Philosophy Without Belief

Zach Barnett

Brown University
zachary_barnett@brown.edu

Should we believe our controversial philosophical views? Recently, several authors have argued from broadly conciliationist premises that we should not. If they are right, we philosophers face a dilemma: if we believe our views, we are irrational; if we do not, we are not sincere in holding them. This paper offers a way out, proposing an attitude we can rationally take toward our views that can support sincerity of the appropriate sort. We should arrive at our views via a certain sort of 'insulated' reasoning—that is, reasoning that involves setting aside certain higher-order worries, such as those provided by disagreement—when we investigate philosophical questions.

1. Introduction

Here is what seems to be a fact about our discipline: some of us really believe the controversial philosophical views we advocate.[1] Some of us really believe that it can sometimes be rational to have inconsistent beliefs, that seemingly vague predicates must have precise application conditions, or that a person would survive if each of her brain cells were replaced with an artificial functional duplicate.

Here is another fact about our discipline: there is widespread disagreement among philosophers surrounding these issues.[2] Given certain assumptions about the nature of these philosophical disagreements, and given certain assumptions about the epistemic import of disagreement more generally, one might come to doubt that our controversial philosophical beliefs are rational—insofar as we have them. Indeed, numerous authors have developed arguments along these lines.[3] The details of their

---

[1] DeRose (forthcoming) argues that we do not genuinely believe our controversial views in philosophy, offering an intriguing story about what we might be doing instead. While I suspect that at least some of us do genuinely believe our controversial views, the arguments given here do not depend on any such assumption.

[2] The Bourget and Chalmers (2013) survey asks philosophers for their views on thirty questions that are taken to be central to the field. For virtually all these, we do not observe anything like consensus.

[3] See Brennan (2010), Christensen (2014), Fumerton (2010), Goldberg (2009, 2013a), Kornblith (2010, 2013), and Licon (2012).

arguments need not concern us, but it will be useful to examine briefly one argument in outline, which will serve as a representative simplification of what they have said:

> *Conciliationism*: A person is rationally required to withhold belief in the face of disagreement – given that certain conditions are met.[4]
> *Applicability*: Many disagreements in philosophy meet the relevant conditions.
> ───
> *No Rational Belief*: Philosophers are not rational in believing many of their controversial views.

The first premise, *Conciliationism*, enjoys ample precedent.[5] Its strengths and weaknesses have been thoroughly explored. I will not rehearse the debate here. The second premise, *Applicability*, is somewhat less familiar, however, so it may be helpful to see what has been said about it. Here is how Christensen motivates the position:

> I do have good reason to have as much epistemic respect for my philosophical opponents as I have for my philosophical allies and for myself … In some cases, I have specific information about particular people, either on the basis of general knowledge or from reading or talking to the particular epistemologists in question. [...]
>
> But another reason derives from the group nature of philosophical controversy. It seems clear that the groups of people who disagree with me on various philosophical issues are quite differently composed. Many who are on my side of one issue will be on the other side of different issues. With this structural feature of group disagreement in philosophy in mind, it seems clear that it could hardly be rational for me to think that I'm part of some special subgroup of unusually smart, diligent, or honest members of the profession. (Christensen 2014, pp. 146-7)

Kornblith takes a similar view, at least with respect to one specific debate:

> Disagreements within philosophy constitute a particularly interesting case … Consider the debate between internalists and externalists about epistemic justification. I am a committed externalist. I have argued for this position at length and on numerous occasions. [...]
>
> At the same time, I recognize, of course, that there are many philosophers who are equally committed internalists about justification[.] It would be

───

[4] Although there is some debate over how these conditions should be characterized, it suffices for us to note that they typically involve the person's having good reason to consider her disagreer(s) equally trustworthy, with respect to the disputed sort of issue, as she considers herself (and her agreers).

[5] See Feldman and Warfield (2009), or Christensen and Lackey (2013), for helpful discussions of this principle.

reassuring to believe that I have better evidence on this question than those who disagree with me, that I have thought about this issue longer than internalists, or that I am simply smarter than they are, my judgment superior to theirs. It would be reassuring to believe these things, but I don't believe them; they are all manifestly untrue. (Kornblith 2010, p. 31)

In light of these observations, *Applicability*, too, can seem a fairly attractive position.[6]

This paper takes both *Conciliationism* and *Applicability* for granted (along with their consequence, *No Rational Belief*) in order to investigate what sense we can make of philosophy if they are true. If we philosophers—in an effort to be more rational—suddenly decided to withhold belief about all philosophically controversial matters, would the practice of philosophy be in any way diminished? As we will see, there is some cause for concern.[7]

2. The sincere philosopher's dilemma

On the face of it, it is not immediately clear why giving up our philosophical beliefs should lead to any problems. Take competitive debate. It is surely quite common for a debater to defend a view she does not, strictly speaking, believe. And this fact hardly serves to undermine the practice of debate. Why should it be any different in philosophy? Perhaps we arrive at certain views, somehow or other, and then defend them as ably as we can manage—without necessarily believing them to be true. Might this be a reasonable way for philosophy to operate?

I find myself somewhat uncomfortable with this picture. In particular, it seems to me that if philosophy were to operate this way, something important would be missing, namely, the *sincerity* with which we defend our preferred positions—a distinctive kind of sincerity that is often lacking in the context of competitive debate.

The kind of sincerity that I have in mind is a way in which I suspect many philosophers identify with the views they defend. The thought is

---

[6] An opponent of *Applicability* is Grundmann (2013). Grundmann points out that even if my opponent is, in general, as philosophically competent as I am (equally honest, equally intelligent, equally hardworking, etc.), she may not be as reliable as I am with respect to the disputed issue. The observation is a good one. But it is worth noting that these general competencies can still serve as fallible indicators of one's domain-specific reliability. Grundmann does not argue that there is no correlation between these distinct competencies.

[7] The following discussion is indebted to Goldberg (2013b). Goldberg's view will be discussed in detail later.

that, for many of us, our views *seem right* to us in some important sense. When I reflect on the relevant issues, my thinking leads me to certain conclusions. And if I were, for some reason, obligated to defend other views (perhaps because they were assigned to me by some governing body), these other views would not be as sincerely held. In defending the assigned views, I would not necessarily be calling the shots as I see them; my own thinking would not have led to them.

This is supposed to capture, intuitively, what it takes for one's views to be *sincerely held*.[8] My claim is not that philosophers *ought* to hold their views sincerely, but rather, that many philosophers do experience this sincere commitment toward their favoured views and would prefer to be able to continue doing philosophy in this sincere manner. And this may be more than just a personal preference for the feeling of sincerity. It is plausible that for many of us, doing philosophy well—energetically and creatively—comes most naturally when we do sincerely identify with the views we defend.

If this is right, then we 'sincere' philosophers have a potentially serious problem on our hands. There seems to be a tension between this sincerity desideratum, on the one hand, and *No Rational Belief*, on the other. We can put the point as a dilemma:

> *Sincere Philosopher's Dilemma*: Either we philosophers will believe our controversial views or we will not. If we do, then we will be irrational. If we do not, then our views will not be sincerely held.

The main task of this paper is to show how this challenge can be met. But first, the challenge should be strengthened. The worry that gives rise to the challenge is that *belief* is required for the relevant kind of sincerity. But this claim is probably too strong.

