

DUNCAN PRITCHARD

VIRTUE EPISTEMOLOGY AND EPISTEMIC LUCK, REVISITED

ABSTRACT. In this paper I return to an argument that I presented in earlier work to the effect that virtue epistemology is at worst false and at best unmotivated. In the light of recent responses to this argument from such figures as John Greco, Guy Axtell, and Kelly Becker, I here re-state and re-evaluate this argument. In the process the original argument is refined and supplemented in key respects and some of the main charges against it are shown to be unfounded. Nevertheless, I also argue that at least one of the objections to the original argument—due to Becker—may well be on the right lines, and draw some conclusions in this regard.

0. INTRODUCTION

In recent work—see especially Pritchard (2003; 2005*a*, chapter 6)—I have argued for a provocative claim regarding the status of virtue epistemology. In short, it goes as follows: virtue epistemology is at worst false and at best unmotivated, and that reflecting on epistemic luck highlights this fact to us. The arguments I put forward in defence of this claim have received quite a lot attention (more than I expected at any rate), and this has prompted me to think again about the argument. In particular, some of the objections put forward have made me realise that the argument could have been presented in a cleaner fashion to avoid confusion. Moreover, at least one of the objections has made me wonder whether the original contention was right, or at least as secure as I thought. With this in mind, I here re-state the original argument in what is, I hope, a better form, and examine in detail the various objections raised against it.

Before I turn to the argument itself, however, I want to make some clarificatory remarks regarding the target of my argument to save possible confusion later on. My focus is virtue epistemology regarding knowledge, which I take to be any view which holds that knowledge must be defined in terms of the epistemic virtues and cognitive faculties. We can distinguish a strong and a weak version of the virtue-theoretic thesis in this regard. The former claims that knowledge must be *essentially* and *exclusively* understood in terms of the epistemic virtues and cognitive faculties, whilst the weak version claims merely that knowledge must be essentially understood in terms of the epistemic virtues and cognitive faculties, though perhaps not exclusively so. My argument against virtue epistemology is that

strong virtue epistemology is false, and that weak virtue epistemology, while perhaps true, is as it stands undermotivated, or at least not motivated to the extent that proponents of the view typically suppose.

Implicit in this argument is the thought that if virtue epistemology is understood in any way weaker than the weak reading just offered, then it is of little interest (at least when it comes to the task of defining knowledge). If, for example, the claim is just that knowledge can be usefully understood in, in part at least, virtue-theoretic terms, although there are alternative and equally adequate ways available, then I don't think the view is all that distinctive. Indeed, I would go so far as to say that in this sense the view becomes almost entirely uncontroversial.

Of course, one might argue that the virtue epistemological project is about so much more than simply offering a definition of knowledge, and has explored and cast light on issues that other epistemological proposals have tended to disregard. This is certainly correct, and three clear examples which illustrate this spring to mind. The first is that the virtue epistemological movement has inevitably brought about a *rapprochement* between epistemology and ethics, prompting a new wave of cross-fertilisation between these two cores areas of philosophy. The second is pretty much a consequence of the first, in that that the distinctive contribution that those working on virtue epistemology have made to the issue of epistemic value is almost certainly a product of the fact that axiological issues lie at the heart of virtue theory. Finally, the third example concerns how virtue epistemology has usefully shifted part of the focus of epistemological theorising away from the minimal conditions for knowledge—a focus that was prompted by the almost interminable post-Gettier literature—and towards higher-order epistemic standings, like understanding and wisdom.

I wouldn't want to dispute that the effect of virtue epistemology on contemporary epistemological debate in each of these three cases (and others) has been undoubtedly a good thing. That said, however, it remains that the appeal of virtue epistemology is very much limited unless it can contribute to the issue of how to define knowledge, and that issue is my concern here.

One final point of clarification that I would like to make is that I am not denying that explaining how creatures such as us come to have knowledge is often best understood along virtue-theoretic lines—indeed, I think that this is the right way to go. My worry just concerns the supposed need to *define* knowledge in virtue-theoretic terms, which I take to be the

characteristic claim of virtue epistemology (if merely saying that we typically acquire knowledge through the operation of one's epistemic virtues and cognitive faculties makes you a virtue epistemologist, then I suspect we are *all* virtue epistemologists).

With these clarificatory points in mind, we can turn to the argument itself.

1. THE BASIC ARGUMENT APPLIED TO AGENT RELIABILISM

A key motivation for virtue epistemology is its apparent ability to deal with epistemic luck in a more satisfactory way than other views. One can see the *prima facie* attraction of this way of motivating the virtue-theoretic thesis. We all agree, after all, that knowledge is non-lucky true belief, since you can't gain knowledge by luck (think, for example, of the Gettier cases in this regard). One natural explanation of this fact is that knowledge is a kind of achievement, and genuine achievements are not due to luck. Instead, the achievement has to be properly attributable to the agent herself. Hitting the target with one's arrow by luck is of no credit to you; whereas hitting the target through skill is. The latter is a case of success which is attributable to the agent, while the former is not. Knowledge is like hitting the target through skill rather than luck: it is success (i.e., true belief) which is properly attributable to the agent rather than being due to luck.

The import of this line of thought to virtue epistemology is that by defining knowledge in terms of the epistemic virtues and cognitive faculties, such views make the cognitive character of the agent a central part of what it is to know. Accordingly, so the thought runs, virtue epistemologies are in a peculiarly good position to account for what knowledge is because they are in a peculiarly good position to accommodate the idea that knowledge is a cognitive success that is properly attributable to the agent rather than due to luck.

One finds this sort of motivation for virtue epistemology throughout the literature, and if it works then it is indeed an excellent way of motivating the thesis. Perhaps the clearest statement of this way of motivating the virtue-theoretic thesis can be found in the work of John Greco (e.g., 1999) who uses it to motivate a reliabilism-based version of virtue epistemology that he calls agent reliabilism.¹ This view holds that knowledge requires a true belief that is formed through the agent's stable and reliable character traits (i.e., epistemic virtues and cognitive faculties) that make up that agent's "cognitive character". In what follows I will specifically focus on this reliabilist version of virtue epistemology, taking

Greco's agent reliabilism to be representative of reliabilism-based epistemology more generally.² Later on, we'll see how the argument I offer against virtue epistemology fares against a representative non-reliabilism-based virtue epistemology.

