CrossMark

Anti-luck virtue epistemology and epistemic defeat

Duncan Pritchard[1]

Abstract This paper explores how a certain theory of knowledge—known as *anti-luck virtue epistemology*—can account for, and in the process shed light on, the notion of an epistemic defeater. To this end, an overview of the motivations for anti-luck virtue epistemology is offered, along with a taxonomy of different kinds of epistemic defeater. It is then shown how anti-luck virtue epistemology can explain: (i) why certain kinds of putative epistemic defeater are not *bona fide*; (ii) how certain kinds of epistemic defeater are genuine in virtue of exposing the subject to significant levels of epistemic risk; and (iii) how certain kinds of epistemic defeater are genuine in virtue of highlighting how the subject's safe cognitive success does not stand in the appropriate explanatory relationship to her manifestation of relevant cognitive ability.

Keywords Epistemology · Knowledge · Epistemic defeaters · Anti-luck virtue epistemology

1 Introductory remarks

It is incumbent upon any theory of knowledge to explain how it handles cases involving epistemic defeaters. To this end, I will be offering an account of how my preferred theory of knowledge—*anti-luck virtue epistemology*—can incorporate this epistemic notion. In particular, I will be arguing that this proposal is well-placed to account for the various different types of epistemic defeater (and can also explain why some related phenomena do not qualify as genuine cases of epistemic defeat). Although my claim is not the ambitious one that anti-luck virtue epistemology is *uniquely* placed to account

---

✉ Duncan Pritchard
duncan.pritchard@ed.ac.uk

1 University of Edinburgh, Edinburgh, UK

for epistemic defeaters, neither is it the completely unambitious claim that anti-luck virtue epistemology is merely able to account for epistemic defeaters. Instead, I hope to show that this theory of knowledge is able to cast a lot of light on why epistemic defeaters of different kinds function as they do. In this sense, then, anti-luck virtue epistemology does not offer merely an adequate account of epistemic defeat, but rather is also able to put forward an enlightening, and multi-layered, diagnosis of this phenomenon. This suggests that anti-luck virtue epistemology is likely to be on stronger ground in this regard than at least some other competing theories of knowledge.

In Sect. 2, I outline anti-luck virtue epistemology, and motivate the position. In Sect. 3, I offer a taxonomy of different kinds of epistemic defeater. In Sect. 4, I explain how anti-luck virtue epistemology can accommodate, and diagnose, different kinds of epistemic defeater. Finally, in Sect. 5, I offer some concluding remarks.[1]

2 The nature of knowledge

Elsewhere, I have put forward an account of knowledge that involves an interplay between manifestations of cognitive agency and extra-agential factors—this is what I call an *anti-luck virtue epistemology*, for reasons that will shortly become apparent. Anti-luck virtue epistemology is uniquely placed to account for what I call the *epistemic dependence* of knowledge. Epistemic dependence has both a positive and a negative aspect. *Positive epistemic dependence* is when a subject manifests a relatively minimal level of cognitive agency—i.e., of a level that would not normally suffice for knowledge—but where she knows nonetheless due to extra-agential factors. *Negative epistemic dependence*, in contrast, is where a subject manifests a relatively high level of cognitive agency—i.e., of a level that would normally be comfortably enough to ensure knowledge—but where she nonetheless lacks knowledge due to extra-agential factors. (I will be offering examples of both kinds of epistemic dependence in a moment).[2]

The reason why I claim that knowledge exhibits epistemic dependence is that it is in effect answering to two distinct constraints. On the one hand, knowledge essentially involves cognitive success (i.e., true belief) that is attributable to one's manifestation of relevant cognitive agency (this is what makes the knowledge *mine*). On the other hand, knowledge essentially involves the exclusion of high levels of epistemic risk. The upshot of the latter constraint is that when one has knowledge, one is forming one's belief in a way that is *safe*—i.e., such that one's belief could not easily have been false (or as the point is sometimes expressed: it cannot be *veritically lucky*).[3] One

---

[1] Note that, henceforth, I will refer to epistemic defeaters as defeaters *simpliciter*.

[2] For more on the notion of epistemic dependence, see Kallestrup and Pritchard (2011, 2012, 2013) and Pritchard (forthcominga). Note that this notion is rooted in an earlier distinction between intervening and environmental epistemic luck, and an associated critique of robust virtue epistemology—see Pritchard (2009a, b, 2012a) and Pritchard et al. (2010, chap. 2–4).

[3] I will not be getting into the issue of how best to unpack the notion of safety in play here. I discuss this notion (and the related notion of veritic luck), and its place within the broader framework of what I call *anti-luck epistemology*, in a number of places. See, especially, Pritchard (2005, 2007, 2012a, c, 2015a). See also my recent defence of what I refer to as *anti-risk epistemology*, which involves a subtle adaption to the anti-luck epistemology framework, in Pritchard (forthcomingb).

might initially think that whatever condition one places on knowledge that satisfies one of these constraints would inevitably satisfy the other. After all, if one's cognitive success is attributable to one's manifestation of relevant cognitive agency, then why would it be subject to high levels of epistemic risk? And if one's cognitive success is not subject to high levels of epistemic risk, then wouldn't that be because it is attributable to one's manifestation of relevant cognitive agency?

