UC Irvine
UC Irvine Previously Published Works

Title
There Cannot Be Lucky knowledge

Permalink
https://escholarship.org/uc/item/6p94x7d1

ISBN
9780470672099

Author
Pritchard, Duncan

Publication Date
2013-09-10

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

For *Contemporary Debates in Epistemology* (2nd Ed.), (eds.) M. Steup & J. Turri, (Oxford: Blackwell).

THERE CANNOT BE LUCKY KNOWLEDGE

DUNCAN PRITCHARD
University of Edinburgh

1. KNOWLEDGE AND LUCK

It is commonplace in epistemology to hear people say that there cannot be lucky knowledge (call this the *anti-luck intuition*). This claim requires some unpacking, however, since epistemologists have something quite specific in mind when they put forward this thesis. In particular, they are not denying that it might well be a matter of luck that an agent has knowledge. Imagine an agent who a few moments ago narrowly avoided being killed by a sniper (because, say, she happened to glance downwards at a very fortuitous moment). If so, then she is lucky to be alive and thus lucky to be in a position to know anything. Accordingly, if that fortuitous downward glance happened to reveal to her that she is wearing non-matching socks, and she comes to know this proposition on this basis, then it will be a matter of luck that she has this knowledge.

It can be a matter of luck that one has knowledge in other ways too. For example, it can be a matter of luck that one is in possession of the evidence in virtue of which one has knowledge. Imagine an agent who just happens to be standing outside a colleague's door at just the right moment to overhear what her colleague, who *always* speaks the truth, was authoritatively asserting on the phone (i.e., if she had passed the door a moment earlier or a moment later then she wouldn't have heard a thing). Given that she has clearly heard what her colleague was saying—for example, that so-and-so is wearing non-matching socks—and given the fact that what this colleague asserts can be relied upon, then she is in a position to come to know the relevant proposition. Accordingly, it is a matter of luck that she knows that so-and-so is wearing non-matching socks, but it is no less knowledge for that.

Clearly, then, there are some senses in which knowledge can be lucky, and no epistemologist would want to deny this. So what specific claim are epistemologists putting forward when they say that there cannot be lucky knowledge? Well, the particular claim they have in mind is that when one has knowledge then *it cannot be a matter of luck that one's belief is true*.[1]

Notice first that in the two cases of lucky knowledge just given it precisely *isn't* the case that the agent's belief is only luckily true. Given how the first agent formed her belief that she is wearing non-matching socks—i.e., by getting a good look at them at close range—it is not a matter of luck that her belief is true, even though it is a matter of luck that she is in a position to know what she does. The same goes for the second agent. Given how she is forming her belief that so-and-so is wearing non-matching socks—i.e., by clearly overhearing a reliable colleague assert that this is the case—it is not a matter of luck that her belief is true, even though it is a matter of luck that she acquires the evidence that she does which supports this knowledge (i.e., the overheard testimony of her colleague).

Consider now some cases in which it is a matter of luck that the agent's belief is true. First off, a lucky guess.[2] Imagine an agent who, purely on a whim, believes that the next person to come into her office will be wearing non-matching socks, and suppose that this belief turns out to be true. Clearly, she didn't know that the next person to come into her office would be wearing non-matching socks, and the natural explanation of why is that given how she formed her belief (i.e., by guessing) it was just a matter of luck that her belief was true.

Second, consider a standard Gettier-style case. Imagine our agent sees what she takes to be a sniper in the office block across the way, with her rifle trained on her building. Accordingly, she forms the belief that there is a sniper in that office block. Now suppose that this belief is true, but that what our agent is looking at is not a real sniper (one trying to assassinate our non-matching-sock-wearing friend, say), but rather merely a cardboard cut-out of a sniper, left over from the promotional materials of a movie (with the real sniper hidden behind). Does our agent have knowledge of what she believes? Intuitively not, and the natural reason for this is that given how she formed her belief—by looking at the cardboard cut-out of a sniper—it was just a matter of luck that her belief was true. Moreover, notice that she lacks knowledge on account of her lucky true belief even though, in contrast to the lucky guesser we just considered, she has a good reason for believing what she does (after all, she did see something which looked like a sniper in the window of the office building).[3]

Now contrast such cases where lucky true belief undermines knowledge with paradigm cases of knowledge. Imagine, for example, that our agent gets to see that her colleague is wearing non-matching socks at close hand and in optimal viewing conditions. Or imagine that our agent is told

by a reliable informant that her colleague is wearing non-matching socks today. In both cases we have an agent who thereby comes to acquire knowledge. But notice that in both such cases we also have an agent whose true belief is not a matter of luck. If one believes that someone is wearing non-matching socks because one gets to see—at close hand, and in normal circumstances—that this is the case, then it is not a matter of luck that one's belief is true. And if one believes that someone is wearing non-matching socks because one is told that this is the case by a reliable informant, then it is not a matter of luck that one's belief is true.