To see this, assume a Lockean account of belief, according to which outright belief *just is* confidence above a certain threshold, say .75. Let us imagine that I often spend my time working out difficult mathematics problems, replete with tempting pitfalls that frequently trip me up. Over the years, my success rate on these problems is only .74, and I know this fact about my reliability. As a result, when I arrive at an answer to any one of these problems, my confidence in the answer I reach tends not to be quite high enough for belief. Despite my lack of outright belief in my answer, the answer I arrive at still seems right to

---

[8] My choice of the word 'sincere' should not be taken to indicate that philosophers who lack this feeling toward their views are being *insincere*, in some problematic way. I am simply pointing out a way in which many of us identify with the views we defend.

me in an important sense. My own thinking led me to it. And even though I recognize that there is a good chance I erred, overall I regard my answer as more likely to be correct than not. In such a case, I find it natural to say that my commitment to the answer I arrived at is sincere in the relevant sense. If this is right, then outright belief in one's view is not necessary for sincerity.

Perhaps this is right. Even so, it seems to me that the dilemma proponent need not be terribly concerned, for she can reply as follows:

> Perhaps I was too quick in suggesting that outright belief was the only doxastic attitude capable of supporting sincerely held views. A fairly high credence probably can do the trick. But this observation hardly saves the sincere philosopher, for it is doubtful that we can rationally maintain high credences in our controversial views. The same considerations that gave rise to *No Rational Belief* (that is, *Conciliationism* and *Applicability*) are sure to entail a parallel *No Rational High Confidence* principle, which will forbid the high credences present in the alleged counterexample. So here's a more general challenge: tell us specifically which attitude you will take toward controversial views that can get you both rationality and sincerity.

This reply seems to me to be exactly right. The challenge is not simply to demonstrate how to achieve sincerity without belief, but rather, to demonstrate that there is some attitude, or some set of attitudes, which allow for sincere and sensible participation in philosophy. The next section examines one potential answer due to Sanford Goldberg.

3. Philosophical views as speculations

Goldberg has explored nearby territory in a series of recent papers (2013a, 2013b). He defends a version of *No Rational Belief* and so he is concerned with a question similar to the one we are considering. He writes:

> Unless we want to condemn philosophers to widespread unreasonableness (!), we must allow that their doxastic attitude towards contested propositions is, or at any rate can be, something other than that of belief. (Goldberg 2013b, p. 282)

Though Goldberg is not explicitly concerned with allowing that philosophers can sincerely hold their views in the sense discussed in the previous section, he is sensitive to nearby issues, such as the sincerity of philosophical assertion.

Goldberg thinks that there is indeed an attitude that we philosophers can and should rationally take toward our views: '[T]here is an

attitudinal cousin of belief which is reasonable to have even under conditions of systematic disagreement and which captures much, if perhaps not all, of the things that are involved in "having a view" in philosophy' (Goldberg 2013b, p. 284). The relevant state is called 'attitudinal speculation':

> *Speculation*: [O]ne who attitudinally speculates that p regards p as more likely than *not-p*, though also regards the total evidence as stopping short of warranting belief in p. (Goldberg 2013b, p. 283)

Goldberg goes on to suggest that this attitude is what is required for sincere and proper assertion in the context of philosophy. The picture of philosophy being recommended (henceforth, 'the speculation picture') is, I think, a fairly natural one: advocates of Incompatibilism, say, should hold their view to be more likely than its rival; at the same time, they should acknowledge that the total evidence (including evidence from disagreement) does not permit sufficient confidence in Incompatibilism for outright belief.

This picture is most attractive when applied to philosophical issues that divide philosophers into exactly two camps. Goldberg's picture may require refinement, however, in order to handle debates consisting of three or more rival positions. For example, take normative ethics. Oversimplifying dramatically, let us suppose that Consequentialism, Deontology, and Virtue Ethics are equally popular, mutually exclusive views exhausting the plausible options. According to the speculation picture, my being a Deontologist will require me to have a credence in Deontology exceeding .5. There are two potential worries here.

First, it is not clear that such a high level of confidence in Deontology can be rationally maintained on a broadly Conciliationist picture. Admittedly, the version of Conciliationism discussed in the introduction, which dealt with all-or-nothing beliefs, would be silent about this question. But some versions of Conciliationism do place rational constraints on one's level of confidence, and in general, they would demand that one's level of confidence in Deontology be roughly $\frac{1}{3}$ in a case like this.[9]

---

[9] In the case described, the three views were supposed to be about equally popular. So about $\frac{2}{3}$ of my peers reject Deontology and opt for one of the other two views. If I have every reason to think these opponents are, in general, as philosophically reliable as my fellow Deontologists (and myself), then I will not be rationally permitted (from within a broadly Conciliationist framework) to think my side is more likely right than not on this occasion. See, for example, Elga (2007).

Second, one of the most attractive features of the speculation picture—its ability to allow us to favour our own view very slightly—seems to disappear in cases like this one. Suppose that my confidence in Deontology is .4, while my confidence in each alternative is .3. It would seem to make sense to classify me as a Deontologist, but, according to Goldberg's view, this would be a mistake. My confidence in Deontology is, apparently, too low. So the attractive feature of the speculation picture disappears in cases like this.

With the foregoing problems in mind, we might modify Goldberg's view as follows:

> *Speculation**: One who attitudinally speculates that *p* regards *p* as the likeliest option (given some set of options), though also regards the total evidence as stopping short of warranting belief in *p*.

This amended version seems to capture the spirit of Goldberg's proposal nicely, allowing us to lean slightly toward our preferred positions even when there are multiple incompatible ones on offer.[10] Can the revised account provide an answer to our dilemma? Specifically, can *Speculation** be the doxastic attitude underlying our philosophical commitments?

4. Obstacles for the speculation account

In assessing his own account, Goldberg points out that the attitude of speculation may not be sufficient for having a view in philosophy, since proponents of a philosophical view are 'typically more motivated to persist in defence of the view when challenged, than is one who merely speculates that p' (Goldberg 2013b, p. 284). In response, he suggests that *Speculation* should be taken to be necessary condition for having a philosophical view, not a sufficient one (Goldberg 2013b, p. 284).

But there is reason to worry that *Speculation* (and even *Speculation**), may not be a necessary condition, either. There seem to be cases in which one can sensibly have a view, despite regarding it

---

[10] It is also worth noting that *Speculation** incurs a problem of individuation from which *Speculation* is immune. Suppose that my confidence in Consequentialism is .4 and that my confidence in each of the others is .3. Since I regard Consequentialism as likelier than the other options, it seems clear that I do take the attitude specified by *Speculation** toward Consequentialism. But we might carve the options up differently: if instead we say that there are two views on the table—Consequentialism and non-Consequentialism—then I cannot be said to take the attitude specified by *Speculation** toward Consequentialism after all. We can set this difficulty aside, however, for I will suggest that both versions of the speculation picture are susceptible to a more pressing problem.

as *less likely*, all things considered, than some rival view. Anticipating this complaint, Goldberg asks whether it ever makes sense for one to defend a view one regards as a 'long shot'. Ultimately, though, he suggests that there is something 'slightly perverse' about one's holding a view even when one does not think that the view will, in the end, be better-supported by the total evidence (Goldberg 2013b, p. 283 fn. 5). While I share the intuition, thinking about certain cases suggests to me that this is not as problematic as it might seem. Consider an analogy:

> *Logic Team*: You are on a five-player logic team. The team is to be given a logic problem with possible answers *p* and *not-p*. There is one minute allotted for each player to work out the problem alone followed by a ten-second voting phase, during which the team members vote *one by one*. The answer favoured by a majority of your team is submitted.
>
>   You arrive at *p*. During the voting phase, Vi, who is generally more reliable than you are on problems like these, votes first, for *not-p*. You are next. Which way should you vote?

