

ANTI-LUCK (TOO WEAK) VIRTUE EPISTEMOLOGY

Fernando Broncano-Berrocal

(published in *Erkenntnis*)

Duncan Pritchard (*et al.* 2010, 2012a, 2012b) has recently shifted his view from an account that primarily understands knowledge in terms of the safety principle (a view mainly developed in his 2005 monograph on epistemic luck) to an ‘impure’ variety of virtue epistemology, which combines the safety principle with a weakened virtue-theoretic condition, a view that he calls *anti-luck virtue epistemology*.¹ This paper argues that anti-luck virtue epistemology is not sufficient for knowledge. Section 1 explains why Pritchard thinks that safety is not sufficient for knowledge. Section 2 explains why he adopts a weakened virtue-theoretic condition. Section 3 interprets anti-luck virtue epistemology as the conjunction of safety and the virtue-theoretic condition, and presents some counterexamples. Section 4 interprets anti-luck virtue epistemology as a mixed condition of safety and the virtue-theoretic condition, and presents some counterexamples. Section 5 raises a problem for any virtue-theoretic account of knowledge: the direction of fit problem.

1 The insufficiency of safety

The core idea of the safety principle is that one knows a proposition only if one’s true belief in that proposition could not easily have been false, where this is usually cashed out in terms of possible worlds: safe beliefs match the truth across a large enough range of close possible worlds.² A belief, however, can be safe without being known, as the following case shows, and this is the main reason Pritchard supplements the safety principle with a virtue-theoretic condition:

¹ I borrow the terminology from Kelp (2012): ‘pure’ virtue epistemology is any view that holds that knowledge is true belief (sufficiently) *due to* intellectual virtue; ‘impure’ virtue epistemology is any account that requires additional non-virtue-theoretic conditions (e.g., modal principles like safety).

² See Broncano-Berrocal (*forthcoming*) for a recent defense of the necessity of the safety principle.

TEMP

Imagine that our agent—let’s call him ‘Temp’—forms his beliefs about the temperature in his room by consulting a thermometer on the wall. Unbeknownst to Temp, however, the thermometer is broken and is fluctuating randomly within a given range. Nonetheless, Temp never forms a false belief about the temperature by consulting this thermometer since there is a person hidden in the room, next to the thermostat, whose job it is to ensure that whenever Temp consults the thermometer the *temperature* in the room corresponds to the reading on the thermometer (Pritchard *et al.* 2010: 48-49; emphasis mine).

Temp’s beliefs about the temperature match the truth in the actual and in all close possible worlds, because the hidden agent is so diligent that she would never fail to modify the room’s temperature. However, although Temp’s beliefs are safe, they do not intuitively amount to knowledge because, unbeknownst to Temp, the world has been changed to make true the content of his beliefs.³ Pritchard offers the following diagnosis of the case:

[W]hat is wrong with Temp’s beliefs is that they exhibit the *wrong direction of fit* with the facts, for while his beliefs formed on this basis are guaranteed to be true, their correctness has nothing to do with Temp’s abilities and everything to do with some feature external to his cognitive agency (Pritchard 2012a: 260; emphasis mine).

In addition, Pritchard’s particular reason for thinking that safety is not sufficient for knowledge is this:

[The safety principle] will simply demand a match between belief and fact in appropriate counterfactual cases, but a cognitive ability requires far more than this—viz., it requires an appropriate direction of fit between belief and fact (Pritchard 2012a: 272).⁴

³ It is important to keep in mind that the hidden agent modifies the temperature of the room and not the values shown on the screen of the thermometer. If the agent manipulated the thermometer instead of the temperature, the intuition that there is no knowledge would not be so strong. In fact, one could argue that the case, so understood, would be analogous to a case of testimonial knowledge.

⁴ Sherrilyn Roush (2005: 121-2) makes a related albeit different point “[t]hough we want safe rather than unsafe beliefs, safety is surely not what makes a true belief knowledge, since knowledge is a matter of responsiveness to the way the world is and safety makes a demand in the opposite direction”.

I will accept Pritchard's diagnosis of TEMP as an assumption of my argument against his view. In particular, I will take for granted the following implication:

DIRECTION OF FIT: *S* does not know that *p* if *S*'s belief that *p* exhibits the wrong direction of fit with the truth-maker of *p*.

I will also assume that the safety principle is unable to guarantee the direction of fit with the facts that known beliefs have.

2 The need of a virtue-theoretic condition

Pritchard's way to remedy the insufficiency of safety and to ensure the appropriate direction of fit between belief and fact is to resort to virtue epistemology. Following Kelp (2012), the core thesis of virtue epistemology can be glossed as follows:

VIRTUE EPISTEMOLOGY: *S* knows that *p* only if *S* truly believes that *p* and *S*'s cognitive success (i.e., *S*'s believing truly) is (sufficiently) due to/because of cognitive competence or ability.

Reliabilist virtue epistemologists conceive cognitive abilities as stable dispositions to form true beliefs (e.g., Greco 2000: 210-7; Sosa 2010: 465-7). These dispositions are reliable with respect to fields of propositions (e.g., propositions about colors) and to ranges of appropriate circumstances, which are typically understood as sets of external conditions of a range of environments (e.g., good light conditions) plus the internal conditions of the agent in those environments (e.g., being in good shape, sober, etc.).⁵

The real dispute among virtue epistemologists is on how to understand the technical 'because of' relation in VIRTUE EPISTEMOLOGY. Although some regard it as primitive (e.g., Zagzebski 1996), virtue epistemologists are divided into two camps (here I follow again Kelp [2012]): those who interpret it in terms of *manifestation* of competence (e.g., Kelp 2012; Sosa 2011; Turri 2011) and those who understand the relation in terms of *explanatory salience* or

⁵ Reliability so conceived corresponds to what is commonly known as *global reliability*. A globally reliable belief-forming method is a method that is truth-conducive with respect to a specified *range* of propositions and circumstances (the relevant truth-conduciveness might be understood in statistical or in modal terms). *Local reliability*, by contrast, is defined solely in terms of the modal profile of a single belief in certain proposition when formed in specific circumstances. In this way, the safety principle might be understood as a sort of local reliability requirement: safety is a matter of a particular belief matching the truth across close possible worlds, where the relevant world order is fixed in terms of similarity to the actual specific circumstances.

credibility (e.g., Greco 2010).⁶ In this section, I will focus on the latter reading, as it is the one advocated by Pritchard. In section 5, I will discuss the former.

The general idea concerning the explanatory/credibility reading is that *S* knows that *p* only if *S*'s cognitive success is *explained* (with some degree of salience) by the exercise of her cognitive abilities. Alternatively, authors who endorse this reading think that an equivalent way to put it is this: *S* knows that *p* only if *S*'s cognitive success is (to some degree) *credible* (or *attributable*) to cognitive ability. The explanatory/credibility interpretation leads to two views on knowledge, the difference between which lies in the degree of explanatory salience or credibility required for knowledge:

FULL CREDIBILITY: *S* knows that *p* only if *S*'s cognitive success is *primarily* credible to the exercise of *S*'s cognitive abilities. (Alternatively: only if the exercise of *S*'s cognitive abilities is the *most salient* factor in the total set of factors that explain *S*'s cognitive success).

PARTIAL CREDIBILITY: *S* knows that *p* only if *S*'s cognitive success is *partly* or *to a significant degree* credible to the exercise of *S*'s cognitive abilities. (Alternatively: only if the exercise of *S*'s cognitive abilities is a *salient* factor in the total set of factors that explain *S*'s cognitive success).

