

VIRTUE EPISTEMOLOGY AND EPISTEMIC LUCK

DUNCAN PRITCHARD

ABSTRACT: The recent movement towards virtue-theoretic treatments of epistemological concepts can be understood in terms of the desire to eliminate epistemic luck. Significantly, however, it is argued that the two main varieties of virtue epistemology are responding to *different* types of epistemic luck. In particular, whilst proponents of reliabilism-based virtue theories have been focusing on the problem of what I call “veritic” epistemic luck, non-reliabilism-based virtue theories have instead been concerned with a very different type of epistemic luck, what I call “reflective” epistemic luck. It is argued that, *prima facie* at least, both forms of epistemic luck need to be responded to by any adequate epistemological theory. The problem, however, is that one can best eliminate veritic epistemic luck by adducing a so-called safety-based epistemological theory that need not be allied to a virtue-based account, and there is *no* fully adequate way of eliminating reflective epistemic luck. I thus conclude that this raises a fundamental difficulty for virtue-based epistemological theories, on either construal.

Keywords: epistemology, Gettier, luck, reliabilism, responsibilism, virtues.

Introduction

Much of the recent work in epistemology has been concerned with the possibility of offering virtue-theoretic accounts of our key epistemic concepts – in particular, justification and knowledge.¹ In essence, such views mark a break with traditional epistemological analyses by defining the key epistemic concepts in terms of the epistemic virtues and/or faculties of the agent (rather than *vice versa*).² As we shall see, however, this basic conception of virtue theory is open to a number of divergent interpretations, and two interpretations in particular that seem almost diametrically

¹ For some of the recent work on virtue epistemology, see the essays in the anthologies edited by Axtell (2000), Fairweather and Zagzebski (2001), DePaul and Zagzebski (2002), and Steup (2001) and the articles in this collection. See also the excellent survey article by Axtell (1997).

² A slightly weaker virtue-based thesis might hold that the epistemic virtues merely present the best *criteria* for determining positive epistemic status, although it is difficult to identify anyone who clearly endorses such a view. Since the objections that I raise against virtue theory in this article are equally applicable on either interpretation, for ease of expression I shall focus on the stronger and more standard rendering of the position.

opposed in terms of what they are trying to achieve in advancing this type of thesis. I argue that the key to understanding this divergence is to recognise that whilst both these types of account of virtue theory regard the elimination of epistemic luck as being a fundamental adequacy condition on their theory, each is focusing upon a *different* species of epistemic luck.³ Simply noting this fact, however, does not suffice to resolve the debate, since, as we shall see, there are further complications at issue here.

In section 1 I describe how the first renderings of virtue epistemology that appeared in the contemporary literature were primarily understood in terms of extant versions of reliabilism. In section 2 I argue that these reliabilism-based versions of virtue epistemology are responding to the problem of what I call “veritic” epistemic luck. Section 3 offers a critique of these reliabilism-based virtue theories on the grounds that there are non-virtue-based reliabilist theories which are equally adept at eliminating veritic luck and that, in any case, the best accounts of knowledge available which respond to the problem of veritic epistemic luck make no essential mention of the epistemic virtues at all. Section 4 introduces the problem of what I call “reflective” epistemic luck, and it explains how one can understand non-reliabilism-based responsibilist virtue epistemological theories in terms of the desire to eliminate this variety of epistemic luck. I also argue, however, that there is no fully satisfactory response to the problem of reflective epistemic luck, and thus that this form of virtue epistemology is in danger of defining knowledge in such a way as to make it unattainable. Section 5 offers some concluding remarks.

1. From Simple Process Reliabilism to Agent Reliabilism

For our purposes, we can take simple process reliabilism to consist, in essence, in the thesis that one has knowledge if, and only if, one forms one’s true belief via a reliable process, where this is a process that ensures a high ratio of true beliefs relative to false beliefs (we shall consider later how a more elaborated version of process reliabilism might go). Such a view is, of course, usually associated with the work of Alvin Goldman.⁴ In terms of rather “low-grade” knowledge, such as basic perceptual knowledge, this is a fairly plausible account. Intuitively, such beliefs count as knowledge just so long as they are formed in reliable ways, which, as the products of our perceptual faculties in normal circumstances, we would

³ Although others have commented on the relationship between virtue epistemology and epistemic luck, no-one, so far as I am aware at any rate, offers anything like the account I present here. For the best discussions in this regard, see Zagzebski (1996, *passim*), Axtell (2002) and Greco (2002).

⁴ See, for example, Goldman (1976; 1979; 1986). See also Armstrong (1973) for a contemporaneous account of a similar theory, and Talbot (1990) for a recent defence of a view of this sort.

expect them to be. Nothing more seems to be necessary to knowledge than meeting such a condition.

However the details are to be spelt out, it should be clear that simple process reliabilism is a paradigm example of an epistemologically externalist thesis, in that it identifies epistemic conditions which are necessary and sufficient for knowledge possession but does not further demand that the agent should be in a position to know, by reflection alone, that these conditions have obtained. More precisely, we can define epistemological internalism as the view that internalist justification is necessary for knowledge, and further define internalist justification as follows:

For all agents, ϕ , an agent's belief in a proposition, ϕ , is internalistically justified if, and only if, the facts which determine that justification are knowable by the agent via reflection alone (that is, through a priori reasoning, introspection of her own mental states, or memory of knowledge gained via either of these means).⁵

I shall henceforth take epistemological externalism about knowledge to consist in the denial of epistemological internalism about knowledge, and thus in the rejection of the thesis that internalist justification is necessary for knowledge. Of course, any characterisation of so contentious a distinction is bound to be at least partly stipulative, but the advantage of this formulation is that it captures what it is about simple process reliabilism that makes it an externalist thesis without committing proponents of simple process reliabilism to an unnecessarily austere rendering of the view. On this model, simple process reliabilists are externalists in that they allow that agents might have knowledge simply by employing reliable processes. It is not further demanded that agents should possess internalist justification for their beliefs. But neither is it excluded that internalist justification might play a fundamental – albeit, perforce, *secondary* – role in such a theory, perhaps as being an important and useful (though inessential) property for one's beliefs about certain matters to have.⁶

As a representative of the externalist position, simple process reliabilism faces the usual kinds of objections levelled against epistemological theses of

⁵ In both this characterisation of epistemological internalism and reflective access, I follow the account given by Pryor (2001) in his extremely useful survey of recent trends in epistemology.

⁶ For a similar conception of the epistemological externalism/internalism distinction, see Brandom (1998). For discussion of the contrast between internalist and externalist notions of justification, see Goldman (1988). Note that a further advantage of this way of construing the externalism/internalism distinction is that it does not commit internalists to an unnecessarily strong rendering of their position either. Internalists can perfectly well allow that there should be external conditions on knowledge over and above the truth condition. The crux for them will only be that the agents in question have internalist justification for what they believe. For more on the epistemological externalism/internalism distinction, see the essays collected in the excellent anthology edited by Kornblith (2001).

this sort, and we shall consider a sample objection of this type below. Those who see in simple process reliabilism the basics for a more developed reliabilist theory understood along virtue-theoretic lines are not concerned about its association with externalism, however, since the revised reliabilist view that they envisage is also an externalist thesis. Instead, their dissatisfaction with simple process reliabilism relates to how it is prone to certain key types of counterexample. Of the main proponents of such a view, John Greco is, perhaps, the most explicit about this, arguing that his reliabilism-based version of virtue epistemology – what he calls “agent” reliabilism – can be motivated by how it is able to respond, in a reliabilist fashion, to various basic problems facing the simple process reliabilist view.⁷ Accordingly, we shall here focus on his view, but the conclusions that are offered should equally apply (*mutatis mutandis*) to other views of this general sort, especially that propounded by Ernest Sosa (e.g., 1985; 1991; 1993).

