A Counterexample to the Robust and Anti-luck Virtue Epistemologies

Leo K. C. Cheung

Department of Philosophy

The Chinese University of Hong Kong

This paper begins with an introduction to four different prominent versions of process reliabilism about knowledge, namely, the simple K-reliabilism, John Greco’s robust virtue epistemology, Duncan Pritchard’s anti-luck virtue epistemology, and the Goldmanian K-reliabilism. Its main objective is to offer a counterexample, namely, the fewer barn façades case, to the anti-luck virtue epistemology, and thus to the robust virtue epistemology and the simple K-reliabilism as well. Since there are already other counterexamples, like Keith Lehrer’s Mr. Truetemp case, to the Goldmanian K-reliabilism, the discovery of a counterexample to the other three prominent versions of K-reliabilism is a strong attack on K-reliabilism in general.

1. Four different versions of process reliabilism about knowledge

Let me begin by introducing four different versions of process reliabilism about knowledge; hereafter “K-reliabilism”. The simplest version may be called “The simple K-reliabilism”, which can be formulated neatly as follows:
S knows $p$ if and only if the true belief $p$ is acquired via a reliable belief-forming process.

The right hand side of the biconditional may be called “The reliability condition”. We may follow Kelly Becker (2009) to hold that a belief-forming process is “reliable”, or “truth-conductive”, if it is a process that “produces either mostly true beliefs or a high ratio of true to false beliefs.” This characterization of the notion of reliability suffices to serve the purposes of this paper.

It is well-known that counterexamples to the simple K-reliabilism can be found easily. One of them is Lehrer’s Mr. Truetemp case:

[Mr. Truetemp] “Suppose a person, whom we shall name Mr. Truetemp, undergoes brain surgery by an experimental surgeon who invents a small device which is both a very accurate thermometer and a computational device [a tempucomp] capable of generating thoughts… Assume that the tempucomp is very reliable, and so his thoughts are correct temperature thoughts. All told, this is a reliable belief-forming process. Now imagine, finally, that he has no idea that the tempucomp has been inserted in his brain, is only slightly puzzled about why he thinks so obsessively about the temperature, but never checks a thermometer to determine whether these thoughts about the temperature are correct. He accepts them unreflectively, another effect of the tempucomp. Thus, he thinks and accepts that the temperature is 104 degrees.” (Lehrer 1990: 163-4)
The popular intuition here is that Mr. Truetemp does not know that the temperature is 104 degrees. However, Mr. Truetemp acquires the relevant true belief via a reliable belief-forming process. This is a counterexample to the sufficiency of the reliability condition, and thus to the simple K-reliabilism.

Greco’s (2006: 30) response to the Mr. Truetemp case is basically that because the cognitive success of Mr. Truetemp, that is, his acquiring the true belief, is not through the exercise of a cognitive ability, he does not know that the temperature is 104 degrees, even though the true belief is reliably produced. Greco’s point seems to be that not all reliable belief-forming processes support knowledge attribution. Only those reliable belief-forming processes produced by the exercise of the agent’s cognitive abilities, which make up the agent’s cognitive character, support knowledge attribution. This is a version of agent reliabilism, and is also a version of, what Pritchard (2009a: 22, 2010a) calls, robust virtue epistemology. Following Greco’s earlier writings (2003: 116, 2009: 17-8, 2010:71-2) and the more recent writing (2012), his version of the robust virtue epistemology can be formulated as follows:

(2) \( S \) knows \( p \) if and only if \( S \)’s acquiring the true belief \( p \) is primarily\(^1\) attributable\(^2\) to cognitive ability.

The right hand side of the biconditional may be called “The ability condition”, which is also the reliability condition of it as a version of agent reliabilism. Greco’s robust virtue epistemology can be formulated as follows:

\( S \) knows \( p \) if and only if \( S \)’s acquiring the true belief \( p \) is primarily\(^1\) attributable\(^2\) to cognitive ability.

\(^1\) Pritchard (2010b) emphasizes that the creditability, or, rather, attributability (see the next footnote), here must be primary. For, otherwise, the relevant robust virtue epistemology cannot explain the intuition of ignorance in Gettier cases. I follow Pritchard to put “primarily” here.