FULL CREDIBILITY is most prominently endorsed by Greco (2010) (he takes the condition to be not only necessary but also sufficient for knowledge).⁷ PARTIAL CREDIBILITY is the epistemic condition that Pritchard chooses to remedy the insufficiency for knowledge of the safety principle (as we will see, Pritchard's definition of knowledge includes a virtue-theoretic clause that he explicitly understands in terms of PARTIAL CREDIBILITY). I will argue that neither the conjunction of safety and PARTIAL CREDIBILITY, nor a non-conjunctive combination of both is sufficient for knowledge. Before that, and to understand Pritchard's virtue-theoretic condition properly, I will clarify some issues concerning the explanatory/credibility reading of the 'because of' relation.

To begin with, some commentators (e.g., Lackey 2007, 2009; Greco 2003; Riggs 2002, 2009; Vaesen 2011) think that the so-called *credit view of knowledge* (roughly, the view that if *S* knows that *p*, *S* deserves credit for truly believing that *p*) is equivalent to the view that if *S* knows

⁶ See Greco (2012) and Turri (2011) for relevant discussion on the 'because of' relation.

⁷ It must be pointed out that Greco has recently changed his view on knowledge. See Greco (2012).

that p , S 's cognitive success is creditable to cognitive ability. However, this equivalence is slightly misleading, since the notion of creditworthiness is not equivalent to the notion of creditability. As Pritchard (2012a: 264, fn. 26) correctly points out, there is a difference between one's cognitive success being *creditable* to one's cognitive abilities and one's cognitive success being *of credit*. For example, it is certainly creditable to S 's memory skills having memorized all the names that appear on the phone book, but such a cognitive success is not of credit if S ought to have used her extraordinary memory abilities to pursue more desirable goals (e.g., to pass a potential life-changing exam). In general, while attributions of epistemic credit seem to be sensitive to the type of inquiry one is pursuing or the kind of activity one is engaged in, creditability to cognitive ability is just a matter of whether the exercise of one's cognitive abilities explains or not one's cognitive success (not a matter of what type of inquiry one is pursuing). Accordingly, Pritchard's PARTIAL CREDITABILITY should not be considered an instance of the so-called credit view of knowledge.

Second, the notion of *cognitive success* means believing truly, which means in turn that the exercise of cognitive ability might explain two things: why the agent forms the belief that p and why the agent comes to believe the truth about p (i.e., why the belief exists and why the agent gets things right). As Sosa correctly notes (2007: 95), one thing is to explain why certain entity exists and another thing is to explain why such an entity has certain property. We can therefore expect (1) cases in which an agent does not know because the *existence* of her belief is not explained by the exercise of the agent's cognitive abilities and (2) cases in which an agent does not know because the fact that she *gets things right* (or, more metaphorically, the fact that her belief 'hits' the truth) is not explained by the intervention of her cognitive abilities.

Examples of (1) are scenarios in which someone hypnotizes a person to implant in her some belief or in which an evil neuroscientist electrochemically stimulates someone's brain to make her believe certain things. Independently of whether such beliefs are true or false, it seems that these are cases in which agents do not know. The following condition explains why: S knows that p only if the existence of S 's belief that p is *fully* creditable to or explained by the exercise of her cognitive abilities. This condition is entailed by FULL CREDITABILITY. Advocates of

PARTIAL CREDITABILITY should endorse it too, as it does not seem possible for an agent to know a proposition if the *existence* of her belief is only partially creditable to her cognitive abilities.⁸

Examples of (2) will be given in a moment. First note that the considerations just offered indicate that the difference between FULL CREDITABILITY and PARTIAL CREDITABILITY lies in the other aspect of the notion of cognitive success, namely in how much creditable to cognitive ability is the fact that the agent comes to believe the truth. Accordingly, FULL CREDITABILITY should be understood as the condition that *S* knows that *p* only if *S*'s coming to believe the truth about *p* (i.e., *S*'s getting things right) is *fully* (or *primarily*) creditable to the exercise of her cognitive abilities. PARTIAL CREDITABILITY, on the other hand, should be understood as the condition that that *S* knows that *p* only if *S*'s coming to believe the truth about *p* is *partly* (or *to a significant degree*) creditable to the exercise of her cognitive abilities. The difference between coming to believe the truth partly and fully because of cognitive ability becomes evident if we consider cases of testimonial knowledge.

According to Pritchard, if *A* knows that *p* via the testimony of *B*, *A*'s cognitive success (viz., *A*'s coming to believe the truth about *p*) is *not primarily creditable* to her cognitive abilities (although the existence of her belief is), because the exercise of *B*'s cognitive abilities is

⁸ All cases of belief formation that is only partially creditable to cognitive ability seem to have the following structure: *S* forms the belief that *p* *partly* because of the exercise of cognitive ability and *partly* because of some feature external to *S*'s cognitive agency (e.g., the hypnotizer's intervention or the neuroscientist's manipulation). It does not seem that we would ascribe knowledge in cases with such a structure. Still, one might think that this does not show that in order for *S* to know that *p* the existence of *S*'s belief that *p* must be *fully* creditable to her cognitive abilities. One will surely imagine counterexamples in which an agent comes to know a proposition by means of the manipulation of an external device (e.g., a notebook). On closer inspection, however, cases of that sort are not counterexamples to the aforementioned condition. For one thing, if it is correct to say that the device is a proper part of the agent's cognitive system (if the conditions for extended cognition are satisfied and the extended cognition hypothesis is true), then the existence of the agent's belief is fully creditable to cognitive ability (i.e., to her biological capabilities and to the relevant external resources (the notebook) that are *coupled* and *integrated* with them). For another, if the device is not a proper part of the agent's cognitive system, then the existence of her belief is fully creditable to the agent's *competent* manipulation of the device. In my opinion, Pritchard's anti-luck virtue epistemology is in need of some extra clause that explicitly rules out cases in which the existence of the agent's belief is not fully creditable to her cognitive abilities as cases of knowledge. However, I am not sure whether he would be willing to include such a clause. See Pritchard (2010) for relevant discussion on this particular point. See also Vaesen (2011) for further discussion.

certainly a *salient* part of what explains why *A* gets to the truth about *p*. But this is not to say that *B*'s cognitive abilities are the *most salient* factor in the explanation of why *A* gets things right. The reason adduced is that when we know by testimony we do not do it blindly, but we typically display a competence for selecting reliable informants (e.g., one that prevents us from trusting children, people who look like tourists or ostensibly crazy people, but that might make us trust expert liars). In this way, when *A* acquires testimonial knowledge from *B*, *A*'s coming to believe the truth is *partly* (i.e., to a significant degree) creditable to her cognitive abilities and *partly* (i.e., to a significant degree) creditable to *B*'s, which shows, according to Pritchard, that FULL CREDITABILITY is too restrictive to account for this kind of knowledge. By contrast, the more permissive PARTIAL CREDITABILITY nicely accommodates cases of testimonial knowledge.

In addition, cases in which the fact that an agent gets things right is not explained by the exercise of her cognitive abilities, i.e., cases with the structure of (2), are ruled out as cases of knowledge by PARTIAL CREDITABILITY. Cases of this sort might have two different sub-structures: (2.1) cases in which the agent forms a belief through a belief-forming process that is considered a cognitive ability but in which the exercise of such a cognitive ability does not explain (not even saliently) why the agent comes to believe the truth and (2.2) cases in which the fact that the agent gets things right is not explained by the exercise of cognitive ability because the relevant belief-forming process does not count as a cognitive ability. Examples of (2.2) are cases in which true beliefs arise out of brain lesions or devices capable of generating thoughts that are not yet integrated with the rest of the agent's cognitive dispositions. PARTIAL CREDITABILITY rules out this kind of cases as cases of knowledge because the cognitive ability clause is not satisfied.

On the other hand, an excellent example of (2.1) is TEMP, a case in which the safety condition holds. In that case, the fact that Temp comes to believe the correct value of the temperature is *solely* explained by the fact that the hidden agent changes the temperature to match the readings of the broken thermometer. In other words, the fact that Temp gets things right is in *no way* creditable to the exercise of his cognitive abilities. PARTIAL CREDITABILITY rules out the case as a case of knowledge because the creditability clause is not satisfied.