Given a broadly Conciliationist view, it is not rational for you to regard *p* as more likely than its negation, after seeing Vi vote. But there is, I think, still pressure on you to vote for the answer you arrived at, rather than the one you now regard as most likely to be correct.

We can illustrate this by adding a bit more information to the case. Suppose that Vi's reliability is .9, that the reliability of each other team member is .75, that each team member is statistically independent of each other, and that the team is aware of this information. If everyone were to defer to Vi during voting, the team would perform sub-optimally in the long run.[11] So in this collaborative truth-seeking context, there is nothing troublesome about 'defending' a view while thinking that it is more likely incorrect than not. More generally, we can see that, in this context at least, one's all-things-considered confidence is no sure guide to what view one should put forward as one's own.

Does this point carry over to philosophy? Perhaps. Within a broadly Conciliationist framework, how popular a position is (among some group of trustworthy evaluators) partly determines how much confidence one should have in that position, all things considered. But it seems doubtful that the philosophical popularity of a view should have

---

[11] Following this strategy, the team's reliability would just be Vi's reliability: .90. If each team member votes without deferring, the team's reliability can be shown to be considerably higher: approximately 0.93.

much impact on whether a given philosopher should hold that view herself. Consider the following example:

> *Turning Tide*: Physicalism seems right to Pat. She finds the arguments for Physicalism to be persuasive; she is unmoved by the objections. And at present, Physicalism is the most popular view among philosophers of mind/metaphysics. On balance, she considers herself a Physicalist.
>
> Later, the philosophical tide turns in favour of Dualism. Perhaps new arguments are devised; perhaps the familiar objections to Physicalism simply gain traction. In any case, Pat remains unimpressed. She does not find the new arguments for Dualism to be particularly strong, and the old objections continue to seem as defective to her as they always have.

What should Pat's view be, in a case like this? If Pat is a Conciliationist who happens to regard other philosophers of mind and metaphysics as generally trustworthy about philosophical matters (which we'll suppose she is), her all-things-considered confidence in Physicalism may well decrease as Dualism becomes the dominant view, perhaps dipping below .5. But what seems strange is that once Pat's all-things-considered confidence in Physicalism falls low enough, and once her all-things-considered confidence in Dualism rises high enough, she should stop being a Physicalist (and perhaps become a Dualist)—solely on the basis of its popularity, *and despite the fact that, when she thinks about the relevant arguments and objections, Physicalism still seems more plausible to her.* Perhaps there is a special role for *one's own* consideration of the issues to play in contexts like these.

5. Disagreement-insulated inclination

Thinking about the preceding examples suggests a different approach altogether: as a philosopher, my views should be informed by only *some* of the evidence. In particular, I should set aside the evidence I get from the agreement and disagreement of other philosophers in thinking the issues through. The views that strike me as correct, with this disagreement evidence set aside, are the views I should hold. Of course, the evidence I get from agreement and disagreement remains epistemically relevant to my all-things-considered *beliefs*. But for the

purposes of the larger project of which I am a part, in order to arrive at *my views*, I should reason as if it is not.[12]

A good way to get a handle on the proposal is to think about how one typically reacts to a perceptual illusion, such as the one below.
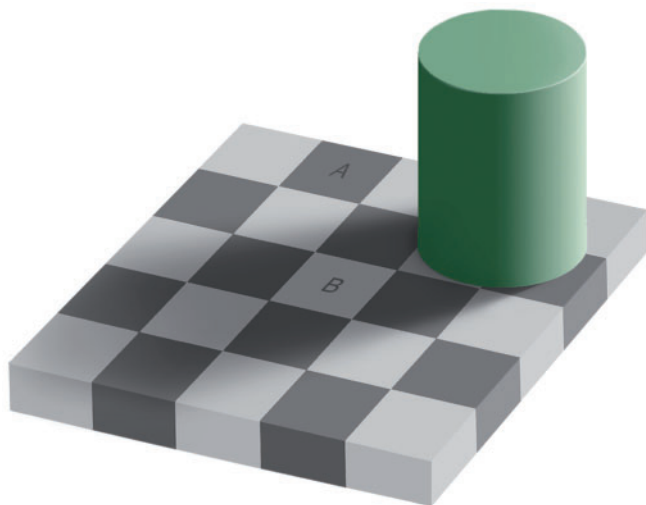
The Checker Shadow Illusion – Edward H. Adelson (1995)

Viewers almost always incorrectly judge the lettered regions to be different in shade. Importantly, the apparent discrepancy between these identically shaded regions tends to remain even after the viewer has become convinced of their constancy. The viewer continues to have the seeming or inclination, but does not endorse it.[13]

An analogous phenomenon occurs when one gains evidence that one's own reasoning about a given topic is likely to be defective in some way. Consider a case involving judgment-distorting drugs:[14]

Deducing While Intoxicated: Basil works through a non-trivial logic problem and comes to believe *p*. She then learns that, before she

---

[12] While this paper is about philosophy, the idea may have broader application to other collaborative, truth-seeking disciplines, such as in the *Logic Team* example.

[13] The seeming prompted by this illusion may involve *alief*, a representational mental state that can conflict with one's explicit beliefs. See Gendler (2008). For the purposes of this paper, the operative attitude will be *conditional belief*, not alief.

[14] See Christensen (2007, 2010, 2011, 2016) for discussion of similar examples.

attempted to solve the problem, she ingested a drug that impinges on one's deductive reasoning skills. It causes ordinarily reliable thinkers to make errors when considering certain logic problems (such as the one just tried) at least half the time. She rereads the problem and finds herself inclined to reason as before: the information given still seems to her to imply *p*. But she refrains from endorsing this seeming and suspends belief.

In the story, Basil is, in some sense, *inclined* to accept a certain claim as true, but opts not to endorse the inclination because of evidence that the mechanisms that produced it may be epistemically defective in some way. This evidence about one's own cognitive capacities is widely known as 'higher-order evidence' (evidence about one's ability to evaluate evidence). Notice that the 'seemings' or 'inclinations' that persist despite what is learned are, in some way, insensitive to this higher-order evidence. In some sense, one can retain the ability to see things *as if* the higher-order evidence were not there, or were not relevant.

But how is this observation relevant to philosophy? Evidence from disagreement (and agreement) is thought to provide higher-order evidence, too.[15] So the suggestion, to put it roughly, is this: philosophers should favour the views that seem right to them, ignoring certain bits of higher-order evidence (including evidence from disagreement/ agreement). David Chalmers helpfully characterizes a related idea:[16]

> [A] level-crossing principle … is a principle by which one's higher-order beliefs about one's cognitive capacity are used to restrain one's first-order beliefs about a subject matter. […] We can imagine a cognizer—call him Achilles—who is at least sometimes insensitive to this sort of level-crossing principle. On occasion, Achilles goes into the mode of insulated cognition. When in this mode, Achilles goes where first-order theoretical reasoning takes him, entirely insulated from higher-order beliefs about his cognitive capacity. He might acquire evidence that he is unreliable about mathematics, and thereby come to believe 'I am unreliable about arithmetic', but

---

[15] See Kelly (2005) and Christensen (2007) for influential early discussions that take this viewpoint.

