has reason to *vote* for *p* while being more confident of *not-p*.

[2] For insightful discussion, see Mayo-Wilson, Zollman, and Danks (2011), Hazlett (2016), Dellsén (2018, 2020), and Skipper and Steglich-Petersen (2021). Relatedly, see Hong and Page (2004), and see Thompson (2014) and Singer (2019) for discussion of how to interpret their findings.

[3] Numerous authors have argued for a related, but importantly different, thesis. Specifically, some have argued that *non-deferential belief* is socially valuable, and that this fact gives us reason to reject deferential epistemology. See Elgin (2010, 2012, 2018), Hazlett (2016), Lougheed (2020, pp. 61–109), Dellsén (2020), Walker (forthcoming) for arguments along these lines. See Kraay (2019), Lougheed (2019), Matheson (2022, 2024*a*, 2024*b*), and Skipper (2023) for discussion. This paper can be seen as responding to this concern about deferential epistemology, arguing that the goods associated with 'thinking for oneself' are compatible with deferential epistemology. Non-deferential inclinations and deferential beliefs can coexist. See Nguyen (2019) for a defense of a related idea in the domain of aesthetic engagement.

Second, inclinations carry information about independent judgments of individuals and, for this reason, must be accounted for when aggregating and updating on the opinions of others. Third, inclinations can underpin sincere and sensible participation within philosophy despite widespread disagreement. Fourth, inclinations bear a tight connection to understanding in a way that ordinary belief does not. And fifth, awareness of the inclination-belief distinction enables us to respond to a recent challenge purporting to show that thinking critically, or 'thinking for oneself' typically reduces expected accuracy and hence should not be encouraged in education. After examining what inclinations are and how they can be useful to us, we will explore whether there are rational constraints on inclination, which will occasion us to consider a sixth role, pertaining to a phenomenon which might be called 'post-conciliation disagreement.'

2   What Is Inclination?

Let's return to the logic team example, for a moment. During downtime between contests, the players would sometimes discuss how to formulate the voting policy by which they should abide. At first, a simple and straightforward policy was proposed.

> Policy 1: Vote for the proposition you believe is correct.

But sometimes, they realized, a player might not have a firm belief one way or another. So they clarified the policy to invoke one's level of certainty or credence.

> Policy 2: Vote for the proposition of which you are most certain.

For a while, it stood as the team's official policy. But after reflecting on the case considered earlier, the players realized that this policy, too, was inadequate. In that case, Ash judged Violet's answer of *not-p* to be most likely, yet as we noted, she could have good reason to vote for *p*.[4]

In response, the team endeavored to find a better policy, one which would advise players to 'vote for themselves' rather than be swayed by the judgments of their (sometimes more reliable) teammates. The first

---

[4] Admittedly, in some circumstances, Ash really should vote for Violet's answer. For example, if Violet's reliability is far higher than Ash's (e.g. 99% vs. 60%), then it will be optimal for Ash to defer to Violet. We will explore some of these issues later on. For now, it's worth noting that, under a fairly realistic specification of the case (see fn. 1), it is best for Ash to vote for *p*, her original answer.

such suggestion was, in effect, to apply Policy 2—but at an earlier time, before any illicit answer-sharing took place.

> Policy 3: Vote for the proposition of which you were *initially* most certain.

This policy fares better; it handles the test case considered before. But upon further reflection, the team realized that this policy didn't quite capture what they were after either.

To see the difficulty with Policy 3, imagine that Ash saw Violet's answer *in advance*, before trying the problem herself. So Ash starts by peeking at *not-p* on Violet's page. Next, Ash works things through for herself, and to her surprise, derives *p*. Given this story, Ash would start out very confident of *not-p*. After arriving at *p* herself, Ash might become less confident of *not-p*, but still favor *not-p* as the most likely answer overall (since Violet is the superior logician). If things play out this way, Policy 3 would advise Ash to vote for *not-p*, the proposition of which she was initially most certain. But this is the wrong result. The same considerations which supported Ash's voting for *p* in the original case apply equally here. A different approach was needed.

Let's distinguish two types of evidence someone like Ash could have in logic team situations. On the one hand, there's the *testimonial evidence*—that is, the evidence provided by the opinions of Ash's teammates about whether *p*.[5] On the other hand, there's the *core evidence*—that is, the evidence provided by the logic problem itself, along with whatever else is relevant. In the cases we've considered, these two types of evidence pull Ash in opposite directions. When Ash thinks about the core evidence alone, she finds herself pulled toward *p*. When she thinks about all the evidence, especially the testimonial evidence from Violet, she is pulled toward *not-p*.

These observations enable us to outline a better voting policy.

> Policy 4: Vote for the proposition of which you are most certain, based on your own assessment of the core evidence alone.

Policy 4 issues the right advice in both of our problematic cases. Whether Ash learns of Violet's answer before or after confronting the problem

---

[5] 'Testimonial' is arguably a misleading term. Arguably, it is odd to say that Ash receives *testimony* from Violet. After all, Ash looked at Violet's answer page without her knowledge, so it's not clear that Violet has *testified* to anything—at least not willingly. In any event, the label is not important here. It will suffice to think of 'testimonial evidence' as consisting in the opinions of others.

herself, Policy 4 will advise Ash to vote for 'her own' answer of *p*. So the team adopted Policy 4 and lived happily ever after.

For us, the story doesn't end here. The particular attitude the team identified—judgment based on one's own assessment of the core evidence alone—has other noteworthy uses. As was previewed earlier, we can use the term *inclinations* to refer to these sorts of judgments.[6] And as we'll soon see, inclinations are relevant in a variety of contexts—when jurors hear a trial, when scientists study the climate, or even when students of philosophy grapple with the mind-body problem. Thinking closely about inclination will enable us to learn several valuable lessons, ranging from the epistemological to the pedagogical.

3   Preliminary Worries

Before putting inclination to work, though, we should be more critical about what, exactly, inclination is. Ash and her friends didn't provide us with a very precise characterization of this epistemic state. Their policy recommends voting on the basis of 'one's own assessment of the core evidence.' But what is *that*? Is it simply a belief about what the core evidence supports? Is it a counterfactual belief—a belief one would have had, if the core evidence were one's only evidence? Or is it something else entirely? Next, we'll consider a few preliminary worries one might have about inclination, which will ultimately help to answer the preceding questions.

3.1 *Inclination as Ordinary Belief?*

> Worry 1: You want to distinguish inclination and belief. But why isn't one's inclination just a certain kind of belief—namely, *a belief about what the core evidence supports*?

It is tempting to think that inclination can be understood as nothing more than an ordinary belief about what the core evidence supports. But this

---

[6] Others have discussed nearby attitudes. Sliwa and Horowitz (2015) refer to a person's 'first-order attitude,' Schoenfield (2015) uses the term '*judgment*,' Skipper and Steglich-Petersen (2021) use the term '*first-order judgment*,' Worsnip (2023) refers to a '*personal take*,' and Skipper (2023) refers to what one '*judges to be the case, independently of other people's influence*.' These authors are discussing something in the vicinity of what I'm calling inclinations. (And the two most recent pieces cited above make the attitude a focal point, though they emphasize distinct roles from the six identified here.) I favor the term 'inclination' because I find that it best captures the phenomenology: a person may find herself inclined toward *p*, but ultimately she must decide whether to trust this inclination or defer to the judgment of others.

temptation should be resisted. In the original scenario, Ash's 'inclination' is supposed to be *p*. Nonetheless, Ash may (and perhaps should) believe that the core evidence probably supports *not-p*. Why? Because Ash can reason as follows:

> *This is a logic problem. So the core evidence (i.e. the problem itself) certainly supports the correct answer, whatever it is. Since my star teammate, Violet, arrived at* not-p, *the correct answer is probably* not-p. *Thus, the core evidence probably supports* not-p (*even though I'm getting* p *when I try it myself*).