The problems that Greco is primarily concerned with are those that involve what he calls “strange and fleeting processes.” Different examples of such processes raise different issues. First, consider the case where a strange and fleeting process is reliable but where the reliability in question is due to the world’s tracking the agent’s beliefs rather than vice versa. For example, suppose that there is a benevolent demon who ensures that every time our protagonist forms a belief the world is adjusted to make it such that the belief is true. Clearly, this would be a highly reliable way of forming beliefs, since it would never fail to result in a true belief. Nevertheless, our intuition in such a case is that the agent lacks knowledge because her reliably formed true beliefs do not reflect a cognitive achievement on her part at all. Greco describes just such an example:

René thinks he can beat the roulette tables with a system he has devised. Reasoning according to the Gambler’s Fallacy, he believes that numbers which have not come up for long strings are more likely to come up next. However, unlike Descartes’ demon victim, our René has a demon helper. Acting as a kind of epistemic guardian, the demon arranges reality so as to make the belief come out as true. Given the ever present interventions of the helpful demon, René’s belief forming process is highly reliable. But this is because the world is made to conform to René’s beliefs, rather than because René’s beliefs conform to the world. (Greco 1999, 286)

In order to ensure that the reliable process at issue here is strange and fleeting, we need to stipulate further that in most near-by possible worlds the demon does not intervene in this way (perhaps because he does not exist in these worlds).⁸

⁷ See, for example, Greco (1999; 2000, chapters 7 and 8).

⁸ Strictly speaking, then, it is the *reliability* of the process, rather than the process itself, that is strange and fleeting. Greco himself is not always clear about this point. As we shall see below, however, this issue is important because it leaves open a possible response to “counterexamples” of this sort that an exponent of a more refined version of process reliabilism can make.

Second, consider the case where a strange and fleeting process is reliable, but where the reliability is due to a “malfunction” on the part of the agent. An example of this sort, offered by Alvin Plantinga (1988; 1993a), concerns a rare brain lesion that causes the victim to believe that he has a brain lesion. Greco (1999, 285; 2000, 175) quotes Plantinga, who describes this scenario as follows:

Suppose . . . that S suffers from this sort of disorder and accordingly believes that he suffers from a brain lesion. Add that he has no evidence at all for this belief: no symptoms of which he is aware, no testimony on the part of physicians or other expert witnesses, nothing. (Add if you like, that he has much evidence *against* it; but then add also that the malfunction induced by the lesion makes it impossible for him to take appropriate account of this evidence). Then the relevant [cognitive process] will certainly be reliable; but the resulting belief – that he has a brain lesion – will have little by way of warrant for S. (Plantinga 1993a, 199)

Again, in order to ensure that the reliability of this process is strange and fleeting in the relevant respect, we need to stipulate further that in most near-by possible worlds this brain lesion would not result in the agent’s forming a belief that he has the brain lesion (and thus would not result in a true belief). We have a strong intuition that there is something epistemically amiss about forming true beliefs via malfunctions in this way, even where those malfunctions happen to support a process of forming beliefs that is reliable.⁹

The basic idea behind the kind of agent reliabilism advanced by Greco – and similar such theories advanced by Sosa, the later Goldman (1993), and others¹⁰ – is to restrict the range of reliable processes that can support knowledge in a principled way by focusing upon the stable and successful dispositions of the agent that are relevant to knowledge and justification – her intellectual *faculties*, if you will. For example (the details need not detain us here), our faculty of sight, if working properly and applied in the right conditions, is a highly reliable way of forming true beliefs about the world. In contrast, the reliability that might attach itself to a cognitive malfunction will not count as knowledge producing on this view, because

⁹ It is significant, of course, that in neither of these examples does the agent have counterevidence which would lead him to believe that the process in question is unreliable. This is because most extant versions of process reliabilism incorporate some extra condition to the effect that if the agent has such a belief then he lacks justification for that belief (and thus knowledge of what is believed). Goldman (1986) is a prime example here. He further adds that the agent should not be in a mental state (such as being in the possession of relevant evidence) from which reliable processes would lead him to conclude that the process at issue is unreliable.

¹⁰ Of course, there are important differences between these views, but they are not relevant to the current discussion. A related view in this respect is Plantinga’s (1988; 1993b; 1993c) “proper functionalism,” although Plantinga has explicitly resisted the description of his view as a virtue-based theory. For an overview of proposals of this sort, see Axtell (1997, especially §2).

such a cognitive process is not one of our natural faculties and this means that it lacks the kind of dispositional stability required. Similarly, forming beliefs in a way that is reliably successful for reasons that are independent of one's cognitive endeavours, as is the case in the "benevolent demon" example, is also ruled out, because this reliable process does not concern a faculty of the agent at all, in any sense of that term, and so is not even in the ballpark to be considered a stable and successful intellectual faculty. The shift from simple process reliabilism to some form of faculty or agent reliabilism thus enables reliabilists to meet two core objections to their view.

2. Veritic Epistemic Luck

One can regard this transition from simple process reliabilism to agent reliabilism as motivated in terms of the need to eliminate a certain form of luck regarding the truth of what one believes. We shall call this type of luck "veritic" luck, and characterise it as follows:¹¹

For all agents, ϕ , the truth of an agent's belief in a proposition, ϕ , is veritically lucky if, and only if, the agent's belief that ϕ is true in the actual world but false in nearly all near-by possible worlds in which the belief is formed in the same manner as in the actual world.

For example, take my entirely ungrounded guess that Lucky Lass will be the winner of today's derby. Suppose, further, that I go on to form the belief that this horse will win, and that the belief turns out to be true (that it's a "lucky guess").¹² Intuitively, to say that this is a lucky guess is to say that although it turned out to be correct, in nearly all near-by possible worlds it would have been false. The belief that results from the guess is thus veritically lucky in the sense just described. It is clearly the task of a theory of knowledge to disallow beliefs formed in this kind of fashion from enjoying the privileged status of being instances of knowledge. If one meets all the relevant epistemic conditions set down by the favoured epistemological theory and one's belief is still veritically lucky, however, then what we have here is a case of veritic *epistemic* luck and thus, prima facie at least, a counterexample to the theory of knowledge in question.¹³ It is

¹¹ I borrow the term from Engel (1992), who argues for a similar characterisation of epistemic luck, although he specifically puts the point in terms of evidence. For further discussion of Engel's view, see Hall (1994), Harper (1996), Vahid (2001), Axtell (2002) and Pritchard (2003; cf. Pritchard and Smith 2002).

¹² It should be noted that a guess that p rarely does result in a belief that p , though we need not concern ourselves with this complication here.

¹³ Of course, there will no doubt be a lot more involved in the notion of luck than this, though the details need not detain us here. I discuss the notion of luck at more length in Smith and Pritchard 2002, and offer a fuller analysis of veritic epistemic luck (and the "reflective" epistemic luck that I discuss below) in Pritchard 2003. For more on the notion of luck, see Rescher (1995).

this type of luck that is famously at issue in the types of counterexamples to the classical tripartite account of knowledge associated with Edmund Gettier (1963).

Consider the following example adapted from one given (though in a different regard) by Bertrand Russell (1948, 170–71). Our protagonist comes downstairs every morning about the same time and looks at the time on the old clock in her hall. The clock has been a highly reliable timepiece now for many years, and she has no reason for thinking that it is faulty this morning. The clock tells her that it is 8:22, and it is indeed usually around about 8:20 A.M. that she comes downstairs. Furthermore, the clock is right, because it *is* 8:22 A.M. Nevertheless, unbeknownst to the agent, the clock has stopped. Given that, intuitively, one cannot acquire knowledge of the time by looking at a stopped clock, the agent in this example lacks knowledge, even though she has a justified belief (and thus, seemingly, has met all the relevant epistemic conditions in this regard) and even though her belief is in fact true.