On closer inspection, however, it becomes clear that these two constraints come apart. In particular, it could be that the safety of one's cognitive success might have absolutely nothing to do with one's own manifestation of cognitive agency, but exclusively down to factors in one's cognitive environment. For example, imagine that one's cognitive environment is such that no matter how badly one is forming one's beliefs about a certain subject matter, these beliefs are nonetheless bound to be correct. One's cognitive success is thus immune to epistemic risk, but there is no manifestation of relevant cognitive ability on display, and hence no knowledge either. Going in the other direction, it is also possible for a subject's cognitive success to be attributable to her manifestation of relevant cognitive ability and yet nonetheless subject to a high degree of epistemic risk. This is because manifestations of cognitive ability can be *modally fragile*, in the sense that although in the actual world nothing impeded them, they would have been impeded in close possible worlds. (Indeed, this is not a point that is specific to manifestations of cognitive ability, as it applies to all manifestations of ability, cognitive or otherwise).

It follows that if knowledge answers to both these constraints, then it thereby answers to them in distinct ways. Indeed, I claim that the right way to respond to the distinctness of these two constraints is to treat knowledge as involving an interplay between manifestations of cognitive ability and extra-agential factors. This is why we get an anti-luck virtue epistemology, where the 'anti-luck' part concerns the exclusion of epistemic risk, and the 'virtue' part concerns the manifestation of relevant cognitive ability. According to anti-luck virtue epistemology, knowledge is safe cognitive success that is significantly attributable to the subject's manifestation of relevant cognitive ability. The safety condition does the job of ruling out high levels of epistemic risk. Moreover, since the safety of one's belief has to be significantly attributable to one's manifestation of relevant cognitive ability, then this excludes the possibility of gaining knowledge in ways that don't involve manifestations of relevant cognitive ability. Notice, however, that the explanatory relationship between one's safe cognitive success and one's manifestation of relevant cognitive ability only needs to be significant. It is not demanded, for example—as some proponents of a strong form of virtue epistemology known as *robust virtue epistemology* demand—that one's cognitive success should be primarily or overarchingly attributable to one's manifestation of relevant cognitive ability.[4] This means that in the right cognitive environments, knowledge can be relatively easy to acquire.

Anti-luck virtue epistemology is uniquely placed to accommodate epistemic dependence. Consider first positive epistemic dependence. Sometimes knowledge is very easy to acquire, as in straightforward cases of testimonial knowledge when one largely

---

[4] For some key defences of versions of robust virtue epistemology, see Zagzebski (1996), Sosa (2007, 2009, 2015), and Greco (2009).

gains one's belief through trust of one's informant. While in such cases one is manifesting a significant level of cognitive ability—one wouldn't ask just anyone, for example, or believe whatever one is told (gullibility is never a route to knowledge)—one could hardly claim that the safety of one's cognitive success is primarily down to one's own manifestation of relevant cognitive ability (as opposed to the cognitive ability of one's knowledgeable informant). Anti-luck virtue epistemology can account for what is happening in such cases. One can, after all, only gain testimonial knowledge in this fashion in very epistemically friendly cognitive environments—largely trusting the word of another is not a route to knowledge in cognitive environments in which, for example, most potential informants are out to trick you. If that's right, however, then the exclusion of epistemic risk is effectively done by the cognitive environment itself. No wonder, then, that a relatively low level of manifestation of relevant cognitive agency can in these conditions suffice for knowledge. In contrast, in cases where the cognitive environment is not so epistemically friendly, a higher level of manifestation of relevant cognitive agency will be required (i.e., less trusting and more scrutiny of both the informant and what she says).

Negative epistemic dependence can also be accounted for. If even a very high level of manifestation of relevant cognitive ability can be consistent with one's cognitive success being a cognitive failure in close possible worlds, then there will no particular level of manifestation of relevant cognitive ability which by itself can guarantee knowledge. We can illustrate this point with what is known as a *twin earth argument*. Imagine two counterpart agents who are microphysical duplicates with identical causal histories. They occupy identical causal environments, one of them on earth and the other on twin earth. Moreover, their normal causal environments—i.e., the causal environments that they usually interact with—are also identical. Suppose each agent now manifests relevant cognitive ability and in the process forms a true belief. Given everything that is common to both agents, it follows that the extent to which each agent's cognitive success is attributable to her manifestation of relevant cognitive ability is exactly the same. After all, in every respect that could possibly have a bearing on this matter, the conditions for both of them are identical. Even so, we could imagine that there is something different about twin earth when compared with earth, in that the agent's true belief on twin earth could nonetheless easily have been false, due to factors purely concerned with the modal environment in which this belief was formed. It follows that although the degree of manifestation of relevant cognitive ability is the same in both cases, only one of the beliefs is safe and hence excludes high levels of epistemic risk. Accordingly, while knowledge is gained by the agent on earth, the agent on twin earth lacks knowledge. This is just as anti-luck virtue epistemology predicts, since on this view even a manifestation of a very high level of relevant cognitive agency will not suffice for knowledge when the belief so formed is unsafe.[5]

The upshot of the foregoing is that we need a theory of knowledge which can allow for an interdependence between agential and non-agential factors in a subject's cognitive success, and that's exactly what anti-luck virtue epistemology offers. Knowledge excludes high levels of epistemic risk, which is why safety is a necessary component.