There is thus a strong *prima facie* case that knowledge excludes luck in just this sense: if one has knowledge, then it cannot be a matter of luck that one's belief is true. Call this the *anti-luck condition* on knowledge. But how, exactly, should we understand the anti-luck condition? The devil, as always in philosophy, is in the detail.

2. UNPACKING THE ANTI-LUCK CONDITION I: SENSITIVITY

One way of understanding the anti-luck condition is in terms of the requirement that knowledge entails *sensitive* true belief. In order for a true belief to be sensitive, it must be such that, had what the agent believed been false, she wouldn't have believed it.[4] Two remarks are in order regarding how we should read this claim.

The first is to note that this is a *modal* claim, in the sense that in applying it we need to consider what is going on in relevant possible worlds. In particular, in assessing whether a true belief is sensitive we need to consider not only what the agent believes in the actual world (where her belief is true), but also what she believes in that possible world (or worlds) which is similar to the actual world in every other respect except that the target proposition (the proposition truly believed in the actual world) is false.[5] Call this the *closest not-p world* (where *p* is the target proposition which is believed by the agent in the actual world).

The second remark modifies the first. For we need to qualify our understanding of the closest not-*p* worlds by considering only those closest not-*p* worlds in which the agent forms her belief about the target proposition in the same way as in the actual world. That is, what we are interested in is not merely whether a certain belief is sensitive, but rather whether that belief is sensitive, *given how it is formed*.[6]

We can illustrate these two points by considering how sensitivity, so understood, can deal with cases in which an agent lacks knowledge due to the epistemic luck in play. Take first the case of a lucky guess that was given above, where the agent believes as a result of a guess that the next person to come into her office will be wearing non-matching socks. Clearly this is an insensitive

belief, since in the closest not-*p* world—i.e., the world in which the next person to enter the room is not wearing non-matching socks, but everything else, including the agent's basis for the belief (i.e., guessing), is the same—our agent will form a false belief.

Now consider the Gettier-style case we gave in which the agent sees what seems to be a sniper in the office block window and on this basis forms the belief, true as it happens, that there is a sniper in the office block that she is looking at. Again, this belief is insensitive, since in the closest not-*p* world—i.e., the world in which there is no sniper in that building, but everything else, including the agent's basis for the belief (i.e., the promotional cardboard cut-out of a sniper that she sees in the window), is the same—our agent will form a false belief.

Clear-cut cases of knowledge also seem to involve a sensitive belief. Imagine, for example, that our agent gets to see that her colleague is wearing non-matching socks at close hand and in optimal viewing conditions. In such a case we have an agent who thereby comes to acquire knowledge, and who is also forming a sensitive belief. For given the way that she is forming her belief, in the closest not-*p* world—i.e., the world in which her colleague is not wearing non-matching socks, but everything else remains the same (including the agent's perceptual basis for her belief)—the agent will no longer believe that her colleague is wearing non-matching socks for the simple reason that she will see that she isn't.

Sensitivity can also be used to shed light on another problem in epistemology, known as the *lottery problem*. Imagine a fair lottery with astronomical odds—a billion-to-one, say—which has been drawn, and imagine two counterpart agents who both own a single lottery ticket. Suppose that our first agent hasn't heard the result yet but believes that she has lost because she has reflected on the very long odds involved in her winning; and suppose that our second agent also believes that she has lost, but does so purely because she has just read the results printed in a reliable newspaper.[7] Here's the thing. Intuitively, the first agent doesn't know that she has lost while the second agent does. But this is puzzling, since the odds in favour of the first agent being right are astronomically in her favour. Indeed, even though reliable newspapers are very careful when it comes to printing lottery numbers (for obvious reasons), nonetheless the probability that these results have been misprinted is surely higher than the astronomical probability that a ticket should win this lottery. So how then can it be that the second agent has knowledge but the first agent doesn't?

The thesis that knowledge requires sensitivity offers us a very attractive way of dealing with the lottery problem, for notice that while the second agent's belief is sensitive, the first agent's belief is not. The closest not-*p* world is where the agent concerned is in possession of a winning lottery ticket. Crucially, though, while this will be a world in which the reliable newspaper prints

the winning result, it will continue to be a world in which the odds in question overwhelming suggest that one has lost. Thus, if one forms one's belief about whether one has lost on the basis of the odds concerned, then one will form a false belief in this world, but if one forms one's belief by consulting a reliable newspaper then one will form a true belief. It is in this sense, claim sensitivity theorists, that the first agent's belief that she has lost the lottery is only luckily true, even though the odds are massively in her favour when compared with the second agent who, by consulting the reliable newspaper, really does know that she has lost.