\(^2\) Greco uses the term “creditable” in his earlier writings. But in his more recent writings like “A (different) virtue epistemology” (2012: 1, footnote 1), he deliberately refrains from talking about “creditability”. Instead, he now uses the rather similar notion of attribution, or that of attributability.
epistemology accounts for the lack of knowledge in the Mr. Truetemp case in this way: The relevant true belief is not acquired via the exercise of any of Mr. Truetemp’s cognitive abilities, if any, and thus the relevant belief-forming process does not support knowledge attribution.

Pritchard (2006) criticizes Greco’s robust virtue epistemology and its response to the Mr. Truetemp case at two points. One of them concerns us here. Pritchard claims that Greco’s “appealing to the agent’s cognitive character won’t suffice to provide one with a fully-fledged theory of knowledge” (2006: 38). In (2006) and some of his more recent writings, including (2008a, 2008b, 2012), Pritchard argues that Greco’s robust virtue epistemology fails to account for those Gettier-style cases involving luck, especially environmental luck. This amounts to be attacking the sufficiency of the ability condition in (2). One of the alleged counterexamples Pritchard is fond of using is the famous Ginet-Goldman “barn façades” case. In what follows, this version of the barn façades case will be considered:

[Barn façades] “Henry is driving in the countryside and sees a barn ahead in clear view. On this basis he believes that the object he sees is a barn. Unknown to Henry, however, the area is dotted with barn façades that are indistinguishable from real barns from the road. However, Henry happens to be looking at the one real barn in the area.” (Greco 2009: 21)
Pritchard shares the popular intuition that Henry in the barn façades case does not know that the object he sees is a barn. (This intuition is adopted in this paper.) He thinks the lack of knowledge in this case is because of the fact that the cognitive success is just a matter of luck. That is, he appeals to what he calls “the anti-luck intuition” or “the intuition that when one knows, one’s cognitive success… is not a matter of luck” (Pritchard 2012: 1).

According to Pritchard (2012), obviously, Henry acquires the true belief that the object he sees is barn through the exercise of Henry’s cognitive ability. The cognitive ability is the same one that he would exercise to acquire true beliefs in an environmentally friendly context; for example, where all the barn façades were replaced by real barns. Thus, in the barn façades case, the true belief is clearly attributable to Henry’s cognitive ability. However, the adopted intuition is the lack of knowledge. Therefore, the barn façades case is a counterexample to the sufficiency of the ability condition, and thus to Greco’s robust virtue epistemology as well.

Pritchard (2012) thinks that the barn façades case shows that an anti-luck condition is also needed as a necessary condition for knowledge, and that the anti-luck condition is best characterized by the safety condition. Here, we may adopt the following definition of the safety condition drawn from Timothy Williamson (2000: 147):

An agent S’s belief is safe if and only if her belief could not easily have been false in a similar case.

Others like Ernst Sosa and Stephen Hetherington share the different intuition that the barn façades case involves knowledge. Hetherington (1999, 2001) holds that Henry very failably knows that the object he sees is a barn, while Sosa (2007) thinks the barn façades case involves animal knowledge. But this intuition will not be considered in this paper.
Or, this one from Pritchard:⁴

S’s belief is safe *iff* in most near-by possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world the belief continues to be true. (Pritchard 2007: 281)

Unlike the viewpoints in some of his earlier papers, Pritchard points out in (2012) that the safety condition is not a sufficient condition for knowledge.⁵ Nevertheless, he claims that the ability condition and the safety condition are jointly sufficient for knowledge. This is his “anti-luck virtue epistemology”, and I would like to formulate it as follows:

(3) S knows p if and only if S’s acquiring the true belief p is primarily attributable to cognitive ability, and the belief that p is safe.

Just like Greco’s robust virtue epistemology, Pritchard’s anti-luck virtue epistemology is a version of process K-reliabilism. Moreover, Prichard’s anti-luck virtue epistemology can account for the barn façades case easily. For the safety condition is violated, that is, the belief is unsafe, and this results in the lack of knowledge. Moreover, since Greco’s robust virtue epistemology can account for the Mr. Truetemp case, Pritchard’s anti-luck virtue epistemology, of course, can account for it as well.

---

⁴ In a paper more recent than (2007), Pritchard gives a more sophisticated definition: “S’s belief is safe *iff* in most near-by possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, and in all very close near-by possible worlds in which S continues to form her belief about the target proposition in the same way as in the actual world, the belief continues to be true” (Pritchard 2009b: 34). The definition in (2007), however, suffices to serve the purposes of this paper.