If Ash reasons this way, her belief about what the core evidence supports (*not-p*) will not align with the judgment she obtains, when she assesses the core evidence herself (*p*). So inclination should not be analyzed as an ordinary belief about evidential support relations.

3.2 Inclination as Counterfactual Belief?

> Worry 2: Okay, I concede that inclinations aren't just beliefs about what the core evidence supports. But now I'm thinking that inclinations can be analyzed counterfactually: Your inclination is what you *would* believe, if the core evidence *were* all you had.

This counterfactual analysis comes closer to capturing the target phenomenon. But it can go wrong. Suppose that Ash's team clarified their voting policy along the lines suggested above.

> Counterfactual Policy: Vote for the answer that you would have obtained, if the core evidence were the only evidence you had.

This policy won't always give the right advice. To see this, let's further develop the logic team scenario.

Suppose that Ash always looks at the answers of her teammates except when she is the only team member present, in which case she experiences severe anxiety and comes up with nothing. Relative to this background, the counterfactual policy would advise Ash to leave her answer sheet blank. After all, Ash would be able to reason as follows:

> *If the core evidence were all I had, that would mean I'm solving alone. And on days when I solve problems alone, I get anxious and come up with nothing. So I guess the policy recommends that I leave this question unanswered.*[7]

Of course, Ash should not leave the question unanswered. Instead, she should vote for the answer she obtained herself, *p*. So inclinations should not be analyzed counterfactually—nor should the team's voting policy be formulated that way.

3.3 Inclination as Rational Counterfactual Belief?

> Worry 3: I see what's wrong with the counterfactual approach. But can it be repaired by imposing a 'no-irrationality' clause? Imagine that the team's voting policy advises Ash to vote for the answer she would have obtained if (*i*) the core evidence were the only evidence Ash had, and (*ii*) Ash reasoned perfectly. Wouldn't this modified policy be immune from the problems facing the simpler counterfactual view?

Yes, it would. But it faces a different problem. In effect, the modified policy recommends that Ash simply vote for whichever proposition the core evidence in fact supports. (After all, if Ash had the core evidence and reasoned perfectly from it, then, of course, she would obtain the answer that this evidence supports.) So then the proposal would enjoin Ash to vote for whichever proposition she thinks that the evidence supports. This proposal turns out to be equivalent to the one addressed in §3.1. And as we saw, this proposal leads Ash to defer to Violet in the original scenario, and so it doesn't track what we're after.

3.4 Inclination as Conditional Judgment

> Worry 4: I agree that the previous proposals are inadequate. So what is the right view, then? Are you saying that inclination is a genuinely new attitude that can't be understood in terms of something more familiar?

No, certainly not. Inclination is not meant to be fundamentally new. Inclinations are a species of *conditional judgments*—judgments based on evidence which differs from the evidence the agent in fact has.

Typically, conditional judgments involve making an additional stipulation or supposition: '*Supposing Harris wins Ohio,* she will win the presidency.' But there are also *inverse* suppositions, which involve hypothetically *removing* something from the current stock of evidence. For example, a trial jury may be instructed to set aside a piece of evidence in its deliberations—if, say, video footage were collected without a search warrant. In such a case, jurors would have to make a judgment based upon the admissible evidence alone (video aside), and this judgment could differ from what they believe, all things considered (video

included). Inclination is similar. Ash makes a judgment based on the core evidence alone (testimony aside), and this judgment can differ from what she believes, all things considered (testimony included).

That's the rough idea. But it is difficult to provide a precise characterization of 'inverse conditional beliefs.' Exactly how to characterize them may depend upon how we characterize beliefs, more generally.

For illustration, let's assume the crude view that beliefs are simply betting dispositions. Under this assumption, conditional beliefs are simply dispositions to take certain *conditional bets*—bets where money changes hands only if a specified condition is met. For example, I might bet on Harris's winning the presidency conditional on her winning Ohio. Under this bet, if Harris loses Ohio, nothing happens; if Harris wins Ohio, however, then there will be a payout—I will win money if Harris wins the election; I will lose money if she loses the election.

Inclinations, then, on the betting account of belief, would be analyzed as dispositions to take certain conditional bets—bets where money would only change hands if somehow the testimonial evidence were lost. In the logic team scenario, Ash's inclination that p would correspond to a disposition to accept a conditional bet on p, a bet which triggers a payout only if the evidence provided by Violet's disagreement is lost.[8]

Admittedly, the idea of losing evidence is somewhat obscure and controversial within epistemology. But there is some precedent for it in discussions of the problem of old evidence[9] and the Sleeping Beauty Paradox.[10] Up to now, the idea of inverse conditional judgment has been a

---

[8] This sounds similar to the counterfactual account discussed earlier. And it is similar. But it's subtly different. Whereas the counterfactual picture understands inclinations as what one would be disposed to bet on if the testimonial evidence were lost, the current proposal understands inclinations as what one is disposed to bet on *now*, if offered a bet that goes into effect only if the testimonial evidence is lost. So the commitment is one that the agent possesses now, rather than being on the agent is counterfactually disposed to acquire.

[9] Traditionally, many analyze evidential support in terms of probability-raising. That is, many hold that we can estimate just how strongly a piece of evidence E supports a proposition p by comparing the current probability of p with the probability of p conditional on gaining E. (If gaining E causes p's probability to rise, then E supports p.) But what if E is already part of one's evidence? Here, the natural thing to do is to compare the current probability of p with the probability of p conditional on *losing E*. (If losing E causes p's probability to fall, then E supports p.) To say this, of course, we need to be able to make sense of judgment conditional on losing evidence. See Glymour (1980), Howson (1991), and Joyce (1999) for discussion.

[10] Michael Titelbaum's textbook (2022, p. 390-391) proposes a rational constraint

relatively minor player in epistemology, featuring occasionally in a handful of mostly technical debates. But we'll soon see how inclination—which is a certain type of inverse conditional judgment—turns out to be useful in several unappreciated ways.

3.5 Changing One's Inclination

> Worry 5: Won't the inverse conditional judgment proposal run into some problems of its own? Specifically, I'm wondering what happens when disagreement prompts an agent to revise some of her background beliefs, leading to a change in the agent's verdict on the question at hand. If I'm understanding correctly, this process would also result in a change to the agent's inclinations. Isn't that the opposite of what is desired? Inclinations were supposed to be disagreement-resistant!

The worrier is right that disagreement can, in principle, cause a person's inclinations to change. But the worrier is wrong that it will happen whenever disagreement changes a person's beliefs via a change in their background beliefs. Disagreement can only change a person's inclinations when it causes her to revise her original reasoning in a way that is epistemically independent of the disagreement. A pair of cases will help to illustrate how this works.

We'll start with a case where disagreement changes Ash's background beliefs, but not her inclination.

> Tautology: Imagine that Ash's team is asked whether a proposition p is a tautology. Ash builds a truth-table and finds that p is true on every line. She is confident that her truth-table is accurate. What she's less sure of is the meaning of the word *tautology*. Her recollection is that *tautology* means *logical truth* and so believes p to be a tautology. But when Ash sees that Violet disagrees, she decides that her recollection is likely mistaken, and that *tautology* is more likely a synonym for *contradiction*. What is Ash's inclination now?[11]

One might conclude that Ash's inclination has changed. After all, Ash now thinks that *tautology* and *contradiction* are probably synonyms. So when she does the reasoning herself, she sees that p is true on every line of its truth-table. And so she can obtain—herself—the result that p is likely

---

[11] Thanks to an anonymous reviewer at *Mind* for suggesting a similar example.

not a tautology. Isn't this Ash's new inclination?

