EPISTEMIC VIRTUE, APPROPRIATENESS, 
AND THE ENVIRONMENTAL LUCK PROBLEM

Fernando Broncano-Berrocal

LOGOS - Universitat de Girona

Draft for the COGITO Seminar, November 10 2010

Abstract

This paper analyzes the notion of epistemic virtue as used by virtue reliabilists. In particular, it discusses the dispositional nature of epistemic virtues in connection with metaphysical debates. It also discusses the relativization of epistemic virtues to external and internal conditions and implements a way of understanding the not much discussed notion of appropriateness. Finally, it explains how virtue reliabilists fail to exclude environmental luck and sketches a possible line of response.

I Theoretical Unity: A Virtue of Virtue Epistemology

§ 1 Foundationalism vs. Coherentism. Along all these years one of the main concerns of Ernest Sosa has been that of pursuing the unity among the different views concerning justification, and also the necessity of putting together the theory of knowledge and the theory of justification. Back in the eighties, he introduced for the first time the notion of virtue in epistemology with the aim of ending the persistent confrontation between the foundationalist and the coherentist parties.\(^1\) As in ethics, epistemic virtues are agent-based excellences to attain some good, which in the realm of the epistemic is truth above all. So the idea is that virtuous agents are virtuous because they possess stable dispositions-innate or habits- to attain truths. How the introduction of epistemic virtues could help to soften the struggle between the foundationalist and the coherentist?

At the intuitive level, no one doubts that our ordinary beliefs are interrelated and it is from this interrelation that their justificatory status comes. For example, my belief that the philosophy library will open tomorrow morning may be supported by my belief that university libraries normally open in the morning while students are attending courses, by my belief that tomorrow is not a holiday, and so on. At the theoretical level, how one conceives the kind of interrelation that a belief must have with others in order for it to be justified?

\(^1\)See, for instance, Sosa 1980 and 1985.
may lead to very different ways of conceiving the nature of knowledge. The two main positions, foundationalism and coherentism, are described by Sosa (1991) with the help of two enlightening metaphors. Foundationalists conceive a body of knowledge as if it were a pyramid. In the same way as two floors of a pyramid are tied by a non-symmetric relation of physical support, two pieces of knowledge are tied by a non-symmetric relation of epistemic support. At the bottom of the structure lies a series of indubitable beliefs (e.g. the belief that I exist) or, alternatively, a set of fundamental epistemic principles (e.g. the principle that if F is a sensible characteristic, the belief that something is F is justified if it is based on a visual experience of something with F in conditions that are normal with respect to F, Sosa 1991: 188). Coherentists, on the other hand, favor a quite different picture of the nature of knowledge. For them, a body of knowledge is like “a raft that floats free of any anchor or tie. Repairs must be made afloat, and though no part is untouchable, we must stand on some in order to replace or repair others. Not every part can go at once” (Sosa 1991: 169). The resulting view is that the justification that a belief has comes from its being part of a coherent system of beliefs (coherence may be put in very different terms, though: logical, explanatory, probabilistic).

Sosa reconciles the two views by placing the notion of epistemic virtue right in the center of the explanation of what it takes for a belief to be justified and known. The idea is that a belief is justified only if it arises out of epistemic virtue (and only if it is justified it is known). The point is that notion of epistemic virtue is broad enough so as to embrace into the resulting conception of epistemic justification the fundamental intuitions both of foundationalism and coherentism. On the one hand, the virtue epistemologist’s notion of justification takes into account foundationalist principles insofar as many virtues are dispositions to form true beliefs on the basis of sensory experience (besides, virtue epistemology can easily defend that self-evident beliefs like the belief that I exist are the product of epistemic virtue). Coherence, on the other hand, is seen as a manifestation of epistemic virtue: attaining coherence is surely an epistemic good by itself. Even if we are not willing to consider coherence an epistemic good per se, a virtue epistemologist could argue that coherence is a manifestation of epistemic virtue insofar as coherent doxastic systems increase the reliability of belief-formation processes.

§2 Externalism vs. internalism. Externalism concerning justification rejects that the factors that justify an agent’s belief need to be (a) internal, (b) accessible. As Sosa remarks, externalists allow that contingent external features determine “(1) the causal origins of the subject’s beliefs and inference patterns, and (2) how reliably truth-conducive such causal determination of beliefs and inference patterns is generally in our actual world” (Sosa et al. 2003: 159). What matters for externalist justification is then that the justificatory factors are sufficiently truth-conducive. Virtue epistemology fittingly accommodates the intuitions of the externalist camp, since epistemic virtues are conceived (at least by some virtue epistemologists) as reliable dispositions to attain truths.