⁵ In his paper “Anti-luck Virtue Epistemology” (2012: 15-6), Pritchard uses the Temp case to show that there are safe beliefs which do not constitute knowledge. Indeed, it is arguable that the Mr. Truetemp case discussed here is also a counterexample to the sufficiency of the safety condition, provided extended cognition is not allowed. But I won’t go into the details here.
Lastly, one should not overlook the version of K-reliabilism that Alvin Goldman puts forward in his paper “Discrimination and Perceptual Knowledge” (1976), where he offers the barn façades case for the first time. Goldman holds that in order for a belief to count as knowledge, there must be a reliable cognitive process, and he explains the notion of reliability here as follows:

To be reliable, a cognitive mechanism must enable a person to discriminate or differentiate between incompatible states of affairs. (Goldman 1976: 771)

Explicitly, Goldman claims that,

I suggest that a person is said to know that \( p \) just in case he distinguishes or discriminates the truth of \( p \) from relevant alternatives. (Goldman 1976: 772)

Goldman emphasizes the same points in his latter discussions on reliabilism (2008, 1998). The version of K-reliabilism Goldman holds can be formulated as follows:

(4) \( S \) knows \( p \) if and only if the true belief \( p \) is acquired via a reliable belief-forming process, and “its mode of acquisition must rule out all serious or ‘relevant’ alternatives in which the belief would be false.” (Goldman 1998)

Let us call (4) “the Goldmanian K-reliabilisim”. The right hand side of the biconditional is, of course, its reliability condition. Goldman explains that, in the barn façades case, “[t]he presence of the facsimiles does not ‘create’ the possibility that the object Henry sees is a facsimile… What the presence of the facsimiles does is make this possibility relevant; or
it makes us *consider* it relevant” (Goldman 1976: 775). Therefore, since the relevant alternatives are not eliminated, the reliability condition in (4) is violated; and this explains the lack of knowledge in the barn façade case nicely.

However, although the Goldmanian K-reliabilism accounts for the barn façades case successfully, the Mr. Truetemp case is actually a counterexample to the sufficiency of its reliability condition and thus to it. In the scenario of the Mr. Truetemp case, there is a reliable belief-forming process. However, there are *no* relevant alternatives in this case. It is in this sense that the mode of acquisition can be said to have been ruled out all relevant alternatives in which the belief would be false. So the reliability condition of the Goldmanian K-reliabilism is satisfied, and yet the Mr. Truetemp case is a case of ignorance. Therefore, the Mr. Truetemp case is a counterexample to the sufficiency of the reliability condition, and thus to the Goldmanian K-reliabilism.

2. **Recognitional and non-recognitional cognitive abilities**

The main objective of this paper is to offer a counterexample to the robust and anti-luck virtue epistemologies (and thus to the simple K-reliabilism as well). To achieve this, it is important to investigate the nature of the various different cognitive abilities playing the essential role of the acquisition of those true beliefs in cases like the barn façades case, as well as normal cases like those where all the barn façades are replaced by real barns. So let us consider this case:

[Real barns] The scenario is exactly the same as that in the barn façades case, except that all of the barn façades are now replaced by real barns.
Our intuition here is, of course, that Henry in the real barns case knows that the object he sees is a barn. Also, in this case, there is a belief-forming process. The belief-forming process is produced by the exercise of Henry’s cognitive ability—that is, the ability to form beliefs or to make judgments. It should be noted that different cognitive abilities can be at play here, and that, in the real barns case, it has not been stipulated which cognitive ability is playing the relevant role. Let me explain this in detail.

In the real barns case, Henry may exercise his recognitional ability to form the true belief that the object he sees is a barn. In other words, knowledge attribution is possible in this case, because Henry recognizes the object as a barn. More precisely, Henry can recognize the object as a barn, because he can distinguish the barn from all other barn-like objects or, more precisely, barn-like non-barns, in the relevant context or environment. So, although it is difficult to define the notion of recognitional ability, at least it can be characterized in terms of the notion of discriminational ability, as follows:\(^6\)

\[(5) \quad S \text{ recognizes an object as an } X \text{ if and only if } S \text{ distinguishes it from all other } X\text{-like non-}X\text{s, if any, in the relevant context or environment.}\]

The ability to recognize an object as an X may be called “recognitional ability”, while the ability to distinguish an object from all other X-like non-Xs “discriminational ability”.