---

[5] For more on the epistemic twin earth argument, see Kallestrup and Pritchard (2011).

But we also need a certain explanatory relationship to obtain between one's safe cognitive success and one's manifestation of relevant cognitive agency. Finally, we need to understand this explanatory relationship such that it can, in epistemically propitious conditions, impose a relatively weak demand on cognitive agency even though, in epistemically hazardous conditions, the demands on cognitive agency are austere. Anti-luck virtue epistemology fits bill on all these fronts.[6]

3 Varieties of epistemic defeat

Although it is common in epistemology to talk about epistemic defeaters as a single general kind, there are in fact lots of different types of defeater. One question is what it is that the defeater is defeating. For example, is the defeater simply depriving one of knowledge (which is compatible with the idea that one's belief continues to enjoy a positive epistemic standing that falls short of knowledge), or does it deprive one's belief of any positive epistemic standing?[7] Given that the role of defeaters with regard to a particular theory of knowledge is our concern here, we will set this question to one side for our purposes by stipulating that a defeater *at least* does the former, while leaving it open whether it also does the latter too. In particular, the idea is that, absent the defeater, the agent would have had knowledge.[8]

One axis along which we can distinguish defeaters is whether they are *undercutting* or *overriding*. In order to capture the distinction, we need to keep apart the proposition that is being believed (and which one takes oneself to know) and the sustaining basis for that belief. An undercutting defeater will be a consideration that counts against the reliability of the basis in question, to the extent that one can no longer know on this basis alone (at least so long as the defeater remains undefeated).[9] Where the basis is perceptual, for example, it could be that the defeater is that one's perceptual faculties are not functioning as they ought (e.g., because one has unwittingly ingested mind-altering drugs). In contrast, an overriding defeater does not attack the basis for belief, but rather gives one sufficient counterevidence against the truth of the target

---

[6] I have defended anti-luck virtue epistemology—and in the process offered further considerations in its favour, including considering a range of putative counterexamples—in a number of places. See especially Pritchard et al. (2010, chap. 1–4) and Pritchard (2012a, forthcomingc).

[7] Related to this point, one sometimes sees commentators running together the notions of defeat and counterevidence. While they are clearly related, it important to keep them apart, since counterevidence is a less demanding notion. That is, in principle one could be in possession of counterevidence and yet it fails to suffice to defeat the epistemic standing in question (knowledge, say). Note also that while I only contrast knowledge with a general positive epistemic standing here, other contrasts are possible (e.g., one could contrast a positive epistemic standing sufficient for justification with a general positive epistemic standing, and so on).

[8] The point of this formulation is that we don't want to understand knowledge-defeaters in such a way that one must have previously had the knowledge prior to it being defeated (even though sometimes this is of course the case), since that would be unduly restrictive.

[9] Note the emphasis on the *sustaining* basis for the belief. This is important since the process which originally gave rise to the belief might not be the one that currently sustains it. When we are discussing undercutting defeaters it is important that our focus is on the sustaining basis rather than the original basis, since grounds for thinking that the original basis is unreliable may have no defeating power if this basis has subsequently come apart from the sustaining basis.

proposition to ensure that one can no longer know this proposition on the same basis as before so long as this defeater is undefeated. So, for example, imagine one takes oneself to see that there is a vase before one, and yet one is sincerely told by one's friends in the vicinity—whom one has no reason to distrust, and who are just as well-placed to see the object in question as one is—that they can see no vase there. In such a case, at least the primary focus of the defeater is not on the reliability of the way that one formed one's belief, but rather on the truth of the proposition in question.[10]

I say 'primary focus' to highlight the fact that the distinction between undercutting and overriding defeaters is not sharp. After all, as the case just given illustrates, the excellent reasons one is given to doubt that there is a vase right before one are also, thereby, excellent reasons for doubting the reliability of one's perception. Overriding defeaters thus have a tendency to collapse into undercutting defeaters. Nonetheless, it is useful to keep them apart, particularly since the same is not true going in the other direction (or, at least, not true to the same extent anyway). That is, there isn't the same general tendency for undercutting defeaters to collapse into overriding defeaters. That one has excellent reasons for thinking that a particular belief-forming process is unreliable need in itself have no bearing at all on the truth of the belief formed via that process.[11]