[16] Others have made independent reference to an idea like this as well. Schoenfield (2014, pp. 2-3) defines your 'judgment' as 'the proposition you regard, or would regard as most likely to be correct on the basis of the first-order evidence alone'. Horowitz and Sliwa (2015) also make use of an idea in this vicinity in their discussion of one's 'first order attitude'. While these attitudes are close to the one I will rely on, the first-order/higher-order distinction turns out not to be the right distinction for the purposes of this paper.

he will go on drawing conclusions about arithmetic all the same. (Chalmers 2012, pp. 103-4)

This idea of 'insulated reasoning' will be useful. The thought is that in philosophy (and in other collaborative truth-seeking contexts), we should try to reason in a way that is insulated from certain evidence, including the evidence we get from disagreement, in determining our views. For reasons we will discuss, it will turn out that we do not want to be insulated from *all* higher-order evidence, as Achilles is in Chalmers' example. But before we can discuss which evidence we will want to wall ourselves off from, we will need to say a bit about what 'walling off' or 'insulation' amounts to. As it stands, it is unclear from Chalmers' discussion whether insulated reasoning is supposed to be something that we humans ever achieve or are even capable of.

The relevant sort of reasoning is of a quite familiar variety: conditional reasoning. We reason conditionally when we reason *as if*, or *supposing*, or *on the condition that* our evidence is different from how it actually is.[17] Conditional reasoning can be divided into *additive* and *subtractive* varieties. In additive cases, we introduce a supposition over and above whatever evidence we already have, and then reason accordingly (for example, 'Supposing we get to the DMV by ten, I think it's likely that we'll be out of there by noon'). Subtractive cases are less common. In such cases, we focus only on a subset of our evidence, ignoring or bracketing some of the evidence we already have, reasoning as if it were not there.[18] Trial jurors are sometimes expected to engage in this kind of reasoning when a piece of evidence is deemed inadmissible.[19] Suppose, for example, that unambiguous video footage of a defendant's crime was collected without a search warrant. A juror who becomes aware of the video might well be instructed to assess the likelihood that the defendant is guilty—setting the video evidence

---

[17] There is perhaps some impulse to understand claims of the form 'Conditional on evidence *E*, I think *p*' counterfactually: 'Were I to acquire evidence *E*, I would think *p*' This impulse should be resisted. An example from Horowitz (2015) can be used to illustrate this. Imagine that Ivan becomes severely irrational whenever he believes that spiders are nearby. But let us add, realistically, that Ivan can suppose that there are spiders nearby without any significant problems. What Ivan does think, on the supposition that there are spiders nearby, may turn out to be quite different from what he would think, were he to learn that spiders are nearby.

[18] It is worth noting that subtractive conditional reasoning is, formally, more problematic than the additive conditional reasoning. We understand conditionalization; can we make sense of 'de-conditionalization'? See Joyce (1999) and Christensen (1999).

[19] Thanks to an anonymous referee for suggesting this example.

aside altogether. Of course, the verdict she reaches may differ markedly from her all-things-considered opinion about the guilt of the defendant.

In both additive and subtractive cases, a person evaluates which way some body of evidence (which may or may not be the evidence she in fact possesses) seems to her to point. For ease of expression, if a certain body of evidence *E* seems to me to support *p*, we will say that I am *inclined*, on *E*, toward *p*.

We can now put the driving thought this way: when someone is doing philosophy with the aim of determining her philosophical views, she should not be evaluating all the evidence she has. Instead, she should be focusing on a special subset of this evidence. Her views should be determined by her inclinations on this evidence (that is, by the way this evidence seems to her to point). What is the special subset of evidence? Provisionally, we will say that the relevant subset is: all the evidence minus that from disagreement and agreement of fellow philosophers. So, on this picture, a person's philosophical views should be her *disagreement-insulated inclinations* (that is, the positions that seem to her supported by the evidence other than that from agreement and disagreement of her peers). To see how this should work, we can apply it to the *Turning Tide* example that caused some trouble for the speculation picture.

At the end of the *Turning Tide* example, the available arguments and evidence seemed to Pat to support Physicalism, but the field was dominated by Dualists. On the current proposal, in arriving at her philosophical views, Pat is supposed to think about all the evidence except for the evidence from disagreement and agreement. This body of evidence includes the arguments and objections with which she is familiar, and which, on balance, seem to her to support Physicalism. In other words, Pat's disagreement-insulated inclination is toward Physicalism. This makes her a Physicalist.

This is not to say that she should *believe* Physicalism to be true, or even that she should be more confident of Physicalism than she is of its negation. If she had to bet on one of the two positions, it would be wiser for her to bet on Dualism, given its popularity among philosophers whom she has good reason to respect. So her all-things-considered confidence in Physicalism may be rather low, since it is sensitive to *all* her evidence, which includes the evidence from disagreement. But the key point is that, as a member of the philosophical community, one's holding of a view and one's all-things-considered

level of confidence in the truth of that view can come dramatically apart.

6. Answering the dilemma

We began this paper seeking an attitude that would enable us to practise philosophy in a sincere, rational way. Disagreement-insulated inclination is an appealing candidate for the elusive attitude we sought. This section explores how well the attitude fares with respect to the sincerity and rationality desiderata, and discusses the role that these inclinations should play more broadly.

Sincerity: Can disagreement-insulated inclination support sincerely held philosophical views? The Physicalism example discussed at the end of the previous section provides some reason to think so. For Pat does seem to be a sincere Physicalist—both before and after Dualism becomes the dominant view. It should be recognized, however, that there is one way in which Pat's commitment is less than fully sincere: she regards Dualism as more likely to be true, all things considered. Nonetheless, it seems clear that Pat's commitment to Physicalism remains sincerely held, in an important sense. When we characterized the sincerity desideratum initially, we wanted it to be the case that the views we hold would trace back to our own consideration of the issues. We wanted our views to *seem right* to us, in some important sense. And Pat's commitment to Physicalism meets these conditions with flying colours. So, at least in this case, disagreement-insulated inclination seems well suited to support the desired sort of sincerity.

Of course, there is a real question about *why* disagreement-insulated inclination is able to support sincerity. After all, disagreement-insulated inclination is a special case of subtractive conditional reasoning (that is, reasoning that involves bracketing some of one's evidence). And subtractive conditional reasoning does not generally foster sincerity of the relevant sort: think of the juror who judges the defendant to be innocent only because she bracketed overwhelming video evidence ruled inadmissible during her deliberation. This juror is inclined, on the admissible evidence, toward the conclusion that the defendant is innocent. But presumably this juror's inclination would not be sincere, in the relevant sense. Why should we expect that, as a rule, our philosophical inclinations will be more sincere than the juror's inclination in this example?[20]

---

[20] Thanks to an anonymous referee for posing this question.

In thinking about this, it is important to note that the view being considered is decidedly not that any old use of subtractive conditional reasoning will deliver sincere inclinations. Given a body of evidence, one can take any chunk of it and set it aside in one's reasoning. When one does this, the inclinations one has will certainly not always 'seem right' in any important sense at all. It will matter crucially which evidence one sets aside. But, then, what makes disagreement evidence so special? Why does bracketing evidence from disagreement tend to result in seemings with which we sincerely identify?

As it happens, this property is not unique to disagreement evidence: it is a feature of higher-order evidence more generally. Very often, bracketing higher-order evidence will tend to facilitate sincere inclinations of the relevant kind. Recall the example of Basil and the judgment-distorting drug. Basil arrives at p by reasoning through the logic problem directly. In the end, she declines to trust the output of this reasoning, because of information about the likely effects of the drug she took. All things considered, Basil might well regard p and *not-p* as equally likely to be correct. But even if she regards both propositions as equally likely, she does not necessarily harbour identical attitudes toward the two propositions. Importantly, Basil still has access to her direct thinking about this question, which led her to p. As a result, if asked to defend p, Basil will have something sincere and substantial to say—even if it turns out to be incorrect. If asked to defend *not-p*, she may not have much to say at all—and if she were to try, it might well feel to her as though she is trying to trick her audience. The key point is that a certain kind of sincerity remains connected to Basil's first-order judgment, even when she admits that the first-order judgment is likely to be mistaken.