To summarize, Pritchard's adoption of PARTIAL CREDITABILITY is motivated by the fact that it is weak enough to accommodate cases of testimonial knowledge while it is strong enough to rule out cases with the structure of (2.1) and (2.2) as cases of knowledge. The safety principle is still needed to explain why there is no knowledge in some other cases; for example, cases in

which one competently and successfully identifies an object (e.g., a Picasso painting) but in which one could easily have misidentified a replica of that object (e.g., a fake Picasso painting). The resulting account of knowledge combines safety with PARTIAL CREDITABILITY in the following way:

ANTI-LUCK VIRTUE EPISTEMOLOGY (ALVE): *S* knows that *p* if and only if *S*'s safe true belief that *p* is the product of her relevant cognitive abilities (such that her safe cognitive success is to a significant degree creditable to her cognitive agency) (Pritchard 2012a: 273).

In the next two sections, I will show that the conditions of this definition are not jointly sufficient for knowledge.⁹

3 The insufficiency of ALVE (I)

The first way of interpreting ALVE is as the conjunction of safety and PARTIAL CREDITABILITY.

More specifically, according to this interpretation ALVE says that *S* knows that *p* if and only if:

- (1) *p* is true,
- (2) *S* believes that *p*,
- (3) *S*'s cognitive success (i.e., *S*'s believing *truly*) is partly or to a significant degree creditable to *S*'s cognitive abilities,
- (4) *S*'s belief that *p* is safe.

This is the reading of ALVE that Pritchard seems to have in mind: he describes anti-luck virtue epistemology as a view “according to which knowledge is essentially safe true belief *plus* a further epistemic condition (an ‘ability’ condition), which handles the ability intuition” (Pritchard 2012b: 184, fn. 19; emphasis mine). The ability condition he refers to is PARTIAL CREDITABILITY and the ability intuition that this condition satisfies is the intuition that

⁹ In his 2010 book, Pritchard (*et al.* 2010: 58-59) takes a liberal view on whether ALVE is a *reductive account of knowledge* (i.e., on whether ALVE analyzes knowledge in terms of necessary and sufficient conditions or, as Pritchard puts it, on whether it involves a “‘decomposition’ of knowledge into its component parts where these parts can be specified in a knowledge-independent fashion” (Pritchard *et al.* 2010: 58)). One might argue that, to the extent that ALVE is not considered a reductive account of knowledge, it is pointless to argue that the theory is insufficient for knowledge. More recently, however, Pritchard (2012a) has explicitly presented ALVE as a theory that fulfills the remit of the analytical project in epistemology, whose goal is no other than “to offer an adequate definition of knowledge, one that [is] informative, noncircular, and which [can] suitably accommodate our salient epistemological intuitions” (Pritchard 2012a: 247). That is, Pritchard not only considers ALVE a reductive account of knowledge, but he also intends with it to put the analytical project back in business.

“knowledge reflects ability, in the sense that when an agent has knowledge, then her cognitive success (i.e., her true belief) is to some significant degree creditable to her cognitive agency (i.e., her exercise of cognitive abilities)” (Pritchard 2012b: 182). Safety, on the other hand, is aimed at satisfying the so-called anti-luck intuition, i.e., the intuition that when one knows, it is not by luck that one comes to believe the truth.¹⁰ Pritchard interprets these two intuitions as imposing two independent constraints on one’s theory of knowledge and, consequently, the epistemic conditions aimed at satisfying them need to be independent too, in the sense that there must be at least one type of case that one condition rules out as a case of knowledge that cannot be ruled out by the other condition, and *vice versa*. For example, safety holds in TEMP; PARTIAL CREDITABILITY does not. PARTIAL CREDITABILITY holds in some Gettier-style cases (see below); safety does not.¹¹ That being so, it is then more congruent with Pritchard’s dual approach to knowledge to interpret ALVE as entailing a conjunction of two epistemic conditions than to understand it as entailing a single condition that merges both requirements in a non-conjunctive way.

Nevertheless, I am not particularly interested in discussing which interpretation of ALVE is the one that Pritchard really has in mind. Pritchard’s formulation of ALVE is consistent with two readings (ALVE as the conjunction of safety and PARTIAL CREDITABILITY and ALVE as a non-conjunctive combination of both). Then, my objections against ALVE can be understood in conditional terms: if the former is the correct interpretation of ALVE, then one objection applies; if the latter is the correct interpretation, then another objection applies. As we will see,

¹⁰ The thesis or intuition that *S* knows that *p* only if *S*’s belief that *p* does not ‘hit’ the truth by luck is part of the orthodoxy in epistemology. I assume this thesis for the sake of the argument, as it is one of Pritchard’s core assumptions. Nevertheless, see Baumann (2012) for a provocative challenge to what most epistemologists consider a platitude.

¹¹ See Carter (2011) for an analysis of Pritchard’s requirement that the two epistemic conditions aimed at satisfying the two ‘master’ intuitions must be independent of each other and for a reconstruction of the argument that Pritchard provides to motivate ALVE (the independence requirement is one of the premises of such an argument). Carter attacks Pritchard’s independence requirement (and hence Pritchard’s argument for ALVE) by disputing one of the core assumptions on which it relies: the thesis that cognitive achievements are compatible with certain type of knowledge-undermining luck. My argument against ALVE is complementary to Carter’s: if Carter is right, ALVE is unmotivated; if I am right, ALVE is insufficient for knowledge. If we are both right, ALVE is in serious trouble.

the conclusion will be the same in either of the two cases: Pritchard's account is not sufficient for knowledge.

Before presenting a couple of counterexamples to the first interpretation of ALVE, let me discuss an objection made by Kelp (2012). According to Kelp, the contribution of Temp's cognitive abilities to his cognitive success parallels the contribution of the cognitive abilities of a person who comes to know the position of a landmark by asking for directions to a stranger, who gives the information correctly. Kelp's point is that if cognitive success is to a significant degree creditable to cognitive ability in the latter case (as Pritchard maintains), so should be in TEMP, and if cognitive success is not to a significant degree creditable to cognitive ability in TEMP, it should not be creditable in the testimony case either. The upshot is, according to Kelp, that while Pritchard can explain each case individually, he cannot explain the two cases in conjunction.

However, there is reason to doubt that the contribution of cognitive ability to cognitive success is the same in both cases. The randomly fluctuating thermometer in TEMP is not analogous to the reliable informant in the testimony case. To be analogous, the informant should give directions at random, some of them probably making no sense. In view of this disanalogy, it is not clear whether there would be testimonial knowledge if the informant uttered directions at random and the person believed the information.

We can try a better objection to this interpretation of Pritchard's theory. Consider *A*, a competent speaker who has a generally reliable ability to select good informants such that, for each informant and each situation individually, the ability monitors (perhaps on the basis of stereotypes) those features of the informant and the situation that indicate that the informant is reliable (e.g., if one wants to know some directions, the ability prevents one from trusting people who look like tourists). Pritchard accepts that if an agent has such an ability, she can know by testimony, because the ability, so displayed, partially contributes to the explanation of the agent's cognitive success. Let *F* be the set of features detected by the recipient of testimony that indicate her that her interlocutor is reliable. Consider now the following case:

ACTOR

A, new in the city, wants to know the position of a landmark. She looks for a good informant and selects *B*, a person who looks like a police officer, on the basis of *F* (*F* may include, for example, those features that make people trust the police in general).

However, *B* happens to be an actor cleverly disguised as a police officer whose intention is to give wrong directions to outsiders (*F* indicates reliability but *B* is not reliable). It also happens that *whenever B* is about to lie about the landmark's location, a benevolent Genie moves the landmark so that *B* speaks truly. Consequently *A*'s beliefs about the landmark are true.