Another distinction that we can apply to defeaters concerns whether or not they are *misleading*. Consider an overriding defeater to one's belief that p. If, in fact, one's belief that p is true, then it follows that this defeater is misleading, in that it falsely suggests not-p. Or consider an undercutting defeater to one's belief that p formed on a particular basis M. If, in fact, M is a reliable belief-forming process, then it follows that this defeater is misleading, in that it falsely suggest that M is unreliable. That a defeater is misleading doesn't in itself deprive it of its power to defeat. Consider a concrete case. Imagine that one is observing the frontage of a barn in entirely normal conditions, and so forms a belief on this basis that there is a barn in front of one. Now suppose that one is confronted with a local who (falsely) maintains that one is in barn façade county, where none of the barns are real. One is thus presented with an undercutting defeater, in that one's perceptual basis for believing that there is a barn there (which only takes in the barn frontage) is no longer a reliable means of forming beliefs about whether there is a genuine barn present. But this defeater is no less effective at undermining knowledge if we stipulate that it is misleading—i.e., that the local is in fact making this story up, and there are no barn façades around.[12]

---

[10] For a classic discussion of the distinction between undercutting and overriding defeaters, see Pollock (1970). (Note though that at this time he is calling defeaters 'excluders' and classifying them as 'type I' and 'type II'. Moreover, in later work—such as Pollock (1974)—although he uses the nomenclature of an undercutting defeater, he refers to overriding defeaters as 'rebutting defeaters').

[11] Of course, there are certainly *some* cases where undercutting defeaters collapse into overriding defeaters. If, for example, I discover that your belief-forming process is so unreliable that it is *guaranteed* to deliver false beliefs, then it follows that I have an overriding defeater as regards any belief produced by this process. My point is just that merely discovering that a belief-forming process has failed to clear the relevant threshold of reliability is not in itself generally a good reason to doubt the truth of any belief formed via this process. Rather, one will normally require further grounds (such as grounds for thinking that the process in question is massively unreliable). I am grateful to an anonymous referee from *Synthese* for pressing me on this point.

[12] That said, misleading defeaters tend to be less epistemically deleterious than non-misleading defeaters to the extent that they are easier to in turn defeat. Once one has taken a look behind the barn frontage and

Note that so long as we restrict our attention to knowledge-defeaters—i.e., defeaters that suffice to deprive one of knowledge that one would have otherwise had—then it follows that all defeaters, whether undercutting or overriding, are misleading defeaters. For if we now remove the defeater what we are left with is knowledge. And yet if one knows that *p*, then it must minimally be the case that *p* is true and that it was formed in a reliable fashion. It follows that the defeater in question must be a misleading defeater.[13] Note that the same isn't true if one's focus is on other epistemic standings. Take the notion of doxastic justification, for example, as classically conceived (i.e., such that it involves the possession of good, reflectively accessible, support reasons). On standard views, one can have a justified belief that *p* even when *p* is false and one's belief that *p* is formed on an unreliable basis. Thus, it is perfectly possible for one to have non-misleading defeaters of both an overriding and undercutting variety as regards doxastic justification. In any case, for our purposes we can treat all defeaters as misleading, which means in effect that we can set this distinction to one side.

A further axis along which we can distinguish defeaters that I want to draw attention to concerns whether or not they are *normative*. There are different ways of drawing this distinction, but my focus will be on those defeaters of which one is unaware but which one (epistemically) ought to be aware of.[14] A couple of comments are in order here. First, notice that we are concerned with defeaters of a certain kind *that one is unaware of*. This is because I don't think it matters all that much whether one ought to be aware of a defeater that one is in fact aware of. This is because once one is aware of a defeater one is rationally bound to respond to it, regardless of whether it is a defeater that one ought to have been aware of (e.g., perhaps it is just by fluke that one came across it).[15] Second, notice that this is the *epistemic* ought that is in play here. What I mean by this is that the defeater is something that one ought to be aware of from a specifically epistemic (rather than, say, ethical) point of view. Perhaps the best way of getting a handle on this notion is to imagine what a good—i.e., intellectually virtuous, to a reasonable (though not exceptional) level—inquirer would be aware of in the same conditions. Part of the rationale behind normative defeat is that agents should

---

Footnote 12 continued

seen that there is a genuine barn, for example, then this particular defeater is in turn defeated. Had it been a non-misleading defeater, in contrast, then clearly one would not be able to defeat the defeater in this fashion.

[13] This feature of knowledge has led to a puzzle, sometimes known as the *paradox of dogmatism*. Roughly, if one takes oneself to know that *p*, then one should regard all evidence that suggests that one doesn't know that *p* (including undercutting and overriding defeaters) as misleading. But doesn't that mean that one should reject all counterevidence in advance of knowing what it is? And, if so, isn't that just a licence for dogmatism? For two important discussions of this paradox, see Harman (1973) and Kripke (2011). See also Sorensen (1988, 2011, Sect. 6.2). I offer my own explanation of why the paradox is illusory in Pritchard (2015b, pp. 210–211). My focus there is on how this paradox affects an account of perceptual knowledge known as *epistemological disjunctivism*—see Pritchard (2012b) for more details about this proposal—a view that seems particularly susceptible to the paradox. Accordingly, insofar as my treatment of the paradox works for epistemological disjunctivism, it ought to be applicable to (plausible) accounts of knowledge more generally.