Of course, the internalist contender does not agree with this way of conceiv-
ing epistemic justification. For her, justificatory factors must be, if not accessible, at least internal. Internal factors may be reflectively accessible states or mental states (Poston 2008). Reflectively accessible states are of course mental states, but the distinction is made because there might be mental states that are not accessible reflectively and, therefore, the resulting form of internalism would reject the necessity of an accessibility requirement. In any case, the idea that underlies all forms of epistemic internalism is that a kind of internal rationality is necessary for knowledge (Sosa et al. 2003: 162). Even when beliefs are causally disconnected from the facts (remember Descartes and his Demon), they must be rationally adequate from a subjective point of view. Sosa (1991), aware of the epistemic import of the internalist concerns, distinguished between animal knowledge and reflective knowledge to cover with the former the externalist intuitions while with the latter the internalist ones.

One has *animal knowledge* about one’s environment, one’s past, and one’s own experience if one’s judgments and beliefs about these are direct responses to their impact - e.g., through perception or memory - with little or no benefit of reflection or understanding.

One has *reflective knowledge* if one’s judgment or belief manifests not only such direct response to the fact known but also understanding of its place in a wider whole that includes one’s belief and knowledge of it and how these come about (Sosa 1991: 240).

We may agree or not with Sosa’s distinction between the two types of knowledge. The point is that the virtue epistemologist, by requiring that any type of justification required for knowledge comes from epistemic virtue, makes plausible the view that both externalist and internalist types of justification are necessary for knowledge. The merit of virtue epistemology is even greater if we consider that, by doing so, one single theory unifies the theories of knowledge and justification under the same conceptual schema.

II What Are Epistemic Virtues?

§3 Responsibilities vs. reliabilists. Code (1984) introduces for the first time the distinction between responsibilist and reliabilist virtue epistemologies, a distinction that has subsequently become a standard. Why we should shift from a reliability-focused virtue epistemology to a virtue epistemology based on epistemic responsibility? Code thinks that reliable knowers are simply passive, though accurate, recorders of experience and holds accordingly that “to be intellectually virtuous is not just to have a good score in terms of cognitive endeavors that come out right” (Code 1984: 41). For Code, to be intellectually virtuous is a sort of pro-active position: it is “a matter of orientation toward the world, and toward oneself as knowledge-seeker in the world” (Code 1984: 44). The concept of epistemic responsibility accounts better than the concept of reliability for this active nature of knowers, because a knower can only be judged epistemically responsible if she is an agent. Thus, while reliabilists conceive epistemic virtues as
reliable belief-forming dispositions, responsibilists are willing to include into the list of epistemic virtues the set of aristotelian intellectual virtues: intelligence, prudence, intellectual honesty, open-mindness, etc.

§4 Sosa’s competences. Here, I will only focus on discussing virtue reliabilist theories of knowledge. What counts as epistemic virtue under the reliabilist conception? Two characteristics are fundamental in the reliabilist conception: (1) the dispositional nature of epistemic virtues; (2) their relativization to environments. Sosa (1991: 284) offers the following definition of epistemic virtue:

\[ S \text{ has an intellectual virtue } V(C, F) \text{ relative to environment } E \text{ if and only if } S \text{ has an inner nature } I \text{ such that if } \]

(i) \( S \text{ is in } E \text{ and has } I, \)
(ii) \( P \text{ is a proposition in field } F, \)
(iii) \( S \text{ is in conditions } C \text{ with respect to } P, \)
(iv) \( S \text{ believes or disbelieves } P, \text{ then } \)
(v) \( S \text{ is very likely right with regard to } P \) (Sosa 1991: 284).

Consider for instance the proposition \( P \) that this before me is white and round. Given \( P \), Sosa explains with more detail what \( S, I, E \) and \( C \) could be:

I might involve possession of good eyes and a good nervous system including a brain in good order; \( E \) might include the surface of the earth with its relevant properties, within the parameters of variation experienced by humans over the centuries, or anyhow by subject \( S \) within his or her lifetime or within a certain more recent stretch of it; \( F \) might be a field of propositions specifying the colors or shapes of an object before \( S \) up to a certain level of determination and complexity (say greenness and squareness, but not chartreuseness or chiliagonicity); and \( C \) might be the conditions of \( S \)’s seeing such an object in good light at arm’s length and without obstructions (Sosa 1991: 139).

Sosa thinks that epistemic virtues, so conceived, are dispositions of an agent to perform well, what he calls competences. More recently, he explains how epistemic virtues share the structure of dispositions, which according to Sosa involves three elements (Sosa 2010). Consider as example of disposition the flammability of a match and as example of competence the archery competence. Thus, dispositions/competences have:

a. A constitutional element, the “seat” of the disposition/competence (e.g. the powdery head/the morphological basis of the archer’s skill).

\[ ^2 \text{As Sosa explains, some epistemic virtues may not be relativized to environments (virtues that can be exercised in the armchair).} \]
b. A condition intrinsic to the host entity (e.g. being dry/being awake and sober).

c. A situation within certain limits (e.g. being in oxygen/enough light, normal wings).

Accordingly, Sosa distinguishes three levels of competence:

Level-1: (a) The constitutional competence or skill.

Level-2: (a + b) The inner competence, “which combines not only a relevant constitution but also an appropriate condition” (Sosa 2010, 465).

