

ARTICLE

# Luck and Reasons

Spencer Paulson 

Northwestern University, Evanston, IL, USA

Email: [spencerpaulson2023@u.northwestern.edu](mailto:spencerpaulson2023@u.northwestern.edu)

(Received 15 May 2022; revised 7 September 2022; accepted 31 January 2023)

## Abstract

In this paper, I will present a problem for reductive accounts of knowledge-undermining epistemic luck. By “reductive” I mean accounts that try to analyze epistemic luck in non-epistemic terms. I will begin by briefly considering Jennifer Lackey’s (2006) criticism of Duncan Pritchard’s (2005) safety-based account of epistemic luck. I will further develop her objection to Pritchard by drawing on the defeasible-reasoning tradition. I will then show that her objection to safety-based accounts is an instance of a more general problem with reductive accounts of epistemic luck. In short, they face a dilemma: they can either fail to vindicate the intuitive verdicts about cases or they can illicitly appeal to the epistemic vocabulary they are trying to reduce. The upshot is that we can only understand epistemic luck in terms of the assessment of the subject’s reasons and we can’t give a reductive account of that.

**Keywords:** Epistemic luck; safety; virtue epistemology; reasons-first; fake barns

## 1. Introduction

Can we analyze the kind of epistemic luck that precludes knowledge in non-epistemic terms? By “non-epistemic,” I mean terms that can be understood prior to understanding epistemic evaluation. Safety theories<sup>1</sup>, for example, are non-epistemic. The intent of these theories is to explain the kind of luck that interests epistemologists by appeal to a counterfactual analysis that can be understood without first understanding epistemic assessment or the vocabulary proprietary to it (i.e., rationality, evidence, defeat, etc.). Some success-from-ability theories also fit the bill.<sup>2</sup> According to these theories, knowledge-undermining luck arises when a subject forms a true belief by exercising a cognitive ability of the right sort, but the fact that the subject believes truly is not explained (sufficiently or in the right way) by the fact that the agent formed the belief by exercising that ability.

<sup>1</sup>See Pritchard (2005, 2007) and Kelp (2013), for example. The worries I raise here will apply just as much to safe-methods views (e.g., Hawthorne 2004) as traditional safety-theories. For this reason, I will use “safety” to refer to a genus of which both are species.

<sup>2</sup>See Sosa (1991, 2007), Greco (2003, 2010, 2012) and Turri (2011). These theorists might not think of themselves as giving an alternative to the safety theory of luck but rather giving an account that entails safety and offers further benefits besides. Even so, if my argument succeeds, they will be shown to fail in these ambitions as well.

I will start by considering the safety account. I will look at Duncan Pritchard's (2005) version of it and Jennifer Lackey's (2006) criticism of it. The purpose of doing this will be to try to extract a general lesson from the counterexamples she offers and the recipe she uses for generating them. The lesson, I will argue, is that safety theories only give us the correct result if our assessment of the relevant counterfactuals is guided by an explanatorily prior assessment of the subject's reasons. This spells trouble for the reductive ambitions of the safety-theorist. I will then turn to success-from-ability theories and argue that reductive versions are plagued by the same problem. That is, our intuitions about cases are only vindicated insofar as our explanations of why beliefs turned out true are guided by an explanatorily prior assessment of the subject's reasons for belief. In both cases, the assessment of the subject's reasons will be the kind familiar from defeasibility analyses of knowledge.<sup>3</sup>

The upshot will be that the problems for non-epistemic analyses aren't generated by features that could be Chisholmed away. There is no way to show this conclusively without considering every possible non-epistemic analysis, but a cumulative case will be made that gives us grounds for skepticism about the tenability of the reductive project. This will give some (strong but inconclusive) support for a reasons-first<sup>4</sup> account of epistemic luck.

## 2. Safety

Duncan Pritchard (2005: Ch. 5) offers the following account of luck generally. For an event, E, to be lucky, it must happen in the actual world but not in a wide class of the nearest possible worlds (Pritchard 2005: 128). The other condition on luck is that E must be significant to the agent concerned (Pritchard 2005: 132). The first condition is the one I will focus on here. The basic idea is that lucky events could have easily failed to happen, and we understand ease of failure counterfactually. This is an account of luck generally. It applies to the epistemic and the non-epistemic alike. To explain veritic epistemic luck<sup>5</sup> (i.e., the kind familiar from Gettier cases), he proposes that the truth of a belief is veritically epistemically lucky just in case it is not safe. A belief is safe just in case it is true in the actual world and could not easily have been false. That is, in nearly all (if not all) of the nearest worlds in which the agent forms the belief in the same way, it is true. Pritchard's account of veritic luck is just the result of taking his general account of luck and applying it to epistemology by making true belief the significant outcome.

In her review of Pritchard's book, Jennifer Lackey (2006) offers a series of criticisms as well as a general recipe for generating them. I will proceed by first considering one of her cases, then considering her recipe, and, finally, adding some details to the recipe in order to make clear the exact nature of the problem she has identified. In the next section, I will apply the lessons learned in this section to reductive success-from-ability accounts and show that a similar objection can be made against them as well. This will give us reason to think that the problem lies not just with safety-based accounts, but rather with reductive accounts generally. In learning about what epistemic luck is not, we will shed some light on what it must be along the way.

<sup>3</sup>See Klein (1980), Pollock (1986) and Moser (1989). See Shope (1983) for a useful overview of the history of defeasible reasoning as a solution to the Gettier (1963) problem.

<sup>4</sup>Here I speak of reasons rather than evidence, though the point could just as well be made in terms of quality of evidence (cf. Klein 2017).

<sup>5</sup>The term is first used in Engel (1992).