[14] See Lackey (2010) for a very useful discussion of normative defeaters.

[15] Notice that I am implicitly taking it for granted here that being aware of a defeater involves being aware of it *qua* defeater. Of course, there may be certain cases where the latter condition isn't met, but I will be setting this complication aside in what follows as it isn't germane to our current discussion.

not gain epistemic benefits by simply failing to be good inquirers. So, for example, that I am generally inattentive to my surroundings ought not to insulate me from epistemic defeat if there is a clear defeater present in my environment (e.g., a sign warning me that I am now in barn façade county).

There is more to be said about defeaters, and that includes more ways of categorising them. But we have enough of a handle on the main types in order to consider how the notion of defeat plays out in an anti-luck virtue epistemology. In particular, notice that the various ways we have just described of distinguishing defeaters are in principle completely independent of each other. That is, in virtue of being, say, a misleading defeater, nothing at all follows regarding whether this is also an undercutting or overriding defeater. And so on for all the other categories. We thus have a reasonably complex way of mapping out different kinds of defeaters.[16]

4 Knowledge, risk, agency and defeat

We are now in a position to see how an anti-luck virtue epistemology accounts for certain kinds of defeater and, relatedly, certain kinds of non-defeater. Let's start with the latter. Here is a question we might ask: why don't non-normative defeaters that one could easily be aware of—but which, in keeping with the normative/non-normative defeater distinction, one is not in fact aware of—undermine knowledge? After all, were one to become aware of these defeaters, then that would undermine one's knowledge, at least temporarily (i.e., until the defeater is defeated). So doesn't the modal closeness of these defeaters, even though they are non-normative, suffice to make them a threat to one's knowledge? In particular, aren't we both exposed to a high level of epistemic risk in such cases, and also exhibiting an insufficient level of relevant cognitive agency (i.e., in not neutralising them)?

The first point to note in this regard is that modally close defeaters—i.e., defeaters which one could easily become aware of—will tend to be normative defeaters by default. After all, if they really are considerations that one could easily become aware of, then it will tend to be the case that they are considerations that the reasonably attentive subject ought to be aware of. For example, if there is a large sign announcing that one is in barn façade county in one's general vicinity, then this is the kind of thing that one ought to be aware of even if one in fact fails to spot it. Even so, there is no entailment here, in that there surely can be modally close defeaters that are nonetheless non-normative. Let's consider a concrete case.

---

[16] For a helpful overview of recent work on epistemic defeaters, see Sudduth (2015). Note that one issue that I have not engaged with here, but which I have discussed extensively elsewhere, is what it takes to turn the mere presentation of an error-possibility into a defeater (or, for that matter, counterevidence). I think this has an important bearing on a number of debates (such as concerning radical scepticism), since epistemologists are often too quick to treat any presentation of an error-possibility as a defeater. I have also not engaged here with a further issue that I think is related, which is how there can be ways of responding to defeaters that are not particularly onerous from an epistemic point of view (again, I think epistemologists are often too quick to impose unduly austere epistemic demands on those defeating defeaters, particularly in the context of certain epistemological debates). For discussion of both of these points, see Pritchard (2010; 2012b, part two, 2015b, part three). See also Carter and Pritchard (2015).

Imagine that one perceives a barn in entirely normal barn-spotting conditions. Unbeknownst to one, however, there is a sign in the vicinity aimed at tricking visitors to the region that announces that one is now in barn façade county. This sign constitutes a misleading undercutting defeater, in that it falsely suggests that one's basis for forming a belief about there being a barn before one (i.e., by simply looking at it) is unreliable. Here is the twist. Suppose that the sign has recently fallen down, and is now completely hidden, such that there could be no reasonable expectation that one ought to spot its presence, even though one could easily spot its presence nonetheless (e.g., one might easily stumble over it). It is thus a non-normative misleading undercutting defeater, albeit a defeater that one could easily have become aware of. Had one been aware of the sign, then one's knowledge would have been defeated, at least temporarily, but since this is not a defeater that one ought to be aware of, it doesn't defeat one's knowledge.