Greco's discussion of epistemic luck focuses on certain standard counterexamples that are often offered against more basic non-virtue-theoretic forms of reliabilism, such as process reliabilism, where these counterexamples are precisely cases in which an agent forms her reliable true belief via luck and hence can't plausibly be thought of as a knower. For example, consider an agent who is forming a belief about what the temperature of the room is by looking at a malfunctioning thermometer the reading on which is randomly fluctuating within a certain temperature band. Suppose further, however, that this way of forming beliefs about the temperature of the room is entirely reliable because, as it happens, there is someone hidden in the room—a 'helper'—who controls the temperature of the room and who ensures that every time our protagonist goes to check the thermometer, the reading it gives is correct.³

The agent in this case thus has true beliefs and, in one sense at least, his beliefs are also reliably formed. Clearly, however, these beliefs do not amount to knowledge, and the explanation that Greco offers for why agents lack knowledge in these cases is that they do not get to the truth via their cognitive characters, but rather due to luck (in this case, the luck that the environment is such that there is someone who is ensuring that one's beliefs are true even despite being formed via looking at a broken thermometer). That is, such beliefs are epistemically lucky, and hence not knowledge, and thus virtue epistemology—in this case in the guise of Greco's agent reliabilism—is needed to rectify the situation.

Agent reliabilism will certainly handle this case because it is uncontroversial that the agent concerned is not forming a true belief via his cognitive character, since the true belief is the result of the helper's intervention. And what goes for the thermometer case is meant to go for others like it where the reliability in question is not knowledge-conducive because it allows an undue degree of epistemic luck, in the sense that the truth of the agent's belief is not due to the agent but rather due to some feature external to the agent.

I don't dispute that virtue epistemology can handle cases of this sort in this way; my concern is rather that such views are unable to deal with all cases of epistemic luck in this fashion. In order to see this, one only needs to note that virtue epistemologists will typically grant that they are unable to deal with (at least some) Gettier cases. Here is a scenario described by Linda Zagzebski that will serve our purposes:⁴

Suppose that Mary has good eyesight, but it is not perfect. It is good enough to allow her to

identify her husband sitting in his usual chair in the living room from a distance of fifteen feet in somewhat dim light [...] Of course, her faculties may not be functioning perfectly, but they are functioning well enough that if she goes on to form the belief **My husband is sitting in the living room**, her belief has enough warrant to constitute knowledge when true and we can assume that it is almost always true [...] Suppose Mary simply misidentifies the chair sitter, who is, we'll suppose, her husband's brother, who looks very much like him [...] We can now easily amend the case as a Gettier example. Mary's husband could be sitting on the other side of the room, unseen by her. (Zagzebski 1996, 285-7, *emphasis in the original*)

Although he later retracted this concession—something which we will discuss further below—Greco did for a long time concede that agent reliabilism is unable to deal with cases like this (see, for example, Greco 2000, 251). After all, the agent in this case is indeed forming her true belief via the operation of the reliable cognitive traits that make up her cognitive character. And yet, because of the distinctive twist that is always involved in Gettier cases, her true belief is not a case of knowledge because it is manifestly substantively due to epistemic luck.

We have, then, a puzzle, in that one primary motivation for virtue epistemology is its ability to offer an account of knowledge that is not susceptible to epistemic luck in the way that other accounts (like process reliabilism) are, and yet virtue epistemologists typically concede that perhaps the most standard concern about epistemic luck in the literature (scepticism aside)—i.e., the epistemic luck at issue in Gettier cases—cannot be dealt with by their view.

On the face of it, this might not seem like such a problem since, after all, *most* theories of knowledge have problems dealing with Gettier cases, just as most theories of knowledge have problems dealing with radical scepticism. Such a situation is not comfortable for the virtue epistemologist, however, because it entails that a strong version of virtue epistemology—i.e., one that defines knowledge *exclusively* in terms of the epistemic virtues and cognitive faculties—is unavailable.

Moreover, the dialectical situation is in fact even worse than this, in that the sort of epistemic condition that one needs to add to one's view in order to deal with the problem of epistemic luck posed by the Gettier cases will in fact also deal with the other kinds of cases of epistemic luck that virtue epistemology is meant to eradicate. Accordingly, this condition does all the work of virtue epistemology in this regard, and more.

In order to see this, consider the fact that the kind of epistemic luck at issue in the Gettier cases—and, indeed, in the other sorts of cases that the virtue epistemologist focuses on—is targeted at the truth of the belief in question—i.e., it is a matter of luck that the belief is true, given the way the case is described. I have argued elsewhere that we can gloss what it

means for an event to be lucky as follows: an event is lucky if it obtains in the actual world but not in a wide class of near-by possible worlds where the relevant initial conditions for that event are the same as in the actual world.⁵ So, for example, a lottery win is a paradigmatic lucky event because while one wins in the actual world, in *most* near-by possible worlds where the relevant initial conditions for that event are the same as in the actual world (where one continues to play, for example, and the lottery continues to be free and fair) one loses. Applied to the kind of epistemic luck that we are interested in, this account of luck gives us the following gloss on what it means for the truth of a belief to be lucky: the truth of a belief is lucky if that belief is true in the actual world but false in a wide class of near-by possible worlds in which the relevant initial conditions for the formation of that belief are the same as in the actual world (i.e., where one forms one's belief in the target proposition in the same way as in the actual world).⁶

This account of epistemic luck is very vague of course—I supply more detail elsewhere, such as in Pritchard (2005*a*, part two)—but it will do for our purposes here, for note that it captures very neatly the kind of epistemic luck at issue in both the Gettier cases and the cases of epistemic luck that Greco was concerned to eradicate with his agent reliabilism, such as the thermometer case. In the Gettier case described above, for example, the agent truly believes in the actual world that her husband is in the sitting room even though, in a wide class of near-by possible worlds where she forms her belief in the same way as in the actual world (by looking across at what appears to be her husband in his chair) her belief will be false. The same goes, *mutatis mutandis*, for the thermometer case. Although the agent happens to form a true belief in the actual world, in a wide class of near-by possible worlds where he forms his belief in the same way as in the actual world (i.e., by using looking at the broken thermometer) his belief will be false because the helper won't be interceding on his behalf to ensure that his belief is true.