I don't think so. I believe that Ash should still judge *p* to be a tautology upon setting Violet's disagreement aside. Setting aside evidence is not mere subtraction. If the proposition being set aside was playing a justificatory role in supporting other beliefs, then when that proposition is removed from the stock of evidence, anything it was supporting should fall along with it. So in the example at hand, when Ash sets aside Violet's disagreement, she thereby sets aside the only evidence she had that *tautology* and *contradiction* were synonyms. All that remains is her original recollection, which suggests that a tautology is a logical truth. This is why her inclination remains that *p* is indeed a tautology.

But now let's turn to a different sort of case—one where Violet's disagreement does initiate a change to Ash's inclinations.

> Double-checking: Ash obtains *p*. She then sees that Violet obtained *not-p*. This prompts Ash to double-check her reasoning, at which point she discovers a mistake. She then works through the problem again from beginning to end, avoiding the mistake, and deriving an answer of *not-p*.

Here, Ash's inclination—her own assessment of the core evidence—has changed. After double-checking, Ash derives *not-p* herself in a way that is entirely independent of Violet's disagreement. Violet's disagreement may have initiated the double-checking process. But Ash is arriving at *not-p* on the basis of the core evidence alone, and her coming to this answer is in no way founded, evidentially, on Violet's belief.

Note the asymmetry between the two cases. In both, Ash comes to believe (or becomes more confident of) Violet's answer. But in only one case does Ash's justification for her new belief refer essentially to the testimonial evidence. In the Tautology case, Ash's belief changes because she comes to suspect that *tautology* is a synonym for *contradiction*. And what justifies this suspicion? Only Violet's disagreement. (After all, Ash's own recollection is the opposite.) But in the double-checking case just considered, Ash's justification for her new belief that *not-p* need make no reference to Violet. Of course, Violet's disagreement is undoubtedly what *caused* Ash's belief to change. But what matters is whether the testimonial evidence features essentially in Ash's justification, not whether the testimonial evidence played a causal role in the acquisition of the judgment.

3.6 What's Special about Inclination?

> Worry 6: I understand the view better now. For any hypothetical body of evidence, we can speak of a person's judgment conditional on it. We can talk about a person's judgments setting aside the testimonial evidence; we can also talk about a person's judgments setting aside evidence from Wikipedia, or setting aside evidence acquired on Tuesdays. But now I fear there's nothing special about inclinations that warrants their being singled out. Is 'inclination' a distinctive mental state? Or is it just a name for one of the countless conditional judgments a person has?

It's more the latter. But singling out inclination is still warranted. One reason it's warranted is that inclination can play a variety of useful roles, as we'll see shortly. But first, I should emphasize: inclinations can also be psychologically salient—a lot more so than, say, your typical conditional judgment.

> Answer Key: You're doing a math problem. You get the answer 3. You check the back of the book. It says that the answer is 4. You double- and triple-check your work. You're still getting 3. You can't see why the answer should be 4. But probably, you conclude, there's an error in your reasoning. Probably, the answer is 4.

This is yet another example of inclination coming apart from belief. Although you believe 4 is probably correct, you're inclined toward 3. When you do the reasoning yourself (setting aside what the book says), you're getting 3, and as a result, this answer seems to be correct in a psychologically gripping sense.

Fundamentally, an inclination is the result of 'your own' reasoning about some question, whether the topic is math, the climate, or the mind-body problem. Next, we'll explore how recognizing inclination can prove fruitful in a variety of philosophical and everyday contexts. Five applications are proposed below (and a sixth is proposed in §5). The following suggestions are offered in a preliminary spirit and, to my mind, warrant further investigation.

4   Applications for Inclination

4.1 Truth-seeking Voting Groups

The logic team case with which we began illustrates our first role. Members of truth-seeking voting groups (such as juries, admissions committees, research panels, and others) may have reason to vote for their inclinations rather than their beliefs when the two attitudes diverge. We have already seen the key idea in action, but it is worth slowing down to justify the claim more fully and to orient the observation with respect to a

background of related insights.

In 1785, Marquis de Condorcet showed that, given certain assumptions, the larger a jury is, the more reliable it is. Which assumptions were needed? First, Condorcet assumed that the jurors were assessing a *binary proposition*, and that the jury would use *majority rule* to reach its verdict. He also assumed that each juror was *individually reliable*—that is, he assumed that each juror had a better than 50% chance of voting for the correct verdict. Finally, and most importantly, he assumed that the jurors' votes were *probabilistically independent* of one another. This sort of independence requires that, for any two jurors, the fact that one voted correctly does not change the probability that the other did the same.

In effect, Condorcet's framework models each juror as a *biased coin*—one that is biased toward voting for the correct verdict. Understood this way, Condorcet's result makes sense. Suppose we have a large collection of coins, all biased toward Heads—each lands Heads ⅔ of the time, say. If we flip only a few of them, a majority of those flipped might well land Tails. But if we flip many such coins, a majority-Tails outcome is far less likely. The more tosses there are, the likelier it is that a majority will land Heads. Analogously, in the case of a jury, provided that the relevant assumptions are in place, the more jurors there are, the likelier it is that a majority will vote for the correct verdict.

The foregoing story helps to illustrate not only why Condorcet's jury theorem is true, but also why voting for one's inclinations—rather than one's beliefs—can be valuable. A notable corollary of Condorcet's jury theorem is that, under the specified conditions, *a group can outperform its most reliable individual member*. (So the logic team may together be more reliable than Violet is alone, even if Violet is the best individual performer.) When the relevant conditions are met, it would clearly be a mistake for a group simply to defer to its best individual performer—such deference would undermine the synergistic effect Condorcet described. When everyone votes for their inclinations rather than their beliefs, the condition of independence is enforced, and Condorcet's synergy remains in effect.

I don't mean to give the impression, though, that the use of inclinations in voting contexts *guarantees* independence of the desired sort. It doesn't. Whether voting for one's inclinations will promote independence (and thereby promote group accuracy) will depend upon the particular features of the case. In the logic team example, to determine whether there is anything to be gained by Ash's voting for her inclination

(rather than simply copying Violet's answer), we would need to assess the frequency with which their errors tend to coincide. (In effect, this tells us how frequently Ash's inclination serves as a 'check' on Violet.) If their errors do tend to coincide (e.g. because Violet only errs on certain advanced problems that Ash does not know how to solve), then Ash is better off simply copying Violet. On the other hand, if their errors do not tend to coincide especially often (e.g. because both players make unavoidable performance errors from time to time, which can happen on any given problem), then Ash's voting for her inclination will foster independence and will be synergistic.

Condorcet's jury theorem is an important result. It asserts that, in a certain sense, a group is epistemically stronger than the best of its parts. But in order to reap this collective strength, a certain form of independence is needed. Voting for one's inclinations sometimes fosters this kind of independence, and so members of truth-seeking voting groups may have reason to vote for their inclinations, rather than their beliefs, when the two attitudes diverge.

This is not to say that everyone serving on an admissions committee or a research panel should always vote for her inclinations. Perhaps in some cases, we should simply identify the most reliable individual member and let them make all of the decisions. But in reasonable and realistic circumstances, voting for one's inclinations rather than one's beliefs can promote collective accuracy, and this fact is worth appreciating.