*Southernmost Barn*: While entering a Midwestern farming community on her cross-country drive, Janice looked at the first barn that she saw, which was on the southernmost end of the field, and formed the corresponding belief “There is a barn”. As it happens, the barn she saw is the only real one, surrounded by barn façades that members of this community have placed in the field in order to make their town appear prosperous. However, as a matter of strict and unwavering policy, the members of this community always place their only real barn on the southernmost end of their land since this is where traffic first enters their town. Moreover, thirty years earlier, Janice had lived in a house on the southernmost end of this field in the precise location of the one real barn. Because of her deep interest in her childhood roots combined with the brief period during which she can safely take her eyes off of her driving, she would invariably have looked at only the particular place in the field where the real barn exists. (Lackey 2006)

Lackey claims, rightly I think, that this is a paradigmatic case of epistemic luck.<sup>6</sup> She also points out that the subject’s belief is safe. A lot would need to change for her to believe falsely forming a belief in that way. This is the point of introducing the “strict and unwavering policy” of the community and the biographical details about the subject only being disposed to look at a very particular spot. Furthermore, if there were any doubt that the subject’s belief is safe, we could simply make the vignette longer and introduce more details to guarantee counterfactual robustness. The key is to make the details sufficient to ground the counterfactual but epistemically irrelevant.

Of course, the counterexample she proposed is not an isolated issue. It is not hard to generate similar ones. Lackey offers us the following recipe for cooking them up:

Still further, numerous other kinds of counter-examples similar to *Southernmost Barn* can be constructed which fail Pritchard’s safety-based view of knowledge. First, choose a paradigmatic Gettier-type case, such as Bertrand Russell’s stopped clock example. Secondly, construct the case so that there is a feature  $x$  such that nearly invariably, when  $x$  is present,  $S$  will form the true belief that  $p$ , and  $p$  will be true only when  $x$  is present. Thirdly, ensure that the connection between  $x$  and  $p$  is entirely non-epistemic, e.g.,  $S$  believes that  $p$  only when what is picked out by  $p$  is pink, and  $p$  will be true only when what is picked out by  $p$  is pink. Fourthly, if there are any residual doubts that the belief is genuinely safe, add further features to guarantee counterfactual robustness across nearby possible worlds. (Lackey 2006)

This is correct, as far as it goes. However, we should try to be more precise about what we mean when we say that the connection between two things is epistemic. I will here go beyond what Lackey explicitly says, although my proposal is in the same spirit. To foreshadow a bit, I will draw on the defeasible reasoning literature to unpack the epistemic/non-epistemic distinction. Apart from the arguments that have

---

<sup>6</sup>She claims that it is a paradigmatic Gettier case (Lackey 2006). If we wish to distinguish classic Gettier cases from post-Gettier cases, we might quibble with this. Nothing important would hang on the outcome of the quibbling. It is generally agreed that post-Gettier cases (such as Ginet’s original fake barn case) are cases of epistemic luck and *Southernmost Barn* resembles Ginet’s original case in all the important respects.

already been made in favor of centering epistemology around defeasible reasoning<sup>7</sup>, we will see that the defeasible reasoning literature earns its keep by enabling us to be more articulate about why some features of *Southernmost Barn* are epistemically relevant while others are not in a way that vindicates the intuitive verdict. Next, we will see that we can use defeasible reasoning to explain why the safety theory gets many cases right, but it gets Lackey-style cases wrong.

I recommend trying to get traction on the epistemic in terms of reasons.<sup>8</sup> Consider *Southernmost Barn* again. The subject has a prima facie reason to think that she is looking at a barn: her perceptual experience of it. However, that reason is defeated by the fact that she is in fake barn county. That is, the fact that she is in fake barn county is a prima facie reason to give up her belief that she is seeing a barn because it undercuts the relation between her perceptual evidence and the truth of her belief. After all, things would look the same to her even if she were looking at one of the many barn facades in the area. The defeating reason is not part of her possessed evidence. So, this reason presents an obstacle to her having knowledge though not to her having justification. This, at any rate, seems plausible. Some varieties of direct realism might say otherwise because they individuate perceptual states in such a way that the perceptual states themselves are different in the good and bad cases.<sup>9</sup> I submit that it would be a bit doctrinaire to deny that perceptual evidence is defeasible (even in the good case). So, I will assume going forward that it is, even if perceptual states themselves are individuated as the direct realist says they are.

It is possible for a subject to have a true belief based on a defeated reason and still have knowledge. Defeaters are prima facie reasons to give up a belief. So, they need not be ultima facie reasons to do so. Defeaters can themselves be defeated. Sometimes they are defeated in such a way as to restore the epistemic potency of the subject's justifying reason. Other times, the defeater-defeater gives the subject new reasons rather than restoring the old ones.<sup>10</sup> *Southernmost Barn* involves the latter sort. The reason to give up the subject's belief is the undercutting defeater mentioned earlier. However, that reason is trumped by the fact that the subject happens to be looking at the one real barn in the area. However, it isn't clear that this restores the epistemic potency of her perceptual reason. Perceptual reasons involve a general capacity for discrimination. However, what is doing the real work here is not her ability to generally discriminate barns from non-barns but rather the fact that she happens to be looking at the one real barn in the area. Insofar as she is basing her belief on a general capacity for discriminating barns from everything else, she is in trouble. She does not have that general capacity in fake barn county. It is true that when she is looking at this one particular real barn, she gets the right answer. Could we say that her perceptual experience is a suitable (for knowledge) reason in rural areas, but not in the narrower

<sup>7</sup>See e.g., Audi (1993), de Almeida and Fett (2016), Klein (2017) and Lehrer (2017).

<sup>8</sup>More specifically, I will do so from within the defeasible reasoning tradition. Some prominent objections to this approach that I will not consider here are those of Feldman (2003), Foley (2012) and Turri (2012). In my view, an ample response to all three can be found in de Almeida and Fett (2016). Lasonen-Aarnio (2010a, 2010b, 2014) and Baker-Hytch and Benton (2015) raise a distinct set of worries I will not be able to address in this paper.

