

“Reliabilism in Philosophy”¹

Sanford C. Goldberg
Department of Philosophy
Crowe 3-179
1880 Campus Drive
Northwestern University
Evanston, IL 60208-2214

s-goldberg@northwestern.edu

DRAFT#2 (comments welcome)

A problem is generated by the following three propositions: (1) Reliability (of some robust sort) is a necessary condition on epistemic justification; (2) On contested matters in philosophy, my beliefs are not reliably-formed (in the relevant sense); (3) some of my beliefs regarding contested matters in philosophy are epistemically justified. I believe that each of the three propositions is defensible, but it appears that they can't all be true. In this paper I explore the nature and scope of the problem, examine and reject some candidate solutions to the problem, compare the problem to a similar problem arising in two other recent discussions, and offer a brief and programmatic assessment of our predicament at the conclusion.

1. The problem

The problem that will occupy me in this paper is the result of three claims, each independently plausible, but which form, or at least appear to form, an incompatible triad. The claims are: that reliability is a necessary condition on epistemic justification; that on contested matters in philosophy, my beliefs are not reliably formed (in the relevant sense); and that some of my beliefs on contested matters in philosophy are epistemically justified. In this initial section I want to give my reasons for thinking that each of the three claims in the triad is true.

1.1 Consider first the claim that reliability is a necessary condition on epistemic justification. Although this claim in the theory of justification² is not uncontroversial, I think it is defensible. Here I will give a quick reason for thinking so, and then go on to argue against what is perhaps the key argument used to resist this thesis based on the so-called New Evil Demon (NED) problem.³ Since much of what I have to say here is well-worn territory, I will try to be quick.

Several considerations support some version of this necessity thesis.

¹ Thanks to the audience at the first annual Midwest Epistemology Workshop (Northwestern University, November 30-December 1, 2007), where this paper was workshopped. Special thanks to EJ Coffman, Bryan Frances, Jennifer Lackey, and Baron Reed for helpful comments on earlier drafts of this paper.

² Although I will be speaking about justification, my main topic here is justified belief (and acceptance). I do not think it will confuse matters to speak of justification *simpliciter*; although where a distinction might be drawn between *a belief's being justified* and *a subject's being justified in believing*, I will attend to it.

³ Owed to Stewart Cohen 1984.

First, there is justification's role in the epistemic assessment of subjects' doxastic states. Two points here are relevant. One is that epistemic assessment has something important to do with the subject's aim at truth. For this reason we might think that 'justification' is a label for how well the subject is doing vis-à-vis that aim. The second is that 'reliability' has counterfactual import: a true belief arrived at through methods that would likely to lead to falsity in a variety of similar circumstances is not reliably formed. We can connect these two points by noting that a subject who forms and/or sustains beliefs through methods that are not reliable will typically perform poorly in her attempts at acquiring truths through her reliance on these methods. From here it seems a short distance to the conclusion that reliability – at least in the sense of being formed and sustained through a reliable method – is a necessary condition on justification.

This argument is quick, and I am under no illusions regarding the impact that it will have. Rather than bolster it, however, I want to take aim at what I suspect is the core argument against the necessity thesis, based on the NED problem. Such an argument, I submit, is misguided.

The NED scenario asks us to consider two subjects, one a normally-embodied (and reliable) perceiver S, the other a BIV doppelgänger of S – call her S* – who is experientially indistinguishable from S but whose perceptual beliefs are invariably false.⁴ If reliability is a necessary condition on justification, then S*'s perceptual beliefs are unjustified. Yet this result has struck many as unacceptable. After all, S and S* are experientially indistinguishable, and form precisely the same beliefs in precisely the same way⁵ – suggesting that they are alike justification-wise. But surely S's perceptual beliefs are justified (since S could be you or me); so if S and S* are alike justification-wise, S*'s perceptual beliefs are justified as well – with the result that reliability is not a necessary condition on justification.

In response it can be (and has been) noted that, while this argument undermines one particular version of the necessity thesis, there are ways of understanding 'reliability' such that if reliability is so understood, the necessity thesis is untouched by the argument. As examples I submit the following distinct versions of the necessity thesis, differing in their respective characterizations of reliability:

AR S's belief that p is justified only if this belief was formed (and sustained) through a method that is reliable in the actual world.

⁴ Most versions of semantic externalism imply that S and S* are not experientially indistinguishable, since (according to these versions) their experiences represent the world as instantiating different properties. To accommodate this idea, let us say that S and S* are *subjectively* experientially indistinguishable: given just her own experiences to go on, neither S nor S* can discriminate her case from the counterfactual case (where the counterfactual case for S is the BIV scenario, and the counterfactual case for S* is the normally-embodied scenario). I will not bother with more of the details.

⁵ Once again, semantic externalism complicates matters here, since most versions of semantic externalism will imply that a lifelong normally-embodied agent and a lifelong BIV do not form the same beliefs. In order to circumvent this problem, let our BIV be recently-*envatted* (and let the claim be that all of the beliefs S* forms formed *after the envatment* and which regard her surroundings *at that time* are false). As such a complication is not relevant to my current concerns, I will disregard it in the text.

DR S's belief that p is justified only if this belief was formed (and sustained) through a method that is reliable in the world in which this method evolved.

TGR S's belief that p is justified only if this belief was formed (and sustained) through a method that is reliable in most of the worlds that are experientially indistinguishable from the world in which S formed the belief that p.

AR (Actual-world Reliabilism) would deliver the right verdict in the NED scenario so long as

(i) our perceptual (and memorial) faculties are reliable in the actual world.

(Non-skeptical epistemological theories will assume as much.)⁶ DR (Developmental Reliabilism) would deliver the right verdict so long as (i) holds along with the claim that

(ii) the (perceptual and memorial) method(s) S* employed in forming and sustaining the belief that p were evolved in the actual world.

(Perhaps it could be argued that if there really were a BIV-only-universe, it would be miraculous if methods of the sort S* is credited with evolved – leading to the conclusion that the thought experiment implicitly trades on methods evolved in the actual world.) And TGR (TransGlobal Reliabilism⁷) would deliver the right verdict in the NED scenario on the even more modest assumption that

(iii) in most of the worlds experientially-indistinguishable from the BIV world, the methods employed by S* are reliable.

(The truth of this assumption, it might be argued, is assured by the facts, first, that these methods are reliable in our world, second, that our world is experientially indistinguishable from the BIV world, and third, that in the relevant reference class – worlds that are experientially indistinguishable from the BIV world – the set of reliability-confirming worlds is larger in size than the set of reliability-disconfirming worlds like the BIV world.)

The foregoing considerations show that, suitably understood, the necessity thesis can withstand the pressure from the NED scenario.⁸ So one who hopes to challenge this thesis must go quite far to do so. It is worth noting the burden that would accrue to any such person. Such a person must say precisely why her proposed notion of justification counts as a notion of *epistemic* justification. After all, most people in epistemology –

⁶ Something like this point seems to inform Sosa's notion of adroitness; see Sosa (2004: 284).

⁷ Owed to Henderson and Horgan (2007).