Anti-luck virtue epistemology can explain what is going on here by noting that in such a case the kind of luck in play is specifically *evidential* rather than veritic, where the former is entirely compatible with one possessing knowledge. That is, it is familiar point of anti-luck/anti-risk epistemology that it can be matter of luck that one has the evidence that one does (i.e., there was a high risk that one lacked this evidence) and yet one can gain knowledge from this evidence nonetheless. The crux of the matter is that our assessments of epistemic luck and risk are explicitly *basis-relative*, as we saw above. That means that when deciding whether there is veritic luck/risk in play, we need to keep fixed what one's evidential basis for that belief is. This is why, for example, one can gain knowledge by luckily overhearing a conversation, even though one could have easily failed to overhear it (i.e., there's a high level of risk that one fails to overhear it). The point is that given that one hears what one does, it may well be that it is not a matter of luck that one forms a true belief on this basis (relatedly, that one forms a false belief on this basis may not be high risk). So anti-luck virtue epistemology has a straightforward way of explaining why modally close non-normative misleading defeaters are not knowledge defeating, since they do not involve knowledge-undermining epistemic luck/risk.[17] Moreover, given that they

---

[17] This general distinction between an epistemically benign form of evidential epistemic risk and an epistemically malignant form of epistemic risk that is veritic can be found in Unger (1968). For a developed account of the distinction and its epistemological ramifications, see Pritchard (2005, chap. 5–6). Note that many putative counterexamples to the necessity of an anti-luck/risk condition like safety for knowledge trade on a failure to take this distinction into account (and, relatedly, a failure to keep the basis for the target belief fixed). Consider, for example, the clock case offered by Bogardus (2014). In this scenario we are asked to imagine the world's most reliable clock which, as it happens, could have easily been disrupted by something in its vicinity (but wasn't). We further imagine our agent just happening to be in the vicinity of the clock when it is functioning perfectly and forming a true belief as a result. Is this belief knowledge? Bogardus claims that it is, but that it is also manifestly unsafe. I agree that it's knowledge, but I dispute that the belief is unsafe. The crux of the matter is that we need to keep the subject's actual evidential basis fixed, and of course his actual evidential basis for his belief is formed by consulting the reliable and unaffected clock. While it is lucky that the subject has this evidential basis (in that there are close possible worlds where it is absent), it is not lucky that she forms a true belief on this basis. Indeed, in all close possible worlds where she continues to enjoy the same evidential basis she continues to form a true belief. (This would thus be a case of a non-normative modally close defeater). See Pritchard (2015a) for a more detailed discussion of why the main putative counterexamples to safety in the literature are defective. See also

clearly do not involve a deficit in one's cognitive performance—since the defeaters are non-normative—hence they do not contravene the virtue-theoretic aspect of anti-luck virtue epistemology either.[18]

In contrast, when it comes to normative defeaters, the reason why they are knowledge-undermining is precisely *because* they highlight an important deficit in one's cognitive performance. In order to see this, we first need to note that not all normative defeaters involve high levels of epistemic risk. Indeed, given that we are focussing on knowledge-defeaters, such that (as noted above) all defeaters are accordingly misleading defeaters, it is simply not possible for any normative defeater relevant for our purposes to involve a high level of epistemic risk. Recall that a knowledge-defeater deprives one of knowledge, such that absent this defeater one would have knowledge. With that in mind, consider what one's basis for belief will be. Since one is by definition unaware of normative defeaters, this basis will accordingly be completely unresponsive to their existence. Hence, it follows that one's basis for belief will be knowledge-conducive, and hence immune to epistemic risk, even though one fails to acquire knowledge on account of the normative defeater. Accordingly, even granted the existence of the normative defeater, there cannot be a significant epistemic risk to the subject's cognitive success, given the basis on which it was formed.

The moral is that normative defeaters are not knowledge-undermining because they prevent one from meeting the safety requirement on knowledge. Instead, what they block is the explanatory link that is required, according to anti-luck virtue epistemology, between one's safe cognitive success and one's manifestation of relevant cognitive agency. We can see this point in action by considering a variation on our last barn façade case. As before, we imagine an agent forming a true belief that there is a barn before her in appropriate cognitive conditions. This time, however, imagine that far from being hidden nearby, the sign announcing that one is in barn façade county is in fact clearly in view, though through inattention our agent fails to spot it. We thus have a normative defeater, in this case of an undercutting variety, in that this is a defeater that one ought to be aware of (but isn't).

If this same scenario were repeated without the presence of the normative defeater, then we would ascribe knowledge, so this prompts the question of what distinguishes this case from the one just offered where the normative defeater is present (and where we do not ascribe knowledge), particularly since both beliefs are not subject to high levels of epistemic risk. The crux of the matter is that only in the case where the normative defeater is lacking would the agent's safe cognitive success be significantly creditable to one's manifestation of relevant cognitive agency. For notice that in the scenario where the normative defeater is present, what ensures that the agent's belief is safe is precisely not her manifestation of relevant cognitive agency but rather her *failure* to manifest relevant cognitive agency. After all, had she properly manifested

---

Footnote 17 continued

Broncano-Berrocal (2014) for further critical discussion of Bogardus's clock case, to which Bogardus and Marxen (2014) is a response.

[18] And notice that it would be irrelevant to respond to this argument by nothing that it only works on misleading modally close non-normative defeaters. After all, as we noted above, insofar as our focus is on knowledge-defeaters, then *all* defeaters are misleading.

her relevant cognitive agency, then she would have spotted the defeater and hence not formed a belief on the target basis (instead, she would have undertaken some further checking).