Condition (4) (the safety requirement) is satisfied because of the diligent intervention of the Genie. In addition, condition (3) (PARTIAL CREDITABILITY) seems to be satisfied too, because there does not seem to be anything in *A*'s cognitive performance that is epistemically defective. If *A* performs badly from a cognitive point of view, her cognitive performance is as 'bad' as in the good case where she trusts a real police officer and acquires knowledge. Moreover, *A* selects her informant on a better basis than *C*, an agent who comes to know the location of a landmark by asking the first adult passerby in a train station. While *C* selects an informant on the basis that the person is an adult (rather than a child), *A* selects an informant on the basis that the person is an adult *and* that he looks like a police officer (police officers and taxi drivers are typically the most reliable informants when it comes to giving directions). Nevertheless, intuition says that *A* does not know the location of the landmark and consequently ALVE is not sufficient for knowledge.

It could be replied that *A* *does* know the location of the landmark, provided that she forms her beliefs on grounds that are at least generally appropriate and *A*'s interlocutor utters truths. That may well suffice for acquiring knowledge via testimony, one may argue, as knowledge may arise whenever the informational source (e.g., an informant or an epistemic instrument) provides true information and one's ability to detect reliable informational sources is itself (by and large) reliable. However, the problem with this reply is that the direction of fit exhibited by *A*'s beliefs seems as wrong as the one exhibited by the target beliefs in TEMP. In this way, it should be explained why ACTOR, in particular, and testimony, in general, are exceptions to DIRECTION OF FIT, which as we saw at the beginning is one of Pritchard's core assumptions (and so is for us).

In any case, we can present another counterexample to ALVE in which this kind of reply is ruled out from the outset, because the case is a Gettier-style case and Gettier subjects do not know (or so is believed by most epistemologists). The counterexample is built on the following case by Greco (2012):

INSPIRED WORKER

Jones believes *that someone in the office owns a Ford*, basing his belief on extensive evidence that his co-worker, Nogot, owns a Ford. But Jones's evidence about Nogot is misleading—Nogot owns no Ford. However, another co-worker in the office, Havit, is inspired by Jones's excellent reasoning about Nogot. So much so that he is shaken out of his long term depression, and goes out and buys a Ford (Greco 2012: 14).

According to Greco, the case illustrates why PARTIAL CREDITABILITY is not sufficient for knowledge. Since Jones's competent reasoning is the cause Havit buys a Ford and Jones's belief that someone in the office owns a Ford is true because Havit buys a Ford, part of the explanation of why Jones gets things right is the deployment of cognitive ability. Yet Jones does not know that someone in the office owns a Ford.

In a sense, the world changes to make true Jones's belief, i.e., the direction of fit exhibited by Jones's belief seems to be the wrong one. The particularity of the case is that the exercise of cognitive ability causes that the proposition believed has a truth-maker. That is to say, unlike in ACTOR, the reason the world changes is the exercise of the agent's cognitive abilities and not some external factor (such as the intervention of benevolent Genies). Assuming that this diagnosis is correct, we can transform INSPIRED WORKER into a counterexample to ALVE by tweaking the case in such a way that the safety condition holds (in INSPIRED WORKER Jones's belief is unsafe):

INSPIRED WORKERS

Jones believes *that someone in the office owns a Ford*, basing his belief on extensive evidence that his co-worker, Nogot, owns a Ford. But Jones's evidence about Nogot is misleading—Nogot owns no Ford. However, unbeknownst to Jones, everyone in the office is inspired by his excellent reasoning about Nogot. So much so that everyone is shaken out of a long term depression, and goes out and buys a car. Curiously, Fords are the only cars that are available for sale on that day.

Jones's actual belief that someone in the office owns a Ford could not easily have been false. After all, if one or two co-workers had refrained from buying a car, the rest of the staff would still have bought one (a Ford). Put it differently, in nearly all (if not all) close possible worlds in which Jones makes the same excellent reasoning about Nogot, it is not the case that everyone in the office does *not* buy a Ford. Therefore, safety holds. If Greco is right in thinking that PARTIAL

CREDITABILITY holds as well, then INSPIRED WORKERS is a counterexample to the sufficiency of ALVE (see a possible line of response below).

4 The insufficiency of ALVE (II)

Another way of interpreting ALVE is as follows. *S* knows that *p* if and only if:

- (1) *p* is true,
- (2) *S* believes that *p*,
- (3) *S*'s *safe* cognitive success (i.e., *S*'s believing that *p* *safely*) is partly (or to a significant degree) creditable to *S*'s cognitive abilities (alternatively: *S*'s *safe* cognitive success is partly (or saliently) explained by the exercise of *S*'s cognitive abilities).

This reading is consistent with the formulation of ALVE quoted earlier. Although this does not seem to be the interpretation of ALVE that Pritchard has in mind, it is worth discussing it because it introduces a novel and powerful idea in the epistemological debate: the requirement that, in order to have propositional knowledge, the *safety* (not just the 'hitting' on truth) of one's beliefs must be *partly* (or *to a significant degree*) creditable to one's cognitive agency. Given Pritchard's own version of the safety principle,¹² this requirement would amount to the condition that, to know that *p*, the exercise of one's cognitive abilities must be a salient factor in the explanation of why the belief that *p* hits the truth in the actual world and in nearly all (if not all) close possible worlds where one forms the belief in the same way as in the actual world.¹³

¹² Pritchard formulates several versions of safety. For example, in (2005: 163) he advocates the following version of the principle: "For all agents, ϕ , if an agent knows a contingent proposition ϕ , then, in nearly all (if not all) nearby possible worlds in which she forms her belief about ϕ in the same way as she forms her belief in the actual world, that agent only believes that ϕ when ϕ is true". Pritchard judges this version of safety to be superior to a weaker version in which the agent's belief must be true in most close possible worlds. Further reflection on the lottery puzzle leads Pritchard to the adoption of another version in which the target belief must be true in most close possible worlds and in all very close nearby worlds in which the agent forms the belief in the same way as in the actual world (see Pritchard 2008). My objections to ALVE apply independently of the version of safety included in the definition of knowledge.

¹³ One might think that this interpretation of anti-luck virtue epistemology aligns Pritchard's theory with Turri's account of knowledge in terms of *amplitude*. Turri (2011) suggests that knowledge may be conceived as ample belief, where an ample belief is a belief that is safe, i.e., that succeeds and that could not easily have failed, such that its safety (not just its success) *manifests* the exercise of cognitive ability. However, thinking that the two accounts are equivalent is a mistake. The reason: the virtue-theoretic conditions on which they are built are

Condition (3) is not met in ACTOR: the fact that *A* comes to believe the truth in most close possible worlds in which she asks *B* for directions is solely explained by the intervention of the benevolent Genie, who relocates the landmark so that *B* speaks truly. In addition, the requirement is not met in the counterexamples to the sufficiency for knowledge of purely virtue-theoretic accounts (cases of knowledge-undermining environmental luck; see section 5). Besides, it is not satisfied in counterexamples to accounts that define knowledge just in terms of safety (e.g., TEMP). Finally, the condition holds in cases of inductive knowledge, which prove that the sensitivity principle is not necessary for knowledge.¹⁴ In sum, ALVE, so interpreted, seems to be in a better position than many of its rivals (viz., than ‘pure’ virtue epistemology and than safety and sensitivity-based theories).