There is a fairly straightforward way of dealing with epistemic luck of this sort—what I have elsewhere called, following Mylan Engal (1992), “veritic” luck—which is to simply stipulate that one's true belief should not be veritically lucky—i.e., should not be such that in a wide class of relevant near-by possible worlds the belief in question is false; or, alternatively, *should be* such that the belief continues to be true in most relevant near-by worlds. The observant reader will have spotted right away that such an anti-luck condition is essentially a version of the safety condition, as defended by Ernest Sosa (1999) and others. This is not surprising, since the whole point of safety has always been to capture the idea that

knowledge involves having a true belief that could not have easily been wrong, and an obvious modal gloss of this claim is precisely that which we just gave as our anti-veritic luck condition.

The trouble is, of course, that an anti-veritic luck condition, like safety, will itself suffice to deal with the sorts of cases of epistemic luck—such as the thermometer case—that the virtue epistemologist cites in order to motivate her view. Moreover, it seems that virtue epistemologists will themselves have to cite a condition of this general sort in order to respond to the Gettier problem anyway, and often do. But if safety is to be part of the view anyway, then why is there any need to appeal to the specific virtue-theoretic aspects of the position in order to deal with other cases of epistemic luck that are neutralised by this anti-luck condition? That is, why not simply opt for a safety-based theory of knowledge and make no essential reference, in one's definition of knowledge at any rate, to the epistemic virtues and cognitive faculties?

Here, then, is the rub, as far as the virtue epistemologist is concerned. Given that one can deal with this kind of epistemic luck by appeal to a condition of this sort—and given that, furthermore, virtue epistemologists will themselves have to appeal to a condition of this sort anyway in order to deal with the Gettier problem, a condition which, note, is not in any obvious way in the spirit of virtue epistemology—then how can consideration of (non-Gettier) cases of veritic luck offer any direct support for virtue epistemology? Indeed, don't such cases motivate, if anything, an anti-luck epistemology defined in terms of some sort of safety condition, rather than an account of knowledge defined exclusively in terms of the epistemic virtues and cognitive faculties? I conclude that one central plank of support for virtue epistemology is in fact, on closer inspection, offering very little support at all.

2. GRECO'S DEFENCE OF AGENT RELIABILISM

As we will see in a moment, we need to complicate this critical line in order to deal with non-reliabilism-based virtue epistemologies, and Zagzebski's non-reliabilism-based virtue epistemology in particular. Before we get on to that issue, however, I want to consider the response that has been offered by Greco (2003; *forthcoming*) to this line of argument.

Greco offers two points in response to my argument. The first is to propose a virtue-theoretic response to the Gettier problem, something which he hasn't tried to do previously. I

will deal with this suggestion in the next section. The second, which I will focus on here, is to argue that safety-based theories of knowledge are fatally flawed.⁷

Essentially, Greco's claim in this regard is that safety is ambiguous in a crucial respect, and that either way one irons out the ambiguity, one is left with a condition that cannot accommodate key cases. In particular, Greco (2003, 263) argues that safety can be read as either *weak* or *strong* safety:

Weak Safety

For all S , φ , S 's belief in a contingent proposition, φ , is weakly safe iff in *most* near-by possible worlds in which S forms the belief that φ in the same way as in the actual world, S 's belief is true.

Strong Safety

For all S , φ , S 's belief in a contingent proposition, φ , is weakly safe iff in *all* near-by possible worlds in which S forms the belief that φ in the same way as in the actual world, S 's belief is true.

The problem with the weak reading of safety, argues Greco, is that it isn't strong enough—it leaves cases of epistemic luck uneliminated that we want eliminated by our theory of knowledge. In contrast, the problem with the strong reading of safety is that it's too strong—it is inconsistent with cases of knowledge that intuitively we want our theory of knowledge to be consistent with.

Greco illustrates the first horn of this dilemma with the lottery case:

S buys a ticket for a lottery in which the chances of winning are ten million to one. A few minutes later, reasoning on the basis of past experience and relevant background knowledge, *S* forms the true belief that she will lose the lottery. Of course her grounds for so believing are merely inductive: it is possible that she buys the winning ticket, although this is extremely unlikely. Greco (*forthcoming*, 1-2; cf. Greco 2003, 266)

Most would agree that agents lack knowledge in the lottery case, and that knowledge is lacked because it is in a sense lucky that the agent's belief is true if it is true. Greco's claim, however, is that using weak safety will force you to grant knowledge to the agent in this case, contrary to intuition. I think Greco is right about this, which is why I reject the weak safety reading of safety.

So if I had to choose between the two readings of safety, then I would opt for the strong reading. Greco thinks that cases like the rubbish chute example show why this reading of safety is unsustainable:

On the way to the elevator *S* drops a trash bag down the garbage chute of her apartment building. A few minutes later, reasoning on the basis of past experience and relevant background knowledge, *S* forms the true belief that the bag is in the basement garbage room. Of course her grounds for so believing are merely inductive: it is possible that the trash bag somehow gets hung

up in the chute, although this is extremely unlikely. Greco (*forthcoming*, 2; cf. Greco 2003, 265-6)

Greco's point about such cases is that we intuitively think that the agent involved does have knowledge, and yet if knowledge demands, in line with strong safety, that one's belief match the fact in all relevant near-by possible worlds, then this doesn't seem to be possible. After all, surely there are quite a few near-by possible worlds in which the garbage doesn't make it to the basement and yet the agent concerned continues to believe that it does (and on the same basis as in the actual world).

I agree with Greco that there is a *prima facie* problem here, though I think it disappears on closer analysis. Even if we are restricted to choosing between strong and weak safety—and I don't think we are, as we will see in a moment—we can still evade this difficulty. This is because if it is indeed unlikely that the bag will snag in this case—and note that this will need to be the situation if the intuition that the agent has knowledge is to hold—then it isn't at all clear that there will be a near-by possible world in which the bag snags and so the agent knows even by the lights of strong safety. In contrast, if we suppose that there is a near-by possible world where the bag snags—such that, for example, there is something in the chute that the bag is nearly snagging on each time it falls—then I think the intuition that the agent has knowledge in this case would subside. So provided that we are clear about the details of the example, then strong safety will predict the right result.