3. UNPACKING THE ANTI-LUCK CONDITION II: SAFETY

Unfortunately, the claim that knowledge entails sensitivity faces some problems. Rather than rehearse them all I will focus here on the most pressing of these difficulties, which is the problem it has with inductive knowledge.[8] Suppose that our agent has spent the entire day with her colleague, and throughout the day she has been continually reminded of the fact that her colleague is wearing non-matching socks. Now suppose that her colleague leaves the office to get the train home. Does our agent know that her colleague is still wearing non-matching socks? Intuitively, she does, since she has an excellent inductive basis for this (true) belief, which is her observations of her colleague throughout the day and her knowledge that she is unlikely to be changing her socks on the train journey home. But here is the problem: her belief is insensitive. For notice that in the closest not-*p* world—i.e., the far-fetched world in which her colleague has been forced to take off her socks, perhaps because she is the victim of an armed sock thief—our agent will continue to believe that her colleague is wearing non-matching socks regardless (after all, her basis for this belief—*viz.*, that she has seen her wearing these socks throughout the day and she has no reason for thinking that she has taken them off—remains constant).[9]

The reason why sensitivity is unable to account for the knowledge present in such cases is that it focuses on the closest not-*p* world and ignores the wide range of close worlds where the agent believes the target proposition and believes truly (after all, in most near-by possible worlds the agent believes that her colleague is wearing non-matching socks and she is wearing non-matching socks). This is one of the main reasons why many in the literature opt for the thesis that knowledge demands not sensitive belief but rather *safe* belief.

The basic idea behind safety is that one has a true belief which could not have easily been false.[10] This is usually cashed-out as the claim that one has a true belief that *p* such that, in close possible worlds, if one continues to form a belief on the same basis as in the actual world, then one's belief continues to be true.

First off, notice that this principle can handle inductive knowledge. For while there is undoubtedly a far-fetched possible world where our agent's colleague has removed her socks while on the train, in all close possible worlds she will retain her socks.[11] Accordingly, our agent will form a belief that her colleague is wearing socks on the train which will be true in all relevant possible worlds, and hence her belief will be safe.

Safety can also deal with Gettier-style cases, for these are characteristically cases in which the agent forms (on the same basis as in the actual world) a false belief in the target proposition in a close possible world. Consider the case of the promotional cardboard cut-out being mistaken for a real sniper that was described above. Although the belief so formed (that there is a sniper in the relevant office block) is true in the actual world, there will be a near-by possible world where the agent will form the very same belief on the very same basis (i.e., by looking at the cardboard cut-out of the sniper), and yet forms a false belief (i.e., the near-by possible world where the cardboard cut-out remains, but the sniper has been delayed and is yet to enter the building).

Safety can also handle the lottery problem, though we need to be a little more precise in how we understand the principle in order to see this. The formulation above talks simply of the agent continuing to form a true belief on the same basis as in the actual world across near-by possible worlds. But this naturally prompts the question of how extensive this range of near-by possible worlds should be, and to what extent, if any, safety is consistent with there being some false beliefs formed within this range of possible worlds.

What is key to answering this question is that safety is capturing an intuition about our tolerance of the risk of error. In the very closest nearby possible worlds we are extremely intolerant when it comes to such epistemic risk, and so would not want to be forming any false beliefs on the target basis. In far-off possible worlds, however, we are extremely tolerant about such epistemic risk, on account of their modal remoteness. In between we have a descending scale of epistemic intolerance, from extreme intolerance to epistemic risk to extreme tolerance. With this point in mind, we need to think of safety as completely excluding false belief in the very closest possible worlds, but becoming increasingly tolerant to such falsity as one moves further away from the actual world.[12]

Safety so construed can handle the lottery problem. Recall that the two agents concerned form their true beliefs that their tickets have lost by, respectively, reflecting on the odds involved and reading the results in a reliable newspaper. What is crucial here is that the remoteness of error is very different in the two cases. All it takes for the first agent to form a false belief is for a few coloured balls to fall in a slightly different configuration. In contrast, what is required for the second agent to form a false belief is a range of mishaps at the newspaper office (e.g., the person inputting the results, despite taking lots of care in doing so, somehow makes a mistake, a mistake

that is not spotted when the various copyeditors, hired and retained for their conscientiousness in such matters, somehow collectively fail to spot the error). This is why the first agent's belief is unsafe, and hence not knowledge, while the second agent's belief is safe and hence is in the market for knowledge.[13]