Pritchard could adopt this reading of ALVE to avoid counterexamples like ACTOR, if the first interpretation is the one that he really has in mind. After all, his formulation of ALVE is consistent with understanding the theory in this way. However, although this reading is better than the conjunctive interpretation, it is not sufficient for knowledge either. To begin with, it seems that condition (3) holds in INSPIRED WORKERS. Condition (3) says that known beliefs have the following modal profile: <true in the actual world and true in nearly all (if not all) close possible worlds in which they are formed in the same way as in the actual world>. More importantly, it says that the exercise of the relevant cognitive abilities must be a salient factor in the total set of factors that explain this modal profile. In INSPIRED WORKERS, the following holds: in nearly all (if not all) close possible worlds in which Jones believes that someone in the office owns a Ford via the same method of belief formation that he uses in the actual world (reasoning), many workers, inspired by that method, buy Fords, and hence the fact Jones comes

clearly different. More specifically, condition (3) of the second interpretation of ALVE is built on PARTIAL CREDITABILITY, whereas the amplitude condition is built on aptness (adeptness in Turri’s terms; see fn. 22). Yet more specifically, PARTIAL CREDITABILITY states an explanatory relation, whereas aptness states a metaphysical relation (see section 5). One consequence of this is that PARTIAL CREDITABILITY is unable to rule out TEMPY as a case of knowledge (the counterexample to the second interpretation of ALVE that I will present in this section), while aptness can arguably dispose of the case. In conclusion, the second interpretation of ALVE might not be the view that Pritchard has in mind, but it is not the view that Turri puts forward either.

¹⁴ The sensitivity principle says that if *S* knows that *p*, if *p* were false, *S* would not believe that *p*. Sosa (1999) and Vogel (2007) present cases that show that one can know that *p* via induction even though one would believe that *p* if *p* were false.

to believe the truth in all those worlds is partly because of his cognitive abilities.¹⁵ Of course, Jones does not know that someone in the office owns a Ford.

In reply, PARTIAL CREDITABILITY could be supplemented with some clause restricting the sort of creditability relation that must hold between the exercise of cognitive ability and the fact that the agent gets to the truth about whether p . In INSPIRED WORKERS, the fact that Jones gets things right, although creditable to his reasoning abilities, is not *directly* creditable to their exercise, in the sense that it is creditable to the causal impact on the world of the way he forms his belief, and namely to the fact that the exercise of cognitive ability produces a truth-maker of the proposition believed (and not a belief that is true in virtue of some already existing truth-maker). To see this more clearly, think of a method of belief formation as a function that takes as input certain piece of evidence or information and that gives as output a belief that is true or false. The fact that an agent comes to believe the truth is not directly creditable to the operation of that function if the function has a side effect, that is, if in addition to returning a value (the belief), it also modifies the state of the world in such a way that the value is true independently of the evidence or the information taken as input. This line of reply needs to be elaborated, but the general idea should be sufficiently clear: by restricting the creditability requirement of PARTIAL CREDITABILITY to something like direct or non-deviant creditability, PARTIAL CREDITABILITY (and ALVE) might be able to explain why there is no knowledge in INSPIRED WORKER(S).

As one can imagine, it is not an easy task to design a counterexample to ALVE. Let us rethink our strategy then. Note that INSPIRED WORKERS is a case originally constructed to call into question PARTIAL CREDITABILITY that I have subsequently modified so that the safety condition holds. We can proceed the other way around, that is, we can start with a case that challenges the sufficiency of safety (i.e., a case in which the modal profile of the target belief is <true in the actual world and true in nearly all (if not all) close possible worlds in which it is formed in the same way as in the actual world>) and then modify it so that the exercise of cognitive ability meets the demands of PARTIAL CREDITABILITY and of condition (3).

A good candidate is TEMP. Temp's beliefs are safe, but their safety is *completely* explained by the manipulation of the hidden agent, so that we cannot credit him with knowledge. For example, if the temperature in the room is 25°C and the broken thermometer

¹⁵ I talk indistinctively of ways of believing and of methods of belief formation.

reads 25.4°C, whenever Temp is about to form the belief that the temperature is 25.4°C by consulting the thermometer, the hidden agent ensures that temperature of the room reaches 25.4°C so that it matches the value shown on the screen. This manipulation takes place in the actual world and in all close possible worlds and, therefore, any version of safety holds.

What we seek next is a case in which the exercise of cognitive ability is a salient factor in the explanation of the safe modal profile of the target belief, while we can still retain the intuition that knowledge is absent because an aspect of the situation, which is beyond the agent's cognitive reach, partially explains, together with her relevant cognitive abilities, the safety of her cognitive success. As in TEMP, this factor will be the manipulation of a hidden person, but this time the manipulated agent (and the case) will be called Tempy:

TEMPY

THE AGENT: Tempy was born with an extraordinary although slightly inaccurate ability to detect the temperature: she can perceive that the temperature of the environment is in a range of $\pm 1^\circ\text{C}$ in a very reliable manner. For example, if the temperature is 25.4°C, she can tell that it is between 24°C and 26°C, but she cannot know that the temperature is exactly 25.4°C.

THE LOCATION: She is in her room, where she does not normally use her detection ability because it requires a lot of effort and concentration. Instead, she usually looks at the digital thermometer on the wall to know the temperature of the room (as in TEMP, the thermometer has always been very reliable in the past and by consulting it the agent has acquired knowledge many times).

THE EPISTEMIC POSITION: This time, Tempy has acquired, via a trustworthy informant (e.g., the heating technician), excellent justification to believe the following *false* proposition: <the thermometer indicates 3°C more than the real temperature>. That is, unless Tempy exercises her detection ability, she will believe that the temperature of the room is 3°C more than what the thermometer says.

THE SPECIAL CIRCUMSTANCES: Unbeknownst to Tempy, the thermostat is broken and fluctuating randomly within a range of $\pm 0.5^\circ\text{C}$ with respect to the real temperature and, as in TEMP, a hidden agent raises or lowers the temperature of the room so as to match the values shown on the screen. This agent can raise or lower the temperature of the room only within the range of fluctuation of the thermometer ($\pm 0.5^\circ\text{C}$).

THE BELIEF-FORMING METHOD: Since Tempy has reasons to distrust the thermometer, she decides to implement the following belief-forming method: first she will use her detection ability to check in which range the real temperature is; then, she will look at the thermometer to check whether it is true that it indicates 3°C more than the real temperature (as she has been told). If this is false, then, given that the thermometer has always been very reliable in the past, she will finally trust the readings of the thermometer.

THE FORMATION OF THE BELIEF: Tempy perceives that the temperature is between 24°C and 26°C (the real temperature is 25.4°C). The broken thermometer says that it is 25°C. As a consequence, Tempy rejects the justification to believe that the temperature is 28°C. Then, by trusting the thermometer's reading, she forms the belief that the temperature is 25°C, but here is the trick: in the meantime, unbeknownst to Tempy, the hidden agent has adjusted the temperature from 25.4°C to 25°C.

THE INTUITION: Tempy's belief that the temperature is 25°C is true, but it does not amount to knowledge.

The case is complex but the explanation of why it is a counterexample to ALVE is simple. Without the detection ability, Tempy (or any other person) would have believed that the temperature is 28°C. Why? Because of the excellent testimonial justification to believe the *false* proposition that the thermometer indicates 3°C more than the real temperature (it actually indicates 25°C). This means that without the detection ability in place the target belief would have had the following modal profile: <false in the actual world and false in all close possible worlds in which it is formed in the same way as in the actual world>.¹⁶ However, given that the use of the detection ability is part of Tempy's belief-forming method, she is in a position to reject her testimonial justification and, consequently, her belief is true in the actual world, but also, and more importantly, in all close possible worlds. That is, Tempy's cognitive abilities *saliently* (but do not completely) explain the safety of her cognitive success. Therefore, condition (3) holds (PARTIAL CREDITABILITY obviously holds as well).

¹⁶ Note that the case is described in such a way that the hidden agent would have adjusted the temperature of the room to 25°C (the reading of the thermometer) regardless of Tempy's evidence or of the reliability or accuracy of her belief-forming method.

The same diagnosis offered by Pritchard of why TEMP is not a case of knowledge applies here. As in TEMP, the direction of fit exhibited by Tempy's belief is the wrong one: although there is correspondence between belief and fact, the world is changed by the hidden agent to match the content of the belief. Following the dictate of DIRECTION OF FIT, Tempy does not know. Therefore, Pritchard's diagnosis of TEMP, fully applicable to TEMPY, indicates that ALVE cannot guarantee that known beliefs have an appropriate direction of fit with the facts. Ironically, the preservation of the appropriate direction of fit was one of Pritchard's main motivations for the adoption of a virtue-theoretic condition in the first place. However, the weak virtue-theoretic condition that Pritchard adopts (PARTIAL CREDITABILITY) is *too* weak.