The type of epistemic luck that is at issue here is veritic in that the problem which an example like this highlights is that, on the classical tripartite model, meeting the relevant epistemic conditions does not eliminate the possibility that one's belief is only luckily true. More precisely, whilst our protagonist has met all the relevant epistemic conditions, and whilst her belief is true, it nevertheless remains that in nearly all near-by possible worlds in which she forms her belief in the same manner as in the actual world (by looking at her clock), she will form a false belief as a result.

As Linda Zagzebski (1994; 1996, §3.1; 1999) points out, what is common to all Gettier-type counterexamples is that they involve a mix of “double luck.” First, there is the “bad” luck that would usually prevent the agent from forming a true belief, even despite the fact that the agent has met all the relevant epistemic conditions (that is, that the agent's normally reliable clock has stopped). Second, there is the “good” luck that cancels out the bad luck by ensuring that, even despite this epistemic misfortune, the agent forms a true belief regardless (that is, that the agent happens to look at the stopped clock when it is telling the right time). Moreover, notice that this double-luck method for constructing Gettier counterexamples is only possible because the epistemic conditions at issue allow the existence of veritic luck in the first place. That is, if these conditions excluded such luck – such that, in so far as one met all the required epistemic conditions and had a true belief then one also had a true belief in nearly all near-by possible worlds where one formed that belief in the same manner as in the actual world – then it would not be possible to be a victim of bad luck in the relevant sense in the first place (that is, bad luck that would ensure that one's belief would *normally* be false).

Like the Gettier counterexamples, the putative counterexamples to the simple process reliabilist account that we have just considered also work by highlighting how this account of knowledge is consistent with veritic

luck. In the benevolent-demon example, for instance, that the agent meets the epistemic conditions of having employed a reliable process does not ensure that his true belief is not lucky. After all, in most near-by possible worlds in which he forms the same belief via the same process (the Gambler's Fallacy), his beliefs will turn out as false because in these worlds there is no demon helping him out. The same goes for the "malfunction" case. Again, although the agent meets the relevant epistemic condition because his belief is formed as a result of what is, in fact, a reliable process of being caused by the brain lesion itself, his true belief is nevertheless lucky in that in most near-by possible worlds this brain lesion would not have resulted in a true belief in this way.

Moreover, these examples are also, in the relevant sense, Gettier-type examples, in that they both involve a kind of double luck. Despite meeting the required epistemic conditions (forming his belief in a reliable fashion), the agent in the benevolent-demon example would normally have formed a false belief, and yet this bad luck is cancelled out by the good fortune of being assisted by a benevolent demon who nevertheless ensures that his belief is true. Similarly, in the malfunction case one would normally expect the agent's belief to be false, and yet this bad luck is cancelled out by the good luck that this particular brain lesion happens to generate true beliefs of the required sort.

Accordingly, one can regard the introduction of stable cognitive dispositions into the reliabilist theory as devices to limit the degree of veritic luck so as to avoid examples like these. For what makes such dispositions relevant in this regard is that they enable the agent to form true beliefs not just in the actual world but also in nearly all near-by possible worlds in which the agent forms her beliefs in the same way as in the actual world. Thus, if the agent's belief in the actual world is formed via a cognitive faculty in this way, then one would expect that the truth of the belief formed will not be a matter of veritic luck.

In responding to these examples, then, the agent reliabilist is trying to eliminate possible ways in which veritic luck can creep into the simple process reliabilist account of knowledge, and in this sense the intuitions that are being appealed to are the same anti-luck intuitions that are at work in the Gettier counterexamples. Agent reliabilism – at least as Greco understands the thesis – can thus be regarded as one way in which one can meet a particular type of species of Gettier-type counterexample that applies to simple process reliabilism by eliminating the veritic luck that is at issue in these examples.

3. Modalized Process Reliabilism and the Safety Principle

Matters are not quite as straightforward as this presentation of the development of agent reliabilism out of simple process reliabilism suggests, however. The reason for this is that there is a more nuanced version of

process reliabilism available that “modalizes” the construal of what counts as a reliable process in such a way that it can deal with these two examples just as adequately as agent reliabilism can. In effect, what the modalized process reliabilist account achieves is to rule out veritic luck to roughly the same extent as agent reliabilism does.

Recall that the putative counterexamples offered against simple reliabilism only worked because it was stipulated that the reliability at issue was strange and fleeting. A crude process reliabilism that simply worked by analysing the reliability of the process at issue in terms of the actual truth-to-falsity ratio of that process is susceptible to such counterexamples because actually reliable processes can nevertheless be strange and fleeting ones, in that their reliability is unstable. For whilst it is true in the benevolent-demon world that our agent always forms true beliefs via the relevant process, in most near-by possible worlds he will (by stipulation) form false beliefs via this method, and thus the “reliability” of the process will disappear. The same goes for the brain-lesion example, at least as it was understood above. For whilst this brain lesion will in the actual world support true beliefs, in most near-by possible worlds this will not be the case. That the reliability at issue is strange and fleeting is thus a result of how it is a reliability that is not sustained in most near-by possible worlds. In contrast, the kind of stable cognitive dispositions that Greco and other agent reliabilists have in mind are defined such that they will retain their reliability in nearly all near-by possible worlds (this is what is meant by calling these dispositions “stable”).

There is no essential reason why process reliabilism should restrict itself to a non-modalized account of reliability, however, and, indeed, there is every reason to think that it *should not* confine itself, given the problems just mentioned coupled with the fact that the very notion of reliability seems to be a partly modal one. Intuitively, a “reliable” thermometer, for example, is not one that just happens to tell the temperature well in the actual world but one that is similarly accurate in nearly all near-by possible worlds as well.¹⁴ Indeed, Goldman himself has expressed the process reliabilist point in counterfactual terms, arguing for what he calls a “normal worlds” reliabilism, which demands that a belief (in any possible world) is justified just in case the process by which it is generated is reliable in normal worlds, where such worlds are in turn understood as those worlds that are consistent with our general beliefs about the actual world.¹⁵

¹⁴ Of course, there is a respectable sense of “reliability” in which the thermometer can be reliable even if it fails to tell the temperature in the actual world or any near-by possible worlds, such as when the environment is radically different from the environment that the thermometer was designed for. In such cases, however, the reliability of the thermometer consists in the fact that, in a world for which it was designed, it tells the temperature in that world and nearly all near-by possible worlds as well.

¹⁵ See Goldman (1986). Goldman (1976) also argues for a counterfactual understanding of the reliabilist position. In recent work, Goldman has moved away from such a position.

On this construal of what counts as a reliable process, the counterexamples at issue above are defused. The Gambler's Fallacy is not a reliable way of forming beliefs in normal worlds, and neither (by stipulation) is the malfunction-based process at issue in the brain-lesion example. Suitably modalized, therefore, process reliabilism is not subject to the critique levelled against it by the agent reliabilist.

Interestingly, what the modalized version of process reliabilism achieves is to restrict the range of processes at issue so that they do not include the ones consistent with veritic luck in the manner of our two examples. In both cases, the veritic luck at issue depends upon the fact that, normally, forming beliefs via the actual manner in which they are formed would lead to falsity rather than truth. By making the notion of reliability at issue relative to a range of normal possible worlds, this variety of process reliabilism ensures that it is not possible for the agent's true belief to meet the relevant epistemic conditions in such a way that, nevertheless, the belief would normally have been false. On this matter, then, modalized process reliabilism is on a par with agent reliabilism. So why should one prefer agent reliabilism over process reliabilism?