One could say an awful lot more about how safety handles particular cases, but given that space is tight I hope that I have said enough to have motivated the thought that safety has advantages over sensitivity, and hence that it is a plausible candidate when it comes to unpacking the intuition that knowledge cannot be lucky. Put simply, the claim that knowledge cannot be lucky, when properly understood, is tantamount to the claim that there cannot be knowledge that *p* which involves an unsafe belief that *p*. Before moving onto some objections to this specific thesis, I want to make two further points.

First, it is often thought that the safety principle has a problem accounting for belief in necessary propositions.[14] One can see how the objection would run. Such propositions are true in all near-by possible worlds, and hence all one needs to do is happen to form a true belief in a necessary proposition in the actual world and—hey presto!—one has a belief which is necessarily safe (since there can necessarily be no near-by possible worlds where one continues to form this belief and believes falsely). Crucially, however, one does not evaluate the safety of a belief by focussing only on near-by possible worlds where the agent continues (on the same basis as in the actual world) to form a belief *in the very same proposition* as in the actual world. Rather, what one is interested in the truth-value of the belief that is formed in near-by possible worlds on the same basis as in the actual world, even when the resulting belief is not of the same proposition.

In order to see this, imagine that someone forms a true belief in a mathematical proposition—that 2+2=4, say—by flipping a coin. Since there is no possible world where the proposition believed is false, there is thus trivially no near-by possible world in which the agent believes *this specific proposition* and believes falsely. But that *doesn't* mean that the belief is thereby safe, and the reason for this is that there are lots of near-by possible worlds where the agent's actual way of forming her belief—i.e., by flipping a coin in order to determine mathematical truths—leads to false belief, such as the possible world where the coin toss prompts her to believe that 2+2=5. The key point here is that in assessing whether a belief that *p* is safe, we are interested in whether the agent forms a belief in the same way in near-by possible worlds and believes falsely, but this is different from being interested in whether the agent forms a belief *that p* in near-by possible worlds and believes falsely.

The second point I want to make about safety is to emphasise that the thesis in play here is specifically the claim that this is a *necessary* condition on knowledge. There is a good reason for this,

which is that aside from the anti-luck intuition we also have a strong intuition that knowledge reflects ability (call this the *ability intuition*), in the sense that when an agent has knowledge then it is to some significant degree the product of the agent's exercise of her cognitive abilities. One can draw this point out by noticing that we can formulate cases where the agent forms a true belief which is clearly not-lucky (i.e., which is clearly safe) and yet which isn't a case of knowledge because the cognitive success in question in no way reflects the agent's cognitive abilities.

Imagine, for example, that an agent was forming beliefs about what colour socks the next person she saw would be wearing purely on the basis of flipping a coin. But imagine now that these beliefs are guaranteed to be true since there is, unbeknownst to our agent, a helpful demon out there whose mission it is to ensure that beliefs formed in just this way are true. Consequently, he makes it his business to see that the socks worn by the next person seen by the agent also correspond to what the agent guessed they would be.

Clearly, one cannot gain knowledge by guesswork in this way. Note, however, that the problem here isn't that the agent's beliefs are only luckily true, since given the interference of the helpful demon they are in fact guaranteed to be true, and hence can't help but be safe. Instead, the problem is that the cognitive success on display in no way reflects the agent's cognitive abilities, but rather is simply the result of the interference of the helpful demon. Put simply, although there is a perfect match-up between belief and fact across the relevant possible worlds (such that the agent always believes truly), there is the wrong direction-of-fit in play for knowledge, in that the facts are changing to fit with what the agent believes rather than the agent's beliefs being responsive to the facts. Given that the problem in play here does not concern knowledge-undermining epistemic luck, and given that the role of safety is to exclude such luck, it should be clear that it is not a failing of safety that it is unable deal with such cases.[15] Rather, what such cases remind us is that there is more to knowledge than safe true belief. In particular, at the very least what is also required is some sort of ability condition on knowledge.[16] As we will see below, this point is important for our purposes since we need to be sure that a putative counterexample to the necessity of safety for knowledge is not in fact trading on something other than the anti-luck intuition.