One might still have the intuition that Tempy knows that the temperature is 25°C degrees, but this thought arises, I think, because the difference between the range of temperature that Tempy can detect ($\pm 1^\circ\text{C}$) and the range of fluctuation of the thermometer ($\pm 0.5^\circ\text{C}$) is not very significant. Accordingly, one may feel that, given the degree of accuracy of Tempy's detection ability, her belief that the temperature is 25°C has an appropriate direction of fit and that, although the world is certainly changed to match the content of the belief (from 25.4°C to 25°C), such a change is not significant enough to reverse its direction of fit with the facts.¹⁷

We can strengthen the intuition that the direction of fit is the wrong one by increasing the difference between the range of temperature that Tempy can detect and the range of fluctuation of the thermometer. For example, suppose that: 1) Tempy is able to perceive that the temperature of the environment is in a range of $\pm 7^\circ\text{C}$ in a very reliable manner; 2) she has justification to believe that the thermometer indicates 10°C more than the real temperature; 3) the range of fluctuation of the thermostat is $\pm 2^\circ\text{C}$ with respect to the real temperature; 4) the temperature in the room is 30°C; 5) the broken thermometer says that it is 28.5°C; 6) unbeknownst to Tempy, the hidden agent lowers the temperature from 30°C to 28.5°C. In this case, Tempy perceives (knows) that the temperature is between 37°C and 23°C (the real temperature is 30°C). Since the broken thermometer says that it is 28.5°C, she rejects the justification to believe that the temperature is 38.5°C and comes to believe (falsely) that the thermometer is working properly. Finally, when she is about to form the belief that the

¹⁷ I am grateful to an anonymous reviewer for urging me to discuss this issue in more detail.

temperature is 28.5°C, the hidden agent adjusts the temperature from 30°C to 28.5° C. Clearly, Tempy does not know that the temperature is 28.5°C.

In what follows, I will evaluate two possible replies. First, it could be argued that Tempy's belief-forming process belongs to the same category as the belief-forming processes of Laurence Bonjour's clairvoyants (Bonjour 1985: 38-41) and Keith Lehrer's Mr. Truetemp (Lehrer 1990: 163-164). Mr. Truetemp, for instance, is an agent with a small device implanted in his brain that can detect with great precision the temperature of the environment, and that is also capable of generating beliefs about the temperature on that basis. Lehrer thinks that although the device is a very reliable way of forming true beliefs about the temperature, Mr. Truetemp's beliefs do not qualify as knowledge. Many epistemologists share this intuition and, consequently, many epistemologists will have the same intuition about TEMPY *if* Tempy's belief-forming process is like Mr. Truetemp's.¹⁸ Pritchard could avail himself of that intuition and argue that Tempy's safe cognitive success is not creditable to cognitive ability because there is no cognitive ability in the first place.

As Lehrer describes him, Mr. Truetemp has no idea that the device has been inserted in his brain and that he is able to detect the temperature. In fact, he is puzzled about why he thinks so obsessively about the temperature. This is sometimes explained in terms of a lack of cognitive integration: Mr. Truetemp's disposition to form beliefs in that way is simply not integrated with the rest of his cognitive dispositions. The following conditions are commonly thought to be necessary for cognitive integration. *S*'s cognitive disposition is cognitively integrated with other of *S*'s cognitive dispositions only if: 1) it is sufficiently stable; 2) it causally interacts with the other dispositions; 3) the beliefs it produces cohere with the beliefs produced by the other dispositions; 4) the whole system of *S*'s cognitive dispositions is disposed to form and maintain a coherent set of beliefs; 5) *S* has taken responsibility for the cognitive disposition; 6) *S* has a reflective perspective over her disposition. There is some disagreement on whether some of these requirements are necessary for cognitive integration or not,¹⁹ but this does not matter for our purposes, as we can modify the details of TEMPY in such a way that Tempy's detection ability satisfies whichever conditions are considered necessary for cognitive integration: from the more

¹⁸ I am grateful to an anonymous reviewer for indicating this possible reply to me.

¹⁹ See Breyer and Greco (2008) and Roberts (2012) for relevant discussion on the notion of cognitive integration.

internalist requirements (coherence, reflective perspective on one's cognitive disposition, awareness of its reliability, etc.) to the more externalist demands (stability and causal interaction requirements, implicit recognition that one is the source of one's beliefs, etc.). By way of illustration, we can stipulate that Tempy's cognitive ability is as cognitively integrated as this other agent's:

[Tempo] is fitted from birth with a highly reliable device which records the ambient temperature and Tempo grows up in a culture where it is taken for granted that one consults one's temperature-recording device in order to form beliefs about the ambient temperature (Pritchard 2010: 146).

Pritchard asks: "Wouldn't we nonetheless straightforwardly regard him as gaining knowledge via this belief-forming process? Moreover, wouldn't we regard Tempo's cognitive success as being to a significant degree creditable to his cognitive agency (...)" (*ibid.*). The same rhetorical question can be asked about Tempy.

The second possible reply is the following: it could be argued that Tempy's abilities saliently explain that she truly believes that the temperature is between 24°C and 26°C and consequently that she does not falsely believe that it is 28°C. However, they do not saliently explain that she truly believes that the temperature is 25°C. In other words, they do not saliently explain that she attains cognitive success (and *a fortiori* that she attains *safe* cognitive success).²⁰

The problem with this reply is that it only works on the assumption that the fact that Tempy manages not to believe falsely that the temperature is 28°C does *not* contribute to the explanation of why she comes to believe truly that the temperature is 25°C. However, as the case is described, Tempy's excellent testimonial justification is a defeater for the belief that the temperature is 25°C. Tempy's beliefs are based on the readings of the thermometer. Therefore, unless Tempy rejects her justification to believe that the thermometer indicates 3°C more than the real temperature, she will never come to believe the truth. In this sense, Tempy's cognitive abilities and namely the belief-forming method that she implements saliently explain why she attains *actual* cognitive success. Of course, her cognitive success is also saliently explained by the intervention of the hidden agent. But since partial contribution to cognitive success is all we need, the case is a counterexample to ALVE (to both interpretations).

²⁰ I am grateful to Chris Kelp for indicating this possible reply to me.

Nevertheless, for those who are not fully convinced yet of the fact that if Tempy had not rejected her good justification to believe that the thermometer indicates 3°C than the real temperature, she would not have believed the truth, we can strengthen the intuition that the rejection of such a good justification saliently (or partly) contributes to her actual cognitive success by slightly modifying the case. Let us stipulate that the hidden agent would intervene only if Tempy exercised her detection ability. Thus, the fact that Tempy exercises her detection ability explains not only the rejection of her justified false belief that the thermometer indicates 3°C more than the real temperature, but also the intervention of the hidden agent. In this way, Tempy's cognitive abilities are a salient factor in the total set of factors that explain her cognitive success. The other salient factor is, of course, the manipulation of the hidden agent, which is precisely what prevents Tempy's belief from being knowledge.

5 The direction of fit problem

This paper has accomplished its aim: to show that Pritchard's anti-luck virtue epistemology is insufficient for knowledge. For the sake of completeness, I will conclude with some thoughts on what is exactly what TEMPY does and does not show and on the reach of the challenge that the case poses. First, the case shows that ALVE is insufficient for knowledge, but it does not show that the whole approach of combining modal conditions (e.g., safety) with virtue-theoretic requirements is misguided. There are other 'impure' developments that can (arguably) explain the case (e.g., Kelp 2012; Turri 2011). Second, TEMPY does not settle the question neither in favor of an interpretation of the 'because of' relation in terms of manifestation of competence (Kelp 2012; Sosa 2011; Turri 2011), nor in favor of an interpretation in terms of creditability/explanatory salience (Greco 2010). Nevertheless, strong virtue-theoretic conditions in terms of creditability/explanatory salience (Greco's FULL CREDITABILITY) seem to be in a better position to explain TEMPY than strong virtue-theoretic conditions in terms of manifestation of competence (Sosa's aptness, Turri's adeptness). We will see in a moment how these strong virtue epistemologies explain the case. First, let me state the challenge that cases like TEMPY present.