This problem becomes further accentuated if one considers how each of the theories fares against the examples just discussed given that we drop the stipulation that the reliability at issue is incidental (and thereby strange *and* fleeting). Consider first the benevolent-demon example. Modalized process reliabilism can accommodate a suitably adapted version of this example just as well as agent reliabilism can. Suppose, for example, that the demon in question necessarily exists and essentially has this property of being benevolent in this way. In such circumstances, it is not an incidental feature of the example that the agent in question forms the true beliefs that he does, since he would form true beliefs via this method in all near-by possible worlds as well. Modalized process reliabilism handles an example like this by claiming that although this process is in fact reliable in the actual world and in most near-by possible worlds, it is not reliable in normal worlds, and so cannot support knowledge or justification. In contrast, the agent reliabilist copes with this sort of example by arguing that the Gambler's Fallacy is just not a virtuous method by which one can gain knowledge and that, in any case, the method at issue here does not reflect a stable and reliable disposition *on the part of the agent* for acquiring true beliefs (the reliability is due to the demon, remember, rather than the agent).

It is when it comes to the brain-lesion example that the accounts come apart, or at least seem to. Suppose that this type of brain lesion *always* produces true beliefs of this sort and that, furthermore, this fact is subsequently discovered by scientific inquiry. On this construal, the brain lesion supports not just a process that mainly results in true beliefs in the actual world but also one that would support such a reliable process in most normal possible worlds in which it occurs as well. What would we then say

about the agents who (lacking this scientific knowledge) believed that they had a brain lesion?

On a radical externalist view of knowledge, the temptation would be to accord them knowledge and, indeed, argue that the brain lesion does not represent a cognitive malfunction at all but rather a bona fide successful cognitive process that had erstwhile gone unnoticed. Moreover, on the supposition that the brain lesion always produces true beliefs, the response employed by the agent reliabilist loses its cogency. For what is now preventing such a process from counting as a stable cognitive faculty?

Reliabilists are used to problems like this, and although there may be qualifications that can be added to the thesis to evade this particular example, it is clear that reliabilists are going to be committed to allowing that *some* type of example of this sort will go through (though perhaps one that is not quite so controversial – we shall consider a possible scenario below). Greco's response to the problem is not so sanguine, however. Instead, he argues that what is lacking in such cases is "subjective justification," which he defines as follows:

A belief *p* is subjectively justified for a person *S* . . . if and only if *S*'s believing *p* is grounded in the cognitive dispositions that *S* manifests when *S* is thinking conscientiously. (Greco 1999, 289; cf. Greco 2000, §7.II)

On the face of it, this seems to rule out the kind of cases under consideration because the belief in question is clearly not grounded in cognitive dispositions that the subject manifests when she is thinking conscientiously. The devil, however, is in the detail concerning what counts as "thinking conscientiously." Greco characterises such thinking in terms of the good hitting of a baseball player, in that one can manifest the relevant dispositions in the appropriate conditions without thereby having beliefs (true or otherwise) about the nature and character of these dispositions (as he neatly puts it, one can be a good baseball player without thereby being a good coach). Accordingly, all that is being demanded here is a doxastic sensitivity to the reliability in question. If this is the case, however, then it seems that Greco's view *does* succumb to brain-lesion-type examples after all, since in such cases agents will, *ex hypothesi*, form their true beliefs in ways that will be sensitive to the reliability in question. At the very least, we need to be given a clearer idea of why agents in such examples are not subjectively justified in forming the beliefs that they do.

Of course, a natural response to this line of objection – one that I take it Greco wants to avoid – is simply to beef up the notion of subjective justification in such a way as to demand that agents have some degree of internalist justification for their beliefs. So understood, brain-lesion-type examples are easily avoided, because agents clearly lack grounds for their belief that would support internalist justification. But note that this move represents not merely a modification of the reliabilist position but rather a

rejection of it, in that it involves incorporating a necessary internalist condition on knowledge and thus entails an endorsement of epistemological internalism.

So, in so far as one remains within a reliabilist framework (and we shall consider in a moment how virtue epistemology is developed when it is not constrained in this way), there is little to tell between an agent reliabilist theory and a modalized process reliabilist position. Crucially, however, neither modalized process reliabilism nor agent reliabilism can exclude all types of veritic luck. In particular, as both these theories stand they have difficulties with the kind of veritic luck at issue in Gettier counterexamples. Since the reliability in question in both cases is consistent with the production of a false belief, it is possible to formulate the kind of double-luck examples that we saw Zagzebski arguing above were characteristic of Gettier counterexamples. The “stopped-clock” example discussed earlier cannot be used here because, at the very least, modalized process reliabilism can meet this problem, since employing a stopped clock to form beliefs is not a reliable way of acquiring true beliefs in normal possible worlds. How agent reliabilism would respond to this example is not altogether clear, but for the sake of parity we shall consider a different example that has application to both agent reliabilism and modalized process reliabilism.

Here is a scenario proposed by Zagzebski that will serve our purposes:¹⁶

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, 285–87; emphasis in the original)

As regards agent reliabilism, we have here the reliable functioning of a faculty that is producing a true belief, but there is nevertheless the kind of veritic double luck involved that prevents the true belief formed from being an instance of knowledge. Moreover, as regards modalized process reliabilism, we also have a process that is reliable in normal possible

¹⁶ Zagzebski was actually directing this example against Plantinga’s related proper-functionalism theory, although the differences between proper functionalism and the agent/faculty reliabilist theory are not important in this regard. Indeed, Greco himself discusses this example and notes that it is just as much a problem for his view as it is for Plantinga’s. See Greco (2000, 251).

worlds and is generating a true belief in the actual world but cannot support knowledge because of the veritic double luck that is in play.

It is far from obvious how one amends the modalized process reliabilist or agent reliabilist accounts to meet veritic luck of this sort (though there have been some notable attempts).¹⁷ Crucially, however, whatever amendment one comes up with will be tantamount to the demand that knowledge is true belief formed in a non-vertically lucky fashion, where this means not only that the belief should be true but also that it remains true in nearly all near-by possible worlds where the belief is formed on the same basis. In this way, one is left with an account of knowledge that can accommodate the example just cited. After all, there will be a number of near-by possible worlds where Mary forms her belief in the same way but where her husband is not in the room and thus where the belief formed is false.

The type of anti-veritic-luck epistemology that is at issue here is one that is based around something like the following “safety” principle for knowledge:

For all agents, ϕ , if an agent knows a contingent proposition, ϕ , then, in nearly all near-by possible worlds in which the agent forms her belief that ϕ in the same way as in the actual world, her belief is true.¹⁸

Like process and faculty reliabilism, any account of knowledge that is solely based on the safety principle will be epistemologically externalist, in that it merely specifies a set of conditions that must be met if the agent is to have knowledge, without further demanding that the agent should also have internalist justification for the target belief. What this principle ensures is that there is no gap between the agent’s meeting the epistemic conditions and forming a true belief in which veritic luck can get a hold, for in so far as the agent meets the epistemic conditions – which includes meeting the safety principle – it cannot be a matter of veritic epistemic luck that her belief is true, since in nearly all near-by possible worlds where she forms her belief on the same basis she will form a true belief as a result. For example, Mary’s belief that her husband is in the room is not knowledge on this view because there are a number of near-by possible worlds where she forms her belief on the same basis and yet what she believes is no longer true.

One can regard the move from a simple process reliabilist view to either a modalized or agent-based reliabilist view as a move towards something

¹⁷ Goldman’s (1967) “causal” theory could be construed as (broadly speaking) a process reliabilist attempt to meet the problem. See also Armstrong (1973). Sosa (1996) offers one influential rendering of the agent reliabilist position that, he claims, can meet the Gettier counterexamples. See also Plantinga (1988).