Higher-order evidence leads one to doubt one's own *direct thinking* about some core body of evidence. But there can often be a kind of seeming produced by that direct thinking—whether or not one trusts it. Even after one acquires evidence that the direct thinking is not reliable, the direct thinking does not tend to vanish; it will still be there, and it may even be correct. And when the direct thinking is still there, we will be in a position to advocate for the conclusion it produced in a way that bears important hallmarks of sincerity: the position will seem right to us in a psychologically gripping way, and we will have something substantial to say in support of it, in which we can see no flaw. In short, because disagreement is a species of higher-order evidence, it seems plausible that bracketing disagreement will tend to

support the kind of sincerity that *No Rational Belief* seemed to place in jeopardy.[21]

Rationality: Let us now turn to rationality. There are two important questions to address here. First, is the attitude of disagreement-insulated inclination rationally assessable? And, second, can we be rational in taking this attitude toward our controversial philosophical views? There is good reason to think that the answer to both questions is 'yes'. Evidence from disagreement is, in general, what precludes us from rationally believing our controversial philosophical views. With the disagreement evidence set aside, being inclined toward a view can indeed be rational—so long as the view one is inclined toward is in fact the view that the remaining evidence supports. A modified version of the judgment-distorting drug case discussed earlier can be used to illustrate the operative rational norm.[22]

> *Deducing While Intoxicated 2*: Basil and Sage both attempt to solve a challenging logic problem with correct answer *p*. Before attempting the problem, each logician learned that she ingested a drug that impinges on one's deductive reasoning skills. During past experimentation, Basil and Sage discovered that they tend to answer challenging logic problems correctly about half the time, after having ingested the drug.
>
> On this occasion, Basil correctly deduces *p*, while Sage loses track of a negation symbol, and arrives at *not-p*. After obtaining these results, both immediately temper their confidence in their respective answers considerably, because they suspect that their logical reasoning faculties were off-kilter. Indeed, despite having reached different answers, both logicians end up with the same level of confidence in *p*: .5.

Basil and Sage ended up at the same place. But intuitively, we want to say that Basil's reasoning was totally rational and that Sage's was not (since Sage's reasoning was based importantly on a mistake).[23] Thinking about this case in terms of inclination can help us to account for this intuition. Even though Basil and Sage have the same attitude toward *p*, all things considered, notice that *setting aside evidence about*

---

[21] Here, one might be tempted to ask: why *only* bracket disagreement evidence in one's philosophical reasoning? Why not just bracket all higher-order evidence? These questions lead to some interesting complexities, which will be addressed in §7.

[22] See Christensen (2016) for discussion of closely related questions.

[23] Kelly (2010, pp. 122-4) discusses a case with a similar structure (though no drugs are involved).

the drug, Basil and Sage are inclined toward different positions. (Basil might say: 'If this drug's a dud, then definitely *p*'. Sage might say: 'If this drug's a dud, then definitely *not-p*'.) And notice that because the non-drug evidence really does support *p*, only Basil's inclination is fully rational. So it does seem to make sense to say that insulated inclinations can be assessed for rationality.

We should think of this situation as analogous to the situation that we philosophers are often in. With respect to our all-things-considered attitude, philosophical disagreement may compel us to be at the same place: agnosticism. But setting the evidence from disagreement and agreement aside, we will still have our inclinations, and these will differ from person to person. Some of us are like Basil, rationally inclined toward the position the relevant evidence supports. Others of us are like Sage, mistakenly inclined toward some other position. So it will not turn out that *all* philosophers who take the attitude of disagreement-insulated inclination are fully rational in holding their views. But so long as one's inclination is directed toward the position that is in fact supported by the relevant evidence (that is, the evidence other than that from disagreement), one *can* be rational in holding that position as one's view. While it may be irrational to *believe* one's controversial philosophical views, it is not necessarily irrational to *be inclined toward them*, setting disagreement aside.

The Role of Inclinations: At this point, we have seen that the prospect of holding controversial philosophical views in a sincere, rational way can survive the threat posed by disagreement—provided that we philosophers take the right kind of attitude toward our controversial views. But if we adhere to this guideline—if the attitude underlying our philosophical commitments is that of disagreement-insulated inclination—one might wonder about the status of any actions based on those commitments.

A person's philosophical commitments are rarely wholly inert; they may lead a person to make assertions or to pursue courses of action. This raises difficult, important questions. If we are supposed to harbour two distinct attitudes toward certain philosophical propositions—our all-things-considered credence and our dispute-insulated view—on which should we rely when we teach our classes, when we advocate for public policies, when we vote, or when we directly influence public policy in other ways?[24]

---

[24] Thanks to the Editors of *Mind* for raising these questions.

Though it may not be possible to answer these questions in a uniform or categorical manner, it will be instructive to consider several different contexts to illustrate how these questions can be answered, and the kinds of considerations that are relevant in thinking through these issues. To start, consider a medical example:

Doctors: Holly is one of fifty doctors who are concerned with a specific illness. These doctors have two distinct responsibilities related to this illness: they treat patients who are suffering from the illness, and they also perform research into how the illness should be treated.

There are two main drugs available for treatment and research: X and Y. Though there is a large body of data available concerning the efficacy of these drugs, the data are not fully conclusive. Forty-five of the doctors are inclined to think that X is more effective, but the other five—including Holly—are inclined to think that Y is more effective. For medical reasons, it is not possible to administer both X and Y to a single patient safely. And, in addition, it is not possible for a doctor to research more than one of these drugs at a time.

Holly knows all this and is trying to determine both how she should treat the patients she sees and where she should focus her research efforts.

Given certain assumptions about the above case, it can seem clear that all the doctors—Holly included—should administer X to their patients. To see this, suppose that Holly does not see herself as any more likely to make accurate assessments of drug efficacy than the other doctors. Indeed, in the past, when there has been an absence of consensus among the doctors about which of two drugs is most efficacious, the larger group has tended to be right nine times out of ten. Given this track record, and given that a great majority of the doctors judged X to be more effective on this occasion, it would seem irresponsible for Holly (or any other doctor possessing the same evidence) to do anything other than administer X to an ailing patient.

At the same time, it may well make good sense for Holly to investigate the efficacy of Y in her research. As a member of the research community, Holly should do whatever will aid the group in its efforts to determine conclusively which drug is most effective. Toward this end, it may not be optimal to have all fifty doctors devoting their research efforts to the same drug—even if that drug is currently the most promising one. Instead, it may be more efficient to have a majority of the doctors researching the most promising drug, with the

rest researching alternatives that still have a decent chance of turning out to be the best.[25] If we add that the doctors, as a rule, tend to produce their best research when they are permitted to investigate whichever drug sincerely seems to them to be the best, then it will make sense for Holly to research Y (even though she will be administering X to her patients).

With this example in mind, we can consider another structurally similar example, in the realm of public policy. Suppose that Gavin is an expert on some issue of public importance and is assessing the merits of a policy, which may soon be enacted. When he thinks about the issue directly, he finds himself inclined toward the view that the policy is likely to have a quite positive impact. But, like Holly, Gavin is in the minority. The vast majority of other experts hold that the policy will, if enacted, have a quite negative impact. In the past, when Gavin has been in the minority, he has tended to be wrong much more often than right. As a result, he adopts two distinct attitudes toward the policy: he harbours a sincere, disagreement-insulated view that the policy would be beneficial, while his all-things-considered opinion is that the policy is more likely to be detrimental. When it is time to take action, on which of these attitudes should Gavin rely? Naturally, the answer to this question will depend on the specific act he is contemplating.