That said, I don't think that we need to choose between strong and weak safety anyway, since filling-out the detail of what an anti-luck epistemology looks like highlights that the right way to formulate safety is in fact somewhere intermediate between these two principles. In order to see this, notice that some events are luckier than others. For example, that I nearly got hit by an accidentally fired bullet that whizzed past my ear is luckier than nearly being hit by a bullet that (with everything else kept fixed) flew by a few feet away. We can accommodate this difference in terms of our account of luck by noting that the range of near-by worlds in which the lucky event (of not being hit by the bullet) fails to obtain—where I get hit—will be greater the luckier the event in question (i.e., greater in the first case than in the second).

With this general point about luck in mind, think again about how best to understand the anti-veritic luck condition on knowledge. I want to suggest that the force of the lottery example is not to make us see that we need to opt for strong safety—i.e., opt for a view which does not allow any relevant near-by possible worlds in which the agent has a false belief in the target proposition. Rather, the import of this case is only this: that there had better not be

any *very close* relevant near-by possible worlds in which the agent has a false belief in the target proposition. After all, the whole point of lottery cases is that the world in which one wins the lottery is just like the actual world, in that hardly anything needs to alter in order to ensure one's success. When it comes to very close relevant near-by possible worlds, that is, I think we are naturally intolerant of any degree of epistemic luck.

Our tolerance increases, however, as we move out into non-very close near-by possible worlds. In particular, with reference to the garbage chute case, suppose that Greco is right that there is some way of reading this example so that there are a few relevant near-by possible worlds in which the bag doesn't make it to the basement and so the agent believes falsely. Given what I said above, such worlds are bound to not be very close worlds, since otherwise the intuition behind the example—that this is a clear case of knowledge—would disappear. Granted this point, however, I think we may well tolerate counterfactual error of *this* sort.

We thus get an intermediate formulation of safety that evades the problem Greco proposes while staying within the spirit of an anti-luck epistemology:

Intermediate Safety

For all S , φ , S 's belief in a contingent proposition, φ , is intermediately safe iff (i) in all very close near-by possible worlds in which S forms the belief that φ in the same way as in the actual world, S 's belief is true; and (ii) in most other near-by possible worlds in which S forms the belief that φ in the same way as in the actual world, S 's belief is true.⁸

With Greco's objection to safety-based theories of knowledge neutralised, there is very little left of this aspect of his response to my argument against virtue epistemology. Everything therefore rests on the other aspect of his argument—the claim that there is a virtue-theoretic response available to the Gettier cases after all.

3. THE VIRTUE-THEORETIC RESPONSE TO GETTIER CASES

Greco is not the first virtue epistemologist to try to offer a virtue-theoretic response to Gettier cases, since Zagzebski, whose view we shall look at in a little more detail in a moment, has also attempted to do this (see, e.g., Zagzebski 1996, §3.2; 1999, §5.B). Moreover, they both aim to resolve the Gettier problem in the same deceptively simple way. Here is the idea: in gaining knowledge what is demanded is not merely that one acquired a belief through one's epistemic virtues and cognitive faculties which was true, but rather that one acquired a belief

which is true *because* it was formed through one's epistemic virtues and cognitive faculties.

On the face of it, this would seem to do the trick. Think, for example, of the Gettier case for standard virtue epistemologies that we saw Zagzebski offering above in which the agent, Mary, is observing her husband's brother across the room and truly believing, on this basis, that her husband is in the room. In this case, while Mary is indeed virtuously forming a true belief, the belief is not true *because* of anything to do with the cognitive character of Mary, since the truth of the belief is instead due to the happenstance that her husband is in the room hidden from view behind his brother.

Although superficially appealing, on closer inspection this proposal starts to unravel. Consider a counterpart to Mary, Mary*, who is in exactly the same situation as Mary except that her true belief has not been Gettiered and hence she has knowledge of what she believes. Suppose, for example, that the husband and his brother have swapped places so that Mary* is indeed looking at her husband. Given that all that is different about the two cases are incidental features of the environment, it is hard to see why there should be any difference in the cognitive characters of Mary and Mary*. But if there is no difference, then how can it be that the truth of Mary*'s belief, but not Mary's, is due to her cognitive character?⁹

Indeed, the natural way to understand what Greco is saying here is as insisting on a true belief formed as a result of a reliable epistemic virtue which also meets the counterfactual condition that the true belief is *sensitive* to the truth—i.e., if the proposition believed had not been true, then the agent would not have believed it in the way that she actually did (i.e., in this case via her cognitive faculties).¹⁰ So in the Gettier case just considered, for example, the true belief, while formed via a cognitive faculty, is not a case of a true belief which is true *because* it was formed via a cognitive faculty because had the proposition believed not been true then Mary would still have believed it in the way that she actually did. That is, in the nearest possible world in which her husband is not in the room, Mary would continue to believe that he is there (because his brother would still be there) and on the same basis as in the actual world (in this case via her cognitive faculties).

Sensitivity will certainly do the trick in this regard, as indeed, would an appeal to safety, since both of them eliminate the veritic luck involved in Gettier cases.¹¹ Notice, however, that if Greco does end up appealing to sensitivity or something like it here then this leaves his position open to precisely the same objection that I levelled against strong virtue epistemology above. In particular, since sensitivity (like safety) is not obviously a virtue-theoretic condition, the fact that virtue epistemologists need to appeal to such a condition in order to deal with Gettier cases indicates that the project of defining knowledge essentially

and *exclusively* in terms of the epistemic virtues is a failure. Moreover, note that it is as a result of reflecting on the need to eliminate epistemic luck that this problem for virtue epistemology is highlighted (as opposed to such reflection providing support for the view).¹²

4. ZAGZEBSKI'S NON-RELIABILISM-BASED VIRTUE EPISTEMOLOGY

Other responses to my argument—in particular, that offered by Guy Axtell (2003; *forthcoming*)¹³—have focused not on its application to reliabilism-based virtue epistemology, but rather on its application to explicitly non-reliabilism-based virtue epistemologies, and Zagzebski's (1996) view in particular. Axtell's papers are rich, and I cannot hope to do justice to *all* the points that he raises here. Instead, I will offer two general points which deal with some of his key objections. I will then offer an outline of the position held by Zagzebski that Axtell is concerned with, and show how my argument against virtue epistemology applies to this view, before returning to make two final observations about Axtell's critique.

To begin with, part of Axtell's critique of my argument seems to proceed by claiming that I fail to recognise the full scope and ambition of non-reliabilist virtue epistemic proposals. Axtell (2003), for example, makes much of the fact that Zagzebski motivates her virtue epistemology by considering the issue of epistemic value, the point being, presumably, that by focussing on the virtue theoretic analysis of knowledge, particularly in the light of the problem of epistemic luck, I fail to recognise this fact.