4. THREE PUTATIVE COUNTEREXAMPLES

A range of putative counterexamples have been put forward to the idea that knowledge entails safety, and we will close by considering a representative sample to explain why they fail to achieve their intended aim.[17] The first is due to Ram Neta and Guy Rohrbaugh (2004):

> WATER: "I am drinking a glass of water which I have just poured from the bottle. Standing next to me is a happy person who has just won the lottery. Had this person lost the lottery, she would have maliciously polluted my water with a tasteless, odorless, colorless toxin. But since she won the lottery, she does no such thing. Nonetheless, she *almost* lost the lottery. Now, I drink the pure, unadulterated water and judge, truly and knowingly, that I am drinking pure, unadulterated water. But the toxin would not have flavored the water, and so had the toxin gone in, I would still have believed falsely that I was drinking pure, unadulterated water. […] Despite the falsity of my belief in the nearby possibility, it seems that, in the actual case, I know that I am drinking pure, unadulterated water." (Neta & Rohrbaugh 2004, 399-400)[18]

My initial reaction to such a case is to hold that it is simply not a case of knowledge. Is it really intuitive that the agent in WATER could gain knowledge that what she is drinking is water even despite the clear epistemic unfriendliness of her environment? After all, she could so very easily have been drinking the toxin rather than the water, and *ex hypothesi* she wouldn't have been able to tell the difference. Indeed, were our agent to discover just how lucky it was that she formed a true belief in this case, then surely she wouldn't ascribe knowledge to herself. But of course, if this isn't a case of knowledge then it can't be a counterexample to the necessity of safety for knowledge.

There is, however, an obvious dialectical drawback to dismissing such a case out-of-hand. After all, it is clear that others have found this example compelling, and so we are in danger of merely trading opposing intuitions here. Fortunately, I think we can diagnose why someone might hold that the protagonist in WATER has knowledge, even though (so says I anyway) she doesn't.

Recall that we noted earlier that when epistemologists say that they want to exclude lucky knowledge they are not meaning to suggest that it can't be a matter of luck that one is in a position to acquire knowledge. As we noted above, it could be merely a matter of luck that one happens to overhear a conversation—a moment earlier or a moment later and one would have missed it—but given that one does get to properly overhear it, one can nonetheless acquire knowledge of what was said as a result (call this example 'OVERHEAR'). Safety captures this point by specifically focussing on the way in which the belief was formed. It may be a matter of luck that one overhears the conversation, but so long as one's hearing is working fine and the voices are loud enough (etc.,) then this will be a way of forming one's beliefs which is safe, and hence knowledge-conducive.

If one thought that WATER was relevantly analogous to OVERHEAR then one might be inclined to ascribe knowledge in this case. There is certainly a surface similarity between the two cases, in that one might think that just as it is a matter of luck that one happens to overhear the conversation in OVERHEAR, so it is a matter of luck that one happens to drink the uncorrupted water in WATER. Crucially, however, there is a key difference. For although it is a matter of luck that the agent in OVERHEAR overhears what she does, nonetheless she is in a great position to epistemically exploit this opportunity (since her hearing is working fine, and so on). But the same

is not true of the agent in WATER. After all, what looks and tastes like water in her environment need not be water. This is why the agent's belief in WATER is unsafe, but the agent's belief in OVERHEAR is safe. I suggest that once we understand the difference between these two types of case then one can see why some epistemologists might be inclined to ascribe knowledge in a case like WATER, even though knowledge isn't in fact possessed by the agent concerned.[19]

A more interesting case is offered by Christoph Kelp (2009), which we can express as follows:

> DEMON: A demon wants our hero—let's call him 'Chris'—to form the belief that the time is 8.22am when he comes down the stairs first thing in the morning (the demon doesn't care whether the belief is true). Since he is a demon, with lots of special powers, he is able to ensure that Chris believes this proposition (e.g., by manipulating the clock). Now suppose that Chris happens to come downstairs that morning at exactly 8.22am, and so forms the belief that the time is 8.22am by looking at the accurate clock at the bottom of the stairs. Accordingly, the demon achieves what he wants without having to do anything.[20]

Kelp's claim is that insofar as the demon doesn't intervene then, given how Chris formed his belief, he gains knowledge. But since the demon will ensure that Chris continues to believe that the time is 8.22am in all near-by possible worlds, even when this is false, Kelp also claims that this belief is nonetheless unsafe.

While I think this example is ingenious, I don't think it works. In particular, I don't at all share the intuition that the agent in DEMON has knowledge. After all, given how Chris formed his belief it is pure luck that this belief happens to be true—had he come downstairs a minute earlier or a minute later then he would have formed a false belief. Indeed, Chris is effectively finding out the time by looking at what is (for him anyway) a stopped clock, since whatever time he comes downstairs the clock will say '8.22am'. But one cannot gain knowledge about the time by consulting a stopped clock, even when one happens to form a true belief!