⁸ Faced with the NED argument, my present strategy has been to endorse the intuition behind the argument (that S and S* are alike, justification-wise) without granting its anti-reliability conclusion. For other responses that share this strategy, though not the details of implementation, see e.g. Bergmann 2006 and Gibbons 2006. (Both Bergmann and Gibbons are interested in defending externalism more generally, rather than a more specific, reliabilist version of externalism.)

including some avowed internalists, such as Larry Bonjour – acknowledge that what demarcates epistemic justification from other forms of justification has to do with the link with truth.⁹ Those who hold that there is no sense in which reliability is a necessary condition on justification must convince us that they have not changed the subject.

I have little doubt that there will still be folks unconvinced by (1). To them I offer the present argument as a conditional one, meant to bring out what happens on the *assumption* that reliability in some sense or other – either AR, or DR, or TGR, or ... – is a necessary condition on justification.

1.2 I now move on to my second claim, which is that (2) when it comes to contested matters in philosophy, my beliefs are not reliably-formed.

Some preliminary comments on the notion of reliability are in order.

First, a belief is reliably formed (in the appropriate sense) when it is formed (and sustained) through a reliable belief-forming method. A belief-forming method is said to be reliable only if it produces a preponderance of true beliefs (relative to the class of all beliefs it produces), when it is employed in normal circumstances by a normally-functioning individual. Further, the normalcy of circumstances is relativized to the belief-forming method itself: normal circumstances for making color judgments are one thing, for making auditory discriminations another. The result is that assessments of reliability are implicitly general: an unreliably-formed belief is one that is formed through a method (or methods) that would not yield a preponderance of true beliefs, even when employed by a normally-functioning individual in (what for that method are) normal circumstances. (We might complicate this further to accommodate the points made above, in connection with the notion of reliability; but for simplicity I will not do so here.) The claim on the table, (2), turns on the claim that the methods I employ in arriving at beliefs on contested matters of philosophy are unreliable in this sense.

In assessing the truth of this claim it would help matters greatly if there were a class of easily-identified belief-forming methods that are standardly employed in belief-fixation in controversial matters of philosophy. Unfortunately, the very topic of philosophical methodology itself would itself appear to be a controversial matter of philosophy. Clearly, reasoning itself, including our competence at determining the relative weights to assign to competing considerations, plays some role. But reasoning merely helps us in the moves we make from one claim to another; how do we arrive at the various claims in the first place? Do we rely on intuitions? semantic competence? imagination? Does it involve a more general sort of competence with theory- or model-construction, or with the interpretation of such? These are vexed matters.

However, we can make some progress simply by noting the diversity of incompatible opinions on contested matters of philosophy. Assume (what is no doubt the case) that among the philosophical views that compete with mine a good many are held by normally-functioning people who are at least as intelligent and philosophically competent as I am, work at least as hard and as carefully as I do, and have the same or better access to the relevant evidence as I have. Under these circumstances it would be nothing short of miraculous if a preponderance of my beliefs on contested philosophical matters are true.

Of course, it is one thing to claim that (taking into account the fact of diversity) there is a low probability that a particular belief of mine on a controversial philosophical

⁹ See Bonjour's contribution to Bonjour and Sosa 2003.

matter is true; it is quite another to claim that the belief in question is unreliably formed. If successful, the foregoing diversity argument establishes the former, low probability claim; but does it establish the latter, unreliability claim? One might think not. For one thing, the reliability of a method has to do with how truth-conducive it is across *some range* of contexts (involving conditions that are normal for its use, and the normal functionality of the subject). Unless we restrict that range to contexts involving philosophical disputations, it would appear that there is no reason to infer from the low probability claim to the unreliability claim. For another, the sense in which there is a low probability of truth for a particular belief of mine on a controversial matter of philosophy has to do with probability *on my evidence* (where this evidence includes the fact of diversity); whereas the sense of “low probability of truth” needed to convict a belief-forming method of unreliability is a more objective sense. Unless there is some way to connect these notions, one cannot establish the unreliability claim by appeal to the low probability claim.

However, two distinct arguments can be used to connect the diversity of opinion within philosophy to the unreliability of beliefs on matters of philosophical controversy. I call the first the *Unreliable Methods Argument*, and the second the *Dangerous Context Argument*. Both aim to comport with the two demands just presented. After presenting these, I briefly present a third argument – the *Undercutting Defeater Argument* – that, though it does not aim to show unreliability in the formation of philosophical belief, might be used nevertheless to establish that the reliability in question is not the sort that can underwrite (*ultima facie*) justification. I begin, however, with the arguments that directly address the issue of reliability itself.

The Unreliable Methods Argument aims to show that the methods employed in philosophy are unreliable even at a first pass. Suppose that on a given philosophical issue under discussion, one of the parties holds a view that is true, and one or more of the parties holds a view that is false. (Assume for simplicity that everyone is cognitively sound.) Then for each of the people in the party that “got it right” we credit the method (or set of methods) she used, and for each of the people in the party or parties that “got it wrong” we blame the method (or set of methods) he used. (In each case we allow the possibility that different people used different methods to reach the same conclusion.) The reliability of a philosophical method can then be characterized in terms of the accuracy of the method, where a rough indication of this would be the number of times it was employed and got things right, divided by the total number of times it was used. Assuming a given method can be, and typically is, used by more than one person and on more than one occasion, we can treat any person’s use of a given method, on any occasion, as relevant to an assessment of that method’s reliability.

There are several reasons to think that this sort of bean-counting will support the hypothesis that the methods used in philosophy are unreliable.

To begin, take any philosophical dispute in which parties to the discussion use the same method(s). In any case of this sort, the fact of diversity gives some grounds for concluding against the reliability of the shared method(s). For in such a case one and the same method (or set of methods) is used by distinct individuals to reach opposing beliefs. Assuming that no single side of the debate has a preponderance of adherents, the result would be that, no matter which of the various beliefs produced by that method in this dispute is true, the method does not produce a preponderance of truths on the disputed

topic. (This is ensured by the fact that a preponderance of people using the method arrived at a false belief on the topic at hand.) In this case, the method (or set of methods) in question is not suitably reliable. Note, too, that this reasoning is perfectly general: it can be applied to *any* case in which the various sides to a dispute all use the same method (or set of methods); and it can be used to establish the unreliability of any method that is used in various disputes under the conditions described. So to the extent that methods are shared by parties to a philosophical dispute, this casts some doubt on the reliability of the method(s) involved; and this doubt is exacerbated the more the method, as used by different folks, yields incompatible results.

The foregoing consideration would go a great distance towards establishing the unreliability of any method that satisfies the following description: (a) the method is widely shared; and (b) it is often the case that, as used by different people, the method gives different (and incompatible) results. If all philosophical methods were like this, the case for the unreliability thesis in (2) would be virtually closed. However, it is unclear to what extent shared methodology is prevalent in philosophical disputation. One's views on this matter will turn on one's views regarding the nature of philosophical methodology.

Happily, making a plausible assumption enables us to address this issue without having to enter an opinion about the nature of specifically philosophical methodology. Although the matter is somewhat controversial, it is plausible to think that the individuation of belief-forming methods should reflect cognitive-psychological natural kinds. Given this assumption, it would appear that there are not all that many methods available. And if there are not all that many methods available, then it is likely that each of the few methods available will be shared by a great many people. Under these conditions, it would be nothing short of incredible if it turned out that, while (taken as a group) the belief-forming methods used in philosophy do not produce a preponderance of true belief, nevertheless there is a philosophical method that is reliable. I confess that I have no proof that this cannot be so; but I think that the sheer unlikelihood of such an outcome makes it not worth taking seriously. (Below, in 2.1, I will consider an objection based on the attempt to identify one such method as reliable; my argument against it will provide a schema for how one might argue against any proposed reliable philosophical method.)