Level-3: (a + b + c) The complete or full competence.

Knowledge requires full competence.

§ 5 Greco’s cognitive abilities. Instead of using the notion of competence, Greco (2010) conceives epistemic virtues as cognitive abilities. Cognitive abilities resemble Sosa’s full competences nevertheless. As competences, cognitive abilities are dispositional entities. They are also relativized to environments and conditions. Along Sosa’s lines, Greco thinks that environments are sets of relatively stable circumstances (e.g. as before, the surface of the earth with its relevant properties and within the parameters of variation experienced by humans over the centuries and the agent in question). By ‘conditions’ he means sets of shifting circumstances within an environment (e.g. the conditions of S’s seeing an object in good light at arm’s length and without obstructions, but also the agent’s internal conditions, like being awake and sober). Greco, therefore, does not distinguish between conditions that are external and internal to the agent, and thus he cannot draw a distinction analogous to Sosa’s between inner and full competence. For Greco, the relevant conditions must only be flexible enough so as to shift within a much bigger set of stable conditions, i.e. the environment. Accordingly, his proposed structure of abilities is as follows (R for ‘result’, C for ‘conditions’, E for environment):

\[
S \text{ has an ability } A \ (R/C) \text{ relative to environment } E = \text{ Across the set of relevantly close worlds } W \text{ where } S \text{ is in } C \text{ and in } E, S \text{ has a high rate of success in achieving } R \text{ (Greco 2010, 77).}
\]

Knowledge requires cognitive ability.

III Epistemic Virtues and the Conditional Analysis of Dispositions

§ 6 The conditional analysis of dispositions. Imagine William Tell, the competent archer, aiming at the target, an apple on the head of his son. He is at the village’s central square, located on planet Earth, where gravity is the usual one. Let this be E, the environment. He is sober, awake, in good shape. Let this be IC, his internal condition. There is no wind, light conditions are good, the distance is
reasonable. Let this be EC, the external conditions. The stimulus for delivering
the arrow could be his intention to shoot and the neural activity that leads to
the relevant body movements. Let this be S. In addition, let R be the result
that he hits the target. Given these specifications it seems that the following
conditional holds:

William Tell has the dispositional property D to respond to stimu-
lus S with response R just in case that were S to obtain, R would
materialize.

This is the standard conditional analysis of dispositions, which Sosa and Greco
(henceforth S&G) seem to be assuming in their respective notions of competence
and cognitive ability. We have seen, though, that they relativize to environments
and conditions. Thus, they seem to assume the following analysis:

William Tell has the dispositional property D to respond to stimulus
S with response R just in case that the archer were in IC and S were
to obtain in E and EC, R would materialize.

Let us test their improved conditional analysis of dispositions/epistemic virtues
with a series of counterexamples to the standard conditional analysis of dispo-
sitions. We will see that in all cases they would appeal in their replies to the
normative notion of appropriateness.

§7 Masked dispositions. Consider, for instance, cases of masked dispositions,
that are cases in which X has a dispositional property D to respond to stimulus
S with response R, S obtains but something prevents R from manifesting. For
example, poison is disposed to kill if ingested, but it is not true that if it were
ingested it would kill if, for instance, an antidote has been administered. The
disposition of the poison, thus, is masked. Imagine an analogous case in which a
competent archer shoots arrows with magnetic tips to a reversed magnetic tar-
get. Despite the skilful shots, arrows never reach their aim (Gundersen 2010).
The competence/cognitive ability of the archer would be thus masked. How
would S&G deal with the case? S&G’s would possibly answer simply by noting
that EC, the external conditions of the archer, are not appropriate.

§8 Finkish dispositions. Let us consider now cases of finkish dispositions. These
are cases in which X possesses a dispositional property D to respond to stimulus
S with response R, S obtains but something causes X to lose D before D is
triggered so that R is not manifested. Imagine a live wire with the disposition
to transmit current when is put in contact with a conductor, so that if it were put
in contact with a conductor, electrical current would flow from the wire to the
conductor. However, an ‘electro-fink’ shuts off current to the wire whenever it
touches a conductor. The wire, thus, loses its disposition. Consider an analogous
case in which a competent archer has his neural network wired up in such a way
that he dies immediately if his brain produces the neural activity required to
form the intention to shoot. If he were to form the intention to shoot at a target,
his archery competence would disappear because we would die and hence the arrow would never hit the target.

What would be S&G’s reaction to a case like this? Certainly, the competence cannot be manifested in any close possible world, because were the archer to shoot, he would lose the competence (and his life) Obviously, we cannot regard such competence as virtuous. Where is the flaw located? The problem must be that either IC or EC, or perhaps both, are flawed. It all depends on whether we think that the fink (i.e. the fact that the archer’s brain is wired up in that tragic way) is constitutive or not of the archer’s inner competence. If it is, the internal condition of the archer is simply and intuitively inappropriate. If it is not part of the constitutional base of the competence, S&G would have to argue that being so wired up is a terribly inappropriate external condition for successfully displaying a *prima facie* healthy inner competence. The problem with this reply (and also with the reply to masked dispositions cases) is that there does not seem to be a principled standard according to which external conditions can be regarded as ‘appropriate’ or ‘inappropriate’ (and the same for internal conditions).