It should be noted that recognitional ability is context-sensitive, while discriminational ability not. If one can distinguish a rabbit from a hare, then she has that

---

\(^6\) If the object is a person, then (4) can be formulated as follows:

\[S \text{ recognizes a person as } X [\text{a proper name}] \text{ if and only if } S \text{ distinguishes her from all other } X\text{-like persons, if any, in the relevant context or environment.}\]
discriminational ability independent of any contextual factors. But, if one can recognize an object as a rabbit, she must be able to distinguish the rabbit from all the rabbit-like non-rabbits, if any, in the relevant context. That’s why recognitional ability is context-sensitive.

In the real barns case, the fact that Henry knows that the object he sees is a barn may be because he recognizes the object as a barn. In other words, it may be because Henry exercises his recognitional ability to produce a belief-forming process which supports knowledge attribution. The above description of the real barns case does not indicate if Henry uses his recognitional ability to acquire the true belief, although that is possible. Of course, when we set up the case, we may stipulate that Henry employs his recognitional ability to acquire the true belief. So let us call the real barns case where, by stipulation, the cognitive ability at place is a recognitional ability “the recognitional real barns case”. Then, in the recognitional real barns case, the cognitive success is attributable to Henry’s recognitional ability.

Let us assume that, in the real barns case, it is possible for Henry to employ one of his other cognitive abilities, which are not recognitional, to acquire the true belief that the object he sees is a barn. (I shall argue later that this is a genuine possibility.) For example, it is possible that, in the real barns case, Henry exercises his perceptual ability to obtain a perception of the object in front of him, and based on this perception, and not on anything else, forms the belief that the object he sees is a barn, independent of any contextual factors. This is an instance of exercising a certain cognitive ability to produce a reliable belief-forming process supporting knowledge attribution in the real barns case, that is, in the unspecified case. Moreover, since this cognitive ability is not context-sensitive, it is not a recognitional ability. Thus, in this case, Henry knows that the object he sees is a barn is not
because he recognizes it as a barn, but because the reliability of the belief-forming process produced by the non-recognitional cognitive ability is strong enough to secure knowledge attribution. (Perhaps, I should remark here that merely reliability, of the relevant belief-forming process, may not be enough to secure knowledge attribution. The reliability has to be strong enough. I will argue for this later.) Let us call the real barns case where, by stipulation, the relevant cognitive ability is a non-recognitional cognitive ability “the non-recognitional real barns case”. It should be noted that the non-recognitional cognitive ability in the non-recognitional real barns case need not be the one just described, as Henry may have other non-recognitional cognitive abilities as well.

I am now going to prove that the aforementioned possibility is a genuine one, and thus that the possibility of the non-recognitional real barns case is also genuine. Consider the following case:

[Single barn façade] The scenario is exactly the same as that in the barn façades case, except that all of the barn façades, save one, are now replaced by real barns. That is, there is one and only one barn façade in the county.

I think the intuition for the single barn façade case is that Henry knows that the object he sees is a barn. The true belief is safe, and is acquired by a belief-forming process produced by the exercise of one of Henry’s cognitive abilities. Also, the belief-forming process is reliable, because it is very likely to produce true beliefs or it “produces either mostly true beliefs or a high ratio of true to false beliefs” (Becker 2009). Furthermore, the reliability of the belief-forming process is so strong that it supports knowledge attribution. This explains the presence of knowledge in the single barn façade case. However, it is important
to note that Henry cannot recognize the object as a barn, because he cannot distinguish the barn from the only barn façade in the context (the relevant county), and this violates the condition in (5). It follows that the cognitive ability producing the true belief, which constitutes knowledge, is not a recognitional ability. This proves that the aforementioned possibility that there be a non-recognitional cognitive ability in the real barns case, that is, the unspecified case, which produces the belief-forming process supporting knowledge attribution, is a genuine possibility. This also proves that the possibility of the non-recognitional real barns case is genuine.