The considerations at play in the Unreliable Methods Argument aim to support the conclusion that the methods used in philosophy are unreliable *tout court*. If this is so, then the beliefs formed through these methods are unreliably formed – which would imply the conclusion I sought to reach, (2). However, the argument has one potential weakness: it needs to assume that there are *distinctly philosophical methods*, that is, methods that are more-or-less only applied to philosophical matters. To see why this assumption needs to be made, suppose this assumption is false. In that case, the fact that the methods used in philosophy are unreliable *in philosophy* no more shows that these methods are unreliable *tout court*, than the fact that the methods used in perceptual belief are unreliable under some conditions (i.e., those that are not ideal to perception) shows that these methods are unreliable *tout court*. On the contrary, in both cases perhaps we are dealing with methods that, though reliable in general, do not reliably yield true belief *in certain restricted contexts* (e.g. contexts of bad lighting; contexts of philosophizing).

Happily, the Unreliable Methods Argument can be modified so as to avoid having to assume the existence of distinctly philosophical methods. The assumption of distinctly philosophical methods was needed in order to warrant the inference, from the fact that a given method is unreliable in the domain of philosophy, to the conclusion that the method is unreliable *tout court*. But perhaps we don't need to appeal to the claim that the methods used in philosophy are unreliable *tout court*, in order to arrive at the conclusion that beliefs on matters of philosophical controversy are unreliably formed. Perhaps we can argue directly from the claim that a given method is unreliable in the domain of philosophy. Our model here might be the very case of perceptual belief formed under conditions not appropriate for the use of the given perceptual modality. Take a vision-based belief formed under conditions of poor lighting. The proper way to describe such a situation is that, while the methods employed in vision-based belief are generally reliable, the particular vision-based beliefs formed under these circumstances are not reliably-formed, for the simple reason that conditions are *epistemically dangerous* relative to the reliance on vision: these are conditions under which reliance on vision results (or would result) in a much higher risk of false belief.

From this a strategy emerges for a new argument to the conclusion that beliefs on matters of philosophical controversy are not suitably reliable. The strategy is to argue for this conclusion from the claim that, whether or not the methods used in philosophical belief-fixation are (generally) reliable, the context of philosophy is itself an epistemically dangerous one relative to the use of those methods. To the extent that we have a real diversity of opinion in controversial matters of philosophy, we have grounds for thinking that the presence of philosophical controversy indicates a situation-type in which reliance on even otherwise-reliable methods is very risky: whether or not the methods used are *generally* reliable, they can be reliably expected to run a high risk of producing falsity under these conditions. And if this is so, beliefs on matters of philosophical controversy are no more reliable than are vision-based beliefs formed under conditions of poor lighting.

I call the foregoing the *Dangerous Context Argument*. It is worth pointing out two of the argument's prospective virtues. One is that it need not assume the existence of distinctly philosophical methods.¹⁰ The other is that the argument is neutral on the very identification of the methods implicated in doing philosophy.¹¹ These potential virtues make the argument one that is worth scrutinizing carefully (something I will do below).

Even so, the Dangerous Context Argument makes what might strike many as a contentious assumption: it assumes that the context of philosophy is suitably like the context of bad lighting, to vindicate the analogy on which the argument is based. This assumption might be resisted on several grounds. Perhaps 'philosophy' is not sufficiently well-defined to demarcate a clear type of context; or perhaps the relation between the context of philosophy and the belief-forming methods involved in fixing philosophical belief is not analogous to the relation between contexts involving bad lighting and the belief-forming methods involved in vision-based belief. I think that these worries can be met, and will try to do so below (section 2.2). In advance of that, however, it is worth noting how the Dangerous Context Argument might be modified so as to avoid these

¹⁰ If the assumption holds, the Dangerous Context Argument reduces to the Unreliable Methods Argument.

¹¹ This is a good thing, since – to repeat something I said above – the issue of the belief-forming methods involved in doing philosophy is itself a matter of some philosophical controversy.

worries, yet remain strong enough to support an interesting thesis regarding the reliability of philosophical belief.

I call the new argument the *Undercutting Defeater Argument*. Suppose (contrary to the Unreliable Methods Argument) that the methods implicated in philosophical belief-fixation are generally reliable. Suppose further (contrary to the Dangerous Context Argument) that the domain of philosophy – or more specifically the domain consisting of matters of philosophical controversy – is not correctly regarded as constituting a dangerous context for the use of these methods. Even so, it seems clear that the fact of diversity – and with it the corresponding diminishment in the probability that one’s belief is true, relative to the total evidence – constitutes an undercutting defeater¹² for the beliefs arrived at through reliance on these methods. We can see this as follows. As the diversity of mutually incompatible views (held by people who may be presumed epistemic equals) on a controversial philosophical topic increases, the subjective probability that a person ought to ascribe to the proposition that her belief on that topic is true should decrease accordingly.¹³ So on any topic on which there is such diversity, it is reasonable for one to suppose that it is not the case that there is a high probability that one’s own belief is true. The fact that this proposition is reasonable (on the total evidence) undercuts the support that would otherwise be provided by one’s reliance on the method in question.¹⁴ Note that this holds *whether or not one’s belief was in fact reliably formed*. If it was reliably formed, the case would be analogous to Bonjour’s (1980) reliable clairvoyant: the belief is formed in a reliable way, but there is a defeater in the form of a reason to suppose that the source itself is *not* reliable – thereby undercutting the support that would otherwise be provided by this *de facto* reliable method.

Now the Undercutting Defeater Argument is importantly different from the two previous arguments. Most importantly, its conclusion is significantly weaker: it does not aim to show that the belief in question is unreliably formed. Rather, it aims to show that, even if the belief is reliably formed, the context is such that the reliability in question does not underwrite justification. This is an important difference.¹⁵ Because the first two arguments aimed to show unreliability in the belief-forming process itself, these arguments, if successfully combined with the necessity thesis, (1), would show that philosophical belief is not even *prima facie* justified; whereas the Undercutting Defeater Argument, if successfully combined with the necessity thesis, would merely show that philosophical belief is not *ultima facie* justified. Even so, I raise the possibility of this sort of argument here, as it would support (a perhaps weakened version of) the thesis I am aiming to defend in this paper: namely, that we have a problem in connection with (1)-(3). (More on this below.)

¹² The distinction between undercutting and rebutting defeaters is owed to Pollock 1986 (especially pp. 38-39).

¹³ It might seem that the issue before us is simply a version of the debate regarding the epistemic status of disagreement. I address this in section 3.

¹⁴ The proposition is thus a normative defeater. For an argument to the effect that epistemic externalists can and should incorporate normative defeaters into their account of justification, see Gibbons 2006. Having noted this paper, I should acknowledge that Gibbons’ form of externalism is very different from my reliabilism here.

¹⁵ Thanks to EJ Coffman for bringing this out, and for the subsequent way of putting the point.