§9 Mimicking dispositions. Cases of mimicking dispositions are cases in which X does not have a dispositional property D to respond to stimulus S with response R, S obtains and something causes R to manifest. Imagine, for instance, an explosive stuck to a rock in such a way that if it were struck, it would shatter. The rock manifests fragility (i.e. it ‘mimics’ fragility), even if it is solid. As Gundersen (2010) usefully notes, cases of mimicking like this are structurally analogous to a case in which an archer shoots at random arrows in all directions but a strong magnet at the tip of the arrows makes them hit the metallic target. In that situation, even if the archer were competent, the hit on the target would not be a manifestation of the archer’s competence, but of the magnet. How would S&G’s treat a case like this?

If the subject were not competent, S&G could simply say that she lacks the inner nature or skill necessary to achieve full archery competence. Alternatively, they could also argue that the external conditions are not appropriate, the same answer that they would give if the archer were competent. In which sense would the conditions be inappropriate? Well, unlike masked disposition cases, here the conditions are ‘too’ friendly, in the sense that no matter how competent one is, the arrows nevertheless hit the target. But many circumstances that help competent agents in their performances are intuitively regarded as ‘appropriate’ (one just has to consider the degree of trust in the speaker that is involved in testimony cases). Which is the threshold between friendly (appropriate) circumstances and ‘too’ friendly (inappropriate) circumstances? Is there a matter of fact about it?

The proponent of the conditional analysis of dispositions seems to be compelled to specify the normal or appropriate conditions in which the relevant subjunctive conditional holds. This is perhaps the reason why S&G, when confronted to all these cases, are forced to appeal to the notion of appropriateness. Appealing to appropriateness is, however, no defect. Appropriateness is a nor-
mative concept that enable us to distinguish those belief-forming dispositions that count as epistemic virtues from those that do not. From an ontological point of view, however, I do no think that it is right to incorporate normative notions in an analysis of dispositions.

§10 A metaphysical digression: appropriateness cuts no ontological ice. There is a deep metaphysical sense according to which the question of whether circumstances are appropriate or inappropriate makes no sense. Consider the following remarks that John Heil (ms) makes concerning a canonical example of disposition, the solubility of salt:

Look more closely at what happens when you stir salt into a glass of water. Certain chemical features of the salt interact with certain chemical features of the water. This interaction is, or appears to be, continuous, not sequential; it is, or appears to be, symmetrical. Both the salt and the water work in concert to yield a certain result: the salt’s being dissolved in the water. One way to understand such cases is to imagine that salt and water possess reciprocal powers or dispositions. The salt’s dissolving is a mutual manifestation of these dispositions. The result is something with new powers, new dispositions capable of further mutual manifestations with further reciprocal disposition partners.

Heil criticizes the conditional analysis for offering a wrong picture of dispositions. According to Heil (2005; ms), the manifestation of a disposition is not triggered by stimuli. He rather thinks that the manifestation of a disposition is always the manifestation of reciprocal disposition partners. Moreover, he thinks that one and the same disposition can manifest itself differently with different reciprocal dispositional partners, as he explains in the following example:

A ball’s sphericity endows it with a power to roll. But it is also, in virtue of being spherical, that the ball has the power to make a concave, circular impression in the carpet, and the power to reflect light so as to look spherical (Heil ms).

Thus, dispositions are not modeled linearly (i.e. a stimulus followed by a manifestation), they are rather intertwined in networks, constantly evolving, massively cooperating and giving rise to a huge number of mutual manifestations. How can we live in such world?

We depend for our existence on stable structures that we inhabit, move about in and on, and deploy. We count on our environment maintaining a high level of stability. Stability requires massive cooperation, the mutual manifestation of countless reciprocal powers to hold things together, to preserve the status quo (Heil ms).

If Heil’s model is correct, then talk about mimicking, finkishness and masking is just perspectival talk. And so it is talk about appropriate and inappropriate
circumstances. Thus, in a strict ontological sense the fact that one forms a belief about an empirical proposition is no more creditable to one’s epistemic virtues than to thousands of dispositions in the world. Fortunately for the virtue epistemologist the notion of epistemic virtue is in good part normative. As said, not every disposition counts as virtuous: one might be a skilled archer but would not count as competent archer unless one is in good internal condition and appropriately situated. Where is the normativity located then? Surely not at the constitutional level (level 1) (that level would be regarded by Heil as the real disposition). Normativity arises at levels 2 (inner competence) and 3 (full competence). Ultimately, it is the notion of appropriateness that enables us to determine whether one has inner and full competence. However, not much is said about how appropriate conditions are to be distinguished from inappropriate ones. Since it is an important topic and not much is said about it, in the next section I will explain how the notion of appropriateness can be made more informative if we give a functional explanation of it. Such explanation will be analogous to the one that Eng (1989) gives to distinguish between deviant and non-deviant causal chains.