<sup>9</sup>For example, McDowell (1994) and Brewer (1999). Some direct realists, such as McDowell, want to accommodate the defeasibility of perceptual evidence. See Ginsborg (2006) for discussion.

<sup>10</sup>To my knowledge, Klein (1980) first made this point. See also Pollock (1986: Appendix) and de Almeida and Fett (2016).

reference class of fake barn county, but it is once again suitable in the yet narrower reference unit class with this particular real barn as its sole member?

It is worth noting that perceptual reasons in general only function against certain background assumptions. For our perceptual experiences to count in favor of believing anything, we must assume that lighting conditions are more or less standard, that we haven't been drugged, etc. It may be true that perceptual experiences can function relative to background assumptions much more fine-grained than the general default ones (perhaps even specifiable using demonstratives and applicable only to a single case). Nonetheless, those are not the background assumptions the subject is actually making. In fake barn cases, we are imagining a subject ignorant of local idiosyncrasies. So, we imagine her making the default background assumptions she makes for perception in general. Those default/general background assumptions are defeated by the fact that she is in fake barn county. Nothing about the particular barn she is looking at restores *those* background assumptions to their knowledge-supporting post: they are false. Since the standing of the subject's reasons depends on her (actual) background assumptions, the standing of her reasons is problematic in such a case.

I am in effect granting the earlier proposal that her perceptual experience is a suitable (for knowledge) reason in rural areas generally, but not in the narrower reference class of fake barn county, but it is once again suitable in the yet narrower reference (unit) class of this particular real barn. So, her foreground perceptual reason is undefeated by the facts and consequently capable of supporting knowledge. However, her background assumptions (to the effect that she is in a normal perceptual environment) are defeated by local idiosyncrasies and never restored later on because they are false. I am not assuming that there can be no knowledge from falsehood (cf. Warfield 2005; Klein 2008). I am assuming that fake barn cases are not cases of knowledge from falsehood.

I am taking the familiar apparatus of defeasible reasoning and then applying it separately to foreground reasons and background assumptions. Her foreground reasons are undefeated by the facts because there are other background assumptions relative to which they could function that are not themselves defeated by the facts (these are the demonstratively specifiable ones only applicable to her particular case). Admittedly, she isn't making those background assumptions and isn't in a position to reasonably do so. However, that doesn't impugn her foreground reasons themselves, only her own epistemic standing in relying on them the way she does. The foreground reasons as such are fine so long as there are some undefeated background assumptions relative to which they could function. Nonetheless, the subject can't know on the basis of them because foreground reasons require background assumptions and the assumptions she is actually making are defeated by the facts.

It is worth contrasting this with a case where a subject sees a barn in ostensibly normal conditions and believes there is one before her but, unbeknownst to her, the *New York Times* just published an article saying that the county she is in is fraught with Potemkin barns. The Times report turns out to be false. The testimony from the *Times* provides an undercutting factual defeater. It undercuts her foreground reason by rebutting her background assumption that she is in normal perceptual conditions. The fact that the reporter made a mistake is a defeater-defeater that restores her perceptual reason by restoring her background assumptions (conditions are normal after all, the defeater is seen to be a "false alarm" at the end of the vignette). Unlike the traditional fake barn case, the original background assumption that conditions are normal is restored and does not need to be replaced by new, more specific

background assumptions. So, unlike the fake barn case, the subject is not a victim of knowledge-undermining luck.<sup>11, 12</sup>

Let us now return to Lackey's point about the epistemic. We have seen that only facts that bear on the quality of the subject's reasons are properly epistemic. This is what explains our intuitions about Lackey-style cases. When I talk about the quality of the subject's reasons, there are two dimensions of assessment that are relevant. One dimension of assessment has to do with the subject's rationality/internalist justification. For her reasons to do well along this dimension they must be *prima facie* reasons undefeated by her total evidence. However, a subject can lack knowledge despite being perfectly rational. This is because her reasons are defeated by facts of which she is unaware. So, the first dimension of assessment involves surviving the crucible of potential defeat against her total evidence. The second involves surviving the crucible of potential defeat against the facts. When I speak of quality of reasons without specifying either dimension of assessment, I mean to talk about the dimensions jointly. That is, if I say there is an issue with the quality of the subject's reasons, I mean that they are lacking along (at least) one of these dimensions.

We can give an account of what goes wrong in Gettier cases in terms of quality of reasons. What cases like *Southernmost Barn* show us is that a subject's reasons can be defeated (and hence incapable of supporting knowledge) but the belief can nonetheless be safe. To generate cases like this, we just have to weave details into the story that are causally significant but not (restoring) defeater-defeaters. That is, we write the story so that it includes a factual defeater (just like in a Gettier case). Then, we can make the truth of the subject's belief counterfactually robust by adding details to the story that have no bearing whatsoever on the quality of her reasons. The details need to be causally efficacious but not (restoring) defeater-defeaters. The fact that the subject happens to have a nostalgic attachment to one particular region of the county has no bearing on the quality of her reasons. It is a brute psychological fact. It does not have any bearing on the question of whether her background assumptions are capable of supporting knowledge. The fact that the community members invariably put the one real barn in this spot has no bearing on the quality of her reasons either. It is causally significant but not epistemically significant in this case. The subject believes truly and in a counterfactually robust way, but it is still a case of luck.