Pritchard is right in thinking that what is epistemically problematic in TEMP is that Temp's beliefs exhibit the wrong direction of fit with the facts. As I argued, Pritchard's diagnosis of TEMP can be extrapolated to TEMPY, ACTOR and, plausibly, INSPIRED WORKER(S). Therefore,

the challenge posed by cases like TEMP, TEMPY, ACTOR and INSPIRED WORKER(S) is that of specifying what is required to preserve the appropriate direction of fit that known beliefs have. Importantly, not only known beliefs exhibit an appropriate direction of fit (to have an appropriate direction of fit just seems a necessary condition for knowledge). Plausibly, in certain cases of knowledge-undermining luck the direction of match between facts and luckily true beliefs is the one that befits knowledge. Consider the following case:

PICASSO

There is a temporary exhibition of Picasso's Blue Period at the local gallery. George, a Picasso enthusiast, thinks that all the works of that period are great and, as soon as he reads about the exhibition, he goes to the gallery. He enters the first exhibition room and starts looking at the paintings that are on his right. He normally proceeds clockwise but for some improbable reason he has decided to start counterclockwise. He looks at the first painting and thinks "The painting I am looking at is a great Picasso painting". Unbeknownst to George, the director of the exhibition has decided to play a joke on a pretentious local art critic who boasts about his vast knowledge of Picasso's work replacing most original Picasso paintings with masterful Picasso fakes that only an expert eye could detect. George is a Picasso enthusiast, but not an expert, and he has looked at the only original Picasso painting in the room.²¹

Many epistemologists would claim that George does not know that the painting he is looking at is a great *Picasso* painting because in most close possible worlds in which he believes the same proposition, that proposition is false. George could easily have started looking at the paintings oacademian his left, and given that he thinks that all the works of the Blue Period are great, he could easily have believed falsely that the paintings on his left are great Picasso paintings. Therefore, it is by luck that George's actual belief is true. Pritchard calls the kind of epistemic luck in PICASSO *environmental veritic luck* because features of the environment would make the target belief fail to hit the truth in close possible worlds *but not in the actual world*.

Crucially, the direction of fit of George's belief seems to be the appropriate one, in the sense that it is as appropriate as the direction fit of an ordinary known perceptual belief (possibly because in both cases there are no non-deviant causal connections between beliefs and facts). In view of this, the *direction of fit problem* is, then, the problem of how to spell out an

²¹ PICASSO is structurally equivalent to Ginet/Goldman's fake barns case (Goldman 1976).

epistemic condition that holds (1) in cases of knowledge *and* (2) in cases of knowledge-undermining environmental luck but that does not hold (3) in cases like TEMP, TEMPY, ACTOR and INSPIRED WORKER(S).

Virtue-theoretic conditions, in general, and strong virtue-theoretic conditions, in particular, seem to be the sort of epistemic requirements best suited to solve the direction of fit problem. Let us start by analyzing whether Greco's FULL CREDITABILITY is able to solve it. FULL CREDITABILITY is satisfied in cases in which an agent knows by means of reliable cognitive faculties such as perception, reasoning or memory. Arguably, it is also satisfied in cases of environmental luck like PICASSO (this is one of the two objections that Pritchard makes against Greco's account of knowledge). In addition, it does not hold in cases like TEMP or TEMPY, because Temp and Tempy's cognitive successes are (respectively) non-creditable and just partly creditable to their cognitive abilities. FULL CREDITABILITY is not satisfied in ACTOR or in INSPIRED WORKER(S) for similar reasons. Therefore, it seems that it is suitable to solve the direction of fit problem. However, as we have seen, FULL CREDITABILITY rules out cases of testimonial knowledge as cases of knowledge (this is the second objection that Pritchard makes against Greco). Therefore, Greco's account fails to solve the direction of fit problem.

Let us evaluate whether strong virtue-theoretic conditions formulated in terms of manifestation of competence can solve the problem. Kelp (2012), Sosa (2007, 2011) and Turri (2011) defend the following condition:

APTNESS: *S* knows that *p* only if *S*'s cognitive success (*S*'s believing *truly* that *p*) *manifests S's cognitive abilities or competences.*²²

²² Kelp's impure virtue-theoretic account defines knowledge as the conjunction of APTNESS and a safety condition. Turri and Sosa's pure virtue-theoretic accounts define knowledge solely in terms of APTNESS. Turri calls his epistemic condition 'adeptness' to differentiate it from Sosa's previous way of understanding the aptness condition, which is less demanding. For example, Sosa suggests APTNESS in his 2007 book (pp. 29, 80), but for the most part of the book he seems to be endorsing something like PARTIAL CREDITABILITY, which he calls 'aptness' (e.g., pp. 94-97). Nevertheless, Sosa explicitly upholds APTNESS in his 2011 book. Therefore, Sosa's actual understanding of aptness (2011) and Turri's adeptness (2011) are the same condition, that is, APTNESS. For an interesting discussion of Turri's view, see Hetherington (2012). Note, in addition, that Turri (2011), in the very same paper in which he defines knowledge solely in terms of adeptness (APTNESS) he proposes an alternative impure virtue-theoretic view (knowledge as ample belief, i.e., belief whose safety manifests competence), a view that entails APTNESS (see fn. 13).

The way these commentators think about knowledge is analogous to the way we think about dispositions. In the same way as salt manifests its solubility when stirred into water, a belief manifests cognitive competence when it hits the truth. Now, salt can only manifest its solubility under some ‘normal’ or ‘appropriate’ conditions, and namely when it is introduced in a liquid with such-and-such characteristics. In the same way, beliefs can only manifest cognitive competence when they hit the truth under some conditions. In particular, the agent’s internal conditions and the external conditions of the environment or the situation must be intuitively appropriate. Only under such conditions, they argue, beliefs can be assessed for aptness.

For example, when we look at objects under good light conditions we can form apt beliefs about those objects. Arguably, we can also form apt beliefs via testimony when our interlocutors are reliable (in ACTOR, for example, the interlocutor is an expert liar). In appropriate circumstances for testimony, our believing the relevant truths manifests our competence for selecting reliable informants (if we have it in the first place, of course). In addition, in cases of environmental luck the target beliefs are also apt because the conditions are considered appropriate and competence is manifested under such conditions. For example, in PICASSO, George looks directly at a real Picasso painting and not to a fake or a hologram, that is, the conditions are as appropriate in the actual world as the conditions in which he comes to know some other proposition by some perceptual channel in entirely normal circumstances. In sum, since APTNESS holds in cases of knowledge and in cases of environmental luck, it seems that it is a suitable condition to solve the direction of fit problem.

However, is APTNESS suitable to explain why the target beliefs in cases like TEMP or TEMPY exhibit the wrong direction of fit? As in ACTOR, the way in which defenders of APTNESS would explain why TEMP and TEMPY are not cases of knowledge would be to point out that the requirement for apt belief that the agent must be appropriately situated is not met. They would argue, for instance, that the unreliability of the thermometer makes the situation inappropriate, so that Temp and Tempy’s cognitive successes do not manifest competence in those specific environments.²³

Intuitive as it may be, appealing *just* to the inappropriateness of the circumstances as a way of explaining problematic cases should be the last resource of any theory of knowledge, as

²³ In the case of Tempy, it seems that her cognitive success *does* manifest her ability to detect the temperature (at least partially). This fact puts even more pressure on accounts based on APTNESS.

one may end up judging that the circumstances are not appropriate whenever a case looks like a counterexample to the theory. Let me illustrate the point with an example. Consider a modification of TEMP in which everything is as in the original case except for the fact that in an interval of time, T_1 , the thermometer is broken and randomly fluctuating within a given range (as in the original case); in the next interval T_2 the thermometer works perfectly well; in T_3 the thermometer is broken and fluctuates again, and so on and so forth. If the intervals are long enough (e.g., one week long) we may presume that when Temp looks at the thermometer in an interval where the thermometer is working well, he knows the temperature of the room. If this is true, advocates of APTNESS are compelled to say that the conditions are appropriate.