Where does this leave us with respect to (2), the claim that my beliefs on matters of philosophical controversy are unreliably formed? If either of the Unreliable Methods Argument or the Dangerous Context Argument were sound, (2) would be established. I am not all that confident in the soundness of the former; I am a bit more confident in the soundness of the latter; and in any case I am even more confident in the soundness of the Undercutting Defeater Argument. Summarizing this, we can say that there is a strong case for thinking that either (2), or some other proposition that could play a similar functional role in the overall argument I am offering here, is true. The other proposition I have in mind is something like this:

(2*) Whatever the reliability of my beliefs in matters of philosophical controversy, none of them are *ultima facie* justified.

Of course, if it is (2*) *rather than* (2) that is true, we would need to modify the set of propositions constituting the problem at the heart of this paper. We might do so by modifying (3) to read that at least some of my beliefs in matters of philosophical controversy are *ultima facie* justified. (Call this claim, (3*.) Depending on how one read ‘justified’ in the original formulation, (3), itself, (3*) may merely make explicit what that formulation was already committed to. In any case in my conclusion I will make clear that whether we formulate the problem in terms of {(1), (2), and (3)}, or in terms of {(1), (2*), and (3*)}, we face what is more or less the same difficulty. Having said this, though, I will move on to consider (3), rather than (3*). (It should be clear that the considerations I offer in support of (3) would offer support to (3*) as well.)

1.3 According to (3), some of my beliefs regarding contested matters in philosophy are epistemically justified. I suppose that this is the most controversial of the claims so far presented. But – call me pigheaded – I still think it’s true.

Suppose (3) is false. Then *none* of my beliefs regarding contested matters in philosophy are epistemically justified. But such a claim carries the suggestion that I am a first-class philosophical hack. And while *you* might be happy to endorse this conclusion, below I will argue that the conclusion generalizes – you’ll get yours in the end, too. (Now who’s laughing?)

The issue before us – whether some of my beliefs regarding contested matters in philosophy are epistemically justified – is somewhat tricky. Nevertheless, several considerations give at least some initial support to the affirmative. I am reasonably intelligent; I have (what passes for) professional training in philosophy and (what passes for) some experience doing philosophy; I devote a decent amount of care, attention, and time to thinking about the philosophical topics that interest me; I have access to, and typically avail myself of, the relevant journals and books where such matters are discussed; I typically have reasons for my philosophical views (available on request!), which I subject to self-scrutiny as well as the scrutiny of others; and at least some of my views can withstand such scrutiny, at least in the sense that these reasons continue to seem compelling (to me) after considering the criticisms. None of this ensures that any of my philosophical beliefs are true; but at the very least they do present an initial case for thinking that some of my beliefs are justified.¹⁶

¹⁶ It is because philosophy by its very nature is a reflective activity – one that requires one to be sensitive to purported reasons to the contrary – that it will typically be true that, if one’s philosophical belief can be

The foregoing considerations might fail to move you to agree. After all, nowhere do I make a case for the *reliability* of my philosophical beliefs. Granted. Even so, the above considerations do suggest the tremendous cost that must be borne by anyone who denies that at least some of my philosophical beliefs are justified. Any such person will have to hold the following: one who is normally-functioning, reasonably intelligent, hard-working, careful, suitably educated and experienced, who has access to and avails himself of the relevant evidence, and who give careful consideration of relevant considerations to the contrary, nevertheless systematically fails to arrive at justified belief in the domain in question. This, I submit, is a curious result: we might be forgiven for thinking that where justified belief is possible, it is just such a person who will attain it. What sort of thing is philosophy anyway, that it attracts reasonably intelligent people into spending their intellectual lives toiling away, only to be systematically thwarted in their efforts at justified belief? (We don't need an evil demon; *philosophy itself* is the culprit.)¹⁷

In sum, while it might be the height of *chutzpa* to hold that I am *reliable* on contested matters of philosophy, it seems much less so to hold that at least some of my contested philosophical beliefs are *justified*. The case for thinking so may be less than conclusive; but it remains a positive case nevertheless.

2. Two objections and some replies

The case I am developing has to do with the problem that arises given three independently plausible but mutually inconsistent claims (1)-(3). In this section I want to consider two objections to the case I have made for (2), the claim that my beliefs in matters of philosophical controversy are unreliably-formed. One of the objections aims to rebut (2) by defending the positive claim that there *is* a reliable method for arriving at philosophical beliefs; if sound this objection would undermine both the Unreliable Methods Argument and the Dangerous Context Argument. I discuss this objection in **2.1**. The other objection aims to undermine my case for (2) by questioning the analogy at the heart of the Dangerous Context Argument; I discuss this objection in **2.2**. It is to be borne in mind, though, that even if I am unsuccessful in rebutting one or both of these objections, there remains the claim in (2*): *it is touched by neither of the objections I am considering.*

2.1 Against (2), it is easy to imagine how some might choose to defend the reliability of the method they use. While I will examine an attempt to defend one such method – the method of common sense – my response to it will indicate how the reliability of any purportedly reliable method might be questioned.

Putting herself in my shoes, the philosophical defender of common sense will be unimpressed with the force of both the Unreliable Methods Argument and the Dangerous Context Argument.¹⁸ “Look,” she will respond, “insofar as one’s philosophical

said to be justified, it can be said to be *ultima facie* justified. (When doing philosophy, few if any of us are like the unreflective chicken sexer.)

¹⁷ Below, in **4**, I will consider a suggestion meant to soften the blow, to the effect that while I might not be epistemically justified in my beliefs on controversial philosophical matters, I am nevertheless justified in some other sense. To anticipate, my reaction is: thanks a lot, but it’s not much solace.

¹⁸ I owe this objection, and the example to follow, to Bryan Frances.

arguments rely on commonsensical premises, and insofar as one employs simple reasoning in reaching one's philosophical conclusions, then, assuming one is competent and otherwise in good cognitive shape, one's philosophical conclusions *will* be highly reliable – even within the domain of matters of philosophical controversy. What is more, one can establish some interesting and controversial philosophical views in this way (or at least views that will be denied by some among the philosophical cognoscenti). E.g., I might establish with Moore that I have hands (I presently see them), and that I *know* this, so that I know that the external world exists, and that radical skepticism is false. The trouble with philosophers, even smart and competent ones, is that they get dazzled by sophisticated arguments, and sometimes forget that commonsense is their reliable route to the truth.”

This objection aims to show that reliance on commonsense in philosophical belief-fixation will yield a preponderance of true belief over false belief, even within domains of philosophical controversy. The objection does not withstand scrutiny.

For one thing, commonsense would appear to deliver too few verdicts. Will it help us settle debates between presentists and four-dimensionalists? Or between rival sides in discussions in political philosophy – let alone philosophy of physics, chemistry, or biology? Will it resolve debates about meaning, reference, or content? About the nature and existence of mental representation? About the proper treatment of vagueness, or the semantic paradoxes? Surely no one side of these debates has a unique claim on common sense. So even if the objection scores points, its application is limited.

In addition it is unclear how many points the objection does score. Consider Russell's (1924/1993) dictum about philosophy: “The point of philosophy is to start with something so simple as not to seem worth stating, and to end with something so paradoxical that no one will believe it.” This dictum suggests that one can have a too-rosy picture of how commonsense figures in philosophy. Even if we concede that it has some role to play – it gives us the so-simple-as-to-be-not-worth-stating starting points – its use in extended chains of reasoning will not always yield commonsensical conclusions. Or so Russell's dictum seems to suggest.