That said, I do think that there is something epistemically laudable about the agent's true belief, in that (given that the demon didn't in fact intervene) it is a cognitive success that it attributable to his cognitive ability and thus his cognitive agency. In this sense, it constitutes a cognitive achievement on the part of the subject, even though it isn't knowledge. Often knowledge and cognitive achievement go hand-in-hand (which I think may explain Kelp's inclination to ascribe knowledge to the agent in DEMON), but what I think cases like this illustrate quite neatly is that they can come apart. In particular, they will come apart in cases where the luckiness of the cognitive success is entirely due to some feature of the modal environment which is absent in the actual world.[21]

The third and final putative counterexample is due to Ian Church (2010):

> VIRUS: "Smith is ill and exhibits a unique set of symptoms, S. Given these symptoms, Dr Jones forms the belief that 'Smith has Virus X', which she deduces from the true proposition that 'Virus X is the only known virus to exhibit S.' What is more, Dr Jones does a blood test which verifies that Smith's body contains antibodies for Virus X, further justifying Jones' belief. Based on the evidence, it is extremely feasible that Smith has Virus X. As it happens, however, Smith's symptoms are in fact due to an unknown virus, Virus Y, which exhibits identical symptoms to Virus X; Smith only exhibits antibodies for Virus X due to an idiosyncratic feature of Smith's particular biochemistry which causes his immune system to maintain high levels of antibodies long past a given infection. Nevertheless, Dr Jones' belief turns out to be true divorced from Smith's symptoms or his blood work, because Smith was infected with Virus X just before meeting with Dr Jones—the infection being so recent that blood work cannot detect it an it is causing no symptoms." (Church 2010, 9)[22]

Notice that this case is different from the other two, in that rather than being a putative case of knowledge where the belief in question is unsafe, it is instead an example of a safe belief which doesn't amount to knowledge. The reason why Church thinks that VIRUS is nonetheless a counterexample to the necessity of safety for knowledge is that he holds that the reason knowledge is lacking in this case is due to the epistemic luck involved. But since it is the job of safety to exclude knowledge-undermining epistemic luck, it follows that this case presents a problem for proponents of safety.

Unfortunately, the case doesn't work. Even if we further add—as Church (2010, 10) himself suggests—that the circumstances of the case are such that Smith is virtually guaranteed to catch virus X just before seeing Dr. Jones, this is only at best a counterexample to a crude form of safety which focuses only on the subject's continued belief *that p* across the relevant possible worlds. Remember the point made above about how a belief in a necessary proposition can be unsafe, even though there is obviously no near-by possible world where the necessary proposition in question is believed falsely. The same applies here. Even if there can be no near-by possible world in which Dr. Jones believes that Smith has virus X and believes falsely—because this proposition is true across all nearby possible worlds—it doesn't follow that the belief is safe, since we also need to consider the other beliefs that Dr. Jones forms in near-by possible worlds on the same basis as in the actual world. Once we remember this, it becomes clear that Dr. Jones will form false beliefs in near-by possible worlds on the same basis as in the actual world. Consider, for example, the close possible world where all that is different is that Smith doesn't happen to maintain a high level of antibodies in his blood and doesn't clearly exhibit the symptoms for virus X. In such a world Dr. Jones would likely form the false belief that Smith *didn't* have virus X, or the false belief that Smith had another virus which he didn't in fact have.

Of course, we can always set a case up in which the agent is guaranteed not just to form a true belief in the actual world, but also to form a true belief on the same basis across all near-by possible worlds. Perhaps the case offered by Church could be reconstructed to do this. But even so, this need be nothing for the defender of the necessity of safety to knowledge to worry about.

For as noted above, such an epistemologist does not claim that safe true belief is sufficient for knowledge, and so it is perfectly compatible with the story they tell that there may be cases of safe true belief which aren't thereby cases of knowledge. Moreover, given that the agent is now guaranteed to have a true belief on the relevant basis across all possible worlds, it seems that what is lacking in such a case is nothing to do with the anti-luck intuition. After all, if one has a true belief in these circumstances then it is surely not a matter of *luck* that one's belief is true.