We can now see why the safety theory has some apparent plausibility as a theory of veritic luck and also diagnose why the plausibility is merely apparent. In the examples Pritchard offers to motivate the theory, the salient features of the vignette are properly epistemic. When he has us consider unsafe beliefs that are not (intuitively) knowledge, the causally salient aspects of the story bear on the quality of the subject's reasons. In classic Gettier cases, for example, the only salient details are the subject's *prima facie* reasons and the facts that defeat them. Our assessment of the relevant counterfactuals will (*ceteris paribus*) be guided by our assessment of the subject's reasons. If her reasons are of the sort required for knowledge, then we will be inclined to think the truth of the subject's belief is counterfactually robust as a result. What

<sup>11</sup>My own view, which I do not have the space to develop in this paper, is that the subject in a fake barn case has their knowledge somewhat but not entirely undermined by luck (Paulson 2023). What matters for my purposes in this paper is that there is *some degree* of epistemic luck in fake barn cases that at least to some degree undermines the subject's knowledge. I succeed so long as I can explain the degree of epistemic luck we find in these cases in terms of how we assess the subject's reasons.

<sup>12</sup>This differs from Harman's (1973) take on this kind of case.

Lackey has shown is that the epistemic and the counterfactual only align the way Pritchard needs them to *ceteris paribus*. When we introduce causally efficacious but epistemically irrelevant noise into the vignette, the counterfactuals come out true despite the subject clearly lacking knowledge. That is, when things are not equal, the counterfactual robustness of the belief is divorced from the quality of the subject's reasons.

In fact, we are now in a position to see why a belief can be irrationally held but nonetheless safe.<sup>13</sup> It is because we can introduce causally efficacious but epistemically irrelevant details into the vignette. That is, we can introduce details that play a causal role but don't bear on the quality of the subject's reasons. This is why the safety theory only gives us the right verdict when the causally efficacious factors also bear on the quality of the subject's reasons. Conversely, when the safety theory gets the right result, it is because the counterfactuals are made true by facts about the quality of the subject's reasons in the actual world. In other words, if we don't already understand which facts are epistemically relevant and restrict the causally efficacious factors in our vignettes accordingly, then we will encounter cases such that the subject's belief is intuitively epistemically lucky but nonetheless safe. This shows that the problem is generated by Pritchard's preferred order of explanation rather than a minor detail that could be Chisholmed away. The defeasible reasoning approach scores points not only because it is not susceptible to Lackey-style counterexample, but also because it enables us to articulate why the safety theory gets cases right when it gets them right and why it gets them wrong when it gets them wrong, as we have just seen.

The assessment of the subject's reasons must be explanatorily prior to the counterfactuals (insofar as the theory is extensionally adequate) because facts about the quality of the subject's reasons determine the truth of the counterfactuals. If we assume the theory of counterfactuals in Lewis (1973),<sup>14</sup> then the point is that the similarity relation on worlds is explanatorily prior to the truth-values of counterfactuals but explanatorily posterior to the assessment of the subject's reasons.<sup>15</sup> If the quality of her reasons is good in the actual world, then worlds in which she bases her belief on them and believes something false will seem remote to us. Quality of reasons judgments guide modal proximity judgments because the quality of the subject's reasons influences how we see the similarity relation. Other things equal, rationally similar worlds are nearby. If the similarity relation is determined by quality of reasons and the counterfactuals in Pritchard's analysis are determined by the similarity relation, then we understand the counterfactuals (partly) in terms of quality of reasons. This, at least, is what happens in the cases where Pritchard gives us the right answer. However, he cannot (given his reductive ambitions) concede that the assessment of the subject's reasons is prior to the theory of counterfactuals. So, he has no principled way of eliminating alternative similarity relations. Lackey exploits this fact. This is just a restatement of what we have already seen, but in the terminology of Lewis' theory of counterfactuals. We have reason to believe it because it is the best explanation of why Pritchard's theory gets many cases right but is susceptible to Lackey-style counterexamples.

We can only state what has gone wrong in Lackey-style cases because we understand the assessment of reasons independently of the counterfactuals that interest Pritchard. It

<sup>13</sup>Thanks to an anonymous referee for pointing out the need to discuss this matter.

<sup>14</sup>Which I am doing only for ease of exposition and because Pritchard does as well.

<sup>15</sup>I am not saying that counterfactuals *in general* are explained in terms of reasons, just that the epistemic counterfactuals relevant to the safety theory either will be explained in terms of reasons or the theory will be extensionally inadequate.

follows that the safety-theorist's order of explanation is what generates the problem. So, it cannot just be Chisholmed away.<sup>16</sup>

There is a worry I should address before proceeding. Those sympathetic to Pritchard will wonder if we can really understand quality of reasons without illicitly relying on a prior understanding of counterfactuals at some point.<sup>17</sup> To address this worry, two hurdles need to be cleared. The account I recommend explains the assessment of reasons by first picking out the *prima facie* good reasons and then using defeasible reasoning to determine whether they are *ultima facie* good reasons. Defeasible reasoning and defeat can easily be shown not to rely on a prior understanding of counterfactuals. The dimensions of assessment involved (i.e., defeat relative to total evidence and defeat relative to the facts) both have to do with rational relations between things in the actual world. The relation between defeater and defeated is a rational relation, so the view does not reduce reasons to something non-epistemic. It also does not rely on counterfactuals since it is actual world defeat that is of interest here and, as I have been arguing, the epistemic counterfactuals that interest Pritchard are explained in terms of rational facts about the actual world rather than vice versa.

However, we could still worry about *prima facie* reasons. In virtue of what are some reasons *prima facie* good and others not? *Prima facie* reasons are epistemic foundations. So, to answer the question I will need to give an adequate foundationalist account. I need to do this eventually, but I cannot in this paper. I can mitigate some worry though by pointing out the following. If some version of internalist foundationalism is tenable, then this worry can be addressed. Internalist foundationalisms do not explain the epistemic potency of the foundations in non-epistemic terms. *A fortiori*, they do not give a reductive counterfactual account of the foundations. So, my conclusions in this paper are conditional on the tenability of some version of internalist foundationalism. It is worth noting in this connection that Pritchard himself is an internalist about justification, he just denies that knowledge entails justification (Pritchard 2005). So, he too is committed to the claims I am assuming here.