Now, if we start reducing the length of the intervals gradually, there is some point where we lose the intuition that Temp knows the temperature of the room. For example, if the intervals were one minute long, we would claim that Temp does not know the temperature (if the intuition does not arise, reduce the intervals, for example, to one second). The problem is that, according to the defenders of APTNESS, the circumstances during the one week long ‘good’ interval are appropriate because the thermometer works well but, at the same time, the circumstances during the one minute long ‘good’ interval are not appropriate *despite* the fact that the thermometer works well. In view of this, it seems that any explanation that they gave at that point of why the circumstances are inappropriate in the short interval case would be *ad hoc*.²⁴

To be clear, I do not claim that Kelp, Sosa or Turri cannot conclude that Temp’s true beliefs about the temperature are inapt (in TEMP or in the modified case). They can draw that conclusion. All that I claim is that such a conclusion seems *ad hoc* in the modified case and that this undermines their strategy of splitting the circumstances into two categories: appropriate and inappropriate. Consequently, I do not think that they have a satisfactory explanation of why the direction of fit of the target beliefs of these cases is the wrong one. This is even more evident once we compare their treatment of the cases to Greco’s clear and straightforward explanation that the right direction of fit between belief and fact can only be preserved if the agent’s

²⁴ For example, some commentators will surely point to the fact that a thermometer that reliably indicates the temperature in the way described does not meet the standards of reliability required for aptness. However, which length of time is enough to meet that standard? Whichever answer one gives should not assume that it is the length of time where we would attribute knowledge to Temp, as one would be begging the question. It does not seem that one can give a satisfactory answer without making such an assumption.

cognitive success is primarily creditable to her cognitive abilities (an explanation that is unsatisfactory for other reasons, as we have seen).²⁵

To conclude, although strong virtue-theoretic conditions seem the best suited to guarantee the appropriate direction of fit with the facts in all the relevant cases, FULL CREDITABILITY and APTNESS are in need of further refinement. Pritchard's anti-luck virtue epistemology does not solve the direction of fit problem either. Whether virtue epistemologists can find a virtue-theoretic condition that can solve the problem and that is not too weak to account for knowledge remains an open question.

Acknowledgments

I am grateful to Fernando Broncano Rodríguez, Chris Kelp, Joan Pagès, Manuel Pérez Otero, as well as the audience of the 2011 European Epistemology Network Meeting held at Lund University (March 2011) for helpful comments.

References

- Baumann, P. (2012). No luck with knowledge? On a dogma of epistemology. *Philosophy and Phenomenological Research*, doi:10.1111/j.1933-1592.2012.00622.x.
- BonJour, L. (1985). *The structure of empirical knowledge*. Cambridge: Harvard University Press.
- Breyer, D. & Greco, J. (2008). Cognitive integration and the ownership of belief: response to Bernecker. *Philosophy and Phenomenological Research*, 76 (1):173–184.
- Broncano-Berrocal, F. (forthcoming) Is safety in danger? *Philosophia*, doi: 10.1007/s11406-013-9467-9.

²⁵ Interestingly, the sensitivity principle (see fn. 14) explains why Temp and Tempy's beliefs exhibit the wrong direction of fit: they are unresponsive to the world, i.e., insensitive. In particular, they are insensitive because both Temp and Tempy would continue to believe the target propositions in the closest possible worlds where the target beliefs are false (worlds where the hidden agents do not intervene). However, as Sosa (1999) and Vogel (2007) show, beliefs that are known by induction are not sensitive. Therefore, either sensitivity does not solve the direction of fit problem or inductive knowledge is an exception to the direction of fit requirement. I leave the question open.

- Carter, A. J. (2011). A problem for Pritchard's anti-luck virtue epistemology. *Erkenntnis*, doi: 10.1007/s10670-011-9315x.
- Goldman, A. I. (1976). Discrimination and perceptual knowledge. *Journal of Philosophy*, 73 (November): 771-791.
- Greco, J. (2000). *Putting skeptics in their place: the nature of skeptical arguments and their role in philosophical inquiry*. Cambridge: Cambridge University Press.
- Greco, J. (2003). Knowledge as credit for true belief. In DePaul, M. & Zagzebski, L. (eds.). *Intellectual virtue: perspectives from ethics and epistemology*. Oxford: Oxford University 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.
- Haddock, A., Millar, A. & Pritchard, D. (2010). *The nature and value of knowledge: three investigations*. Oxford: Oxford University Press.
- Hetherington, S. (2012). The significance of fallibilism within Gettier's challenge: a case study. *Philosophia*, doi:10.1007/s11406-011-9340-7.
- Kelp, C. (2012). Knowledge: the safe-apt view. *Australasian Journal of Philosophy*, doi:10.1080/00048402.2012.673726.
- Lackey, J. (2007). Why we don't deserve credit for everything we know. *Synthese*, 158 (3):345-361.
- Lackey, J. (2009). Knowledge and credit. *Philosophical Studies*, 142 (1):27-42.
- Lehrer, K. (1990). *Theory of knowledge*. Boulder, CO: Westview Press.
- Pritchard, D. (2005). *Epistemic luck*. Oxford: Oxford University Press.
- Pritchard, D. (2008). Knowledge, luck and lotteries. In Hendricks, V. & Pritchard, D. (eds.). *New waves in epistemology*. New York: Palgrave Macmillan
- Pritchard, D. (2010). Cognitive ability and the extended cognition thesis. *Synthese*, 175 (Suppl. 1):133-51.
- Pritchard, D. (2012a). Anti-luck virtue epistemology. *The Journal of Philosophy*, 109 (March):247-279.

- Pritchard, D. (2012b). In defence of modest anti-luck epistemology. In Black, T. & Becker, K (eds.). *The sensitivity principle in epistemology*. Cambridge: Cambridge University Press.
- Riggs, W. D. (2002). Reliability and the value of knowledge. *Philosophy and Phenomenological Research* 64 (1):79-96.
- Riggs, W. D. (2009). Two problems of easy credit. *Synthese* 169 (1):201-216.
- Roberts, T. (2012). Taking responsibility for cognitive extension. *Philosophical Psychology*, 25 (4):491-501.
- Roush, S. (2005). *Tracking truth: knowledge, evidence, and science*. Oxford: Oxford University Press.
- Sosa, E. (1999). How to defeat opposition to Moore. *Philosophical Perspectives*, 13 (s13):137-149.
- Sosa, E. (2007). *A virtue epistemology: apt belief and reflective knowledge, volume 1*. Oxford: Oxford University Press.
- Sosa, E. (2010). How competence matters in epistemology. *Philosophical Perspectives*, 24 (1):465-475.
- Sosa, E. (2011). *Knowing full well*. Princeton, New Jersey: Princeton University Press.
- Turri, J. (2011). Manifest failure: the Gettier problem solved. *Philosophers' Imprint*, 11 (8).
- Vaesen, K. (2011). Knowledge without credit, exhibit 4: extended cognition. *Synthese* 181 (3):515-529.
- Vogel, J. (2007). Subjunctivitis. *Philosophical Studies*, 134 (1):73-88.
- Zagzebski, L. (1996). *Virtues of the mind: an inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge: Cambridge University Press.