¹⁸ Since necessary truths are true in all possible worlds, they present complications for any view of this sort. As these complications are not important to the present discussion, I focus here solely upon contingent propositions.

like a safety-based account, in that the reliability in question becomes tied not just to actual truth-to-falsity ratios but also to what those ratios would be across a range of normal possible worlds (either directly, in the case of modalized process reliabilism, or indirectly, in the case of agent reliabilism). The trick is thus to supplement the basic reliabilist account in such a way as to achieve something like that achieved by the safety principle.

The salient question to ask at this juncture, however, is just why we need a reliabilist account at all if a safety-based theory will do just as well (which will meet the Gettier-type examples, along with all the other examples that play upon the presence of veritic luck). Indeed, this issue is especially pertinent once one notes that the foremost defenders of agent reliabilism, Greco and Sosa, have each used something like a safety principle in order to meet the sceptical challenge. In both cases, the basic idea is that sceptical error possibilities can be dismissed in that they are concerned with far-off possible worlds, so that just so long as the reliability at issue is maintained in near-by possible worlds, this will suffice for knowledge. Accordingly, so long as the actual world is “normal” in the sense that we saw Goldman characterise it above, we can know such anti-sceptical truths as that we are not brains in vats, since in normal worlds our means of determining these kinds of truths are appropriately reliable and thus generate “safe” beliefs.¹⁹

There is no inherent reason, however, why a safety-based account of knowledge should mention the reliability of processes at all, whether those at issue in agent reliabilism or those at issue in the modalized process reliabilist story. Indeed, given the difficulty that reliabilist accounts have in meeting the Gettier counterexamples, the appropriate moral to draw (at least given that we are happy with epistemological externalism) does not seem to be that we need to keep supplementing the reliabilist thesis ad infinitum in order to try to meet the challenge posed by Gettier cases, but rather that we should simply accept that knowledge is, at root, just true belief that meets the safety principle. The job of reliabilism, of either description, would then be to tell the necessary *explanatory* story about how, as the creatures we are in the environment that we find ourselves in, we come to form true beliefs about the world in such a way as to support knowledge. Here one would expect the reliabilist story to be essential, but this project is not the project of discovering necessary and sufficient conditions for the possession of knowledge.²⁰

¹⁹ See Greco (1994; 2000, chapter 8) and Sosa (1999; 2000). Sainsbury (1997), Williamson (2000a; 2000b, chapter 8) and Pritchard (2001a; 2002a; 2002b; 2002d) have all defended versions of the safety-based response to scepticism (the so-called neo-Moorean response to scepticism). Black (2002) offers an account of Nozick’s counterfactual conditions on knowledge that ends up construing them as much like the safety principle. For further discussion of this type of anti-sceptical thesis, and how it relates to the other main anti-sceptical theories in the literature, see Pritchard (2002c).

²⁰ This point is especially germane as regards Greco’s account because his agent reliabilist theory does not entail the safety-based response to the relevant Gettier-type counterexamples that he offers. (Indeed, Greco seems to be aware of this, noting only that his safety-based

So, in so far as it is the elimination of veritic epistemic luck that is motivating the adoption of the agent reliabilist thesis over the process reliabilist thesis, the position is ill motivated as a theory of knowledge for two reasons. First, because a modalized version of process reliabilism can be just as successful in this regard. Second, because, in any case, both agent reliabilism and modalized process reliabilism are subject to a certain form of veritic epistemic luck. Accordingly, if one is to take the challenge posed by such luck seriously, one should prefer the type of safety-based account of knowledge that is specifically defined such that it eliminates veritic epistemic luck. The upshot of this is that in so far as we should advocate a version of virtue epistemology modelled on reliabilism, this should be as an explanatory thesis concerning how creatures such as ourselves come to have safe beliefs, and thus knowledge. This conclusion hits right to the heart of the view, however, since the defining characteristic of virtue epistemology was meant to be the fact that it characterised epistemic concepts in terms of epistemic virtues rather than the other way around. Hence, this conclusion is not simply a *demotion* of the status of virtue epistemology but a straightforward *rejection*.²¹

4. Responsibilist Virtue Epistemology and Reflective Epistemic Luck

Not all virtue-based epistemological theories model themselves on reliabilism in the way that agent reliabilism does, however, so this conclusion

account of relevant alternatives is “grounded” by his version of agent reliabilism.) Accordingly, his agent reliabilist theory seems to be more of an explanation of how cognitive beings come to have safe beliefs (and thus knowledge) rather than a definition of knowledge. See Greco (2000, chapter 8). Foley (1994) also raises the issue of why we need a *virtue-based* epistemology, rather than simply an epistemology that involves an account of the epistemic virtues, though his point is specifically directed at Sosa’s version of virtue-based reliabilism.

²¹ In a recent essay, Greco (2002) has argued that safety alone will not suffice for knowledge because it cannot handle the so-called lottery problem. Suppose I buy a ticket for a free and fair lottery with long odds and form the belief that I shall not win. Now suppose that the draw has been made and this belief has been confirmed, but that I have not been informed of the result yet. Do I *know* that I haven’t won? Intuitively, I do not, but Greco argues that the safety-based theory demands an affirmative answer to this question because my belief is true not only in the actual world but in nearly all near-by possible worlds as well. Thus, safety alone cannot meet our epistemic intuitions, and Greco diagnoses this difficulty by arguing that the problem rests with the fact that safety-based views cannot capture the sense in which knowledge demands responsibility for truth. This argument does not go through, however, for the simple reason that, contrary to first impressions, there *are* near-by possible worlds in which one wins the lottery. After all, the whole attraction of a free and fair lottery is that the possible world in which I win is extremely similar to the actual world in which I do not (just a few coloured balls need to fall in a slightly different configuration). The example misleads us by talking of a low-probability event, thereby giving us the impression that the event only obtains in far-off possible worlds, but possible worlds are ordered in terms of their *similarity* to the actual world, not in terms of their *probability* (though the two notions are of course related). It follows that Greco cannot motivate his position via examples of this sort. I am grateful to John Greco for useful discussion on this point.

does not have an unrestricted application to virtue epistemology as a whole. There are a number of reasons for this schism, but one of the main sources of dissatisfaction with the agent reliabilist account of virtue epistemology has been its commitment to epistemological externalism, a thesis that some epistemologists find counterintuitive, *especially* when it comes to the epistemic virtues. As we shall see, the issue here is that merely offering an account of knowledge that is (largely) free of veritic epistemic luck does not appear to be enough – we also seem to need a theory that eliminates what I shall call “reflective” epistemic luck, at least as regards our knowledge in a wide range of cases.

We have already seen that as externalist theories modalized process reliabilism and agent reliabilism appear obliged to allow that in the “strengthened” brain-lesion case – where, we recall, it is stipulated that the process in question is stable in the relevant respects, and thereby will support true beliefs in nearly all (normal) near-by possible worlds – the agent *can* know that he has a brain lesion. Such cases are standard fare in epistemology these days. The most famous is the so-called chicken-sexer example. In this case we are asked to imagine an agent who has a natural and highly reliable ability to distinguish between male and female chicks, but who has no idea how she is doing this and is not even aware that she is reliable in this respect. (Some even supplement the example by saying that she is mistaken about how she is doing it, thinking that she is seeing or touching something distinctive, when in fact she is guided by her unusually sensitive sense of smell).²²

Externalists are inclined to allow that such agents can indeed have knowledge in such cases, whilst internalists tend to demur. Internalists contend that this cannot be knowledge, because the agent has no reasons to support her beliefs. Externalists respond by arguing, as it were, from the third-person perspective – by looking at what we would say about the epistemic status of the agent’s belief rather than at what she would say about it. This shift in perspective is important in that it distinguishes the issue of what the agent is able properly to claim regarding her epistemic situation from what her epistemic situation actually is.²³ Crucially, however, in so far as it is plausible to contend that the chicken-sexer has knowledge via the employment of her reflectively inaccessible though reliable belief-forming processes, it ought (as we noted earlier) to be plausible to contend that one can gain knowledge via brain lesions just so long as the reliability in question is stable in the relevant respects.