Before proceeding further, two remarks should be made here. First, unlike the case of reliably acquired true belief, which does not necessarily secure knowledge attribution, successfully exercising a recognitional ability always gives rise to knowledge. For to say that the subject recognizes the object as an X simply means that she knows that the object she sees is an X.

Second, it should be noted that whenever there is an indistinguishable non-X in the vicinity or the relevant context, the subject cannot recognize the relevant object as an X, and thus does not know that the object she sees is an X by merely exercising her recognitional ability. It follows that the exercise of a recognitional ability to acquire perceptual knowledge with respect to, say, X is successful, only if there are no indistinguishable non-Xs in the relevant context. So, Henry cannot exercise his recognitional ability in the single barn façade case. But this does not mean that the single barn façade case is a case of ignorance. For there may be, and indeed there are, other cognitive abilities capable of producing the true belief in a way such that knowledge
attribute is supported. The non-recognitional cognitive ability discussed above is exactly one such cognitive ability.

3. \textbf{A counterexample to the robust and anti-luck virtue epistemologies}

I am now going to offer a counterexample to the sufficiency of the conjunction of the ability and safety conditions, and thus to Pritchard’s anti-luck virtue epistemology. Of course, it is also a counterexample to both the simple K-reliabilism and Greco’s robust virtue epistemology. Although it is not a counterexample to the Goldmanian K-reliabilism, the Mr. Truetemp case is already one. Therefore, since three of the prominent versions of K-reliabilism are given a counterexample, and the fourth one, the Goldmanian K-reliabilism, already has another counterexample, the Mr. Truetemp case, this is a strong attack on K-reliabilism in general. (Moreover, as the Mr. Truetemp case is not a counterexample to the robust and virtue epistemologies, the discovery of a counterexample to them can also be seen as supplementary to the argument against K-reliabilism by means of counterexamples like the Mr Truetemp case.)

The strategy to construct a counterexample is as follows:

To find a case of ignorance, where (i) there \textit{is} a reliable belief-forming process produced by a non-recognitional cognitive ability, and (ii) there \textit{isn’t} any other knowledge-attribution-supporting belief-forming process produced by a recognitional ability or one of the other non-recognitional cognitive abilities, if any.

The following is such a case:
[Fewer barn façades] The scenario is exactly the same as that in the non-recongitional real barn case, except that 40% of the barn-like objects in the county are barn façades, and 60% real barns. Also, the barn façades and the real barns are distributed evenly in the county.

In the fewer barn façades case, there is a belief-forming process produced by the exercise of one of Henry’s non-recognitional cognitive abilities, based on perception. (The non-recognitional cognitive ability may be, or may not be, the one described in the above discussion concerning the non-recognitional real barns case.) The belief-forming process is reliable, because it is more likely to produce true beliefs than false beliefs, with the ratio being 6 to 4. So, the cognitive success is attributable to Henry’s cognitive ability. Also, the belief, as the product of the reliable belief-forming process, is safe, because in most, indeed 60%, of the near-by possible worlds in which Henry continues to form his belief in the same way as in the actual world the belief continues to be true. As a result, the fewer barn façades case fulfils both the ability and the safety conditions of the anti-luck virtue epistemology.

However, the intuition for this case should be the lack of knowledge, that is, that Henry does not know that the object he sees is a barn. The reliability of the belief-forming process is not strong enough to secure knowledge attribution. (This also proves the aforementioned claim that merely having reliability, of the relevant belief-forming process, may not be enough to secure knowledge attribution.)

The reason why there is a lack of knowledge in the fewer barn façades case is that, besides the fact that the reliability of the reliable belief-forming process is not strong
enough, there is also a lack of recognitional ability and other non-recognitional cognitive abilities capable of producing strong enough reliable belief-forming processes. For, first, Henry cannot distinguish the barn from 40% of the barn-like non-barns (barn façades) in the relevant context, and thus cannot recognize the object he sees as a barn; and, second, there are no other belief-forming processes produced by any other non-recognitional cognitive abilities in the fewer barn façades case so stipulated.

This shows that the fewer barn façades case is a counterexample to the sufficiency of the conjunction of the ability and safety conditions, and thus to the simple K-reliabilism, and the robust and anti-luck virtue epistemologies.⁷

References


⁷ [Acknowledgement.]


——— (2006). “Greco on Reliabilism and Epistemic Luck.” *Philosophical Studies* 130:

35-45.


Oxford: OUP