### 3. Success-From-Ability

When Stephen Curry makes a three-pointer, the shot is made because of his skill. Curry's skill explains the good outcome. When I make a three-pointer, it is quite often not because of my skill. I make a lot of mistakes and sometimes they happen to off-set each other in just the right way. This suggests a general schema. There are skills (i.e., abilities, competences, virtues) and there are successful outcomes that the exercise of the skill is supposed to bring about. Archery is a skill, hitting the bullseye is a successful outcome. Shooting baskets is a skill, making a basket is a successful outcome. Sometimes the exercise of the skill explains why the successful outcome is obtained. Sometimes the skill is exercised and the successful outcome is obtained, but the latter is not explained by the former.

A number of authors<sup>18</sup> have tried to make sense of epistemic luck in terms of this schema. They argue in one way or another that a subject knows just in case her true belief is formed by exercising an ability or competence and the fact that she exercised

<sup>16</sup>Lackey (2006) says things to this effect more than once, although her discussion of the point is brief and her diagnosis less committal than mine.

<sup>17</sup>Thanks to Sandy Goldberg for raising this concern.

<sup>18</sup>Among them Sosa (1991, 2007), Greco (2003, 2010, 2012) and Turri (2011).



that ability explains why she ended up with a true belief. It is not lucky that the subject's belief is true when she knows. If this schema can shed light on the absence of luck, then it can also shed light on the presence of luck. If the subject's ability does not explain why her belief is true, then the truth of her belief is lucky. The belief is consequently not knowledge.

The basic idea is clear enough. Consider the Nogot/Havit case from Lehrer (1965). The Professor formed the belief that Nogot owns a Ferrari because she saw him driving one (i.e., she exercised her perceptual competence) and he told her he owned it. She existentially generalized and came to believe that someone in her class owns a Ferrari (exercising a competence of deductive reasoning). The resulting belief was true, but only coincidentally. Mr. Havit, another student of hers, happened to really own a Ferrari. Nogot was lying. He was borrowing his rich uncle's Ferrari. So, the professor formed a true belief by exercising her competences, but the explanation of why she believed something true is peripheral to all this. She believed something true because another student of hers happened to own a Ferrari.

The basic idea seems right. The question is whether there is an account of explanation in the offing that can vindicate it without tacitly relying on a prior understanding of quality of reasons. The kind of success-from-ability account I am considering here is supposed to be reductive. It is supposed to explain epistemic luck in non-epistemic terms.<sup>19</sup> This only works if we can understand the relevant sense of the non-epistemic terms without a prior understanding of any epistemic terms. I claim that if we really do understand luck in non-epistemic terms, we will get the wrong verdicts on cases. If we get the right verdicts on cases, we will end up illicitly presupposing the vocabulary of epistemic assessment that is supposed to be getting reduced. In this way, this section mirrors the previous one.

To see if there really is such a dilemma, let us consider some prominent versions of the theory. Ernest Sosa (2007) tells us that the subject knows just in case her belief is accurate (i.e., true) because it is adroit. Adroit beliefs issue from competences (basically skills or abilities). How does the adroitness need to figure into the explanation of the accuracy? Sosa takes a permissive stance here. The adroitness of the belief needs to be an important part of the explanation of its truth. However, it does not need to be the most important. It can just be one of many important factors.

Sosa makes a distinction between a competence explaining why a subject has a belief and a competence explaining why that belief is true to get around the classic Gettier cases. In Lehrer's case from earlier, Sosa's account maintains that the professor's competence only explains why she formed the belief but goes no distance toward explaining why she achieved accuracy in forming it. This is true, but it does not help in post-Gettier cases where the competence explains not just why the subject has a belief, but also why it is an accurate one. In barn façade county, the subject sees a real barn in normal lighting conditions, etc. Her perceptual competence is an important part of the explanation of why it is a true belief: she saw it in standard lighting conditions where things appear as they are. Sosa is aware of this consequence and comfortable biting the bullet (2007: 96 fn. 1). This plays into the hands of my dilemma. I claim that the reductive theorist can either illicitly rely on defeasible reasoning or they will be unable to vindicate intuitive verdicts about cases. Sosa took the second horn. Furthermore, Sosa's permissive stance allows for certain self-fulfilling prophecies to count as knowledge when, intuitively, they are not. Suppose I am raising a child and teaching them to

<sup>19</sup>So, knowledge-first virtue epistemologists (e.g., Kelp 2013; Miracchi 2015) are not the target here.

perform inductive inferences. Suppose the first time they do so competently, I want to reinforce the behavior by rewarding it. Unfortunately, the conclusion they draw is false (despite being well-supported). So, I change the world to make the conclusion true and then congratulate the child for getting it right.<sup>20</sup> The child's competence is an important part of why they believed something true. The fact that the child believes truly is nonetheless epistemically lucky.

John Turri (2011) also takes this horn (regarding fake barns, not self-fulfilling prophecies).<sup>21</sup> He claims that in cases of knowledge, the truth of the subject's belief is a manifestation of her cognitive virtue. In Gettier cases, the subject's true belief is merely caused by the virtue, but not a manifestation of it. He claims that we should rely on our "robust pre-theoretical" (Turri 2011) understanding of manifestation to understand the difference. He is right that we have a solid, pre-theoretical understanding of manifestation vs. mere causation generally. The fact that a safe is fire-resistant might cause me to buy it. Nonetheless, my buying the safe is not a manifestation of its disposition to endure fire. When the building burns and the safe is still standing, that is a manifestation of its disposition to endure fire.