²² For further discussion of the chicken-sexer example, see Foley (1987, 168–69), Lewis (1996), Zagzebski (1996, §2.1 and §4.1), and Brandom (1998). See also the exchange between Sainsbury (1996) and Wright (1996).

²³ This is a key move that the epistemological externalist needs to make if she is to offer the necessary motivation for her view, because whilst the possession conditions for knowledge will closely mirror the propriety conditions for knowledge claims on the internalist view, this will not be so on the externalist view. The externalist is thus able to use this observation in order to explain why certain cases of knowledge, whilst in fact *bona fide*, might nevertheless *seem* suspect. For more on this point, see Pritchard (2001b).

Expecting a consensus to emerge concerning examples like this is almost certainly expecting too much. What we *can* expect to achieve, however, is a consensus that there is *something* lacking about the epistemic status of the agent's beliefs in these examples, even if the further issue of whether or not what is missing is essential to knowledge possession is one that cannot be settled in a way that is acceptable to both externalists and internalists. For even externalists would surely agree that the knowledge possessed by the chicken sexer (if indeed it is knowledge) is knowledge of a very "brute" variety, and that it would be preferable (even if unnecessary, strictly speaking) that the agent have an internalist justification in support of her belief as well.

Notice, however, that one cannot capture the "lack" at issue here in terms of veritic epistemic luck. If the chicken sexer is exhibiting a genuine stable cognitive faculty that generates safe beliefs, then not only will her belief be true in the actual world but, in nearly all near-by possible worlds in which she forms her belief in the same way, she will also continue to form true beliefs as a result. Nevertheless, whilst it is not veritic epistemic luck that is at issue, it does seem to be a matter of luck, *in some sense*, that the agent has knowledge in this case, and this at least partly explains why we are reluctant to ascribe knowledge.

We shall call the luck that is at issue here "reflective" luck, since it concerns not simply the luck that one's belief is true but rather the luck that one's belief is true given what one is able to know by reflection alone. We can express this form of luck as follows:

For all agents, ϕ , the truth of an agent's belief in a proposition, ϕ , is reflectively lucky if, and only if, the agent's belief that ϕ is true in the actual world, but, in nearly all possible worlds consistent with what the agent is able to know by reflection alone, were the agent to form a belief that ϕ , that belief would be false.

As with veritic luck, if one meets all the relevant epistemic conditions set down by the favoured epistemological theory and one's belief is still reflectively lucky, then this is a case of reflective *epistemic* luck.²⁴

The description of the relevant class of possible worlds here deserves comment. These worlds are described in this manner because we are trying to capture the sense in which, whilst it may be the case that the agent's belief is in fact safe such that it matches the truth in near-by possible worlds in which it is formed in the appropriate way, this alone does not ensure that the truth of the belief is not reflectively lucky. Consider again the example just given of the chicken sexer, but this time compare our chicken sexer with an "enlightened" counterpart who not only has this

²⁴ Riggs (1998) offers a distinction between two types of epistemic luck that roughly parallels (though is still importantly different from) that offered here between veritic and reflective epistemic luck. For discussion of this view, see Axtell (2002).

ability but also has the reflective knowledge to back up the beliefs that she is forming about the sex of the chicks (she has, say, true beliefs about how this ability works, and about its relative reliability in different types of environmental conditions). Clearly, neither of these agents has a belief that is veritically lucky, since the safety of their beliefs ensures that in nearly all near-by possible worlds in which they form their beliefs in the same way as in the actual world, their beliefs will be true.²⁵

Nevertheless, in terms of what the agents are able to know by reflection alone, the “unenlightened” agent is lucky to have a true belief in a way that the enlightened agent is not. This is because, unlike the enlightened agent, the unenlightened agent has no knowledge reflectively available to her to indicate that her belief is safe. We can express this point by saying that whilst the enlightened chicken sexer has reflective knowledge that excludes various error possibilities from obtaining, this is not true of the unenlightened chicken sexer. So far as she is able to know in this respect, her beliefs could be radically in error. So whilst in most possible worlds consistent with what she is able to know by reflection alone the enlightened chicken sexer’s belief remains true, this will not be the case for her unenlightened counterpart. Instead, there will be a wide class of worlds that are consistent with what she can know by reflection alone but where she is radically in error, such as those worlds where she simply lacks this ability. The same goes for the brain-lesion example, since here too there are many possible worlds that are consistent with what the agent is able to know by reflection alone but where the belief formed is false.

So whilst externalist epistemological theories eliminate one type of epistemic luck – veritic epistemic luck – they do not eliminate another type of epistemic luck that is reflective. Externalists may well be sanguine about this fact – in the sense that they do not think that reflective epistemic luck is a bar to knowledge possession – but this should not prevent them from being aware that reflective epistemic luck, whilst perhaps consistent with knowledge possession, is at least epistemically *undesirable*. Part of the problem here is that the reliability at issue, since it is not reflectively accessible to the agent, operates in this sense merely at the “sub-personal” level. As a result, it does not seem entirely appropriate to ascribe the cognitive achievement of knowledge possession to such an agent when, in one sense at least, the cognitive ability in question is not due to the *agent* at all.

It is a worry of this sort that has prompted some commentators attracted

²⁵ Though, of course, one would expect the agent who has the reflective knowledge to have a belief that tracks the truth in far more near-by possible worlds than her counterpart who lacks this knowledge because, unlike the unenlightened chicken sexer, she will revise her beliefs in appropriate ways in response to different circumstances. For example, given that she reflectively knows that she has evidence which strongly suggests that her ability is based on smell, she will be cautious about forming beliefs about the sex of chicks via this ability in situations where she is aware that this faculty is not functioning as normal (e.g., when she has a cold).

by the prospects of a virtue-based epistemology to make the move away from reliabilism altogether and adopt an agent-based epistemology that is, in the relevant regard, non-reliabilist. The central idea is that what is lacking from the reliabilist schema, even when modified along agent reliabilist lines, is a fundamental role for the epistemic *responsibility* of the agent. Such views are, to varying extents, approximations to an Aristotelian virtue-theoretic account of knowledge that conceives of intellectual virtues in terms of processes at the personal level – *traits of character* – that agents are, at least in some sense, responsible for (as opposed to the broader, possibly Platonic, reading of “virtue” merely in terms of an excellence [*arete*] that could just as well apply to mere cognitive faculties).²⁶ Zagzebski is, perhaps, the foremost exponent of a view of this sort, a view that is often referred to as “neo-Aristotelian.” Accordingly, I shall focus on her theory here, though much of what I say ought to be equally applicable, *mutatis mutandis*, to any view of this sort.

Zagzebski sees the central problem with reliabilist theories as being their inability to explain what is valuable about knowledge possession. In particular, she argues that the “machine-product” analogy they work with is faulty, in the sense that no matter how reliable a machine is at producing a certain (valuable) product, the reliability of the machine alone will not explain the value of the product. The same is meant to apply to knowledge, in that she argues that if knowledge is to be solely understood in terms of reliability, it is difficult to see why knowledge is such a valuable thing to have. The chicken-sexer example provides support for this contention, in that both externalists and internalists would surely agree that there is *something* more valuable about the knowledge possessed by the enlightened chicken sexer that is missing in the case of her unenlightened counterpart. Plausibly, this difference in value is a direct result of the fact that the enlightened chicken sexer is able to take responsibility for her beliefs in a way that is impossible for the unenlightened chicken sexer.