What is more, not even the biggest proponents of the commonsense method would regard it as infallible. In addition to delivering verdicts that are open to philosophical criticism (as in the situation Russell describes), commonsense may sometimes deliver verdicts that are open to criticism by *others'* commonsense verdicts. (There is often debate about what exactly commonsense has to say on a given topic.) Here one might also add that if commonsense issues in judgments having the phenomenology of intuitiveness or obviousness, the emergence of cross-cultural research on what is found intuitive or obvious suggests more variation, and so more reason to doubt the univocality of commonsense.¹⁹

Given that disagreements can arise within commonsense, that commonsense itself can support the noncommonsensical, and that commonsense often flies in the face of what at least many philosophers regard as reasoned opinion to the contrary – not to mention that it is often silent about matters of philosophical controversy – it would seem that commonsense methodology offers no refuge from my two arguments for the unreliability thesis, (2).

¹⁹ One thinks here of the work of Stephen Stich, Jonathan Weinberg, and others.

2.2 Rather than try to argue *in the affirmative* that there is some reliable philosophical method, one might think to argue *against* my case for thinking that beliefs on matters of philosophical controversy are unreliable. Here I focus my attention on the Dangerous Context Argument, since it does not have to make the controversial assumption of distinctive philosophical methods. The objection I want to focus on questions the analogy at the heart of this argument.

In thinking about the reliability of beliefs on matters of philosophical controversy, the Dangerous Context Argument makes an analogy with vision-based beliefs formed under conditions of bad lighting. This analogy allows the argument to grant (if only for the sake of argument) that the methods implicated in belief-fixation in philosophy, like the methods implicated in vision-based belief, are generally reliable,²⁰ and still maintain that these beliefs, like vision-based beliefs formed under conditions of bad lighting, are unreliably formed. But perhaps it will be said that the analogy is not a good one. So here I would like to do a bit more to defend it.

The part of the analogy that is crucial to my argument pertains to my use of the notion of a ‘dangerous context’ for the use of an otherwise-reliable method. No one will dispute the application of this notion to the case of vision-based belief under conditions of poor lighting. The question will be whether it is appropriate to apply this notion to the case at hand. Suppose that, as a matter of fact, most of the perceptual beliefs that have been formed at exactly noon in the local cafeteria are false. Surely this would not motivate the claim that the context of being in the local cafeteria at noon constitutes a dangerous context. (The category of ‘being a belief formed at exactly noon in the local cafeteria’ is not a projectible one.) Perhaps philosophical contexts are better treated like the context of being in the cafeteria at noon, than like the context of bad lighting.

But this suggestion does not withstand scrutiny. Since the category of ‘being a belief formed at exactly noon in the local cafeteria’ marks an accidental feature of a belief, it would be surprising indeed if future beliefs in this category were false a disproportionately high percentage of the time. If philosophy beliefs were in the same category, we would analogously expect that they should not be false a disproportionately high percentage of the time. But if the history of philosophy is any guide, such an expectation would be dashed: we *would* find a disproportionately high percentage of falsity among future beliefs on matters of philosophical controversy. Indeed, the history of philosophy would appear to furnish the basis for something akin to a “pessimistic induction”: since (it is arguable that) most of the views held by even influential philosophers have not withstood the test of time (where this implies that such views are not true), it is likely that most of the views held by any given philosopher will not withstand the test of time (where this implies that such views are not true). No such inference would be compelling in the case of beliefs formed at noon in the local cafeteria.

It must be acknowledged, of course, that the inference from diversity of opinion to the postulation of a dangerous context is not valid: there are areas where opinions are diverse but there is no temptation to explain this diversity in terms of a ‘dangerous context’. One thinks here of diversity of opinion on matters of taste, or on straightforward matters of fact. Disagreement on matters of taste do not signal the presence of a dangerous context so long as these disagreements do not pertain to how

²⁰ I note that if the claim I have been granting is false, my argument is even stronger, since in that case the route to (2) involves fewer premises than I used above in defending it.

things stand objectively. (I do not enter this dispute here regarding the metaphysics of taste properties.) Disagreements on matters of fact do not signal the presence of a dangerous context so long as there is a plausible alternative explanation for such disagreements – perhaps in terms of the relative epistemic position of the various disputants, their access to the relevant evidence, etc. – but which in any case obviates the need for an appeal to the ‘dangerous context’. Many disagreements are like this. So my argument is not merely that diversity of opinion, in itself, signals a ‘dangerous context’. Rather, there is something special about diversity of opinion *in philosophy*:²¹ it is rarely explicable in terms of the relative epistemic positions of the disputants or their respective access to the relevant evidence, and it is particularly resistant to resolution (again, I cite over two millennia of philosophical disputation). Although I have no knock-down argument to show this, I think that these considerations suggest that philosophical disputes arise owing to something having to do with the subject-matter itself (and with the relevant limitations of the relevant belief-fixing processes). It is because of this that the claim, that matters of philosophical controversy constitute a ‘dangerous context’ in which even otherwise-reliable methods are likely to produce a disproportionately high percentage of falsity, is motivated.

But mightn’t the guiding analogy (with the case of vision-based beliefs under conditions of bad lighting) be vitiated by the fact that, unlike the philosophy case, when it comes to vision-based beliefs, we have a theory that would lead us to expect that such beliefs will be unreliable under the conditions alleged to be dangerous? Now I admit that the postulation of a ‘dangerous context’ is more compelling to the extent that we have an account of why, under conditions of just that kind, the reliance on a particular otherwise-reliable method (or set of such methods) is risky. And I admit further that I have no such theory to offer about what happens to the methods involved in philosophical belief-fixation, when they are employed under conditions of philosophical controversy. But for at least a good many candidate methods, it is easy to imagine how such an account might go. And since the current objection, to score points, would have to be that there couldn’t be any such account, it suffices to offer some speculations – and I admit that they are no more than that – merely to indicate how such an account might go.

Suppose that beliefs on matters of philosophical controversy are (sometimes) reached by employing one’s own semantic competence. (A similar story could be told regarding intuition in place of semantic competence.) The reliability of one’s semantic competence as a source for belief-formation could be seen in wide variety of statements whose truth-values we can recognize merely in virtue of being competent with the language. Consider such statements as that the number two can’t get married or have regrets; that if A is taller than B, and B is taller than C, then A is taller than C; and that it can happen that D loves E but E does not love D. While semantic competence puts one in a position to discern the truth-values of these and a good many other ‘obvious’ truths, nevertheless the method of semantic competence has its limits, and in a good many cases a subject who relies on it will be inclined to accept as true statements others will reject as false. In fact, we might speculate that it is precisely at this point when philosophers enter the picture: to settle disputes that arise under conditions in which the semantic competence of various individuals is compatible with their arriving at different verdicts

²¹ It may not be just philosophy. As I suggest below, the same argument could be run on any intellectually demanding domain where great intellectual achievement is compatible with the prevalence of false belief.

on matters that are not straightforwardly factual. If this is the case, it is clear that the cost of regarding philosophy as in the business of making truth-claims is that matters of philosophical controversy will signal a dangerous context. Again, this is very speculative, and a full assessment will have to await the outcome of relevant empirical work. But my point here is merely to show how it could come to pass that we would arrive at an account that would lead us expect that belief-fixing processes used under conditions of philosophical controversy are at least as likely to lead to error as to truth.