As I noted above, however, I do not deny that many positive things have come out of the virtue theoretic programme in epistemology, and their contribution to the debate about epistemic value is one of them. Moreover, such positive outcomes of the programme obviously do lend the programme some support. Nevertheless, the point remains that unless virtue epistemology is able to offer an analysis of knowledge which is in the spirit of the view, then the epistemological import of the programme is severely limited. Thus, it is essential that virtue epistemologists find a way of responding to my argument.

This brings me to a second aspect of Axtell's critique, which is his claim that virtue epistemologists are somehow not interested in the project of defining knowledge at all. I'm not sure what to make of this point. It is certainly true, as I have noted, that virtue epistemologists are not *solely* concerned with the project of defining knowledge, and that this is to be applauded. Nevertheless, I would argue that a fully-fledged proposal in epistemology

must be able to say something substantive in this respect if it is to be accepted, and its appeal is severely limited if it cannot do this.

Interestingly, in more recent work—especially Axtell (*forthcoming*)—the suggestion seems to not merely be that I have ignored the fact that virtue epistemologists do not exclusively focus on the project of offering an informative and non-circular analysis of knowledge, but rather that I have failed to recognise that virtue epistemology, properly conceived, is not engaged in that project *at all*, even as a secondary part of the programme.

This stronger claim clearly raises some important metaphilosophical issues about what the role of philosophy is, issues that I cannot engage with fully here. I am inclined to think, however, that Axtell is overstating his stance here, since, given that he later goes on in this paper to himself offer an informative and non-circular analysis of knowledge (see §5), he obviously does not think that virtue epistemology should disengage with the project of analysis entirely. Instead, as far as I can tell, the point is rather to treat the details of the analysis to be spelt-out within specific contexts, rather than regarding them as fixed in advance. This sort of appeal to context raises a dilemma for Axtell, however.

To begin with, notice that an anti-luck condition could also be regarded as only spelt-out relative to a specific context, at least on one plausible reading of what makes a context here. The particular way in which a belief is safe is, after all, dependent upon a myriad of contextual factors which go beyond purely psychological facts about the agent. Accordingly, if Axtell just has this kind of context-sensitivity in mind, then there is no challenge here that I should be worried about. If, on the other hand, he has a more robust sense of context-sensitivity in mind, then I think we need to be given much more detail about how it should be understood. Moreover, it is worth noting that such a robust account of context-sensitivity would have the implausible result of wedding virtue epistemology to a further, and rather contentious, epistemological thesis, one that does not obviously seem to be directly related to the main thrust of virtue theoretic programme in epistemology. Either way, this is not a happy result for Axtell's critique.

With these two points in mind, I will now briefly reiterate what I say about the kind of non-reliabilist virtue epistemologies that Axtell is concerned with, taking Zagzebski's view to be representative of the kind.

What is distinctive about Zagzebski's view, in contrast to the kind of reliabilism-based views defended by Greco and others, is that it is a form of epistemic internalism. By epistemic internalism here, I mean any view which demands of a knower that she has good reflectively accessible grounds in favour of her belief, with epistemic externalism being any

view which does not make this demand.¹⁴ Zagzebski's commitment to epistemic internalism in this sense comes out at a number of junctures in her work, but perhaps the best way to highlight her stance in this regard is by considering how she responds to the chicken-sexer case.

The case of the chicken-sexer concerns someone who (so the story goes) reliably forms her beliefs about the sex of the chicks before her because of her sense of smell. As the story is usually told, however, it's also the case that this agent has false beliefs about how she is doing what she is doing—she thinks she's touching something distinctive, for example, even though there is nothing distinctive for her to touch—and that she lacks good reasons for thinking that she is reliable in this regard (perhaps she was told that she was reliable, but told by someone who thought that she wasn't and was trying to deceive her). Can such an agent *know* the sex of the chicks before her?

Characteristically, epistemic externalists say 'yes' to this question, as do virtue epistemologists who adopt a reliabilism-based version of the view, as Zagzebski rightly notes. After all, the agent's beliefs are being formed in a reliable fashion. Moreover, she is forming true beliefs as a result of her cognitive character, since her chicken-sexing faculty is surely part of her cognitive character. It is thus reliable true belief that arises out of a virtuous cognitive character. Furthermore, it hasn't been Gettiered, so there is no difficulty from that direction. Accordingly, on this view at least, it is hard to see what could prevent the belief from counting as an instance of knowledge.

In contrast, Zagzebski, in common with most epistemic internalists, argues that the chicken-sexer lacks knowledge. Where externalist versions of virtue epistemology go wrong, as far as Zagzebski is concerned, is by allowing that an agent might gain knowledge simply via her cognitive faculties. In contrast, she maintains that knowledge in addition requires the epistemic virtues, and it is clear from the text why she maintains this, since without the operation of the virtues the agent would not be in a position to offer good grounds in favour of her belief, as the case of the chicken sexer (who is clearly not exhibiting epistemic virtue) illustrates.

What is important for our purposes, however, is that Zagzebski thinks that the problem with the kind of 'brute' knowledge allowed by externalists—i.e., knowledge which is not supported by good reflectively accessible grounds—is that it is lucky. Indeed, Zagzebski (1996, 39) complains that epistemic externalists are unduly "sanguine" about epistemic luck.

There's clearly *something* right about this thought, but it won't do as it stands. After all, as we have seen, Greco (and he's not alone on this score) is very much concerned to eliminate epistemic luck—indeed, this is an explicit motivation that he offers for advancing virtue epistemology. Moreover, the type of epistemic luck at issue here clearly isn't of the veritic form—it is not the type of epistemic luck that is at issue in Gettier-style case, for example. Given that the chicken-sexer really does have this reliable chicken-sexing faculty, then she doesn't just *happen* to form true beliefs about the target propositions—in relevant near-by possible worlds where she continues to form the beliefs that she forms in the actual world, her beliefs will continue to be true.