IV Appropriateness

§11 An application of Eng’s account of non-deviance to virtue epistemology.

Consider the following case of epistemic masking:

Flop, forms a belief as to whether the light is switched on next door by implementing some acceptable epistemic method—observing the light switch, say. However, a high-tech remote-control mechanism registers Flop’s belief and, unbeknownst to Flop, adjusts the lighting next door such as to contradict Flop’s belief. If Flop comes to believe the light is on, the remote-control mechanism will register this and immediately switch the light off. So when Flop comes to believe p, not p will actually be the case. And if Flop comes to believe not p, it will be the case that p (Gundersen 2010, 359).

As we have seen, about cases like this S&G would say (correctly) that the external conditions are inappropriate for using the acceptable epistemic method. But in which sense are the agent’s circumstances inappropriate? I think that answer here should be straightforward an simple: the circumstances or the external conditions are inappropriate for forming beliefs concerning a field F of propositions by means of cognitive competence C because in those circumstances C would likely form more false beliefs than true ones. As this, however, more than a condition on appropriateness, is a rule of thumb because, as we will see, there are circumstances like those of Goldman’s Fake Barn Case in which beliefs are likely to be false but, intuitively, they are appropriate.

To make the notion of appropriateness more informative, here I would like to focus on cases in which beliefs would likely be false but still the circumstances

3The same idea applies to faulty internal conditions.
are intuitively inappropriate. Let us give some examples. Case 1: Suppose that S sees a green object and comes to believe that the object is green and the causal chain is the ordinary one. Case 2: Suppose that S sees a yellow object and comes to believe that there is a yellow object and the causal chain is the following: “the chain starts with a yellow object, goes through a blue filter, produces a green-object percept, and because of the effect the filter has on the man’s cortical connections, ends with the belief that the object is yellow” (Enç 1989: 253).

Enç argues that what is wrong with cases like Case 2 is that the system (or in our case the competence involved) produces the output in a way it is not supposed to. By this he does not mean that some explanatory relation fails to hold between the cause (the yellow object) and the output (the yellow-belief). What does determine what a system (or a competence) is supposed to do? Enç’s answer is that its function determines it, and the way it is supposed to do is determined by how it has been designed to work. How may a system fail to do what it is its function to do? According to him, in two ways: by malfunctioning or by the environment not cooperating (remember that Sosa’s structure of competences distinguishes between these two -internal and external- levels). In sum, here is the basic idea of how Enç conceives a normal or appropriate causal chain:

The account I propose to develop starts with the basic intuition that a key to understanding what counts as “normal” causal chain is the concept of executing a function - the concept of a system’s doing something that it is supposed to do in the way it is supposed to do it (Enç 1989: 239).

Enç admits that it is easier to apply functional concepts to judge the performances of designed artifacts than of natural systems. To extend his account to cognitive systems of higher organisms he appeals to the theory of natural selection. I will assume that the theory of natural selection can offer some functional explanation of how it is that cognitive competences produce certain outputs (true beliefs) rather than others (false beliefs) in a way they are supposed to. This is a big assumption crying out for elaboration, but the as Enç points out:

If the basic intuition is right, though, then the correct account of non-deviance [i.e. appropriateness] should focus on an explanatory context in which we explain not why this output is produced under these conditions, but rather why the system is capable of producing this type of output under these conditions in this particular way (Enç 1989: 239).

That is, the order of the explanation is reversed. S&G’s focus is on explaining why a true belief is produced by an epistemic virtue under certain conditions that are intuitively regarded as appropriate or inappropriate. I propose to put the focus on explaining why an epistemic virtue is capable of delivering a true belief under certain conditions in the particular way that is supposed to deliver
The shift introduces an important normative element that will subsequently help to distinguish appropriate from inappropriate conditions.

Which could be the measure of appropriateness then? Enç draws an analogy with a biological case to show how can be distinguished normal from deviant causal chains. He explains that some subtropical orchids emit a particular fragrance that attracts a species of bees so as to cross-pollinate. This is the particular normal way in which cross-pollination is supposed to take place. This is so because in an evolutionary process emitting fragrance is and has been beneficial for orchids. By contrast, a deviant way in which pollination could take place would be for instance that someone finds the fragrance of the flower offensive to the extent that he cuts the flower and throws it away in such a way the wind carries the pollen to other orchids.

Now, to tell deviant from non-deviant causal chains Enç proposes to answer sorting questions of the following type: Would the organism (the orchid) doing X, an event type, (producing the fragrance) even if tokens of X did not result in certain consequences to the organism (achieving cross-pollination by attracting the bees)? If the answer is no, then the causal chain is normal. Why has the answer to be negative?