Anti-luck virtue epistemology is thus well-placed to account for normative defeaters. In particular, the fact that normative defeaters are knowledge-defeating highlights the importance of demanding an explanatory relationship not simply between one's cognitive success and one's manifestation of relevant cognitive agency, but rather between one's *safe* cognitive success and one's manifestation of relevant cognitive agency. For notice that in the case where the normative defeater is present one's cognitive success *simpliciter*—i.e., one's cognitive success on this particular occasion—*is* significantly creditable to one's manifestation of relevant cognitive agency. It is only when we switch our attention to the explanatory relationship between one's safe cognitive success and one's manifestation of relevant cognitive agency that the explanation for why normative defeaters are knowledge-undermining becomes apparent.

We now come to actual defeaters—i.e., defeaters that the agent is aware of (unlike normative defeaters, which by our definition the agent is unaware of). We noted above that overriding defeaters had a habit of collapsing into undercutting defeaters, so let's focus on an instance of the latter. We can adapt our previous barn façade case involving the normative defeater of the sign announcing that one is in barn façade county to serve this purpose—we just need to stipulate that the agent comes across the sign. Like all the defeaters we are considering, this is a misleading defeater, and hence it follows, as we noted above, that there is, objectively speaking, no epistemic risk involved in disregarding this defeater and believing regardless. And yet such a defeater undermines knowledge. Is this then just a failure at the level of cognitive agency, as in the case of the normative defeater just described?

While there is an overlap between the two cases, I think the situation is at least slightly different. What this case shares with the case of the normative defeater is that if one disregards the defeater in question—which is the analogue in this case of failing to attend to the defeater that one ought to have attended to—then it is no longer true that one's safe cognitive success is significantly attributable to one's manifestation of relevant cognitive agency. For imagine a subject who sees the sign in question, but who then opts to continue believing that there is a barn before her regardless, without making any effort to defeat this defeater (e.g., by checking behind the frontage of the barn). As the case is described, our subject is assured a safe true belief regardless, but we would not attribute the safety of her belief to her manifestation of relevant cognitive agency any more. On the contrary, the safety of her cognitive success is now *in despite of* her manifestation of relevant cognitive agency, rather than being to any significant degree because of it.

Aside from this commonality between cases involving normative and actual defeaters, there is also a further knowledge-undermining aspect that concerns only the latter case. As we have noted, in neither case would the agent's belief be subject to a high level of epistemic risk were the agent to continue believing the target proposition regardless of the defeater in question. Nonetheless, there is a sense in which the agent who ignores an actual defeater—unlike an agent who is unaware of, and hence unresponsive to, a defeater that they ought to be aware of—is taking an epistemic risk. That is, from her point of view, in disregarding a defeater (which she knows that she

cannot defeat), she is forming her belief in an epistemically risky fashion, such that she could very easily end up with a false belief. This is different from the case of the normative defeater, since of course this agent is unaware of the defeater in question. We thus get an extra layer of epistemic negativity in this case, and hence more reason to regard it as knowledge-undermining from the perspective of anti-luck virtue epistemology.

This captures the idea that there is *more* epistemic fault in this case than in the previous case. That is, there is more epistemic fault involved in disregarding a defeater of which one is aware than simply failing to attend to a defeater which one ought to have attended to, in that the former involves an active disregard for one's epistemic duties. While neither case involves objective epistemic risk (as both defeaters are misleading), and while both cases involve a deficit in one's cognitive performance, only the former involves actively opting to take an epistemic risk.

Across a range of types of defeater, then, we find that anti-luck virtue epistemology is not only generating the right verdict as to whether or not the agent has knowledge, but is also able to offer a compelling diagnostic story about why these defeaters function as they do.

5 Concluding remarks

I noted at the outset that it is incumbent upon the proponent of any theory of knowledge to explain how their view accommodates defeaters, and that has been my primary goal here. I hope I have also demonstrated, however, that the specific way in which anti-luck virtue epistemology involves an interplay between agential and non-agential factors in the production of knowledge puts the view in a particularly good position to diagnose why different kinds of defeater function in the way that they do. In particular, the crux of the matter is that in evaluating whether an agent's cognitive success amounts to knowledge we are evaluating both whether there was a significant epistemic risk in play, and also whether the elimination of this risk is suitably related to the subject's manifestation of relevant cognitive ability. Sometimes defeaters play upon one aspect of this feature of knowledge and sometimes the other (and occasionally both). This is not to say that alternative theories of knowledge cannot also adequately accommodate defeaters (indeed, I'm sure that they can). But the challenge will be whether their accommodation of them offers the kind of sensitivity to the nature of defeaters that we find with anti-luck virtue epistemology, since if that is lacking then this is an additional consideration in favour of this particular theory of knowledge. I leave this as a further question to explore on a different occasion.[19]

References

Bogardus, T. (2014). Knowledge under threat. *Philosophy and Phenomenological Research*, *88*, 289–313.
Bogardus, T., & Marxen, C. (2014). Yes, safety is in danger, *Philosophia*. doi:10.1007/s11406-013-9508-4.