Indeed, the point made above about how we shouldn't expect the safety condition to fully capture our intuitions about the role of cognitive ability in knowledge acquisition are very salient to such a case. For it seems that with the case so construed what has gone awry, epistemically, is not that the agent is forming beliefs such that they could so very easily have been false, but rather that her cognitive success is not appropriately related to her cognitive ability. Put another way, it seems that what is epistemically problematic about such beliefs is not that they fail to satisfy the anti-luck intuition about knowledge, but rather that they fail to satisfy the ability intuition.[23]

REFERENCES

Becker, K. (2007). *Epistemology Modalized*, London: Routledge.
Black, T., *&* Murphy, P. (2007). 'In Defense of Sensitivity', *Synthese* 154, 53-71.
Black, T. (2008). 'Defending a Sensitive Neo-Moorean Invariantism', *New Waves in Epistemology*, (eds.) V. F. Hendricks *&* D. H. Pritchard, 8-27, Basingstoke: Palgrave Macmillan.
Chisholm, R. (1977). *Theory of Knowledge*, (2nd ed.), Englewood Cliffs, NJ: Prentice-Hall.
Church, I. (2010). 'Getting 'Lucky' with Gettier', *European Journal of Philosophy*.
Comesaña, J. (2005). 'Unsafe Knowledge', *Synthese* 146, 393-402.
Dretske, F. (1970). 'Epistemic Operators', *Journal of Philosophy* 67, 1007-23.
—— (1971). 'Conclusive Reasons', *Australasian Journal of Philosophy* 49, 1-22.
Frankfurt, H. (1969). 'Alternate Possibilities and Moral Responsibility', *Journal of Philosophy* 46, 829-39.
Greco, J. (2007). 'Worries About Pritchard's Safety', *Synthese* 158, 299-302.
Hetherington, S. (1998). 'Actually Knowing', *Philosophical Quarterly* 48, 453-69.
—— (2002). *Good Knowledge, Bad Knowledge: On Two Dogmas of Epistemology,* Oxford: Oxford University Press.
Hiller, A., *&* Neta, R. (2007). 'Safety and Epistemic Luck', *Synthese* 158, 303-14.
Kelp, C. (2009). 'Knowledge and Safety', *Journal of Philosophical Research* 34, 21-31.
Luper, S. (1984). 'The Epistemic Predicament', *Australasian Journal of Philosophy* 62, 26-50.
—— (2003). 'Indiscernability Skepticism', *The Skeptics: Contemporary Essays*, (ed.) S. Luper, 183-202, Aldershot: Ashgate.
Madison, B. J. C. (2010). 'Combating Anti-Anti-Luck Epistemology', *Australasian Journal of Philosophy*.
McEvoy, M. (2009). 'The Lottery Puzzle and Pritchard's Safety Analysis of Knowledge', *Journal of Philosophical Research* 34, 10-20.
Nozick, R. (1981). *Philosophical Explanations,* Oxford: Oxford University Press.
Neta, R., *&* Rohrbaugh, G. (2004). 'Luminosity and the Safety of Knowledge', *Pacific Philosophical Quarterly* 85, 396-406.
Pritchard, D. H. (2002). 'Resurrecting the Moorean Response to the Sceptic', *International Journal of Philosophical Studies* 10, 283-307.
—— (2005). *Epistemic Luck*, Oxford: Oxford University Press.
—— (2007). 'Anti-Luck Epistemology', *Synthese* 158, 277-97.
—— (2008). 'Sensitivity, Safety, and Anti-Luck Epistemology', *The Oxford Handbook of Scepticism*, (ed.) J. Greco, 437-55, Oxford: Oxford University Press.
—— (2009). 'Safety-Based Epistemology: Whither Now?', *Journal of Philosophical Research* 34, 33-45.
—— (*Forthcoming*). 'Anti-Luck Virtue Epistemology', *Journal of Philosophy*
Pritchard, D. H., Millar, A., *&* Haddock, A. (2010). *The Nature and Value of Knowledge: Three Investigations,* Oxford: Oxford University Press.
Roush, S. (2005). *Tracking Truth: Knowledge, Evidence and Science*, Oxford: Oxford University Press.
Sainsbury, R. M. (1997). 'Easy Possibilities', *Philosophy and Phenomenological Research* 57, 907-19.
Sosa, E. (1999). 'How to Defeat Opposition to Moore', *Philosophical Perspectives* 13, 141-54.
Unger, P. (1968). 'An Analysis of Factual Knowledge', *Journal of Philosophy* 65, 157-170.
Williamson, T. (2000). *Knowledge and its Limits*, Oxford: Oxford University Press.
Zagzebski, L. (1994). 'The Inescapability of Gettier Problems', *Philosophical Quarterly* 44, 65-73.

NOTES

[1] See Unger (1968) for what is almost certainly the first clear statement of the idea that there is a specific sense in which we want our knowledge to be non-lucky (or 'non-accidental', as he terms it). See Pritchard (2005) for a sustained development of this idea.