In sum, it would seem that the analogy at the heart of the Dangerous Context Argument is defensible. But if it is, then philosophical beliefs on matters of controversy are unreliably formed, precisely at (2) would have it.

2.3 This lands us precisely at the heart of the difficulty: if reliability is a necessary condition on justification, and if my beliefs on contested matters of philosophy are not reliably formed, then it would seem that I cannot be justified in the beliefs that I form on these matters. Or, flipping things around, if I *am* justified in some of the beliefs that I form on these matters, then (assuming that reliability is a necessary condition on justification) it apparently must be the case that on at least some occasions my beliefs on contested matters of philosophy are reliable. Or perhaps the odd man out is the claim about reliability being a necessary condition on justification. At any rate, what I am calling “the difficulty” is that three claims, each independently plausible, appear to form an incompatible triad.

3. The scope of the difficulty

In this section I am to show that the difficulty under discussion is one for a good many people; and that, while it bears on several recent discussions about the epistemic significance of disagreement, it is interestingly different.

First, the problem is one for a good many people. As I have formulated the problem so far, it is framed in the first-person, and in particular regards me. But you can run it on yourself – that is, unless you’ve got some special sort of reliability in matters of philosophical controversy. What is more, the problem can be put in the third-person as well: nothing is lost if instead of being directed at me or you it is directed at ‘any competent philosopher’.

In addition, the problem does not depend on any particular conception of the nature of philosophical claims. Understand such claims as you like: let them express conceptual analyses, or elucidations of the community’s concepts, or the ideology implicit in one’s own concepts, or the grammar of our language, or the presuppositions of our framework, or the output of our faculty of rational intuition, or one’s prejudices, or The argument goes through so long as it is assumed that philosophy aims (perhaps among other things) at saying things that are true.²² And in this connection it is worth bearing in mind that there is more than one way to fail to say something true: one fails to

²² Baron Reed points out (in conversation) that ancient Skeptics aimed not at true philosophical belief, but at avoiding philosophical (and other kinds of) error; and he notes that, for all we know, they were reliable in this aim. I take his point. Even so, they are hardly a model for contemporary philosophizing: few of us have followed their lead, and presumably there is good reason for this. (This said, I confess that my argument could be resisted if, instead of belief, the proper attitude towards the conclusions of one’s own philosophizing ought to be the sort of attitude Van Fraassen regards as appropriate for scientific theory regarding unobservables.)

say something true when one says something false; but one fails to say something true when one fails to say anything at all. I mention this since unintelligibility does seem to be one of philosophy's vocational hazards – and its prevalence only strengthens my argument.

Now it is true that the difficulty will be felt only by those who endorse, or who have some temptation to endorse, the necessity thesis (reliability is necessary for justification). Admittedly, this does not cover all epistemologists: perhaps Fumerton escapes its clutch (a shame, given how fun it is to annoy him); but it should be felt by the vast majority of us. I am not alone among epistemologists in having to face the difficulty.

Though it is related to current work on the epistemic status of disagreement, the issue presently before us differs in orientation from what is found in the disagreement literature. To a rough first approximation, the core issue discussed in the disagreement literature has to do with the *rationality* of sustaining one's (degree of) belief in a proposition, under conditions of disagreement with people one has reason to think of as one's epistemic peers. Our present orientation focuses, not on rationality, but on *reliability*. Considerations pertaining to the reliability of one's belief in a given domain need not have any implications for the rationality of those beliefs. For one thing, many people regard rationality as an epistemically internalist notion, whereas reliability is a paradigmatic example of an epistemically externalist notion. But even waiving this initial point, it would take an argument to connect the issue of reliability with that of rationality.²³ To see this, suppose one were to succeed in showing that it can be rational to retain one's (degree of) belief in a proposition in the face of disagreement with apparent epistemic peers; this would not settle the question whether such a belief is reliably formed and sustained (we might add: in a way that ensures that the belief is *ultima facie* justified). This suffices to make clear that, while the two sorts of issue are connected, they are different in substance.

To be sure, those who regard rationality as constrained by reliability, or who see the debate about disagreement's epistemic significance as a debate about the (reliability-informed) justification of sustained degree of belief in the face of disagreement, will see the present discussion as having a more direct bearing. To those folks I offer the present discussion as one way that the disagreement debate might be developed. (I will not follow this idea further here.)

In addition, the sort of difficulty I am pointing to – one that culminates in the unhappy claim that one's views on contested matters of philosophy are not justified – is distinct from what might seem to be a similar sort of skepticism emerging from an interesting and provocative new book by Bryan Frances. In his (2005), Frances is at pains to establish a conclusion that is at once both more and less ambitious than mine. Frances' conclusion is more ambitious than mine, in that it pertains to a much wider class of agents (not just philosophers, but ordinary folks as well), and to a much wider class of

²³ Objection: my very argument has provided such a connection, since (the objection goes) it cannot be rational to maintain one's degree of belief in the face of acknowledging that one oneself is unreliable in the demarcated set of beliefs. Reply: Either this connection holds, or it does not. If it does not hold, then the objection fails for being based on the false contention that it does hold. But if the connection does hold, then (as I note in what follows) the foregoing argument suggests a novel way to develop the rationality of disagreement literature. Either way, the foregoing argument is not guilty of simply repeating what has been presented in that literature. (Thanks to Baron Reed for raising this point.)

beliefs (not just beliefs regarding contested matters of philosophy, but beliefs on any topic where there is an incompatible hypothesis that is both ‘live’ and not ruled out by the subject in question). But in another respect Frances’ conclusion is less ambitious: his skepticism concerns *knowledge*, not justification. And, although the point is not entirely clear, it seems that, for all his skeptical argument has to say, the beliefs that his argument targets nevertheless continue to enjoy a good deal of positive justification.²⁴ It is precisely such a claim – that beliefs regarding contested matters of philosophy enjoy a good deal of positive justification – that my two arguments for (2) call into question. They do so by questioning whether the methods through which such beliefs are formed make it any more likely than not that the beliefs are true.

Indeed, in a more recent follow-up paper, Frances makes clear that he sees things rather differently from how they are presented here. In “Who am I to argue with David Lewis?” (unpublished ms; but see also the postscript to Frances (2005)), Frances wonders how it can be “epistemically responsible” (his expression) for one to disagree with another whom one recognizes to be one’s epistemic superior on the topic at hand. This is an interesting question; but it is not the question I am presently pursuing. For one thing, I am focusing on reliability rather than responsibility. But more importantly, the difficulty I am discussing applies with virtually equal force to David Lewis as well: for as great a philosopher as he was, I think that, given the nature of philosophy, he was *at best* only marginally more reliable in philosophy than the rest of us, if he was more reliable at all.²⁵

²⁴ I say the point is not entirely clear. Frances does not much talk about justification, and his few remarks are somewhat cryptic. Here is an example, taken from a footnote on p.122 of Frances (2005): “According to live skepticism [the view presented in Frances (2005)], we have true beliefs with loads of positive justification that don’t amount to knowledge, due to some negative justification (or ‘defeaters’). The negative justification (or warrant) is enough to falsify both ‘S knows P’ and ‘S justifiably believes P’ and ‘S warrantably believes P’.” Frances’ view, then, seems to be this: the beliefs in question have “loads of positive justification,” albeit not enough to neutralize the defeaters – with the result that the beliefs in question cannot be called *on-balance justified* (a term I borrow from Peter Graham; see his (forthcoming a and forthcoming b)). My argument against (2) is meant to call into question whether the beliefs even have “loads of positive justification” in the first place.