4.2 Updating on Other's Opinions

A second role concerns judgment aggregation. The thought here is that attention must be paid to inclination when updating on the opinions of others.

In their spare time, Ash and Violet study medicine. Although Violet is the superior logician, when it comes to diagnosing patients, both are very reliable—and equally so. Ash is more humble than Violet, however: When their medical judgments conflict, Violet always continues believing her initial diagnosis, but Ash always suspends her belief.

Today, the two friends see the same patient and arrive at conflicting diagnoses: Ash arrives at *p*; Violet arrives at *not-p*. But upon sharing their assessments, in typical fashion, only Ash suspends her belief.

Given the above information, what should the patient think? Hopefully not the following:

I consulted two equally competent doctors today. One is now agnostic. The other

now believes not-p. *So I suppose* not-p *is the best guess about my condition.*

This is a poor inference. The trouble is that Ash's revised opinion takes account of Violet's judgment, but not vice versa. So in effect, the patient is double-counting Violet's judgment but only single-counting Ash's. For this reason, when updating on outside opinion, we shouldn't simply aggregate the experts' *beliefs*.

A better approach is to aggregate inclinations instead. Ash's inclination is *p*; Violet's inclination is *not-p*. Since both sources are equally competent, neither proposition is favored. And more generally, since inclinations are insensitive to testimony, this approach will be immune to the 'double-counting' issues that this case exemplifies.[12]

Some might object, however, that the advice outlined above—*aggregate inclinations, not beliefs*—isn't always feasible to implement. After all, in many realistic cases, we might know a person's beliefs without knowing their inclinations.

This is a fair point. Even in the medical example considered above, if the patient only learned the respective beliefs of the two doctors at the end of the story, then the patient would be unable to equally weigh the two doctors' respective inclinations, as was advised.

In response, there are a few points worth making, which show that it is still useful to be aware of these issues.

Suppose that you are a doctor, and a patient comes to you seeking a second opinion. If the patient plans simply to weigh your judgment equally against the first doctor's judgment, then it may make sense for you to communicate your inclination rather than your belief, if the two come apart. Otherwise, the first doctor's judgment could end up exercising outsized influence, in shaping the patient's outlook.

Alternatively, suppose that you are a patient seeking a second opinion from a doctor. If you plan to share the first doctor's judgment with the second doctor, it may make sense to you to ask this second doctor to communicate their *independent assessment* (i.e. their inclination). That is the point of a second opinion, after all.

In saying all of this, I don't mean to suggest that most doctors/

---

[12] This is not to suggest that we should simply compute the average of the sources' inclinations. This suggestion doesn't work. A full treatment of updating on the opinions of others would need to take into account the inclinations of the different sources, their respective levels of reliability on the proposition in question, and the degree of probabilistic dependence between their inclinations. See Easwaran et al (2016), Leitgeb (2017), and Gallow (2018) for some recent perspectives on this issue.

patients are typically making these mistakes (though it may happen sometimes). Rather, the aim is to make explicit something that we may implicitly understand already: namely, that when compiling and weighing the judgments of multiple sources, the judgments should be as independent as possible from each other to avoid contaminating the results. And because inclinations are, in a certain sense, independent of others' influence, they are better positioned than beliefs to serve as inputs to the sort of aggregation procedures we've been thinking about. In short, making these observations explicit helps to guard against potential mistakes and fosters a deeper understanding of the epistemology of opinion aggregation.[13]

4.3 Practicing Philosophy

The third role for inclination concerns the practice of philosophy itself. The basic idea is that—just as Ash might sensibly vote for her inclination (*p*) rather than the answer she deems most likely in light of Violet's disagreement (*not-p*)—in philosophy, we should think of ourselves as identifying with and defending our philosophical inclinations rather than the views we deem most likely in light of agreement and disagreement from other philosophers.[14] An example will help to motivate the idea.

I've thought about Newcomb's Problem many times, as carefully as I can. When I reflect on the imagined predicament and all its variations, I find myself more compelled by the case for one-boxing. When I confront arguments for two-boxing, objections spring to mind. At the same time, I know that many in philosophy favor two-boxing—indeed, many of the philosophers and decision theorists I respect most are in this opposing camp. These facts give me significant pause. In light of this disagreement, my all-things-considered judgment is that there's a pretty high chance I'm wrong (though I'd also say there's still a decent chance I'm right). So in short, I have a one-boxing inclination, but due to disagreement, I'm not confident, all things considered, of this view. If I had to bet, I'd probably

---

[13] Does any of this rest on controversial presuppositions about the epistemology of disagreement? No. If two doctors have access to evidence that you lack, all sides of the disagreement debate will agree that your opinion should be informed by weighing the doctors' opinions. And I've argued that inclinations can be helpful to do this. (What *is* controversial is whether the doctors' opinions should still count if you have access to the same evidence that they possess. The preceding section has been neutral on this question.)

[14] Barnett (2019) sketches a version of this idea. For related discussion, see DeRose (2017), Feldman and Conee (2018), Fleisher (2018, 2021*a*), Palmira (2020), and Barnett (forthcoming).

bet on the truth of two-boxing.

It should be clear, I hope, that despite my disagreement-driven doubts, I still should be classified as a one-boxer, in a familiar sense. One-boxing seems right to me, when I think about the question directly. That's also the view I can best defend. Admittedly, my credence in one-boxing is not very high—probably below ½.[15] But my inclination toward one-boxing remains intact. More generally, the views we identify with and defend in philosophy need not be the views of which we are most confident, all things considered. In my own case, what seems to matter is not my all-things-considered credence, but rather, my own assessment of the relevant issues—or in other words, my inclination.

I don't mean to suggest, here, that philosophers do not typically believe their views. I know that many do. Maybe most do. Instead, the thought is just that, when inclinations and beliefs come apart—that is, when disagreement undermines belief but leaves inclination intact (as it does for me, with respect to Newcomb's Problem)—it makes sense to identify one's philosophical view with one's inclination, which can be passionately and sincerely held.

4.3.1 The Puzzle of Philosophical Testimony?

One virtue of this proposal is that it potentially helps us answer what Christopher Ranali has called *the puzzle of philosophical testimony*.[16] Most philosophers are compatibilists about free will.[17] But it would clearly be out of bounds to write a paper defending compatibilism by citing its popularity. On reflection, it isn't immediately obvious why this should be. If the goal of a philosophy paper is to get the reader to *believe* its conclusion, then it's not obvious why a certain kind of evidence—testimonial evidence—should be categorically banned. This is the puzzle.

If we think of inclination, rather than belief, as being the type of commitment underlying our views, we can answer the puzzle. The goal of a philosophical argument is not to get the reader to *believe* its conclusion, but rather, to *incline the reader toward its conclusion*. On this picture, the norm against testimonial evidence is precisely what we'd expect. After all, learning that a given philosopher believes (or does not believe) a given

---

[15] Goldberg (2013*b*) suggests that one's philosophical views are the views one regards as more probable than not. The example discussed above puts pressure on this suggestion.

[16] See Allen (2019) and Ranali (2020).

[17] See Bourget and Chalmers (2014, 2023).

philosophical thesis cannot by itself change one's inclination about that very thesis.[18] And notably, this aim within philosophy of trying to *incline* one's audience toward one's preferred view does not solely apply to written work—it extends to presentations, conversations, or even to a philosopher's own personal evaluation of the various views.