If Gavin is the governor and thereby has the ability to enact or veto this policy entirely on his own, it is clear that he should veto the policy. The case is relevantly similar to the one involving Holly's treatment of her patients. Gavin, like Holly, should pursue the course of action that is likelier to bring about a favourable outcome, given all the information available. It would seem irresponsible to enact unilaterally a policy that the vast majority of experts deem likely harmful.

In contrast, suppose that Gavin and the other experts are all members of a committee tasked with assessing the likely effects of the policy in question. Here, it is important for Gavin to advocate for his insulated viewpoint, describing the direct considerations that incline him to think that the policy is likely to bring about a favourable outcome.

---

[25] See Ch. 8 of Kitcher (1993) for a thorough defence of the epistemic advantages bestowed upon a scientific community by this kind of diversity. One intuitive takeaway of Kitcher's discussion is that communities that rapidly approach consensus can be at a disadvantage, since inevitably they will occasionally converge around a mistaken opinion—which turns out to be quite inefficient. It is better for a community to hedge its bets, delaying convergence until the situation is relatively clear.

© Barnett 2017

If Gavin is wrong, then, with any luck, the rest of the committee will be able to identify direct considerations that outweigh or undermine the ones he raised. Of course, there is some chance that his advocacy will sway the group to favour the policy. As long as the other experts on the committee are generally reliable in making these sorts of direct assessments, this change of opinion ought to be more likely to occur when Gavin happens to be right. If Gavin were to stay quiet in such a case, then the committee would be less likely to correct itself.

Voting contexts raise still different issues. First, consider a jury that votes. One reason for jurors to rely on their disagreement-insulated attitudes in determining how to vote is that doing so will tend to make the group more accurate. As we saw in the Logic Team example, a group can have a higher probability of reaching the right verdict when the group members reason and vote independently than when they all defer to the most reliable individual person in the group.[26] And insulation from disagreement can serve to enforce this kind of independence, thereby making the group more accurate.[27]

Let us turn to citizen voting. Suppose that Gavin is now contemplating whether to vote for or against the policy mentioned above. This is a particularly complex case.[28] On the one hand, there is some intuitive pressure to say that if Gavin really does regard the policy as likely to be detrimental, all things considered, on the basis of expert testimony, then he should vote against the policy. On the other hand, the considerations about group accuracy discussed in the context of jury voting, which count in favour of insulation, also seem applicable in this case. So there is a tension here that cannot be resolved in the abstract, without reference to the particulars of a given case.[29]

---

[26] Condorcet's (1785) jury theorem shows this. The Logic Team is, in a way, an application of the theorem.

[27] Insulation can prevent *herding*—congregating around a particular opinion—which is a natural result of properly accounting for disagreement evidence. See Lackey (2013) for a discussion showing that independence in belief-forming contexts is a tricky issue to resolve.

[28] One layer of complication stems from the fact that the body of voters might not properly be regarded as a 'truth-seeking body' in the way that juries can be. Instead, we might think of voting as a fair way to reconcile our incompatible preferences. On this picture, Gavin is entitled to vote for the policy he would prefer to see enacted, quite apart from its likely consequences for society at large. To circumvent this worry, we can suppose that Gavin would prefer to enact the policy if and only if it is likely to have a positive impact, on the whole. See, for example, Estlund (2008) for discussion of these issues.

[29] For example, one issue that arises here is that of expertise. In citizen voting, a person may not think that her fellow voters are particularly reliable with respect to the issue being

The context of teaching also raises difficult and important issues. In one's efforts to portray a controversial philosophical issue honestly and fairly to one's students, is it better to rely on one's insulated attitude or on one's all-things-considered attitude? It is worth noting that one need not rely on only one of these attitudes in all aspects of one's teaching. For example, a philosopher should not present her controversial views in the same way that an introductory science teacher would present an established scientific theory: she should make clear to the students that there are well-qualified philosophers on both sides. But a teacher's own inclinations can play a role in her teaching: she may tell her students how things look to her, when she considers the matter directly. And she may encourage the students to work out their own views by directly considering the relevant arguments. Similarly, these attitudes may figure in complex ways in other decisions a teacher might make (for example, in deciding what to put on the syllabus, or in designing exams).

There are certainly going to be many other contexts in which it may not always be clear which of one's attitudes one should rely on. This should not trouble us. In a real-world medical context, the treatment of patients and the research into potential treatments may not be cleanly separable (as was supposed in the *Doctors* example). Such a case would make it difficult to determine whether and to what extent any of the doctors should be acting on their insulated views. Real-world cases are messy, and we should not assume that there is bound to be a clean way to handle them. Nonetheless, we can often identify some of the considerations that are relevant to making these kinds of determinations.

Broadly speaking, we have observed two distinct benefits that can be associated with relying on one's disagreement-insulated attitude (rather than on one's all-things-considered attitude): first, a distinctive kind of sincerity can be enabled; second, the efficiency of a deliberative, truth-seeking body can be enhanced due to the way that insulation can foster a valuable sort of cognitive diversity. Which attitude a person may choose to rely on in a given context will depend on which of these benefits can be obtained and on the extent to which she values them in that context. A philosopher may place a high value on sincerity; a doctor seeing patients may not.

Insulated inclination is not a general substitute for belief. It is a separate item in our epistemic toolkit, to be used when the task before us

---

voted on. The strength-in-numbers consideration discussed previously is applicable only if the other voters are sufficiently reliable, with respect to the disputed issue.

© Barnett 2017

makes its use appropriate. I suspect that there is no simple rule precisely describing the conditions that warrant reliance on an insulated attitude. But it does seem that, in the context of doing philosophy, employing a certain kind of insulated reasoning does make good sense.[30]

7. Inclination, insulated from what?

I will conclude by exploring whether disagreement evidence is the only kind of higher-order evidence from which our philosophical reasoning should be insulated—for there is a compelling case to be made that our philosophical reasoning should be insulated from at least some other kinds of higher-order evidence as well.

While disagreement in philosophy can be used to argue persuasively for *No Rational Belief*, it is by no means the only route to this conclusion. Suppose that Erika, a committed Conciliationist, is philosophizing in her office. She discovers a philosophical question that, as far as she knows, has never been investigated by philosophers before. Since there is no disagreement (that she is aware of), her disagreement-insulated inclination and her all-things-considered opinion will match. After some reflection on this new question, she finds that p seems right to her. But then she pauses, thinking about how much confidence she should have in p, given that it seems right. And it is not at all clear that Erika can, all things considered, regard p as likelier than its competitors. Here is one route Erika might take to a more agnostic position on p:[31]

[30] It might be thought that the foregoing picture runs into difficulties when accounting for philosophical assertions—particularly if it is assumed that warranted assertion requires knowledge or rational belief. There are two important observations to make about this worry. First, it should be noted that this is not a special problem for the insulated inclination picture described here; it arises as soon as *No Rational Belief* is assumed. Second, it is worth pointing out that there may be some divergence between philosophical norms of assertion and ordinary norms of assertion. See Goldberg (2013b) and DeRose (forthcoming) for discussions that explore this idea. If norms of assertion are contextually sensitive in this way, it should be possible to generate norms of philosophical assertion that are compatible with the picture outlined in this paper. One such suggestion is that, in the context of philosophy, an assertion is warranted just in case it is based on a rational inclination, insulated from the appropriate evidence.