²⁵ Many folks will resist thinking that reliability considerations have anything to do with Lewis’ greatness in philosophy. I am inclined to agree. It would seem that his greatness consisted rather in the boldness of his vision and the obvious creativity, insight, and intelligence with which he pursued and defended even bizarre theses. These would stand as philosophical virtues even if every thesis for which he argued throughout his career turned out to be false. (It should be clear that this is more grist for my mill!) If nothing else, the present considerations raise an intriguing question about the nature of philosophy as an intellectual endeavor: although most philosophers purport to be in the business of making true statements – or at least we feel embarrassed if we deny this – the nature of evaluation is such that one can fail utterly in this regard and still be regarded as among the very best philosophers. (See Thomas Reid’s quote, below). I can already hear my mother asking: what kind of shady activity is this?

Unfortunately, a new worry arises. If philosophical greatness is consistent with a preponderance of falsity over truth in one’s philosophical beliefs, then philosophical greatness is consistent with (largely or wholly) unjustified philosophical beliefs. Mightn’t this call into question my defense of (3) above, to the effect that some of my philosophical beliefs on controversial matters are justified? After all, if even a genius can fail to have justified philosophical beliefs, why think that a mere philosophical mortal such as myself has justified philosophical beliefs? But two mutually-supplementing points can be made in reply. First, it may well be that genius aids philosophy in some way *other* than via justified claims to truth; perhaps genius’s contribution to the philosophical landscape is to enhance our philosophical imagination – something it can achieve without aiming to expand our knowledge of truths (for which see the introduction of Robert Nozick (2005)). I should be quick to add that this view of philosophical genius need not regard justified philosophical belief as of *no* value, only of lesser value than that of enhanced philosophical

In any case Lewis certainly was not sufficiently reliable to count as (*ultima facie*) justified in his philosophical beliefs. I mean this as no slur on Lewis' work. On the contrary, he worked on some of the hardest issues in M&E: modality, causation, counterfactuals, ontology, vagueness, mind, language, knowledge, etc. These are topics regarding which there is a good deal of controversy and differing opinion, and the players include some who were Lewis' epistemic peers (or near-enough). So it is no strike against Lewis even if he was at best only marginally more reliable than you or I in matters of philosophical dispute; given the nature of the job even that much is an achievement – though not one inoculates him against the foregoing argument.²⁶

It would thus seem that the difficulty I am characterizing is general, in that it should be felt by most epistemologists; has a different orientation than what is typically found in the related discussions about the epistemic significance of disagreement; and is interestingly different from the sort of skepticism that emerges from Bryan Frances' (2005) and more recent work.

4. Possible solutions

The difficulty developed here appears to leave us with three main options: surrender (1), surrender (2), or surrender (3). (One might opt to surrender more than one of these, but you get the idea.) In section 1 I have given my reasons in defense of each of these three claims, taken singly. But given that we appear to face a forced choice, which should be given up? In this penultimate section I make some rather programmatic remarks.

Consider what can be said in defense of the first option. Perhaps it will be said that we should surrender (1) in favor of a reliability-free notion of justification. We need such a notion, it might be contended, to assess belief in any domain where great intellectual achievement is compatible with a preponderance of false opinion. (Perhaps even science fits in here, at least if one thinks that most interesting scientific theories turn out to be false.) On this line of reasoning, the foregoing considerations regarding philosophical belief serve as yet another example illustrating the need for such a notion of justification.

Although the reliabilist in me recoils at the thought, I do not dismiss such a position out of hand; rather, I regard it as a position of last resort, something to endorse only if nothing better presents itself. My reason is simply this: I think that the reliability-informed notion of justification will be needed elsewhere, and I'd like to make do with one notion of epistemic justification, rather than two (or more). To be sure, it may not be possible to make do with just one notion of justification; but the evidence before us does not strike me as sufficiently compelling to draw such a conclusion just yet.

imagination. And second, since few of us are philosophical geniuses, the point in question, that genius need not be a matter of justified philosophical belief, is of limited relevance. To wit: a philosophical genius might offset the badness of massively unjustified philosophical belief by appeal to her role in expanding our philosophical imagination; but the mere mortals among us cannot claim that our massively unjustified philosophical beliefs are redeemed in this way. (Thanks to Baron Reed for pressing me on this.)

²⁶ For a contrasting view of the philosophical genius, compare this from Thomas Reid (himself no mean philosopher): "It is genius, and not the want of it, that adulterates philosophy, and fills it with error and false theory." (Taken from his An Inquiry into the Human Mind.) If Reid is right, Lewis is in even more trouble than you or I!

What, then, of the move to surrender (2)? Such a move strikes me as a non-starter. Why should I think that I am any more reliable in contested matters of philosophy than the likes of David Lewis, Sidney Morgenbesser, Timothy Williamson, Peter van Inwagen²⁷ – or any of the other great philosophers with whose views I have often found myself disagreeing?²⁸ (And I suspect each of these people would say, or would have said, the same with respect to some other set of colleagues.) It seems to me that even casual observation of the philosophical scene suffices to support the view that not even the best among us is particularly reliable in contested philosophical matters. (And recall Reid’s view, cited in a footnote above, to the effect that genius is actually a *hindrance* to truth in philosophy.) (2) is surrendered on pain of wild implausibility.

Suppose that you fail to be moved by this, or by the two arguments I offered on behalf of (2): the Unreliable Methods Argument and the Dangerous Context Argument. Rejecting these arguments will not get you out of the forest: you still will need to confront the Undercutting Defeater Argument. Admittedly, this argument does not establish (2); but it establishes (2*), the claim that even if I am reliable in matters of philosophical controversy, it is not the sort of reliability that underwrites (*ultima facie*) justification. And while the claim that I am not (*ultima facie*) justified in my philosophical beliefs is compatible with the claim that I am *prima facie* justified in my philosophical beliefs, still it is an unhappy thing to have to concede that I am not *ultima facie* justified in (at least some of) these beliefs. For, as I will note in my conclusion, if I am not *ultima facie* justified in my philosophical beliefs, then, given a rather weak assumption about the epistemic norm of assertion, I am not warranted in making any assertions in matters of philosophical controversy – an unhappy consequence.

What, finally, can be said in defense of surrendering (3)? Here the claim might be made that, while my – or your – philosophy beliefs are not *justified*, they are ... well, what? *Reasonable*? This is not much solace, if all it amounts to is conformity to the norms of logic. *Coherent*? There are well-known problems here, especially if coherence is not presumed to have any relation to reliability. *Egocentrically rational*? Again, if this has no connection with reliability, then I wonder why we should care about this.²⁹ *Sufficiently well-defended to impress those in the philosophy conversation*?³⁰ That might earn me or you a small place at the philosophical dinner table, but...