[2] I bracket here the issue of whether a guess can ever be properly considered a belief.

[3] This is a variation on a famous example given by Chisholm (1977, 105). Although most epistemologists grant that agents lack knowledge in Gettier-style cases, there are some detractors. See, especially, Hetherington (1998; 2002, ch. 1). For a very useful recent discussion of Hetherington's view, see Madison (2010).

[4] For the key texts in this regard, see Dretske (1970; 1971) and Nozick (1981). For some recent texts which sympathetically explore the sensitivity principle, see Roush (2005), Becker (2007), Black & Murphy (2007) and Black (2008).

[5] Of course, a change in truth-value of any proposition is bound to have further implications, and hence the difference between the actual world and the target possible world is unlikely to be only that the proposition in question is false. But we can set this complication to one side for our purposes.

[6] For the classic discussion of this constraint on sensitivity, see Nozick (1981, 179ff.). For further discussion, see Pritchard (2008, §3).

[7] If you like, add the stipulation that this second agent doesn't even know how long the odds for this lottery are.

[8] For a more thorough discussion of some of the problems faced by the sensitivity principle, see Pritchard (2008).

[9] This problem for the sensitivity principle is due to Sosa (1999).

[10] Versions of safety-type principles have been offered by a number of authors, including Luper (1984; cf. Luper 2006), Sainsbury (1997), Sosa (1999), Williamson (2000), and Pritchard (2002; 2005; 2007).

[11] And note that it will do no good to say that the actual world is in fact such that in a great deal of near-by possible worlds the colleague would have removed her socks (perhaps because she often likes to do this while on the train, unbeknownst to our agent), since if that's the case then there is no intuition that our agent has inductive knowledge which requires rescue.

[12] Although it would take me too far afield to go into this issue here, there is in fact a strong motivation not just for safety but also for this specific rendering of safety in an epistemological programme that I call *anti-luck epistemology*. Very roughly, this programme involves identifying what luck is, and identifying the specific sense in which knowledge excludes luck, in order to gain an insight into the nature of knowledge. For more details, see Pritchard (2005; 2007; 2008; 2009).

[13] For scepticism about the prospects of a safety-based account of knowledge dealing with the lottery problem, see Greco (2007) and McEvoy (2009). For responses, see (respectively) Pritchard (2007; 2009).

[14] Or, for that matter, any proposition which is true across all near-by possible worlds, even if not necessarily true.

[15] In any case, no formulation of safety could exclude such cases. The reason for this is that a modal principle like safety cannot capture the 'direction-of-fit' between belief and fact that is key to the ability intuition on account of how it simply specifies a match between belief and fact across a range of worlds. For more on this point, see Pritchard, Millar & Haddock (2010, ch. 3) and Pritchard (*forthcoming*).

[16] I have argued elsewhere for a view I call *anti-luck virtue epistemology* according to which knowledge is essentially safe true belief plus a further epistemic condition (an 'ability condition) which handles the ability intuition. See, for example, Pritchard, Millar & Haddock (2010, ch. 3) and Pritchard (*forthcoming*).

[17] Note that these cases are often put forward against particular formulations of the safety principle, but our interest will be whether they work against the specific formulation we offer here.

[18] See also the very similar counterexample to the necessity of safety for knowledge offered in Hiller & Neta (2007, 310-11).

[19] The same diagnosis will apply to the structurally similar, though more complex, 'Halloween party' case offered by Comesaña (2005, 397), which is also meant to be a counterexample to the necessity of safety for knowledge. That said, as Kelp (2009) points out, it isn't at all obvious that Comesaña's example even involves an unsafe belief in the first place.

[20] This is essentially a type of 'Frankfurt-style' example in that what is significant is that the demon *would* have intervened rather than that he did intervene. See Frankfurt (1969).

[21] I develop this point about how knowledge and cognitive achievement come apart in a number of places. See, for example, Pritchard (2009; *forthcoming*) and Pritchard, Millar & Haddock (2010, ch. 2).

[22] This example is a variation on a case originally proposed by Zagzebski (1994, 71), albeit to illustrate a different point.

[23] This paper was written while I was in receipt of a Phillip Leverhulme Prize. My recent thinking about these issues has been informed by conversations with (amongst others) Kelly Becker, Tim Black, Adam Carter, Ian Church, Julien Dutant, Georgi Gardiner, John Greco, Allan Hazlett, Avram Hiller, Jesper Kallestrup, Chris Kelp, Ram Neta and Ernie Sosa. Special thanks go to Stephen Hetherington, Brent Madison and John Turri who offered comments on an earlier version.