Broncano-Berrocal, F. (2014). Is safety in danger? *Philosophia*, *42*, 63–81.

Carter, J. A., & Pritchard, D. H. (2015). Perceptual knowledge and relevant alternatives, *Philosophical Studies*. doi:10.1007/s11098-015-0533-y.

Greco, J. (2009). *Achieving knowledge*. Cambridge: Cambridge University Press.

Harman, G. (1973). *Thought*. Princeton, NJ: Princeton University Press.

Kallestrup, J., & Pritchard, D. H. (2011). Virtue epistemology and epistemic twin earth, *European Journal of Philosophy*. doi:10.1111/j.1468-0378.2011.00495.x.

Kallestrup, J., & Pritchard, D. H. (2012). Robust virtue epistemology and epistemic anti-individualism. *Pacific Philosophical Quarterly*, *93*, 84–103.

Kallestrup, J., & Pritchard, D. H. (2013). Robust virtue epistemology and epistemic dependence, ch. 11. In T. Henning & D. Schweikard (Eds.), *Knowledge, virtue and action: putting epistemic virtues to work*. London: Routledge.

Kripke, S. (2011). Two paradoxes of knowledge. In *Philosophical troubles: collected papers* (Vol. 1, pp. 27–51). Oxford: Oxford University Press.

Lackey, J. (2010). Testimonial knowledge. In S. Bernecker & D. H. Pritchard (Eds.), *Routledge companion to epistemology* (pp. 316–325). New York: Routledge.

Pollock, J. (1970). The structure of epistemic justification. *American Philosophical Quarterly*, *4*, 62–78.

Pollock, J. (1974). *Knowledge and justification*. Princeton, NJ: Princeton University Press.

Pritchard, D. H. (2005). *Epistemic luck*. Oxford: Oxford University Press.

Pritchard, D. H. (2007). Anti-luck epistemology. *Synthese*, *158*, 277–297.

Pritchard, D. H. (2009a). Apt performance and epistemic value. *Philosophical Studies*, *143*, 407–416.

Pritchard, D. H. (2009b). Knowledge, understanding and epistemic value. In A. O'Hear (Ed.), *Epistemology. Royal Institute of Philosophy Lectures* (pp. 19–43). Cambridge: Cambridge University Press.

Pritchard, D. H. (2010). Relevant alternatives perceptual knowledge, and discrimination. *Noûs*, *44*, 245–268.

Pritchard, D. H. (2012a). Anti-luck virtue epistemology. *Journal of Philosophy*, *109*, 247–279.

Pritchard, D. H. (2012b). *Epistemological disjunctivism*. Oxford: Oxford University Press.

Pritchard, D. H. (2012c). In defence of modest anti-luck epistemology. In T. Black & K. Becker (Eds.), *The sensitivity principle in epistemology* (pp. 173–192). Cambridge: Cambridge University Press.

Pritchard, D. H. (2015a). Anti-luck epistemology and the Gettier problem. *Philosophical Studies*, *172*, 93–111.

Pritchard, D. H. (2015b). *Epistemic angst: radical skepticism and the groundlessness of our believing*. Princeton, NJ: Princeton University Press.

Pritchard, D. H. (Forthcominga). Epistemic dependence, *Philosophical Issues*.

Pritchard, D. H. (Forthcomingb). Epistemic risk, *Journal of Philosophy*.

Pritchard, D. H. (Forthcomingc). Knowledge, luck and virtue: resolving the Gettier problem. In C. Almeida, P. Klein & R. Borges (Eds.) *The Gettier problem*. Oxford: Oxford University Press.

Pritchard, D. H., Millar, A., & Haddock, A. (2010). *The nature and value of knowledge: three investigations*. Oxford: Oxford University Press.

Sorensen, R. (1988). *Blindspots*. Oxford: Clarendon.

Sorensen, R. (2011). Epistemic paradoxes. In E. Zalta (Ed) *Stanford encyclopedia of philosophy*. http://plato.stanford.edu/entries/epistemic-paradoxes/.

Sosa, E. (2007). *A virtue epistemology: apt belief and reflective knowledge*. Oxford: Oxford University Press.

Sosa, E. (2009). *Reflective knowledge: apt belief and reflective knowledge*. Oxford: Oxford University Press.

Sosa, E. (2015). *Judgement and agency*. Oxford: Oxford University Press.

Sudduth, M. (2015). Defeaters in epistemology. In J. Fieser & B. Dowden (Eds.) *Internet encyclopedia of philosophy*. http://www.iep.utm.edu/ep-defea/.

Unger, P. (1968). An analysis of factual knowledge. *Journal of Philosophy*, *65*, 157–170.

Zagzebski, L. (1996). *Virtues of the mind: an inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge: Cambridge University Press.