My contention about internalist virtue epistemological theses, like that offered by Zagzebski, was that they are concerned to eliminate not just veritic luck, but also a second type of epistemic luck, what I termed “reflective” luck. In essence, reflective luck is like veritic luck except that the ordering of worlds is not fixed in the usual ‘objective’ way by the facts in the actual world, but is rather fixed in a ‘subjective’ way in line with what the agent has good reason to believe the facts are in the actual world. That is, if an agent has good reason to think that the actual world is such that, for example, she has a certain cognitive ability, then she will retain this ability in most near-by worlds on this subjective ordering. In contrast, if she has no good reason for thinking that she has this ability, then there will be no tendency for near-by worlds to be such that she has this ability.¹⁵

It should be clear that while the chicken-sexer's beliefs are not veritically lucky, they are reflectively lucky, since her lack of good reflectively accessible grounds in favour of her chicken-sexing beliefs will mean that there will be near-by possible worlds on the subjective ordering in which she lacks her ability and so forms her belief on the same basis as in the actual world and yet believes falsely.¹⁶ The advantage of this distinction is that it can explain why someone like Zagzebski claims to be offering the real anti-luck epistemology, even though agent reliabilists like Greco also explicitly offer an anti-luck motivation for the view.

In any case, this feature of Zagzebski's position makes little difference to the central claims of my argument against virtue epistemology. Even with this additional demand on knowledge, it is still the case, as I showed above, that Zagzebski needs to appeal to a further non-virtue theoretic condition on knowledge in order to deal with the Gettier cases, which means that a strong virtue epistemological thesis is unsustainable even for her.

Moreover, the appeal to an anti-veritic luck condition in order to deal with Gettier cases will undermine the anti-luck motivation for the view as a whole, even considered as a

weak virtue epistemic thesis. After all, safety (or something like it), will deal with the veritic luck, and insofar as reflective luck can be eliminated at all (the sceptic would claim otherwise I think¹⁷), the most immediate way to eliminate it is via simply demanding good reflectively accessible grounds on the part of the subject. But note that one can make *that* demand without making any mention of virtue theory at all. It seems then that, just as with agent reliabilism, the anti-luck motivation for the view is not doing the job that it was meant to do.

With this outline of Zagzebski's argument—and my treatment of it—in mind we can return to consider some more specific charges laid down by Axtell. In correspondence, Axtell has usefully summarised his critical line in terms of the following two points. First, that I fail to recognise that virtue epistemological views, like Zagzebski's, are in a sense both epistemically externalist *and* internalist. Second, that I fail to recognize that one could advance both an anti-luck epistemology *and* a virtue epistemology. Neither point hits its target, and it is important to recognise why.

The first misses its target because in the original paper (as in this paper) I explicitly define the externalism and internalism as mutually excluding and exhaustive options, so there is no compatibilist picture available. Moreover, I motivate my drawing the distinction in this way, and show that it corresponds to our standard way of understanding this contrast. Oddly, Axtell doesn't in either of his papers offer an alternative non-standard account of the distinction, and neither does he explain how there could be a compatibilist view available in terms of my account. Accordingly, it isn't clear what to make of his objection. Note that I wouldn't deny that there can be a different ways of understanding the externalism/internalism distinction such that some views end up occupying some sort of middle ground—the cake can surely be sliced in more than one way (though some ways are better than others of course, and I would argue that my way of doing it is the right way to go). Nothing obviously follows from this for my argument though.

The second point is similarly defective. My argument shows that the required anti-luck condition that is needed to deal with Gettier-style cases is not going to be such that it must be understood in a virtue-theoretic fashion. Strong virtue epistemology is thus unavailable. Moreover, if the non-virtue-theoretic anti-luck condition does the work that the virtue-theoretic condition was supposed to do, as I claim, then merely combining the two conditions in the manner of a weak virtue epistemology is not going to be a very satisfactory option to take.

5. BECKER'S DEFENCE OF PROCESS RELIABILISM

There is, however, one line of response that has been made against my argument—due to Kelly Becker (*forthcoming*)—that I do think might have some force, and I want to close by considering this response here.

To begin with, we need to note that Becker's target in his paper is not my attack on virtue epistemology as such. He claims that my argument, if it works, will be just as effective against a process reliabilist view—one that defines knowledge in terms of reliable processes—and it is reliabilism of this sort that he wishes to defend in the light of my argument. Becker may be right that my argument has ramifications for process reliabilism as well as virtue epistemology, especially if that view is understood strongly in line with a strong reading of virtue epistemology such that knowledge is essentially and exclusively defined in terms of reliable processes. After all, if my argument undermines the idea that we must essentially appeal to epistemic virtues and cognitive faculties in our understanding of knowledge by showing that an anti-luck epistemology would do just as well (and be better motivated), then it would seem that this conclusion should 'spill-over' to impact on process reliabilism too. In any case, it strikes me that Becker's defence of process reliabilism may highlight a way in which virtue epistemologists can begin defending their view.

Becker's defence of process reliabilism turns on the observation that in order to get a formulation of safety that works, it is essential that one indexes safety to the belief-forming process actually employed. This is indeed how we formulated safety above, and we also saw that this aspect of the view naturally followed from our formulation of luck which required that the initial conditions for the event in question be fixed across the possible worlds under consideration. Accordingly, when talking about a non-lucky (and thus safe) true belief, it is important that one keeps the process that actually led to the formation of that belief fixed across possible worlds as well.

It is worth noting what odd results one would get if one dropped this restriction on safety. Consider the following example, adapted from one offered by Robert Nozick (1981, 179ff.). A grandmother has a highly reliable ability to tell when her grandson is well just by getting a good look at him (in good light and so forth). On this basis, she forms a true belief that he is well. There is a bug going around, however, which the grandson very nearly succumbed to. Had he fallen ill his parents would have kept him away from his grandmother and told her that he was fine. Furthermore, she would have believed them. Is the

grandmother's belief safe?

According to the version of safety adopted here it is, since given that she has this highly reliable ability to tell that her grandson is well, there won't be any near-by possible worlds where she forms the belief that he is well by this process and the belief is false. Without the restriction to processes, however, safety will not yield this result, since unrestricted it will follow that there *is* a wide class of near-by possible worlds where the grandmother believes that her grandson is well when he isn't (i.e., the ones where she forms her belief about his health via the testimony of her grandson's parents), and so her belief is unsafe. Clearly, however, the grandmother does have a safe belief in this proposition—and has knowledge of it to, for that matter—and so we should prefer the restricted version of safety over an unrestricted version.