We should worry that insofar as we get the intuitively correct results, we are relying on more than just our pre-theoretical understanding of manifestation generally. In the classic Gettier cases, we are inclined to say that the truth of the subject's belief is not a manifestation of the subject's competence, but merely caused by it. Why? Turri says it is just our pre-theoretical understanding of the manifestation relation, which is not further analyzable. However, we notice a general pattern. When there is an issue with the quality of the subject's reasons, we will say that we are dealing with mere causation. When there are no concerns about the quality of reasons, we will say that we are dealing with a case of manifestation. This gives the impression that we partition cases based on our understanding of a distinctively epistemic kind of assessment and it guides our judgments about manifestation when we are dealing with epistemic matters. Our prior understanding of reasons makes us a great deal more articulate about why we intuit manifestation where we do. However, our understanding of manifestation in general does not make us any more articulate about quality of reasons. If we could ground epistemic assessment in a prior understanding of manifestation, then we should be able to *derive* facts about quality of reasons from facts about manifestation. However, we cannot, according to Turri, say much about the nature of manifestation and why it shows up where it does. We just rely on our intuitive understanding of it to partition cases. So, the prospects for this look bleak. In fact, if Turri were to derive facts about manifestation from something more fundamental, he would be giving up on his preferred order of explanation. Recall that according to him the manifestation relation is metaphysically and explanatorily primitive.

However, if we go in the opposite direction, a different picture emerges. We can derive facts about manifestation from facts about quality of reasons. This is just to say that if the subject's reasons do well along the two dimensions mentioned earlier, then the truth of her belief is intuitively attributable to her cognitive abilities and not just caused by them. In Turri's preferred terminology, the subject truly believing is a

<sup>20</sup>This is similar to the "Guardian Angel" case from Greco (2012).

<sup>21</sup>Which is not to say his view is the same as Sosa's, only that they both take the non-standard position that the subject knows in fake barn cases. Turri offers independent argument for reconsidering our intuitions, so it might not be exactly fair to say that he is biting bullets here. Nonetheless, the conclusion is tough for most people to accept.

“manifestation” of those abilities. I submit that the best explanation of the above is that Turri’s order of explanation is the reverse of what it should be. The defeasible reasoning tradition gives us the resources to explain why the subject believing truly is differently attributable to her cognitive capacities in the two cases: they afford her reasons of differing quality. The defeasible reasoning tradition can go a step further using the apparatus of defeat to explain why they differ in quality. Turri’s proposal can’t offer anything comparable. There is a lot we can say about reasons and how they are susceptible to defeat, but according to Turri there is not much we can say about the manifestation relation and why it shows up where it does. So, it holds no promise of helping us better understand defeasible reasoning, whereas defeasible reasoning has sufficiently rich resources to help us better understand it.

The upshot is that Turri gives us the worst of both worlds. In fake barn cases where he gives us a principled answer that really does only depend on our understanding of manifestation generally, we get an intuitively incorrect verdict. When we get the intuitively correct verdict in classic-Gettier cases, the verdict is grounded in an independent understanding of the quality of the subject’s reasons.<sup>22</sup>

John Greco (2003, 2010) offers us a similar picture. Greco holds that a subject knows *p* just in case that subject believes the truth (with respect to *p*) because she formed a belief that *p* by exercising an intellectual virtue (Greco 2010: 71). An intellectual virtue here is understood in the same way as abilities/competences. Greco says that, in general, a full explanation of an event requires many causal factors. However, some of those factors are going to be more salient to us (the explainers) than others. This is a function of their relative abnormality and our practical interests (2003, 2010). The presence of sparks and the presence of oxygen are both causal factors contributing to a warehouse fire. Nonetheless, the sparks are more salient than the oxygen. The presence of oxygen is typical. It is there on the days when there are no fires. The sparks are a different story. Sparks are unusual. So, they are more salient in our explanation. Relatedly, it is difficult for us to control the presence of oxygen but much easier for us to control the presence of sparks. So, we have a practical interest in focusing on the sparks rather than the oxygen in our explanation.

Greco thinks he can use the above to shed light on Gettier cases. Given our practical interests in exchanging information, it makes sense that intellectual virtues have a default salience for us. So, in a normal case of learning through perception, the subject’s perceptual ability will be the most salient part of the explanation of why she believed truly as a result. In Gettier cases, on the other hand, some abnormal events are even more salient. For instance, the fact that Nogot was lying and that, coincidentally, Havit owns a Ferrari are more salient due to their abnormality. Since those are the most salient parts of the explanation of why the professor believed truly, her intellectual virtues have been demoted. So, she does not know. Even though her virtues are a part of the story, they are not a sufficiently salient part of it.

Later, Greco (2012) recognized that some kinds of abnormality matter and others do not. He offers the following example to make this point. Suppose I hear my child shriek and turn around to see what happened. I then see that the cat has a bird in its mouth. The sound of the shriek is abnormal. It doesn’t seem that this abnormality has any bearing on whether I know that the cat has a bird in its mouth, however. Something abnormal got my attention, so in that respect I might be lucky to have acquired the

<sup>22</sup>Turri has more recent work on the topic I cannot go into here. See Carter (2014) and Turri (2016) for what I take to be a good response.

evidence I in fact did through perception. Nonetheless, there is no knowledge-undermining luck here.<sup>23</sup>

Greco does not think that this is the most pressing problem for his view, so he puts it aside after mentioning it briefly. I disagree. I urge that this problem is the same one we keep seeing. If it is, it cannot be solved by Chisholming because it stems from the reductive ambitions of his theory rather than the details.