[31] There are other higher-order routes to agnosticism in philosophy that make no essential reference to actual disagreement. Ballantyne (2014) pursues one such route, which appeals to the could-have-been disagreement of 'counterfactual philosophers' (that is, people who likely would have disagreed with your philosophical views, had they chosen to pursue philosophy). Frances (2016) discusses a different route, which involves being aware of one's past philosophical failings. Still another might flow from thoughts about the difficulty of philosophical questions.

> *Expected Disagreement*: Although there is no disagreement about whether *p* yet, Erika expects there soon to be some. She thinks that this question is exceedingly likely to provoke disagreement from philosophers whom Erika respects (though she may not be able to predict exactly who will be on which side). She decides that she needn't wait for anyone to say the words 'I disagree'. In anticipation of the disagreement, she divides her confidence equally between *p* and *p*'s soon-to-be competitors.

The above reasoning can be transformed into a more general argument for agnosticism about difficult issues in philosophy. So there might well be disagreement-independent reason to refrain from having much confidence in some of one's philosophical views. And if there is, then for certain philosophical questions (such as those which are likely to engender substantial disagreement), disagreement-insulated inclination cannot fully deliver on its promise: it cannot allow us to arrive at firm, rational opinions about how these philosophical questions should be answered.

One way to react is to recognize that 'evidence from expected disagreement', like evidence from actual disagreement, constitutes higher-order evidence of a certain sort. (The knowledge that fellow philosophers are likely to disagree with one about a particular question is some evidence that one's own thinking about this question is likely to be somehow mistaken.) Perhaps the next move is to demand more insulation—insulation from all higher-order evidence. Let's call an inclination based solely on first-order evidence 'fully insulated'.[32] So on this new proposal, our philosophical views should be our fully insulated inclinations. This would allow Erika to have the view that *p*, since *p* seemed correct to her before considering certain higher-order worries. Unfortunately, this revised proposal is too simple. The problem is that sometimes higher-order evidence does not seem to be evidence that philosophers should set aside in their thinking about the issues.

Consider an analogy. Imagine that a member of an admissions committee learns that she is prone to implicit bias: her (perhaps fully insulated) inclination is, often, to regard male applicants as being more deserving of admission than relevantly similar female applicants. This information constitutes higher-order evidence, since it concerns the committee member's ability to evaluate first-order evidence. But it does not seem to be evidence that she should set aside, in

---

[32] Recall that this was the kind of insulated reasoning that Chalmers (2012) and others seemed to have in mind.

arriving at an independent judgment about a given applicant's merit. Intuitively, she should attempt to compensate for her bias to some extent. And this will preclude her being fully insulated.

Let us move to an example from philosophy. One can imagine learning that, in weighing a theory's elegance against a theory's resistance to potential counterexamples, one tends to overvalue one of these virtues relative to the other. This information constitutes higher-order evidence, since it is evidence about one's capacity to evaluate competing philosophical theories. But once someone did discover this about herself, I think it quite natural to think that she would be justified in compensating accordingly in reasoning about these theories. If she knows that she ordinarily tends to overvalue elegance (say), then she might respond by settling on a less elegant theory slightly more often than she otherwise would. This does not seem immediately problematic. If anything, it would be problematic to ignore this information. So it seems clearly acceptable, at least sometimes, not to bracket higher-order evidence in our philosophical thinking.

From what, then, should our philosophical reasoning be insulated? In trying to answer this question, we should think again about why insulation can be useful in the first place. Insulation has in part been motivated by sincerity. But in the previous section, we discussed a second advantage: it fosters cognitive diversity. Plausibly, philosophy progresses most efficiently when various distinct positions are being investigated. We could, perhaps, imagine a variant of the *Logic Team* case, in which each team is trying to find a proof of a theorem as quickly as possible. Other things being equal, the team might well be better off if its members are not all trying the same type of proof. One might think that something similar could be true for philosophy too: if there is something valuable out there to be discovered (an argument, a counterexample, etc.), a group might be on the whole more likely to find it if the group's members pursue many different paths.

Suppose that the above is right: cognitive diversity is indeed desirable in philosophy. This insight can help us to determine which higher-order evidence to set aside in our philosophical reasoning. Specifically, if some higher-order evidence tends to undermine cognitive diversity (when properly accounted for), then that counts in favour of setting it aside. And if the higher-order evidence *guarantees* cognitive uniformity, then that counts strongly in favour of setting it aside. Does higher-order evidence ever tend to breed cognitive uniformity in this way? It can. Evidence from disagreement is the most straightforward case. When two groups of philosophers disagree,

Conciliationism will recommend that both sides suspend judgment—a uniform outcome. Evidence from expected disagreement functions in a similar manner. Like evidence from actual disagreement, it tends to support suspension of judgment. If all philosophers were to account for the expectation of disagreement in their reasoning (even if their reasoning were insulated from actual disagreement), a uniform agnosticism would result—at least with respect to certain difficult issues. So it makes sense to ensure that our philosophical reasoning is insulated from disagreement and expected disagreement (and other kinds of higher-order evidence that would guarantee uniformity, if properly accounted for).

But what about evidence about one's own bias toward elegant theories? It is worth noting that evidence about one's elegance bias is unlike evidence from disagreement in an important respect: correctly compensating for it in no way guarantees uniformity; a disagreement between two philosophers could easily persist, even after one or both of them compensated for an elegance bias afflicting them. At the same time, compensating for a pernicious elegance bias can undermine diversity in at least some cases: if I am the only one afflicted by the bias, then (before compensating) I might be inclined toward certain (particularly elegant) positions that others find less plausible. After compensating, I might come to dismiss those positions, too, making the overall distribution of views more uniform. If diversity is all that counts, then it would seem clear that I should not compensate (that is, my reasoning should be insulated from evidence of this bias).

But diversity is not all that counts. Yes, my bias might lead me to pursue views that no one else would. But since the bias is a pernicious one, it is likely making me less individually accurate. And individual accuracy counts, too. So our concern for diversity should be tempered by our concern for accuracy.

One might worry, though, that a concern for accuracy could be used to motivate taking account of disagreement evidence as well. After all, taking proper account of disagreement evidence also makes one more accurate in the long run. Setting it aside makes one less accurate. (This was especially dramatic in the *Logic Team* example.) In the case of both bias evidence and disagreement evidence, insulating facilitates diversity at the cost of individual expected accuracy. Why insulate in one case but not the other?

To see the difference, we should notice that there are three values we are simultaneously trying to promote: (1) individual expected accuracy, (2) cognitive diversity, and (3) sincerity. In trying to determine

which of our inclinations should underlie our philosophical views, we should consider the extent to which each of these values may be promoted or undermined by a given policy, were everyone to adopt it.

Start with evidence from disagreement. Ensuring that philosophers reason in a way that is insulated from disagreement evidence will hurt everyone's individual expected accuracy (as insulation always does). But it will enable sincerity to a great extent (as we saw in *Turning Tide*), and it will facilitate cognitive diversity to a great extent (for if everyone were to take account of disagreement evidence, *complete uniformity*—suspension of judgment by all parties —would result). So while there is a cost associated with insulating from disagreement evidence, there are substantial benefits (such as the avoidance of complete uniformity).

The case for insulating from bias evidence is weaker. With respect to values (1) and (3), accuracy and sincerity, the case for insulating from bias evidence may be as strong as the case for insulating from disagreement evidence. After all, insulating from bias evidence hurts personal expected accuracy (just as insulating from disagreement evidence does), and insulating from bias evidence may enhance sincerity (since bias evidence does constitute a kind of higher-order evidence). But while insulating from bias evidence may enhance cognitive diversity to some extent (depending on the details of the case), it will not enhance diversity *to the extent* that insulating from disagreement evidence does. Properly accounting for bias evidence would not guarantee uniformity in the way that accounting for disagreement evidence would. Even if every philosopher learned tomorrow that she harboured a bias toward elegant theories and compensated accordingly, it is very doubtful that consensus would result. So there are reasons to favour insulating from disagreement over insulating from biases.