Now Becker's thought is that if it is essential that we index safety to the process through which the belief in question was formed, then we are in effect talking about the safety of belief-forming processes here, rather than the safety of beliefs *simpliciter*. If that's right, though, suggests Becker, then isn't safety really just a way of spelling out a form of process reliabilism, rather than being an alternative to process reliabilism?

I think Becker is on to something here, and that the point he makes is not confined to process reliabilism. For just as the process reliabilist could claim that safety is, properly understood, a modal characterisation of process reliabilism, so too could the virtue epistemologist. Take reliabilism-based virtue epistemology first. The line of thought that I am exploring would hold that a safety-based approach to knowledge which indexes safety to belief-forming processes is simply offering a modal specification of the epistemic virtues and cognitive faculties that make up one's cognitive character. That is, that it is in the nature of one's epistemic virtues and cognitive faculties that they provide one with non-veritically lucky true beliefs in the propositions in question.

The obvious objection that one might raise to this proposal is to argue that there can be true beliefs that arise out of one's good cognitive character which are not safe, and safe true beliefs which do not arise out of one's good cognitive character. Accordingly, since these two notions come apart, safety cannot be plausibly thought to be a way of analysing epistemic virtue.

I think there is a problem here, although I don't think the problem facing this line is quite as bad as one might have initially thought. For example, is it really so clear that a true belief not formed via the cognitive traits that make up an agent's good cognitive character

could nevertheless be safe?

One possibility that immediately springs to mind is that of a true belief that was formed entirely on a whim. A true belief so formed, the idea goes, would not reflect one's good cognitive character, but could nevertheless be safe. On closer inspection, however, such examples start to look suspect. Setting aside for a moment why we might think such a belief could be a safe, is it really so clear that it is possible to form a *belief* on a whim? One can act on a whim, for sure, but that is a very different matter. Remember that beliefs are not occurrent states, but dispositional in character, so the idea of a belief being formed seemingly out of nothing in this way is deeply problematic. In any case, it is problematic enough for a virtue epistemologist to dismiss such cases as bad psychology.

Dismissing the idea that beliefs can be formed on the basis of whims leads us to consider whether we can make sense of the idea of there being *stable* belief-forming processes which gives rise to safe—and thus, presumably, reliable—true beliefs but which are *not* part of that agent's cognitive character? By agent reliabilist lights, at least, I don't see how such a distinction could be motivated. As Greco (2003, 356-7) concedes, even the belief-forming process that Alvin Plantinga (1993, 199) describes (which concerns a brain lesion which reliably gives rise to true beliefs) could count as part of one's good cognitive character if the belief-forming trait in question was a stable enough feature of the agent's cognitive character.

Of course, that's not to deny that there may be those virtue epistemologists, like Zagzebski, who think that there can be stable belief-forming traits on the part of the agent which do not form part of that agent's cognitive character by *her* lights. Indeed, even setting the brain lesion possibility to one side, we have already seen an example in this respect: that of the chicken-sexing belief-forming process employed by the chicken-sexer. As we noted above, the mere successful operation of a reliable cognitive trait will not suffice for knowledge by Zagzebski's lights, since it is essential that the agent also employs her epistemic virtues, more narrowly conceived, and thus has reflectively accessible grounds in favour of her beliefs.

Nevertheless, when it comes to internalist virtue epistemological theories like that offered by Zagzebski, the right question to ask is not whether there can be safe true beliefs which arise as a result of stable cognitive traits on the part of the agent but which do not arise out of the agent's good cognitive character. Instead, the question is rather whether there can be safe true beliefs that meet this description which are not only safe in the sense of not being veritically lucky, but also safe in the sense of not being reflectively lucky either, since only in

this way will we capture the internalism inherent in the view. With the question so posed, however, it is no more plausible to think that safety and good cognitive character come apart on this construal of safety and good cognitive character than it does on the construal of these notions that is in play when we consider reliabilism-based versions of virtue epistemology.

Nevertheless, while the foregoing remarks indicate that there is no obvious problem with this direction of fit between safe true belief and true belief that arises out of good cognitive character, the opposing direction of fit does pose a problem. In particular, it does seem that there are cases where an agent forms a true belief as a result of her good cognitive character, and yet the belief so formed is not safe.

As we noted above, virtue epistemology faces the problem—key to my critique of the view—that it is unable to deal with the Gettier cases, and these cases precisely fit the bill in this respect. Think again of Mary’s true belief that her husband is in the room, formed out of epistemic virtue and yet Gettierized. This is a clear example of a true belief that is formed as a result of good cognitive character, and yet which is not safe.

The prospects for applying Becker’s proposal to virtue epistemology thus hinge on whether virtue epistemologists are able to offer the response to the Gettier problem that they claim to be able to offer. I’m sceptical on this score as I have indicated above, but at least it is now clear, I take it, just how much weight the virtue epistemologist’s response to the Gettier problem carries, since it is only with an answer to this problem in hand that it can adequately evade my argument against it.¹⁸

REFERENCES

- Axtell, G. (2001). ‘Epistemic Luck in Light of the Virtues’, *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*, (eds.) A. Fairweather & L. Zagzebski, 158-77, Oxford: Oxford University Press.
- (2003). ‘*Felix Culpa*: Luck in Ethics and Epistemology’, *Metaphilosophy* 34, 331-52.
- (Forthcoming). ‘*Felix Culpa* Revisited: Two For the Show’, *Synthese*.
- Becker, K. (Forthcoming), ‘Reliabilism and Safety’, *Metaphilosophy*.
- Engel, M. (1992). ‘Is Epistemic Luck Compatible with Knowledge?’, *The Southern Journal of Philosophy* 30, 59-75.
- Greco, J. (1999). ‘Agent Reliabilism’, *Philosophical Perspectives* 13, 273-96.
- (2000). *Putting Sceptics in Their Place: The Nature of Skeptical Arguments and Their Role in Philosophical Inquiry*, Cambridge: Cambridge University Press.
- (2003). ‘Virtue and Luck, Epistemic and Otherwise’, *Metaphilosophy* 34, 353-66.
- (Forthcoming). ‘Worries about Pritchard’s Safety’, *Synthese*.
- Nozick, R. (1981). *Philosophical Explanations*, Oxford: Oxford University Press.