4.3.2   The Self-Undermining Problem for Conciliationism

I've suggested that philosophically legitimate arguments aim to *incline* the audience toward a particular view. And we've seen how this proposal can help to explain why certain arguments from popularity (e.g. 'compatibilism is popular; therefore compatibilism is likely true') are philosophically illegitimate. But interestingly, the proposed picture may rule out a range of arguments that some have earnestly advanced. Is this a problem for the picture I'm proposing? Or are these authors violating an accepted norm? Let's look.

The example I have in mind is the *self-undermining problem*[19] for conciliationism, originally due to Alvin Plantinga (1999).[20] According to conciliationism, the fact that a trustworthy person disagrees with you counts as some evidence that you are wrong—even if this person has the same evidence you have, and even if your original judgment was justified and correct. This is a substantive epistemological thesis, and it's not overwhelmingly popular—there are respected epistemologists on both sides of the issue. As a result, some have observed that, by their own lights, conciliationists should reduce confidence in their view. More pointedly, it can be argued that, at present, conciliationism recommends that we not believe *it*! In other words: conciliationism *self-undermines*.

This objection strikes many as forceful. But upon closer inspection, it runs afoul of the prohibition canvassed earlier. The objection from self-undermining just is an argument from popularity. Earlier, it was argued that a philosophically legitimate argument aims to incline its

---

[18] Though see §6 for one potential complication.

[19] Two distinct objections to conciliationism fall under this heading, as Christensen (2020) and Fleisher (2021*b*) point out. The first alleges that conciliationism can't presently be rationally believed by its adherents. This is essentially Plantinga's worry (though Plantinga's discussion predated the term 'conciliationism'), and it's the one we'll discuss below. The second objection alleges that conciliationism issues incompatible demands and so is incoherent. See Bogardus (2009), Elga (2010), Pittard (2015), Christensen (2020), and Fleisher (2021*b*) for various attempts to solve this theoretical problem.

[20] In addition to Plantinga, Sampson (2018) and Dixon (2021) have appealed to the self-undermining objection in response to various challenges from disagreement. See Fleisher (2021*b*) for a response to this line of objection.

audience, rather than to engender belief. Arguments from popularity cannot do this. They can only engender belief. And the argument above is a textbook case of an argument from popularity.

To be clear, 'arguments from popularity' can still be sound arguments. Indeed, the argument presented above could still be sound. The argument may succeed in showing that conciliationists working within anglophone philosophy in the year 2023 are not rational to be confident of their view. But if—as I've argued—'having a view' in philosophy is a matter of inclination rather than belief, then the argument should not occasion any conciliationist to change their view. This argument is purely about the import of the testimonial evidence, and inclination is a function of how one evaluates the core evidence. If a person was inclined toward conciliationism before confronting this argument, they will almost certainly continue to be so after confronting it.

Admittedly, it's still a noteworthy fact that, if conciliationism is true, few controversial philosophical theses can rationally be believed by their adherents.[21] But this fact should not be seen as threatening to those views on their first-order merits. And the same can be said about conciliationism itself.[22]

4.4 Connections to Understanding
Fourth, inclination seems to be intimately related to understanding in a way that ordinary belief is not.

4.4.1 Might Accurate Inclinations Be Necessary for Understanding?
There is controversy surrounding the nature of the relationship between testimony and understanding. While some have assumed that testimony does not typically confer understanding, others have called this

---

[21] For various arguments to this effect, see Goldberg (2009, 2013*a*), Brennan (2010), Fumerton (2010), Kornblith (2010, 2013), Licon (2012), Christensen (2014).

[22] Fleisher's (2021*b*) defense of conciliationism from the problem of self-undermining is similar to that presented above. Fleisher argues that we need not believe our controversial views in philosophy; rather, we need only *endorse* them, which roughly amounts to committing to act as if they're true, for the good of the enterprise of which we are a part. Since one need not believe a view to make such a commitment, the problem of self-undermining won't prevent any conciliationists from endorsing their view. In this way, Fleisher's defense is fairly limited. He shows only that conciliationists can consistently accept conciliationism for the purposes of inquiry. I'm asserting something more contentious: the objection from self-undermining violates a norm against arguments from popularity and is therefore no better than an argument that makes the case for a given philosophical view—such as compatibilism or the correspondence theory of truth—on the basis of an observed consensus among practitioners in the field.

assumption into question.[23] But most will agree that in many cases, someone can come to be justified in believing a proposition via testimony without understanding why the proposition is true.[24] This fact suggests that inclination and understanding may be closely related.

The expert consensus in economics, I'm told, is that it can be beneficial for nations to run large trade deficits. This has always struck me as odd. When I reflect on the question, it seems to me that, just as an individual household should avoid falling into debt, so too should the United States. But economists have examined this issue far more carefully than I, and they are largely in agreement. Apparently, my untutored inclination is incorrect. That's okay—all things considered, I believe the economists are right: trade deficits can be beneficial.

This belief of mine may well be true. It may well be justified. It may even constitute knowledge. But I lack understanding. If indeed trade deficits can be beneficial, I don't understand why this is the case. Why not? Well, for one, my own consideration of the issues (my inclination) is thoroughly confused. When I think through the economic question myself, there's something I'm missing or not fully appreciating, which inclines me to the wrong conclusion. This error could certainly be corrected—if, say, an economist were to walk me patiently through the prevailing rationale step by step. If this happened, then perhaps I would come to understand—but this process would be liable to change my inclination as well.[25] In light of these observations, it seems worth exploring further whether having an accurate inclination is necessary for understanding. At the very least, accurate inclination seems more plausible as a necessary condition on understanding than is true belief.

4.4.2 Might Epistemically Perfect Inclinations Be Sufficient for Understanding?

We've seen why having a poor inclination may be incompatible with understanding. But there may be reason to think that something stronger

---

[23] For some examples of those who reject this assumption see Boyd (2017, 2020), Croce (2020), and ch. 7 of Greco (2020).

[24] Many authors have observed that, in some sense, 'thinking for oneself'—rather than simply deferring to authorities—can sometimes be required for understanding. For relevant discussion, see Zagzebski (2007), Hills (2009), Sliwa (2012), Hazlett (2016), Nguyen (2018), and Matheson (2022, 2024a, 2024b).

[25] Why? My inclination—my judgment based on the core evidence alone—is insensitive to the economic consensus about the proposition in question, but it is sensitive to other information, sometimes conveyed via testimony, such as a comprehensive first-order explanation proffered by an economist.

is true—that having an *epistemically perfect* inclination could actually be sufficient for understanding. An example will help to motivate this thought.[26]

> Cicadas: Asked why a certain species of cicada emerges for their mating season every 13 years, Ash proposes the following: 'Cicada predators tend to emerge more regularly—every 2, 3, or 4 years. Cicada populations with prime periods tend to avoid coinciding with their predators. So this trait was selected for.' In fact, Ash's explanation is correct. But she learns that Violet, her esteemed colleague, disagrees. Violet tells Ash that she discovered an even better explanation of the data (but doesn't say what that alternate explanation is). As a result, Ash suspends belief about her proposal, but it remains the best explanation she can think of.

Ash doesn't have knowledge or justified belief in this case—after all, she doesn't even believe her theory to be true. All she has, it seems, is an epistemically perfect inclination.[27] But arguably, Ash does seem to have a certain kind of understanding here—or, at least, an important ingredient of understanding.[28]

To defend this idea, it will be useful to draw on Allison Hills' (2016) account of understanding. Hills analyzes understanding as a certain sort of *know-how*. She argues that understanding why something is true brings with it certain abilities. Roughly, she suggests that, if a person understands why *p*, then she'll be able to...