This is not to offer a clean-cut rule telling us precisely when to insulate and when not to. But it is not obvious that any such rule should be given. I am tempted to think that we philosophers might have some flexibility in certain cases, due to the fact that philosophers value sincerity to varying degrees. Suppose for the moment that both of the first two values (individual expected accuracy and cognitive diversity) tell in favour of accounting for one's elegance bias (that is, not reasoning in a way that is insulated from it). Still, I think that it might well make sense for a philosopher to adopt the view corresponding to her inclination insulated from the bias. Perhaps this philosopher is better able to defend her views when she finds herself sincerely committed to them. Or perhaps she finds that she

is most creative when she argues on behalf of views that 'seem right' to her in the relevant sense. Even if, other things being equal, the community might prefer (on accuracy and diversity grounds) to have a philosopher defending view A rather than view B, it could turn out to be more valuable to the philosophical community to have a creative and persuasive advocate of B rather than a lacklustre advocate of A. So while there do seem to be some general observations we can make (for example, insulating from disagreement is typically a good idea), we may not be able to delineate the bounds of insulation in philosophy with perfect precision. Indeed, there may well be cases in which the philosopher has a choice: it is acceptable to insulate, and also acceptable not to. Since we are trying to promote several different values that can diverge, this permissiveness is not a problem—in fact, it is just what we should expect.

8. Conclusion

We began by assuming that philosophical belief in controversial claims was irrational, in order to see what sense we could make of philosophy if this were so. We saw that the 'sincere' philosopher—the philosopher who holds views that seem correct to her—faces a dilemma: either she may believe her views irrationally, or else she must abandon an important kind of sincerity underlying her philosophically controversial commitments. This paper has proposed that, if careful, the sincere philosopher can retain most of what she might want. In thinking about philosophical questions, the sincere philosopher should set some of her evidence aside—including the evidence provided by disagreement. She can sincerely and rationally advocate the views that she is inclined toward with this evidence set aside. Though she may not believe her controversial views, all things considered, she can hold views that seem correct to her in an important sense.

This is the view I am putting forward for your consideration. I fully expect to encounter dissenting opinion, and I confess that this expectation prevents me from having high confidence in the proposal, all things considered. Nonetheless, I can say sincerely that it seems to me to be correct.[33]

References

Ballantyne, Nathan 2014: 'Counterfactual Philosophers', *Philosophy and Phenomenological Research* 88 (2): pp. 368–387.

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

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

Chalmers, David 2012: 'Insulated Idealization and the Problem of Self-Doubt', *Constructing the World*. Oxford University Press: pp. 101–7.

Christensen, David 1999: 'Measuring Confirmation', *Journal of Philosophy*, 96: pp. 437–1.

—— 2007: 'Epistemology of Disagreement: The Good News', *Philosophical Review* 116: pp. 187–217.

—— 2010: 'Higher-Order Evidence', *Philosophy and Phenomenological Research* 81 (1): pp. 185–215.

—— 2011: 'Disagreement, Question-Begging and Epistemic Self-Criticism', *Philosophers' Imprint* 11 (6): pp. 1–22

—— 2013: 'Epistemic Modesty Defended', in Christensen and Lackey, pp. 77–97.

—— 2014: 'Disagreement and Public Controversy', in J. Lackey (ed.) *Essays in Collective Epistemology*. Oxford University Press: pp. 143–63.

—— 2016: 'Disagreement, Drugs, etc.: From Accuracy to Akrasia', *Episteme* 13 (4): pp. 397–422.

Christensen, David and Jennifer Lackey, (eds.) 2013: *Disagreement: New Essays*. Oxford University Press.

Condorcet, Marie Jean Antoine Nicolas de Caritat, Marquis de 1785: 'Essay on the Application of Mathematics to the Theory of Decision-Making', in Baker, Keith Michael (transl. and ed.) *Condorcet: Selected Writings*. Bobbs-Merril, 1976, pp. 33–70.

DeRose, Keith forthcoming: 'Appendix C: Do I Even Know Any of This to Be True?: Some Thoughts about Knowledge, Belief, and Assertion in Philosophical Settings and Other Knowledge

Deserts', *The Appearance of Ignorance: Knowledge, Skepticism, and Context*, Vol. 2. Oxford University Press.

Elga, Adam 2007: 'Reflection and Disagreement', *Noûs* 41 (3): pp. 478–502.

Estlund, David 2008: *Democratic Authority: A Philosophical Framework*. Princeton University Press.

Feldman, Richard and Ted Warfield, (eds.) 2010: *Disagreement*. Oxford University Press.

Frances, Bryan 2016: 'Worrisome Skepticism about Philosophy', *Episteme* 13 (3): pp. 289–303.

Fumerton, Richard 2010: 'You Can't Trust a Philosopher', in Feldman and Warfield: pp. 91–111.

Gendler, Tamar 2008: 'Alief and Belief', *Journal of Philosophy* 105 (10): pp. 634–63.

Goldberg, Sanford 2009: 'Reliabilism in Philosophy', *Philosophical Studies* 124 (1): pp. 105–17.

—— 2013a: 'Disagreement, Defeat, and Assertion', in Christensen and Lackey: pp. 167–89.

—— 2013b: 'Defending Philosophy in the Face of Systematic Disagreement' in Diego E. Machuca (ed.) *Disagreement and Skepticism*. Routledge: pp. 277–94.

Grundmann, Thomas 2013: 'Doubts about Philosophy? The Alleged Challenge from Disagreement' in Tim Henning & David Schweikard (eds.), *Knowledge, Virtue, and Action: Essays on Putting Epistemic Virtues to Work*. Routledge: pp. 72–98.

Horowitz, Sophie 2015: 'Predictably Misleading Evidence', Conference on Rational Belief and Higher Order Evidence, Brandeis University.

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

Kelly, Thomas 2005: 'The Epistemic Significance of Disagreement' in J. Hawthorne & T. Gendler (eds.), *Oxford Studies in Epistemology*, Vol. 1. Oxford University Press. pp. 167–96.

—— 2010: 'Peer Disagreement and Higher Order Evidence', in Feldman and Warfield: pp. 111–74.

Kitcher, Phillip 1993: *The Advancement of Science: Science Without Legend, Objectivity Without Illusions*. Oxford University Press.

Kornblith, Hilary 2010: 'Belief in the Face of Controversy', in Feldman and Warfield: pp. 29–52.

——2013: 'Is Philosophical Knowledge Possible?' in Diego E. Machuca (ed.) *Disagreement and Skepticism.* Routledge: pp. 260–76.

Lackey, Jennifer 2013: 'Disagreement and Belief Dependence: Why Numbers Matter' in Christensen and Lackey: pp. 243–68.

Licon, Jimmy Alfonso 2012: 'Sceptical Thoughts on Philosophical Expertise'. *Logos and Episteme* 3 (3): pp. 449–58.

Schoenfield, Miriam 2014: 'A Dilemma for Calibrationism'. *Philosophy and Phenomenological Research* 90 (2): pp. 1–31.

Sliwa, Paulina and Sophie Horowitz, 2015: 'Respecting *All* the Evidence'. *Philosophical Studies* 172 (11): pp. 2835–58.

© Barnett 2017