- Plantinga, A. (1993). *Warrant and Proper Function*, New York: Oxford University Press.
- Pritchard, D. H. (2003). 'Virtue Epistemology and Epistemic Luck', *Metaphilosophy* 34, 106-30.
- (2005a). *Epistemic Luck*, Oxford: Oxford University Press.
- (2005b). 'Scepticism, Epistemic Luck and Epistemic Angst', *Australasian Journal of Philosophy* 83, 185-206.
- (2006). 'Knowledge, Luck, and Lotteries', *New Waves in Epistemology*, (eds.) V. F. Hendricks & D. H. Pritchard, Aldershot: Ashgate.
- Sosa, E. (1991). 'Intellectual Virtue in Perspective', *Knowledge in Perspective: Selected Essays in Epistemology*, Cambridge, UK: Cambridge University Press.
- (1999). 'How to Defeat Opposition to Moore', *Philosophical Perspectives* 13, 141-54.
- Umbers, R. (2005). 'Non-Random Knowledge in Virtue Epistemology', *manuscript*.
- Zagzebski, L. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*, Cambridge, UK: Cambridge University Press.
- (1999). 'What is Knowledge?', *Epistemology*, (eds.) J. Greco & E. Sosa, 92-116, Oxford: Blackwell.

NOTES

¹ For two other examples of commentators who use epistemic luck to motivate their virtue epistemology, see Zagzebski (1996, *passim*) and Axtell (2001).

² Very roughly, one could class Sosa (1991) and Plantinga (1993) as offering agent reliabilist theses, though I don't doubt that they would resist this characterisation of their views (Plantinga in particular).

³ This is not an example that Greco (1999) offers, but it is structurally similar to the evil demon case that he describes (see 286). I have chosen to use this example solely because it has the advantage of not appealing to anything like an evil demon to make the required point.

⁴ Zagzebski was actually directing this example against Plantinga's (1993) related 'proper functionalism' theory, although the differences between proper functionalism and agent reliabilism are not important in this regard. Indeed, Greco (2000, 251) himself discusses this example and notes that it is just as much a problem for his view as for Plantinga's.

⁵ Note that the worlds are here being ordered in the usual way in terms of their similarity to the actual world. For more on this account of luck (including what needs to be added in order to deal with some problem cases), and its application to epistemology, see Pritchard (2005a, part two).

⁶ For obvious reasons, this account will only apply to beliefs in contingent propositions. In order to keep matters as simple as possible, in what follows I will take it as given that the target proposition is always fully contingent (i.e., not necessary in *any* sense, whether nomological, physical, metaphysical, etc.).

⁷ Notice that it is essential that Greco's second point is allied with the first if it is to do any major work undermining my argument, since without it his response will only *at best* weaken that part of my argument which is directed at whether weak virtue epistemology is adequately motivated. After all, the correctness of this second point leaves my claim that strong virtue epistemology is false completely untouched.

⁸ For more on this intermediate conception of safety, see Pritchard (2006).

⁹ Notice that I'm not suggesting here that cognitive character needs to be understood in such a way that it is completely independent of all environmental factors (indeed, I think this is false). The point is rather that the changes in the environment that we encounter as we move from the example of Mary to that of Mary* seem to have no impact at all on either agent's cognitive character.

¹⁰ Sensitivity is the main modal condition imposed on knowledge by Nozick (1981, chapter 3).

¹¹ That said, safety is in fact much more effective at eliminating epistemic luck more generally; it is also better motivated and not subject to the same kinds of problems. For more discussion of the relative merits of safety and sensitivity, see Pritchard (2005a, §6.3).

¹² Interestingly, in her discussion of this 'because of' clause, Zagzebski (1999, §5.B) explicitly grants that a natural gloss of what this condition demands can be captured by the sensitivity condition, and that this is probably the best account available of this condition. She goes on to argue, however, that strictly speaking the gloss cannot be right. The reason she gives for this is that such counterfactual analyses cannot capture other uses of the 'because of' clause, such as the claim that X is a bachelor because X is unmarried. The reason the analysis doesn't work in this case, however, is clearly that there are no relevant counterfactuals. The failure of the

analysis in this case therefore has no obvious implications for uses of the clause in cases where there are counterfactuals, as in the cases that we are interested here where we are trying to offer an analysis of knowledge of contingent propositions.

¹³ A third response to my argument that takes this general line can be found in Umbers (2005), but since this is an unpublished manuscript, I won't comment on the specifics of this paper here.

¹⁴ Note that, confusingly, Zagzebski often refers to her view as a form of epistemic externalism, on the grounds that epistemic virtues as she understands them are of their nature reliable. That is, Zagzebski treats internalist views as being those positions which deny that there are any external epistemic conditions on knowledge, with externalist views demanding such conditions (perhaps in conjunction with an internal epistemic condition, like a justification condition, internalistically conceived). While Zagzebski's view is no doubt more 'externalist' than an extreme form of epistemic internalism which held that knowledge was *simply* true beliefs backed up by good reflectively accessible grounds, it is still an internalist thesis by the lights of the way that this distinction is drawn here (which is, I believe, a standard way of drawing this distinction). In any case, notice that the way Zagzebski draws the internalist/externalist distinction is such that hardly anyone in the literature endorses internalism (at least post-Gettier), since hardly anyone thinks that true belief that meets an internal epistemic condition will suffice for knowledge. Since the distinction doesn't capture the one side of the debate, it doesn't capture the debate.

¹⁵ For more on this distinction, see Pritchard (2005a, part two; 2005b).

¹⁶ Notice that one also needs to individuate the process in which the belief was formed differently on the objective and subjective orderings. In the case of the former, one individuates it in the usual way in terms of the actual belief-forming process used, while in the case of the latter one needs to individuate it in terms of the process that the agent *thinks* was used. For more discussion of this point, see Pritchard (2005a, chapter 6).

¹⁷ I relate reflective luck to the sceptical problem in Pritchard (2005a, chapters 8-9; 2005b).

¹⁸ Thanks to Guy Axtell, Kelly Becker, John Greco, and Richard Umbers. An earlier version of this paper was given as part of an invited symposium at the 2006 Pacific APA in Portland, Oregon. Thanks to my commentator—Jim Montmarquet—to the other participants in this session—Robert Roberts and Jay Wood—to the chair—Abrol Fairweather—and to the organiser—Heather Battaly. Finally, thanks also to the audience, especially Jason Baehr, Mylan Engal, Richard Greene, and Ken Lucey.