Suppose we look at all those possible worlds which are exactly like ours except that the fragrance fails to attract the bees. The chances are that in some of these worlds, the scent is retained because the feature is genetically linked to some other feature that is functional; in others it is retained either by chance, or because, in terms of ecological economy, it is too “expensive” for the plant to eliminate the mechanism responsible for the production of the fragrance. However, knowing certain facts about evolution, we know that there will be some possible worlds in the set of nearest possible worlds in which the species goes extinct and some worlds in which the species survives (e.g. by self-pollination) but the fragrance atrophies. The existence of the latter type of worlds yields the negative answer to the sorting question (Enç 1989: 245).

Now consider this other question: Would the flower be producing this particular fragrance even if the fragrance did not have the effect of achieving cross-pollination by offending me? The answer is obviously yes, so the causal chain that leads to cross-pollination is deviant. Let us reconsider the cases of perception mentioned at the beginning of the section. Suppose that S sees a green object and comes to believe that the object is green. Enç asks us to suppose that there is some event type, E, between the production of the percept and the production of the belief that is describable at the level of specificity appropriate to the Functional Organization of the cognitive system (p.253). Would the chain of events leading to E from the production of the percept be the same even if E did not result in the belief that the object is green? The answer is no, so the causal chain is normal. Why?

For, one presumes that the evolution of the cognitive system has involved a process of selection, and that this process has favored
those features of the cognitive system which make the system capable of bringing about those events between the percept and E. These features have been selected partly because they enable the system to bring about the kind of events that are causally sufficient for producing a belief that corresponds to the percept (Enç 1989: 253-254).

Reconsider now the other case in which S forms the belief that there is a yellow object on the basis of a causal chain that starts with a yellow object, goes through a blue filter, produces a green-object percept, and because of the effect the filter has on the S’s cortical connections, ends with the belief that the object is yellow. Would the chain of events leading to E from the production of the percept be the same even if E did not result in the belief that the object is green? The answer would be yes because “the capacity to produce a green-object percept under these conditions is independent of whether or not this percept causes the belief that the object is yellow” (p. 253).

What conditions are appropriate for the reliable display of an epistemic virtue? Conditions that pass the sorting test, that is, conditions under which the causal chains that lead to the target true beliefs are normal. Of course, even when conditions are appropriate in this sense one may form a false belief, but that is compatible with the reliability requirement that the notion of epistemic virtue incorporates.

In conclusion, I have proposed two criteria for judging whether the circumstances are appropriate or not. Given field of propositions F and some epistemic virtue EV: 1) a condition C is appropriate only if one would likely form true beliefs about F in C (remember this is more a rule of thumb than a condition). 2) If one would likely form true beliefs of propositions of F in C, then C is appropriate if the causal chains that lead to the target beliefs are normal in the way that EV is supposed to.

Is the proposed interpretation of appropriateness complete? Consider Pritchard’s Temp case:

Temp forms his beliefs about the temperature in the room by consulting a thermometer. His beliefs, so formed, are highly reliable, in that any belief he forms on this basis will always be correct. Moreover, he has no reason for thinking that there is anything amiss with this thermometer. But the thermometer is in fact broken, and is fluctuating randomly within a given range. Unbeknownst to Temp, there is an agent hidden in the room who is in control of the thermostat whose job it is to ensure that every time Temp consults the thermometer the ‘reading’ on the thermometer corresponds to the temperature in the room (Pritchard forthcoming).

Temp would likely form true beliefs in this scenario and his beliefs are not the result of a deviant causal chain. The problem with Temp’s beliefs is that there is no causal connection between them and the temperature of the room. Why? Because Temp’s beliefs are formed on the basis of seeing the readings of
the broken thermometer, but the temperature of the room does not cause the
mercury line of the thermometer to grow or decrease (because, say, it is blocked)
and therefore there is no causal connection between Temp’s beliefs formed on
the basis of his reading of the thermometer and the temperature of the room.

We can easily regard these circumstances as inappropriate if we apply the test
of the sorting questions. Would the chain of events that result in a thermometer
reading be the same even if the chain did not result in the correct reading
(and as result the relevant belief)? In a normal case where a thermometer
works properly the chain of events resulting in a thermometer reading is the
temperature making the mercury bar grow or decrease so as to reach certain
value in the scale. Would this chain of events that result in a thermometer
reading be the same even if the chain did not result in a correct reading? No. A
thermometer is designed in such a way that it would not give rise to that chain
unless the chain resulted in the correct reading. If the reading were not correct
would be, for example, because of the intervention of someone manipulating the
bar by different physical means.

In the Temp case the chain of events resulting in the thermometer reading
is some random physical process (the thermometer is fluctuating randomly).
Would this chain of events that result in a thermometer reading be the same even
if the chain did not resulted in a correct reading? Yes. The very same random
physical process could have produced any other result. Forming beliefs about
the temperature of the environment by looking at the readings of a thermometer
is a method of knowing that counts as virtuous only in circumstances in which
thermometers would pass the sorting test. However, Temp’s circumstances do
not pass it.