Still, mightn’t we soften the blow of denying (3) by introducing some other notion of justification? Perhaps it will be said that, while none of my controversial philosophical beliefs are *epistemically* justified, they (or at least some of them) might be (to coin a new notion) *philosophically* justified.³¹ Call a belief *philosophically justified* when it is formed through processes that are intellectual virtues, and it rests on reasoning that passes some threshold level of philosophical adequacy. The trouble with this suggestion is that it fails to distinguish philosophical achievement from that of any other intellectually-demanding activity – intellectually-demanding games, for example. Take

²⁷ Or, if you think there is no interesting connection between greatness in philosophy and reliability in one’s philosophical views, replace these names with the names of merely competent philosophers.

²⁸ Compare the attitude expressed throughout Frances (2005) and (unpublished manuscript).

²⁹ In a land of perpetually terrible soil and little water, how much value would there be in a gardner’s keeping his feeble plants as best he can?

³⁰ With a wink to the late Richard Rorty.

³¹ Thanks to Bryan Frances for impressing this upon me, and for introducing the notion of *philosophical justification*.

chess. Success in chess requires some degree of intellectual skill as well as a competence in the game itself. As a result we might coin the notion of *chess-justification* to designate a sort of justification that applies to those moves that are formed through processes that are intellectual virtues, and that are informed by reasoning that passes some threshold level of chess adequacy. If a move you make is chess-justified, you have a reason to be proud, to be sure. But – and please don't be offended – *it's only a game*. Nothing serious hangs on it. In particular, nothing in the way of our attempts to attain truth hangs on it. Now, on the assumption that the best that can be said of a controversial philosophical belief is that it is philosophically justified, the same sort of thing can be said there: when you attain philosophically justified belief on controversial matters, you have reason to be proud, to be sure, but nothing in the way of our attempts to attain truth hangs on it.³² This strikes me as profoundly unsatisfying. I want philosophical belief to be held to a higher standard than one that can be used for intellectually-challenging games. In particular, I want it to be held to the standards employed in other truth-directed inquiries. The analogy of *philosophical justification* with *chess-justification* suggests that the former is not the sort that is wanted.

The drawbacks of giving up (3), then, are akin to the drawbacks of giving up (1). I really would like to think that there is a single notion of justification that can be applied both to (epistemically virtuous) quotidian belief, and to (epistemically virtuous) belief in domains like philosophy, where acquiring truths requires some intellectual effort, and may even be elusive. There may be no such notion; but the foregoing does not seem to me to be sufficient grounds to draw such a conclusion (or to be satisfied with some second-class notion of justification).

5. Conclusion

I'm at a loss. I really want to hold onto each of (1)-(3), but they do strike me as forming an incompatible triad. The odd man out would appear to be (3), but this leaves me – us, really – in an unhappy position.³³ After all, on the prevalent and plausible assumption that assertion has an epistemic norm, it would not be very controversial to

³² Although I have sometimes heard it said by the cynics among us, I stop short of calling philosophy *only a game*.

³³ Is *this conclusion itself* the result of philosophical reasoning, and so susceptible to the very line of argument I am running throughout this paper? I am inclined to think that it is, but that this does not undermine the point the paper wishes to make. It should be borne in mind that the position I am taking in this paper is one of presenting a difficulty without claiming to solve it. The paper itself aims to present us with something like a methodological problem – one the solution to which (if there is one) will require reflecting on philosophical practice and on the evaluative tools we use to assess truth-claiming practices generally. Suppose then that the claim, that the use of the methods we bring to bear in doing philosophy are known (or at least justifiably believed) to yield at least as many falsehoods as truths, is itself a philosophical claim, and so one for which we lack justification (by the very argument I have been offering). In that case our situation is even worse than advertised. Whereas I was claiming to make difficulty for the hypothesis that some philosophical beliefs on contested matters are justified, the present difficulty bears on our ability to make sense of philosophical practice more generally: it would now appear that we cannot reach justified belief regarding that practice. This is more worrisome than the problem I aimed to develop. Thus while I find the present worry to be worrisome indeed, far from undermining the present paper's aim, it actually advances that aim – albeit in terms different that are slightly different from the ones I have been using. (Thanks again to Baron Reed for raising this issue.)

assume further that this norm will imply that an asserter S is warranted in asserting p, only if S justifiably believes p (or perhaps: only if p is justified for S³⁴). To be sure, assertion's norm might require *more* than this – as it would if it requires that asserters be knowers. But most likely it will require *at least* this much.³⁵ If so, neither you nor I are warranted in making assertions on controversial matters of philosophy.³⁶ This will require either a wholesale reconfiguration of the practice of philosophy,³⁷ or else a recognition that when you and I do philosophy – when we give talks, for example – we are systematically behaving in an unwarranted fashion. This is not a happy situation. Any help?

Work Cited

- Bergmann, M. 2006: Justification Without Awareness. (Oxford: Oxford University Press.)
- Bonjour, L. 1980: “Externalist Theories of Empirical Knowledge.” *Midwest Studies in Philosophy* 5.
- Bonjour, L. and Sosa, E. 2003: Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues. (Oxford: Blackwell).
- Cohen, S. 1984: “Justification and Truth.” *Philosophical Studies* 46: 279-295.
- Frances, B. 2005: Scepticism Comes Alive. (Oxford: OUP.)
- Frances, B. Unpublished Manuscript: “Who am I to argue with David Lewis?”
- Gibbons, J. 2006: “Access Externalism.” *Mind* 115:457, 19-39.
- Graham, P. Forthcoming a: “Testimonial Justification: Inferential or Non-Inferential?” *Philosophical Quarterly*.

³⁴ See Lackey 2007.

³⁵ Consider those who are hostile to the view that only knowledge warrants assertion. I have heard such theorists argue that such a view jeopardizes warranted philosophical assertion. My claim is that such people are susceptible to a *tu quoque*. They are susceptible in this way so long as they hold that the rule of assertion entails the (on most views, weaker) thesis that justification warrants assertion.

³⁶ Am I warranted in asserting that we are not warranted in making assertions on controversial matters of philosophy? A critic of my argument might think not. But I reply that if my critic is right, she wins the battle (defeats my argument) but loses the more general war being waged here (the need to acknowledge a problem regarding the practice and assessment of philosophy). See footnote 35 for discussion.

³⁷ Perhaps the ancient skeptics were right after all, at least in practice: we have no business claiming philosophical knowledge. Though neither we nor they should actually *say* this. We and they will just have to resist every attempt by other philosophers to make knowledge claims. Is *this* a happy position? (With thanks to Baron Reed.)

- Graham, P. Forthcoming b: "Theorizing Justification." In O'Rourke, Campbell, and Silverstein, eds. Contemporary Topics of Philosophy 5: Knowledge and Skepticism. (Cambridge: MIT Press).
- Henderson, D. and Horgan, T. 2007: "The Ins and Outs of Transglobal Reliabilism." In S. Goldberg, ed., Internalism and Externalism in Semantics and Epistemology (Oxford: Oxford University Press).
- Lackey, J. 2007: "Norms of Assertion." *Nous* 41:4, 594-626.
- Nozick, R. 2005: Invariances. (Cambridge: Harvard University Press.)
- Pollock, J. 1986: Contemporary Theories of Knowledge. (Totowa, NJ: Rowman and Littlefield.)
- Senor, T. 1996: "The Prima/Ultima Facie Justification Distinction in Epistemology." *Philosophy and Phenomenological Research* 56:3, 551-66.
- Sosa, E. 2004: "Replies." In J. Greco, ed., *Ernest Sosa and His Critics*. (Oxford: Blackwell), pp. 275-325.