- ○ ... explain why *p* obtains.
- ○ ... defend the explanation from objections.
- ○ ... refute proposed alternative explanations of *p*.

The details of Hills' account can be debated. In particular, it is controversial whether we should see these abilities as constitutive of understanding, or as merely correlated with understanding. But for our

---

[26] This example is adapted from the literature on mathematical explanation in science—see Baker (2009).

[27] Finnur Dellsén (2017) proposes an interesting alternative example of understanding without justification or belief. In Dellsén's example, a researcher does not believe her theory, nor is she inclined to think it is true, but she accepts it in Cohen's (1992) sense. To my mind, the case is most forceful if the agent is inclined toward a true theory for the right reasons and only stops short of believing the theory on account of disagreement. So I agree with Dellsén, and I think the appeal to inclinations dramatically strengthens his case.

[28] On top of being true, we can suppose that Ash's inclination is 'epistemically perfect' in that it manifests an ideally rational response to the core evidence.

purposes, it suffices to note that someone with an epistemically perfect inclination would be likely to possess the foregoing abilities.

One might ask: Why think that Ash would possess these abilities? After all, Ash lacks belief in her proposal, and this lack of belief could well preclude her performing the stated tasks. If someone asked whether anyone could explain why these cicadas have prime periods, we would not expect Ash to volunteer.

In response, though, it should be noted that it is possible for a person to have abilities she believes she lacks. If Ash had not ridden a bike in many years, she might come to believe that she no longer had the ability. Accordingly, she might sincerely report that she could not ride a bike if asked. Still, this is compatible with her having the ability to ride a bike (if, say, it were true that she would succeed if she tried). In the cicada example, something analogous could well be true. If Ash were to try to defend the prime-number hypothesis, it's plausible that she'd succeed, despite her lack of belief.

None of this is meant to show that Ash definitively *has* the relevant abilities; rather, it shows only that Ash's lack of belief is not, by itself, a decisive indication that she lacks the abilities outlined by Hills. And this indicates that the relationship between inclination and understanding deserves further scrutiny.

4.5 *Defending Critical Thinking*

In a provocative paper, Michael Huemer (2005) argues that critical thinking is 'epistemically irresponsible.' Huemer suggests that, although it is often held up as an intellectual ideal, critical thinking (or 'thinking for yourself') actually reduces one's accuracy and should not generally be encouraged. In its place, *deference to the appropriate experts* is proposed as a more accuracy-conducive ideal. (Of course, identifying those experts requires at least some critical thought, and Huemer recognizes this.[29]) Awareness of the inclination-belief distinction can help us to see what's right and wrong with Huemer's argument, and it can offer us a more nuanced way of thinking about accuracy, critical thinking, and their relationship.

Huemer's argument is straightforward enough: When people think for themselves, they often get the wrong answers. If they trusted experts instead, they would more often get the right answers. And it's better to be right than wrong.

---

[29] See Goldman (2001), Anderson (2011), Brennan (2020).

There is clearly something right about Huemer's argument. When it comes to climate change or the efficacy of vaccines, many are too willing to trust their own judgment over that of the experts. But shouldn't it be possible to advocate critical thinking in a way that doesn't encourage people to trust themselves over the experts? The inclination-belief distinction turns out to be useful here.

Here's one way to develop the thought, as it might function in a classroom setting. Students could be taught the difference between a *personal inclination*, and a *final opinion*. Accordingly, students would be encouraged to exercise critical thinking in developing and maintaining their own personal inclinations, by looking directly at the core evidence surrounding issues of interest to them. At the same time, students would be reminded that their own inclinations are not always going to be perfectly accurate. And in light of this, students could be taught to base their final opinions mostly on what the experts (the people who have studied the issues the hardest) believe. These two norms are clearly compatible and can be taught side by side.[30]

A skeptic of critical thinking might ask, though: Why encourage students to maintain their own inclinations? What is gained by promoting this (relatively) inaccurate form of thought? What good are inclinations, if they're less accurate than the student's own deference-based beliefs?

In response, I can offer some brief and speculative suggestions. Why encourage students to think 'critically' in developing their own inclinations? Doing so can foster understanding; it can build one's argumentative skills; it can enable students to explain things to others without having to appeal to authority; and it puts students in a better position to have a well-supported view in case they are not able to rely on expert testimony. In most of one's life, one can't simply rely on expert opinion, so it's important to be a good thinker, not just a good testimony-receiver. For these reasons, while we should defer to experts when expert testimony is available, there is still ample reason to value and encourage critical thinking among our students and one another.

5   Normative Constraints on Inclination

So far, we have seen five roles that inclinations are well suited to play. No intricate argumentation or philosophical gymnastics were required to

---

[30] This point is related to Matheson's (2022, 2024*a*, 2024*b*) observation that the virtues of epistemic autonomy and intellectual humility need not be seen as rivals—it is possible to possess and cultivate both virtues at once. Indeed, the distinction between inclination and belief may actually be useful in explaining just how these virtues can be compatible.

make the case for the usefulness of inclination in these different contexts. The reader may not agree that inclination is up to the task of playing all five of these roles in exactly the manner I've set forth. But the discussion makes it apparent that the inclination/belief distinction is a useful one, both within philosophy and more broadly. Next, we'll examine a sixth role for inclinations, which will invite careful reflection on what it means for an inclination to be rational.

Suppose that Ash is solving logic problems yet again, this time with her equally competent friend Jet. Ash and Jet encounter a difficult problem and obtain different answers: Ash obtains q, and Jet obtains *not-q*. Upon learning of their disagreement, both friends reduce confidence in their respective judgments, ultimately arriving at the same credence in q: ½.

Do Ash and Jet now agree? In a sense, they do. If we asked them to *pronounce* on q, or to *guess* whether q, or to *bet* on q, they'd feel equally ambivalent. Still, there is something akin to disagreement which might persist here. This can be illustrated in a few ways.

Despite their having equal credence in q, there likely are phenomenological differences: It might still seem to Ash that q; and it might seem to Jet that *not-q*. Ash might have before her mind the reasoning that led to her original conclusion, and likewise for Jet. If they were to discuss the problem, Ash would likely have something to say on behalf of q, and Jet would likely have something to say on behalf of *not-q*. These facts are plausibly explained by the fact that Ash and Jet harbor divergent inclinations—despite having equal credences. So inclinations can help us to account for the sense in which something akin to a genuine disagreement can persist, even when the 'disagreeing' parties have the very same credence.

The payoff is not purely descriptive, though. Suppose that, in fact, Ash's original reasoning was perfect, while Jet made a mistake. Once they share their answers and converge on the same credence in q (½), it might be thought that, rationally speaking, the earlier asymmetry washes out. But even after converging, Ash and Jet still harbor divergent inclinations. And plausibly, their respective inclinations are not equally rational. If Ash is inclined toward q for the right reasons, while Jet is inclined toward *not-q* on the basis of a mistake, then Ash's inclination toward q could be rational, while Jet's inclination toward *not-q* would not be.

Some might hesitate to regard Jet as less than fully rational merely for having a mistaken inclination. Are inclinations really rationally assessable in a manner similar to beliefs? Yes.

One way to illustrate the irrationality of Jet's inclination is to think about how Jet's beliefs would change, if—for whatever reason—Jet stopped trusting Ash. If, for example, Jet became convinced that his friend's answer resulted from a random guess, then, he would resume believing his original inclination of *not-q*—the wrong answer. This is an irrational disposition. And notably, it's a disposition that Ash lacks. If Ash became convinced that Jet had randomly guessed, then presumably, she would resume believing her inclination of *q*, the correct answer. So there is a rational asymmetry between Ash and Jet, and this asymmetry can be explained by appeal to the (ir)rationality of their inclinations.[31]

These observations are closely related to what Adam Elga (2007) labeled the *problem of self-trust*.