V The Environmental Luck Problem

S&G think that knowledge is true belief because of epistemic virtue. This means
that S knows that p if and only if the fact that S’s belief hits the truth is com-
pletely creditable to S’s cognitive abilities/competences. In Sosa’s terms, a belief
that fulfills that requirement is an *apt* belief. S&G’s definition of knowledge rules
out standard Gettier-style cases, cases of reliable strange and fleeting processes,
and so on. However, they are in trouble when it comes to exclude environmental
luck, i.e. luck in which the accidental factor that makes the belief luckily true
is merely potential. A classic case of environmental luck is the well-known Fake
Barns case:

Henry is driving in the countryside with his son. For the boy’s
edification Henry identifies various objects on the landscape as they
come into view. “That’s a cow,” says Henry, “That’s a tractor,”
“That’s a silo,” “That’s a barn,” etc. Henry has no doubt about the
identity of these objects; in particular, he has no doubt that the last
mentioned object is a barn, which indeed it is. Each of the identified
objects has features characteristic of its type. Moreover, each object
is fully in view, Henry has excellent eyesight, and he has enough
time to look at them reasonably carefully, since there is little tracks
to distract him.

Suppose we are told that, unknown to Henry, the district he has just
entered is full of papier-mâché facsimiles of barns. These facsimiles
look from the road exactly like barns, but are really just facades,
without back walls or interiors, quite incapable of being used as
barns. They are so cleverly constructed that travelers invariably
mistake them for barns. Having just entered the district, Henry has
not encountered any facsimiles; the object he sees is a genuine barn.
But if the object on that site were a facsimile, Henry would mistake
it for a barn. Given this new information, we would be strongly
inclined to withdraw the claim that Henry knows the object is a
barn (Goldman 1976: 772-773).

The problem for S&G is that the fact that Henry’s belief hits the truth is
completely creditable to the exercise of his cognitive abilities. That is, Henry’s
true belief is because of epistemic virtue. After all, he does not make any
cognitive mistake. However, the intuition that everyone shares about this case
is that the modal proximity of error makes hard to attribute knowledge to Henry.

One possible response to the case is to argue that Henry lacks the relevant
discriminatory capacity. Greco (2010), for instance, defends that Henry lacks
the ability to discriminate between barns and non-barns in that particular en-
vironment, although in normal environments he has it. This is however highly
controversial, as Greco recognizes, because it implies that Henry loses his nor-
mal capacity when he enters the Fake Barns district. Moreover, the response
involves a very fine-grained relativization to environments but, as Pritchard ar-
gues, we tend to understand what constitutes a suitable environment in a very
coarse-grained fashion: “in crediting you with the ability to play the piano, for
example, we are not thereby supposing that you can play the piano underwater”
(Pritchard et al 2010: 37). To show how intuitive Greco’s response is, Pritchard
ask us to imagine a case in which a pianist is happily playing in an empty cham-
er which in most nearby possible worlds is full of water. Does she lack the
capacity to play piano? Greco would have to give an affirmative answer, since
she could very easily be underwater.

On the other hand, for Sosa Henry’s conditions are appropriate. What is
inappropriate according to Sosa is to claim that Henry loses his complete com-
petence on entering the Fake Barn district. He could easily lose it, but he does
not de facto. Even if a competence is fragile, this does not affect the aptness
of the resulting beliefs, and that is the reason why Sosa is happy to credit
Henry with knowledge, but only of the animal kind. What Henry really lacks
is reflective knowledge. However, it is highly unintuitive to credit Henry with
knowledge. But even if we grant that Henry has knowledge in the animal way,
Sosa’s response does not work.

Kornblith (2009) has recently explained why. In his 2007 book Sosa conceives
reflective knowledge as apt belief aptly noted. According to Kornblith, reflective
knowledge so conceived is animal knowledge twice over: animal knowledge of a fact and animal knowledge that one has animal knowledge of such fact. This is an important shift in Sosa’s conception of reflective knowledge with respect to that of 1991, which involved a reflective stance or perspective over one’s state of knowledge acquired by the exercise of one’s first-order capacities. Therefore, in this new conception the internalist intuitions are put aside.

Given these clarifications, Kornblith argues that it is incorrect to claim that in cases like Fake Barns subjects have animal knowledge but lack reflective knowledge. Why? Take for instance the Fake Barns case: Henry allegedly has animal knowledge that that object is a barn; Henry also believes that he has knowledge that that object is a barn. Kornblith ask us to suppose that the subject (the old experienced Henry) has a competence to judge whether he has animal knowledge that an object is a barn. To add more force to Kornblith’s claim imagine that Henry is a kind of skeptic about barns, so every time he sees one he wonders whether he knows that the object is barn, to that extent that he has developed a very accurate capacity to deliver judgments about the state of his knowledge about barns. This being the case, it seems that Henry has reflective knowledge.