I can best make this point by making a comparison with the earlier discussion of Pritchard. The problem for Pritchard is that the truth of the subject's belief can be as counterfactually robust as we like (relative to its actual basis) and still be the result of knowledge-undermining luck so long as the counterfactual robustness is the result of factors that have no bearing on the quality of the subject's reasons. Someone could form a true belief about the number of a winning lottery ticket by just wishfully thinking about her favorite numbers. With enough ingenuity, we can make the fact that these are her favorite numbers as counterfactually robust as we want. We can make the fact that those numbers were on the winning ticket for this drawing as counterfactually robust as we want (e.g., by telling a story about the number generator's algorithm and the previous states of the machine). We can tell a story about why she would only have formed that belief about the winning ticket for this drawing by making that particular date significant to her in a way that could not have been different unless the whole course of her life had also been different. None of this changes the fact that she made a lucky guess. And we can change the example a bit and make it a Gettier case. The key point is that we can make the truth of her belief counterfactually robust *in the wrong way* for knowledge. We can make the truth of the belief modally secure and lucky, so long as we off-set the defeaters with causally compensating but epistemically irrelevant factors.

The same problem arises in a different form for Greco. Abnormality, as such, does not allow us to distinguish between epistemically relevant abnormality and epistemically irrelevant abnormality. No amount of Chisholming will change this. The reason is that Greco wants a reductive account of epistemic luck. He wants to explain it in non-epistemic terms. This forces him to traffic in terms such as "abnormality." Abnormality, as such, is not a distinctly epistemic property. So, it suits his reductive purposes. However, for the same reason, it does not enable us to distinguish between the kind of abnormality that matters for epistemology and the kinds that do not. It seems clear enough why the abnormal shriek did not undermine my knowledge: it did not amount to a defeater! However, he can't say that. Abnormalities as such are too blunt an instrument to help us distinguish between cases when you are lucky to have the evidence you do and cases where it is a matter of luck that your belief (formed for the reasons it was) is true. Greco cannot pick out the right class of abnormalities without shirking his reductive ambitions.

Greco is right that some events have a default salience in our explanations of why a subject believes truly and that this salience can be overridden by abnormal factors. He is right that this is a general fact about explanation and in no way specific to epistemology. However, he is wrong to think that this gives him everything he needs for a reductive account of epistemic luck. The reason is that in order to understand the relevant sense (for epistemology) of *default salience* and *abnormality*, you must first understand defeasible reasoning. Some reasons give us prima facie justification for a belief. If we end up with a true belief based on them, they have a default salience in the explanation

<sup>23</sup>It is of the same general sort as Nozick's (1981) "Jesse James" case. Cf. Engel's (1992) novelist case.

of why the subject believed something true. Nonetheless, the justification (and consequently the salience) can be overridden by a defeater. The default salience of her *prima facie* reason is trumped by the greater salience of the defeater.

In some sense, defeaters are abnormal. That is why *prima facie* reasons are good reasons, at first glance. They do the trick unless something unusual happens. However, to specify the kind of unusual event that matters, you must have recourse to the vocabulary of defeat. This is because we are interested in a specifically epistemic kind of abnormality. Defeaters are only salient to us insofar as we are interested in epistemology. Insofar as we are interested in epistemology, other kinds of abnormality don't matter at all.

It is also worth noting that defeaters are *prima facie* (not *ultima facie*) reasons to give up a belief. So, something can be abnormal in the epistemically relevant sense of "abnormal" without undermining knowledge. This happens when the defeater is itself defeated in such a way as to restore the subject's original justification (cf. Lehrer and Paxson's (1968) "Tom Grabit" case). So, for an account like Greco's to work, we would not just need to find a way to delineate the relevant sense of "abnormal" without using any epistemic vocabulary. We would also have to explain why certain iterated abnormalities undermine knowledge and others do not. We would have to do this without mentioning the rational relations between the abnormalities. Yet the rational relations are the only ones that matter.<sup>24</sup>

Greco (2012) has offered us a revised account. He now says that the truth of a subject's belief is attributable to her abilities just in case her abilities contribute to believing the truth in a way that could regularly serve the interests of the community. The problem with Gettier cases is that crucial aspects of the explanation are too idiosyncratic to be consistently exploitable.

I see no reason why we can't generate recurring Lackey-cases. Suppose I (for complicated biographical reasons) suffer a compulsion to look at the same clock every day at noon. My circadian rhythm is acute, sub-personal processes cause me to look at exactly noon even though I have not yet formed the belief it is noon. The clock has long been stopped at 12 and (for complicated historical reasons of which I am ignorant) this could not change without much else also changing. When I base my belief that it is noon on what I see looking at the clock, I am Gettiered. I am lucky to get the time right, but I do not know the time. Nonetheless, I'm always right and after forming the belief that it's noon, I walk through town and tell everyone the time. If they could set their watches to Kant's walks, they could set their watches to mine. The classic Gettier cases have a one-off character. However, that is an accidental property. Perhaps the townspeople could not come to know the time on the basis of my testimony. They could consistently get the right answer, despite being lucky to do so. Their reasons would be defeated. My testimony is undercut by the fact that the clock I based my belief on was stopped. None of the biographical idiosyncrasies need restore my reasons for belief to ensure that I always check the clock at the same time. The possibility of regularly exploiting a source of information only requires that certain regularities be in place between the state of the source and the state of whatever the information is about. Those regularities may or may not have any bearing on the quality of the subject's reasons.

---

<sup>24</sup>This also puts pressure on related accounts that depart from Greco's by maintaining that success from ability is a matter of degree, such as that of Carter (2014). The problem isn't that Greco ignores gradations of success from ability. Rather, it is the order of explanation to which he is committed.

#### 4. Conclusion

I began this paper by drawing out the implications of Lackey's criticism of Pritchard's (2005) safety-based account of veritic epistemic luck. I then showed that the problem she identified is an instance of a more general problem that plagues other reductive accounts of epistemic luck. The general lesson is that reductive accounts of epistemic luck only deliver the intuitively correct verdicts about cases when they tacitly (and illicitly) rely on a prior understanding of the quality of the subject's reasons. I showed how this allows us to diagnose the shortcomings of a number of well-known accounts in the success-from-ability tradition.