Zagzebski argues that the way to deal with this problem is to abandon the machine-product analogy and focus instead upon the motivational states of the agent in much the same way as we are inclined to do in the moral case. Simply being able to act in such a way as to produce good consequences reliably is not sufficient for being accorded a high moral status for one’s character. Rather, it is necessary that one also have the right motivational states that forge the required connection between one’s inner psychological states and one’s resultant actions. Similarly, in the epistemic case, it is important that the agent should be motivated by the right epistemic virtues and thus that the reliability is brought about in the right way.

²⁶ For further discussion of this contrast between reliabilism-based and responsibility-based virtue epistemologies, see Axtell (1997; 1998), although note that Axtell allows some virtue theories to count as responsibility even though they allow the possibility that the cognitive processes in question might operate at the sub-personal level. In particular, Axtell (1997, 13) cites Greco as being an example of a responsibility virtue epistemologist.

Zagzebski, therefore, understands beliefs on the model of actions, and in doing so, she brings the deontic theoretical machinery to bear on the epistemic character of the agent.²⁷

The problems regarding veritic luck facing process reliabilism thus lead some epistemologists to opt for an agent-centred version of reliabilism, but the further problems that remain spur others, such as Zagzebski, to go further and develop an epistemology entirely focused upon the epistemic acts of the agent. In doing so, Zagzebski moves away from epistemological externalism by arguing that it is essential to knowledge possession that the agent be able to know by reflection alone those facts that determine the justification for her belief. Consider the following passage:

The value of knowledge is connected with the components of intellectual virtue. . . . It is in part because knowledge arises out of acts of intellectual virtue that it is an achievement in a way that mere true belief cannot be. The value of the truth obtained by a reliable process in the absence of any conscious awareness of a connection between the behaviour of the agent and the truth he thereby acquires is no better than the value of the lucky guess. (Zagzebski 1996, 303–4)

If this model of knowledge is to rule out cases like those of the chicken sexers as being genuine instances of knowledge (on the grounds that an instance of knowledge of this sort would be no better than a “lucky guess”), then we need to interpret the “conscious awareness” of the reliability at issue here in terms of the agent having evidence reflectively available to her in such a way as to support internalist justification of the target belief. In effect, then, Zagzebski is demanding that reliability alone is not enough and that what is further needed is *internalist* knowledge, knowledge that meets the internalist rubric as defined above. Moreover, it is important to note that Zagzebski is here explicitly making the point that the problem with an account of knowledge construed externalistically solely in terms of reliability (or, *a fortiori*, safety) is that it does not eliminate epistemic luck. Since, as we have seen, safety-based accounts – of which modalized process reliabilism and agent reliabilism are approximations – *do* succeed in eliminating veritic epistemic luck, and given that Zagzebski clearly believes that the type of epistemic luck she has in mind is dealt with via an appeal to reflective knowledge, we can safely assume that it is specifically *reflective* epistemic luck that is being referred to here. The problem with agent reliabilist versions of virtue epistemological theories

²⁷ See Zagzebski (1996, especially part 3) for a full account of her position. She expounds upon what she calls the “value problem” facing externalist theories of knowledge at more length in her essay in this collection. For discussion of the issues that she raises, see Greco (2002), Riggs (2002) and Percival’s essay in this collection. For other versions of the virtue epistemology thesis that have more in common with Zagzebski’s account than agent reliabilism, see Code (1984; 1987), Montmarquet (1987; 1993), Kvanvig (1992) and Hookway (1994).

according to Zagzebski is thus that, unlike the neo-Aristotelian responsibilist view that she endorses, they fail to eliminate reflective epistemic luck. One can thus understand the move from a reliabilism-based virtue theory to a responsibilist virtue theory in terms of how the latter is concerned to eliminate not just veritic epistemic luck but also reflective epistemic luck.

The problem with understanding this type of virtue epistemology as being motivated by a desire to eliminate reflective epistemic luck is that *no* epistemological theory can meet this challenge. This is highlighted for us by the radical sceptic via her use of radical sceptical error scenarios (such as that one might be a brain in a vat being “fed” one’s experiences). It is part of the point of these sceptical error possibilities that they are entirely consistent with what one is able to know by reflection alone, even in the best case. Given this fact, however, this means that it is part of what we might call the “epistemic condition” that one’s knowledge is always in this sense subject to reflective epistemic luck. Accordingly, if one demands that one’s theory of knowledge should eliminate this type of luck, one is directly playing into the hands of the sceptic.

If the responsibilist view is to work, then, it must be willing to countenance at least one variety of reflective epistemic luck, despite protestations to the contrary by those who advance such a position. The difficulty facing this form of virtue theory now is to explain just what it offers over and above a mere epistemologically internalist account of knowledge that combines a safety-based view with the demand for internalist justification. Recall that the responsibilist position is meant to be a further refinement of an agent reliabilist thesis that eliminates the problem of veritic epistemic luck. Given that agent reliabilism can be regarded as an approximation to the kind of safety-based view that does fully succeed in eliminating veritic epistemic luck, any viable responsibilist theory is going to have to incorporate a safety-type principle. We have already seen that a non-virtue safety-based account seems, *prima facie* at least, to present the best way of dealing with the problem of veritic epistemic luck. The challenge facing responsibilist virtue theories is thus to supplement this type of principle in such a way as to eliminate non-sceptical cases of reflective epistemic luck whilst also explaining why appeal to the epistemic virtues is essential to one’s theory of knowledge. As it stands, however, there is no reason to prefer this responsibilist view over a safety-based internalist theory of knowledge that, whilst conceding that it cannot meet sceptical cases of reflective epistemic luck, nevertheless incorporates a role for responsibility by demanding internalist justification and eliminates veritic epistemic luck by including a safety condition.

Note that I am not denying here that the responsibilist account of the epistemic virtues may well be an important feature of any fully fledged internalist theory of knowledge. Rather, I am simply questioning the key virtue-theoretic claim that one cannot account for knowledge and the other

key epistemic concepts *without* appeal to the virtues. If veritic epistemic luck is the problem, then a safety-based account of knowledge is the answer, and this need make no essential reference to the epistemic virtues. Similarly, if reflective epistemic luck is a supplementary problem in this regard and one is entitled to set aside the radical sceptical challenge, then a safety-based account of knowledge coupled with the demand for internalist justification is the answer, and, again, this need make no essential reference to the epistemic virtues. Far from motivating the adoption of virtue-based theories, consideration of the role of epistemic luck, in both its most salient forms, actually reveals that virtue epistemology is curiously ill motivated as it stands.

5. Concluding Remarks

Not one of the considerations adduced here is meant to constitute decisive grounds for rejection of the virtue-epistemological project. Aside from anything else, the particular views that I have focused upon do not exhaust the theoretical possibilities as regards the two main types of virtue epistemology, and there may be, in any case, other types of virtue epistemology available that do not fall into either of these camps. Nevertheless, the considerations adduced here should suffice to prompt proponents of virtue epistemology to reflect further on just what it is that is motivating their adoption of this thesis. In particular, the play with epistemic luck that is common amongst defenders of this type of account of knowledge masks important distinctions concerning what is involved in epistemic luck that are substantively relevant to the plausibility of the position. Despite the pessimism that is expressed here, I am actually quietly optimistic about the prospects for virtue epistemology. Until it is clear what question virtue epistemology is an answer to, however, and that it does constitute an answer to this question, this optimism needs to remain on ice.²⁸

Department of Philosophy
University of Stirling
Stirling FK9 4LA
United Kingdom
d.h.pritchard@stir.ac.uk

References

Armstrong, David M. 1973. *Belief, Truth and Knowledge*. Cambridge: Cambridge University Press.

²⁸ Thanks to Guy Axtell, Michael Brady and John Greco, and to the Leverhulme Trust for the award of a Special Research Fellowship to undertake work in this area.