Now, here comes the key point in Kornblith’s argumentation against Sosa: the condition that Sosa requires for a meta-competence like Henry’s in order for it to yield knowledge is that it must not be excessively liable to yield a falsehood when exercised in its appropriate conditions. According to Sosa, that requirement would not be met due to the presence of fake barns. Kornblith immediately argues that if the presence of an error possibility is what makes the meta-judgement easily false then it should equally make easily false the first-order belief. Therefore, either Henry has both animal and reflective knowledge or neither of them. The argument does not only prove that cases like Fake Barns are counterexamples to Sosa’s theory, but also undermines his notion of reflective knowledge.

At this point, the virtue epistemologist has two possible options. The first would be to incorporate a modal condition on knowledge. This strategy is adopted, for instance, by Pritchard (2010; forthcoming), who gives a definition of knowledge that mixes the safety principle with a virtue-theoretic condition. The second possible line of response is the one suggested next.

§13 Virtue epistemology and metacognition. When we learn we use strategies to self-regulate and improve our learning. When we remember we use capabilities and strategies to aid our memory. When we play chess, when we reason diagrammatically (by means of visual representations) or when we design, we use constructive perception, the capacity to infer conceptual ideas from perception (Suwa & Tversky 2003). All these cognitive activities involve metacognitive processes.

In 1990 Nelson and Narens published a seminal paper that prompted much research on metacognitive processes. In the paper they proposed a model according to which metacognition is a capacity to evaluate and control one’s cognitive processes. They based their model on the assumption that the brain, insofar
as it is a regulator of survival, proceeds by forming a model of its environment. Then, they proposed the idea that any metacognitive system contains two interrelated levels: a meta-level and the object level. Between the object level and the meta-level flows information. The meta-level has a model of the situation and of the object-level and the information that flows from it to the object-level exerts controls by either changing the state of the object-level process or by changing the object-level itself. Possible actions that this control could produce are: (1) to initiate an action; (2) to continue and action; (3) to terminate an action. On the other hand, the information that flows from the object-level to the meta-level monitors the latter by changing the state of the meta-level's model of the situation (and this includes also no change in the state of the meta-level). They propose a useful analogy to understand the model better. A metacognitive system, they say, is like a telephone handset. The meta-level is like speaking into the telephone handset in such a way that it modifies whatever or whoever is on the other end of the line (i.e. the object-level). The object-level is like listening to the the telephone handset in the sense that one gets information from whatever or whoever is on the other end of the line (i.e. the meta-level).

Other theorists describe metacognition as ‘thinking about thinking’, that is, a metarepresentational capacity to self-ascribe mental states. As Arango-Muñoz (2011: 72) explains, “such self-ascription depends on a more general mindreading capacity consisting in an inferential capacity to attribute mental states in order to interpret and rationalize other people’s behavior”. By contrast, the model initiated by Nelson and Narens does not require the possession of psychological concepts or of a mindreading capacity, and what is more important, it does not require reflexive thinking, but ‘metacognitive’ feelings. Arango-Muñoz illustrate the point with this example:

A subject confronted with a cognitive problem such as a multiplication task has to select a cognitive strategy to solve it. She has to choose either to remember the answer (in the case of a familiar problem) or to calculate. This decision does not seem to be based on a reflexive process considering all the possible alternatives and the pros and cons of each one (a maximizing and metarepresentational process), but on a feeling that affords or makes salient one of the possible strategies (Arango-Muñoz 2011: 76).

Koriat (2000) and Arango-Muñoz (2011) suggest that the two ways of conceiving metacognition are not incompatible, but they are describing two different levels of complexity: (1) a high-level, which is self-interpretative, that delivers judgments and rationalizes one’s behavior and mental states, and thus involves conceptual content and an inferential structure; (2) a low-level that controls one’s behavior and mental activities by delivering non-reflexive epistemic feelings.

I think that a possible strategy for dealing with the environmental luck problem is to introduce in our theory of knowledge a condition that requires metacognitive control of the reliability of the agent’s cognitive processeses. A manifestation of such control would be the capacity to selectively apply one’s
belief-forming process. To exclude environmental luck would be, then, a matter of having control over the conditions in the sense that the agent must be able to refrain from using her belief-forming capacities when they are likely to yield to false beliefs.

Does this imply some kind of reflective knowledge? Note that the metacognitive control may be caused by different things. Sometimes it is a piece of background knowledge that prevents the subject from forming false beliefs. This is why one is tempted to say that environmental luck can be excluded only with reflective knowledge: having background knowledge of the conditions of the environment enables the agent to deliver justified judgments of the epistemic status of her beliefs about the environment and of the reliability of her belief-forming processes. However, sometimes it is an epistemic feeling that tells the subject to stop forming beliefs in the way he is doing. This may be all that is needed to exclude environmental luck. It is hard to put the label ‘reflective knowledge’ to these epistemic feelings, since they are clearly non-reflective. I think, therefore, that we should keep the term ‘reflective knowledge’ to refer to the kind of internalist knowledge we acquire by judging the epistemic status of our mental states. Anyhow, it seems that metacognition can offer a line of response to the environmental luck problem, a line of response that does not involve modal principles.

References

URL:http://web.me.com/jonathandjacobs/Putting_Powers_to_Work/Papers_files/Heil_Causing.pdf