#### References

- Audi R. (1993). *The Structure of Justification*. Cambridge: Cambridge University Press.
- Baker-Hytch M. and Benton M. (2015). 'Defeatism Defeated.' *Philosophical Perspectives* 29(1), 40–66.
- Brewer B. (1999). *Perception and Reason*. Oxford: Oxford University Press.
- Carter J.A. (2014). 'Robust Virtue Epistemology as Anti-Luck Epistemology: A New Solution.' *Pacific Philosophical Quarterly* 97, 140–55.
- de Almeida C. and Fett J.R. (2016). 'Defeasibility and Gettierization: A Reminder.' *Australasian Journal of Philosophy* 94(1), 152–69.
- Engel M. (1992). 'Is Epistemic Luck Compatible with Knowledge?' *Southern Journal of Philosophy* 30(2), 59–75.
- Feldman R. (2003). *Epistemology*. Upper Saddle River, NJ: Pearson.
- Foley R. (2012). *When Is True Belief Knowledge?* Princeton, NJ: Princeton University Press.
- Gettier E. (1963). 'Is Justified True Belief Knowledge?' *Analysis* 23, 121–3.
- Ginsborg H. (2006). 'Reasons for Belief.' *Philosophy and Phenomenological Research* 72(2), 286–318.
- Greco J. (2003). 'Knowledge as Credit for True Belief.' In M. DePaul and L. Zagzebski (eds), *Intellectual Virtue: Perspectives from Ethics and Epistemology*, pp. 111–34. Oxford: Clarendon Press.
- Greco J. (2010). *Achieving Knowledge: A Virtue Theoretic Account of Epistemic Normativity*. Cambridge: Cambridge University Press.
- Greco J. (2012). 'A (Different) Virtue Epistemology.' *Philosophy and Phenomenological Research* 85(1), 1–26.
- Harman G. (1973). *Thought*. Princeton, NJ: Princeton University Press.
- Hawthorne J. (2004). *Knowledge and Lotteries*. Oxford: Oxford University Press.
- Kelp C. (2013). 'Knowledge: The Safe-Apt View.' *Australasian Journal of Philosophy* 91(2), 265–78.
- Klein P. (1980). 'Misleading Evidence and the Restoration of Justification.' *Philosophical Studies* 37, 81–9.
- Klein P. (2008). 'Useful False Beliefs.' In Q. Smith (ed.), *Epistemology: New Essays*, pp. 25–62. Oxford: Oxford University Press.
- Klein P. (2017). 'The Nature of Knowledge.' In R. Borges, C. de Almeida and P. Klein (eds), *Explaining Knowledge: New Essays on the Gettier Problem*, pp. 35–56. Oxford: Oxford University Press.
- Lackey J. (2006). 'Critical Study: Pritchard's *Epistemic Luck*.' *Philosophical Quarterly* 56(3), 284–9.
- Lasonen-Aarnio M. (2010a). 'Is There a Viable Account of Well-Founded Belief?' *Erkenntnis* 72(2), 205–31.
- Lasonen-Aarnio M. (2010b). 'Unreasonable Knowledge.' *Philosophical Perspectives* 24(1), 1–21.
- Lasonen-Aarnio M. (2014). 'Higher-order Evidence and the Limits of Defeat.' *Philosophy and Phenomenological Research* 88(2), 314–45.
- Lehrer K. (1965). 'Knowledge, Truth and Evidence.' *Analysis* 25, 168–75.
- Lehrer K. (2017). 'Defeasible Reasoning and Representation: The Lesson of Gettier.' In R. Borges, C. de Almeida and P. Klein (eds), *Explaining Knowledge: New Essays on the Gettier Problem*, pp. 169–79. Oxford: Oxford University Press.
- Lehrer K. and Paxson T. (1968). 'Knowledge: Undefeated Justified True Belief.' *Journal of Philosophy* 66, 225–37.
- Lewis D. (1973). *Counterfactuals*. Chichester: John Wiley & Sons.
- McDowell J. (1994). *Mind and World*. Cambridge, MA: Harvard University Press.
- Miracchi L. (2015). 'Competence to Know.' *Philosophical Studies* 172, 29–56.
- Moser P.K. (1989). *Knowledge and Evidence*. Cambridge: Cambridge University Press.

- Nozick R.** (1981). *Philosophical Explanations*. Cambridge, MA: Harvard University Press.
- Paulson S.** (2023). 'First-Class and Coach-Class Knowledge.' *Episteme*. doi: 10.1017/epi.2023.5.
- Pollock J.** (1986). *Contemporary Theories of Knowledge*. 1st Edition. Totowa, NJ: Rowman & Littlefield.
- Pritchard D.** (2005). *Epistemic Luck*. Oxford: Oxford University Press.
- Pritchard D.** (2007). 'Anti-Luck Epistemology.' *Synthese* **158**(3), 277–97.
- Shope R.** (1983). *The Analysis of Knowing: A Decade of Research*. Princeton, NJ: Princeton University Press.
- Sosa E.** (1991). *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge: Cambridge University Press.
- Sosa E.** (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge*. Volume 1. Oxford: Clarendon Press.
- Turri J.** (2011). 'Manifest Failure: The Gettier Problem Solved.' *Philosophers' Imprint* **11**(8), 1–11.
- Turri J.** (2012). 'In Gettier's Wake.' In S. Hetherington (ed.), *Epistemology: The Key Thinkers*, pp. 214–29. London: Continuum.
- Turri J.** (2016). 'Knowledge as Achievement, More or Less.' In M. Vargas (ed.), *Performance Epistemology*, pp. 124–34. Oxford: Oxford University Press.
- Warfield T.** (2005). 'Knowledge from Falsehood.' *Philosophical Perspectives* **19**, 405–16.

**Spencer Paulson** is a PhD candidate in the Philosophy department at Northwestern University. He primarily works on Epistemology and the Philosophy of Mind.