> Suppose that a great fraction of those you count as your epistemic peers agree with you on some issue. Then the equal weight view says: stick with your initial assessment. Great! Except that the reason for sticking to that assessment has very little to do with *your own evaluation of the common stock of evidence*, and very much to do with the fraction of your peers who agree with you. Shouldn't *your own careful consideration of the issue* count for more than 1/100th, even if there are 99 people you count as epistemic peers? If not, then one might just as well form views on controversial matters by simply sending out a poll... It is implausible that rationality requires you to give *your own consideration of the issue* such a minor role. (484–485, my emphasis)

The role carved out for inclination here is clear. If I'm one of 100 experts on an issue, my own consideration of that issue (i.e. my inclination) can seem to get washed out by the overall consensus. If they all independently arrive at *p*, then I should be confident of *p*, regardless of what my inclination may be. But this doesn't mean that my own judgment disappears, or that it loses its normative significance. According to the picture advanced in this section, when I examine an issue directly, rationality does require me to be inclined toward the best-supported view—even if, at the end of the day, I decline to trust that inclination. So in short, even when we are required to defer to others, we needn't worry that our own thinking falls out of the picture. It is still there, and it is still subject to rational requirements.

6   Conclusion

---

[31] The asymmetry described above is a counterfactual one. Alternatively, we could appeal to conditional bets that Jet would (now) be disposed to accept: for example, he might accept a bet on *not-q* conditional on losing Ash's disagreement.

We will conclude with a final challenging case. Suppose that you are a researcher studying how a certain historical figure should be interpreted. You interpret him as having believed *p*—an interpretation of his work not shared by your peers. Though you find their disagreement disconcerting, you still harbor a strong inclination toward this bold interpretation, which explains your continued identification with and advocacy for this view.

As it happens, one of your reasons for attributing *p* to this historical figure is that his contemporaries seem to have thought he believed it. But as you reflect upon this fact, it occurs to you that the opinions of the historical figure's contemporaries seem to constitute testimonial evidence. After all, his contemporaries are just other people, other people who happen to share your interpretation. If that's right, then—given the picture developed in this paper—it would seem to follow that the opinions of this historical figure's contemporaries should be set aside when doing history of philosophy. But that looks implausible. Isn't it legitimate to take into account the opinions of the historical figure's contemporaries?

Yes it is. This is clearly legitimate—and particularly so if the figure's contemporaries are believed to have possessed evidence we lack today, such as direct conversations with the figure in question. This case shows that, in doing history of philosophy at least, testimonial evidence shouldn't always be set aside. And although this critique is focused on the 'practicing philosophy' role, I don't doubt that challenging cases can be cooked up for other roles as well.

In response, it should be said that the goal of the preceding sections was not to advance six fully developed theories, each of which asserts that inclination is both necessary and sufficient for something. Rather, the more modest goal was to show that there are six areas where inclinations seem relevant and useful—and indeed more useful than belief. While challenging cases, such as the history of philosophy example above, do show that more work is needed to properly describe the relationship between inclination and some of the roles proposed, these cases do not undermine the paper's more modest objective.

The overall lesson can be summarized perhaps less contentiously, without referring to 'inclination' at all. It can be beneficial for members of voting bodies to set aside the opinions of their peers in arriving at their vote. In philosophy, it can be helpful to set aside the popularity of a view in assessing its first-order merits. And in thinking critically about the controversial issues of the day, it can be worthwhile to set aside the expert consensus in arriving at your own personal take—for the purpose of

honing your argumentative skills and for gaining a deeper understanding of the world. In short, it can be useful to set aside the opinions of others, so as to isolate the component of your thinking that is distinctively and independently your own. And regardless of what we call it, I'm inclined to think that this pattern is worth recognizing. [32]

References

Allen, Keith (2019): 'Should We Believe Philosophical Claims on Testimony?' *Proceedings of the Aristotelian Society* 119: pp. 105–125.

Anderson, Elizabeth (2011): 'Democracy, Public Policy, and Lay Assessments of Scientific Testimony,' *Episteme* 8: pp. 144–164.

Baker, Alan (2009): 'Mathematical Explanations in Science,' *British Journal for the Philosophy of Science* 60: pp. 611–633.

Barnett, Zach (2019): 'Philosophy Without Belief,' *Mind* 128: pp. 109–138.

— (forthcoming): 'It Seems to Me I'm Right, But Probably I'm Not: Philosophical Partisanship with Humility,' in *Attitude in Philosophy* (eds. S. Goldberg and M. Walker). Oxford University Press.

Bogardus, Tomas (2009): 'A Vindication of the Equal-Weight View,' *Episteme* 6: pp. 324–335.

Bourget, David and David Chalmers (2014): 'What Do Philosophers Believe?' *Philosophical Studies* 170: pp. 465–500.

— (2023): 'Philosophers on Philosophy: The 2020 PhilPapers Survey,' *Philosophers' Imprint* 23: pp. 1–53.

Boyd, Kenneth (2017): 'Testifying Understanding,' *Episteme* 14: pp. 103–127.

— (2020): 'Moral Understanding and Cooperative Testimony,' *Canadian Journal of Philosophy* 50: pp. 18–33.

Brennan, Jason (2010): 'Skepticism about Philosophy,' *Ratio* 23: pp. 1–16.

Brennan, Johnny (2020): 'Can Novices Trust Themselves to Choose Trustworthy Experts? Reasons for (Reserved) Optimism,' *Social Epistemology* 34: pp. 227–240.

Christensen, David (1999): 'Measuring Confirmation,' *Journal of Philosophy* 96: pp. 437–461.

Christensen, David (2014): 'Disagreement and Public Controversy,' in *Essays in Collective Epistemology* (ed. J. Lackey). Oxford University Press: pp. 143–163.

Christensen, David (2020): 'Akratic (Epistemic) Modesty,' *Philosophical Studies* 178: pp. 2191–2214.

Cohen, L. Jonathan (1992): *An Essay on Belief and Acceptance.* Clarendon Press, Oxford.

Croce, Michel (2020): 'Moral Understanding, Testimony, and Moral Exemplarity,' *Ethical Theory and Moral Practice* 23: pp. 373–389.

Dellsén, Finnur (2017): 'Understanding without Justification or Belief,' *Ratio* 30: pp. 239–254.

— (2018): 'When Expert Disagreement Supports the Consensus,' *Australasian Journal of Philosophy* 96: pp. 142–156.

— (2020): 'The Epistemic Value of Expert Autonomy,' *Philosophy and Phenomenological Research* 100: pp. 344–361.

DeRose, Keith (2017): *The Appearance of Ignorance: Knowledge, Skepticism, and Context*, Vol. 2. Oxford University Press.

Dixon, Jonathan (2021): 'Moral Disagreement Scepticism Leveled,' *Ratio* 34: pp. 203–216.

Easwaran, Kenny, Luke Fenton-Glynn, Christopher Hitchcock and Joel Velasco (2016): 'Updating on the Credences of Others: Agreement, Disagreement, and Synergy,' *Philosophers' Imprint* 16: pp. 1–39.

Elga, Adam (2007): 'Reflection and Disagreement,' *Noûs* 41: pp. 478–502.

— (2010): 'How to Disagree about How to Disagree,' in *Disagreement* (eds. R. Feldman and T. Warfield). Oxford University Press.