- Axtell, Guy. 1997. "Recent Work on Virtue Epistemology." *American Philosophical Quarterly* 34: 1–26.
- . 1998. "The Role of the Intellectual Virtues in the Reunification of Epistemology." *Monist* 81: 488–508.
- Axtell, Guy. 2001. "Epistemic Luck in Light of the Virtues." In *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*, edited by A. Fairweather and L. Zagzebski. Oxford: Oxford University Press.
- , editor. 2000. *Knowledge, Belief and Character: Readings in Virtue Epistemology*. Lanham, Md.: Rowman and Littlefield.
- Black, Tim. 2002. "A Moorean Response to Brain-in-a-Vat Scepticism." *Australasian Journal of Philosophy* 80: 148–63.
- Brandom, Robert. 1998. "Insights and Blindspots of Reliabilism." *Monist* 81: 371–92.
- Code, Lorraine. 1984. "Toward a "Responsibilist" Epistemology." *Philosophy and Phenomenological Research* 44: 29–50.
- . 1987. *Epistemic Responsibility*. Hanover, N.H.: University Press of New England and Brown University Press.
- DePaul, Michael, and Linda Zagzebski, editors. 2002. *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Engel, Mylan. 1992. "Is Epistemic Luck Compatible with Knowledge?" *Southern Journal of Philosophy* 30: 59–75.
- Fairweather, Abrol, and Linda Zagzebski, editors. 2001. *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*. Oxford: Oxford University Press.
- Foley, Richard. 1987. *A Theory of Epistemic Rationality*. Cambridge, Mass.: Harvard University Press.
- . 1994. "Sosa's Epistemology." *Philosophical Issues* 5: 42–58.
- Gettier, Edmund. 1963. "Is Justified True Belief Knowledge?" *Analysis* 23: 121–23.
- Goldman, Alvin. 1967. "A Causal Theory of Knowing." *Journal of Philosophy* 64: 355–72.
- . 1976. "Discrimination and Perceptual Knowledge." *Journal of Philosophy* 73: 771–91.
- . 1979. "What Is Justified Belief?" In *Justification and Knowledge*, edited by G. S. Pappas, 1–23. Dordrecht: D. Reidel.
- . 1986. *Epistemology and Cognition*. Cambridge, Mass.: Harvard University Press.
- . 1988. "Strong and Weak Justification." *Philosophical Perspectives* 2: 51–69.
- . 1993. "Epistemic Folkways and Scientific Epistemology." *Philosophical Issues* 3: 271–84.
- Greco, John. 1993. "Virtues and Vices of Virtue Epistemology." *Canadian Journal of Philosophy* 23: 413–32.

- . 1994. "Virtue Epistemology and the Relevant Sense of Relevant Possibility." *Southern Journal of Philosophy* 32: 61–77.
- . 1999. "Agent Reliabilism." *Philosophical Perspectives* 13: 273–96.
- . 2000. *Putting Sceptics in Their Place: The Nature of Sceptical Arguments and Their Role in Philosophical Inquiry*. Cambridge: Cambridge University Press.
- . 2002. "Knowledge as Credit for True Belief." In *Intellectual Virtue: Perspectives from Ethics and Epistemology*, edited by Michael DePaul and Linda Zagzebski. Oxford: Oxford University Press.
- Hall, Barbara J. 1994. "On Epistemic Luck." *Southern Journal of Philosophy* 32: 79–84.
- Harper, William. 1996. "Knowledge and Luck." *Southern Journal of Philosophy* 34: 273–83.
- Hookway, Christopher. 1994. "Cognitive Virtues and Epistemic Evaluations." *International Journal of Philosophical Studies* 2: 211–27.
- Kornblith, Hilary, editor. 2001. *Epistemology: Internalism and Externalism*. Oxford: Basil Blackwell.
- Kvanvig, John. 1992. *The Intellectual Virtues and the Life of the Mind: On the Place of Virtues in Epistemology*. Savage, Md.: Rowman and Littlefield.
- Lewis, David. 1996. "Elusive Knowledge." *Australasian Journal of Philosophy* 74: 549–67.
- Montmarquet, James. 1987. "Epistemic Virtue." *Mind* 96: 487–97.
- . 1993. *Epistemic Virtue and Doxastic Responsibility*. Lanham, Md.: Rowman and Littlefield.
- Plantinga, Alvin. 1988. "Positive Epistemic Status and Proper Function." *Philosophical Perspectives* 2: 1–50.
- . 1993a. *Warrant and Proper Function*. New York: Oxford University Press.
- . 1993b. "Why We Need Proper Function." *Noûs* 27: 66–82.
- Pritchard, Duncan H. 2001a. "Contextualism, Scepticism, and the Problem of Epistemic Descent." *Dialectica* 55: 327–49.
- . 2001b. "Radical Scepticism, Epistemological Externalism, and 'Hinge' Propositions." *Wittgenstein-Studien*: 81–105.
- . 2002a. "McKinsey Paradoxes, Radical Scepticism, and the Transmission of Knowledge across Known Entailments." *Synthese* 130: 279–302.
- . 2002b. "Radical Scepticism, Epistemological Externalism and Closure." *Theoria* 69: 129–61.
- . 2002c. "Recent Work on Radical Scepticism." *American Philosophical Quarterly* 39: 215–57.
- . 2002d. "Resurrecting the Moorean Response to Scepticism." *International Journal of Philosophical Studies* 10: 283–307.
- . 2003. "Epistemic Luck." *Journal of Philosophical Research* 28.

- Pritchard, Duncan H., and Matthew Smith. 2002. "The Philosophy and Psychology of Luck." Typescript.
- Pryor, James. 2001. "Highlights of Recent Epistemology." *British Journal for the Philosophy of Science* 52: 95–124.
- Rescher, Nicholas. 1995. *Luck: The Brilliant Randomness of Everyday Life*. New York: Farrar, Straus, and Giroux.
- Riggs, Wayne. 1998. "What Are the 'Chances' of Being Justified?" *Monist* 81: 452–72.
- . 2002. "Reliability and the Value of Knowledge." *Philosophy and Phenomenological Research* 64: 79–96.
- Russell, Bertrand. 1948. *Human Knowledge: Its Scope and Its Limits*, London: George Allen and Unwin.
- Sainsbury, Mark. 1996. "Crispin Wright: *Truth and Objectivity*." *Philosophy and Phenomenological Research* 56: 899–904.
- Sosa, Ernest. 1985. "Knowledge and Intellectual Virtue." *Monist* 68: 224–45.
- . 1991. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge: Cambridge University Press.
- . 1993. "Proper Functionalism and Virtue Epistemology." *Noûs* 27: 51–65.
- . 1999. "How to Defeat Opposition to Moore." *Philosophical Perspectives* 13: 141–54.
- . 2000. "Skepticism and Contextualism." *Philosophical Issues* 10: 1–18.
- Steup, Mathias, editor. 2002. *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford: Oxford University Press.
- Talbott, William J. 1990. *The Reliability of the Cognitive Mechanism*. New York: Garland Publishing.
- Vahid, Hamid. 2001. "Knowledge and Varieties of Epistemic Luck." *Dialectica* 55: 351–62.
- Williamson, Timothy. 2000a. "Scepticism and Evidence." *Philosophy and Phenomenological Research* 60: 613–28.
- . 2000b. *Knowledge and Its Limits*. Oxford: Oxford University Press.
- Wright, Crispin. 1996. "Response to Commentators." *Philosophy and Phenomenological Research* 56: 911–41.
- Zagzebski, Linda. 1994. "The Inescapability of Gettier Problems." *Philosophical Quarterly* 44: 65–73.
- . 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.
- . 1999. "What Is Knowledge?" In *Epistemology*, edited by J. Greco and E. Sosa, 92–116. Oxford: Basil Blackwell.