Elgin, Catherine (2010): 'Persistent Disagreement,' in *Disagreement* (eds. R. Feldman & T. Warfield). Oxford University Press: pp. 53–68.

— (2012): 'Begging to Differ,' *The Philosopher's Magazine*. Fourth Quarter: pp. 77–82.

— (2018): 'Reasonable Disagreement,' in *Voicing Dissent: The Ethics and Epistemology of Making Disagreement Public* (ed. C. R. Johnson). Routledge: pp. 10–21.

Feldman, Richard and Earl Conee (2018): 'Between Belief and Disbelief,' *Believing in Accordance with the Evidence: New Essays on Evidentialism* (ed. K. McCain). Springer Verlag: pp. 71–89.

Fitelson, Branden (2000): 'A Bayesian Account of Independent Evidence with Applications,' *Philosophy of Science* 68: pp. 123–140.

Fleisher, Will (2018): 'Rational Endorsement,' *Philosophical Studies* 175: pp. 2649–2675.

— (2021a): 'Endorsement and Assertion,' *Noûs* 55: pp. 363–384.

— (2021b): 'How to Endorse Conciliationism,' *Synthese* 198: pp. 9913–9939.

Fumerton, Richard (2010): 'You Can't Trust a Philosopher,' in *Disagreement* (eds. R. Feldman and T. Warfield): pp. 91–111.

Gallow, Dmitri (2018): 'No One Can Serve Two Epistemic Masters,' *Journal of Philosophy* 175: pp. 2389–2398.

Glymour, Clark (1980): *Theory and Evidence*. Princeton University Press.

Goldberg, Sanford (2009): 'Reliabilism in Philosophy,' *Philosophical Studies* 124: pp. 105–117.

— (2013a): 'Disagreement, Defeat, and Assertion,' in *Disagreement: New Essays* (eds. D. Christensen and J. Lackey). Oxford University Press: pp. 167–189.

— (2013b): 'Defending Philosophy in the Face of Systematic Disagreement,' in *Disagreement and Skepticism* (ed. D. Machuca). Routledge: pp. 277–294.

Goldman, Alvin (2007): 'Experts: Which Ones Should You Trust?,' *Philosophy and Phenomenological Research* 63: pp. 185–185.

Greco, John (2020): *The Transmission of Knowledge*. Cambridge University Press.

Hazlett, Allan (2016): 'The Social Value of Non-Deferential Belief,' *Australasian Journal of Philosophy* 94: pp. 131–151.

Hills, Allison (2009): 'Moral Testimony and Moral Epistemology,' *Ethics* 120: pp. 94–127.

— (2016): 'Understanding Why,' *Noûs* 49: pp. 661–688.

Hong, Lu and Scott Page (2004): 'Groups of Diverse Problem Solvers Can Outperform Groups of High-ability Problem Solvers,' *Proceedings of the National Academy of Sciences* 101: pp. 16385–16389.

Howson, Colin (1991): 'The "Old Evidence" Problem,' *British Journal for the Philosophy of Science* 42: pp. 547–555.

Huemer, Michael (2005): 'Is Critical Thinking Epistemically Responsible?' *Metaphilosophy* 36: pp. 522–531.

Joyce, James (1999): *The Foundations of Causal Decision Theory*. Cambridge University Press.

Kornblith, Hilary (2010): 'Belief in the Face of Controversy,' in *Disagreement* (eds. R. Feldman and T. Warfield): pp. 29–52.

— (2013): 'Is Philosophical Knowledge Possible?,' in *Disagreement and Skepticism* (ed. D. Machuca). Routledge: pp. 260–276.

Kraay, Klaas (2019): 'Elgin's Community-oriented Steadfastness,' *Synthese* 198: pp. 4985–5008.

Leitgeb, Hannes (2017): 'Imaging All the People,' *Episteme* 14: pp. 463–479.

Licon, Jimmy Alfonso (2012): 'Skeptical Thoughts on Philosophical Expertise,' *Logos and Episteme* 3: pp. 449–458.

Lougheed, Kirk (2019): 'Catherine Elgin on Peerhood and the Epistemic Benefits of Disagreement,' *Synthese* 198: pp. 3183–3202.

— (2020): *The Epistemic Benefits of Disagreement*. Springer.

Matheson, Jonathan (2022): 'The Virtue of Epistemic Autonomy,' in *Epistemic Autonomy* (eds. J. Matheson and K. Lougheed). Routledge: pp. 173–194.

— (2024a): 'Why Think for Yourself?' *Episteme* 21: pp. 320–338.

— (2024b): 'Epistemic Autonomy and Intellectual Humility: Mutually Supporting Virtues,' *Social Epistemology* 38: pp. 318–330.

Mayo-Wilson, Conor, Kevin J. S. Zollman, and David Danks (2011): 'The Independence Thesis: When Individual and Social Epistemology Diverge,' *Philosophy of Science* 78: pp. 653–677.

Nguyen, C. Thi (2019): 'Autonomy and Aesthetic Engagement,' *Mind* 516: pp. 1127–1156.

Palmira, Michele (2020): 'Inquiry and the Doxastic Attitudes,' *Synthese* 197: pp. 4947–4973.

Pittard, John (2015): 'Resolute Conciliationism,' *The Philosophical Quarterly* 65: pp. 442–463.

Plantinga, Alvin (1999): 'Pluralism: A Defense of Religious Exclusivism,' in *The Philosophical Challenge of Religious Diversity* (eds. K. Meeker and P. Quinn). Oxford University Press: pp. 172–192.

Ranalli, Christopher (2020): 'The Puzzle of Philosophical Testimony,' *European Journal of Philosophy* 28: pp. 142–163.

Sampson, Eric (2019): 'The Self-Undermining Arguments from Disagreement,' *Oxford Studies in Metaethics* 14: pp. 23–46.

Schoenfeld, Miriam (2015): 'A Dilemma for Calibrationism,' *Philosophy and Phenomenological Research* 91: pp. 425–455.

Singer, Daniel J. (2018): 'Diversity, Not Randomness, Trumps Ability,' *Philosophy of Science* 86: pp. 1178–1191.

Skipper, Mattias and Asbjørn Steglich-Petersen (2021): 'When Conciliation

Frustrates the Epistemic Priorities of Groups,' in *The Epistemology of Group Disagreement* (eds. F. Broncano-Berrocal and J. A. Carter). Routledge.

Skipper, Mattias (2023): 'Wise Groups and Humble Persons: The Best of Both Worlds?' *Asian Journal of Philosophy* 2: pp. 1–10.

Sliwa, Paulina (2012): 'In Defense of Moral Testimony,' *Philosophical Studies* 158: pp. 175–195.

Sliwa, Paulina and Sophie Horowitz (2015): 'Respecting All the Evidence,' *Philosophical Studies* 172: pp. 2835–2858.

Thompson, Abigail (2014): 'Does Diversity Trump Ability?,' *Notices of the American Mathematical Society* 61: pp. 1024–1030.

Titelbaum, Michael (2022): *Fundamentals of Bayesian Epistemology*. Oxford University Press.

Walker, Mark (forthcoming): 'Socratic Ignorance, Skeptical-dogmatism, and Self-Refutation,' *Southwest Philosophical Studies*.

Worsnip, Alex (2023): 'Compromising with the Uncompromising: Political Disagreement under Asymmetric Compliance,' *Journal of Political Philosophy* 31: pp. 337–357.

Zagzebski, Linda (2007): 'Ethical and Epistemic Egoism and the Ideal of Autonomy,' *Episteme* 4: pp. 252–263.