

Colby



Colby Quarterly

Volume 27
Issue 2 *June*

Article 6

June 1991

Externalism and Epistemic Responsibility

Jennifer Armstrong

Follow this and additional works at: <https://digitalcommons.colby.edu/cq>

Recommended Citation

Colby Quarterly, Volume 27, no.2, June 1991, p.88-99

This Article is brought to you for free and open access by Digital Commons @ Colby. It has been accepted for inclusion in Colby Quarterly by an authorized editor of Digital Commons @ Colby.

Externalism and Epistemic Responsibility

by JENNIFER ARMSTRONG

ACCORDING to Lawrence Bonjour, an adequate theory of empirical knowledge must provide two things: (1) an account of epistemic justification, and (2) a metajustification for the proposed account showing the proposed standards to be truth-conducive (Bonjour, p. 9). His attempt to use coherence in the long run as the latter exposes what I take to be one of the central problems in contemporary epistemology: “hooking up” justified belief with knowledge. In thinking about this problem and the various attempts Bonjour and others have made to overcome it, it occurs to me that this difficulty is remarkably like a parallel problem in utilitarian normative ethics. This paper is an attempt to draw the parallel convincingly and to suggest that similar responses to those made by the utilitarian ethicist are available to the epistemologist.

The Ethical Line

Total Utility Utilitarianism (TUU): Act A is morally right if A produces as much or more total utility than any alternative to A.

UTILITY here is a term of generic intrinsic value. Exactly what thing or things are intrinsically valuable is left unspecified, but a state of affairs has positive utility to the degree that it includes what is intrinsically good and negative utility to the degree that it includes what is intrinsically bad. One can properly talk about what is produced by an act A in terms of states of affairs (complexes of the consequences of A and the consequences of acts performed prior to and concurrent with A) or more directly in terms of the total utility of those states of affairs. The above definition is expressed in the latter terms but could as acceptably be recast in the former.

There are two assumptions which must be made by the utilitarian in using utility as a measure of the intrinsic value of states of affairs: (1) that there are objective, measurable (in utils, for example) amounts of positive and negative utility in any given state of affairs, and (2) that the units of positive and negative utility are commensurable. To determine the total utility of a state of affairs, S, its positive and negative utils are summed; this sum (total utility) can then be compared to the total utility of other states of affairs. According to TUU, that act is right which produces a state of affairs the total utility of which is lower than that of states of affairs produced by no alternative act to A. Any act which produces less total utility than any of its alternatives is simply wrong.

This theory about what acts are right has paved the way for the serious objection that, if TUU is true, agents are obligated to act according to something which they in principle cannot know: the future consequences of acts. This I will call the impracticability objection. Take, for example, the following case:

On their daily walks, baby Hitler's nurse regularly pushes little Adolf's perambulator across an old wooden bridge which stretches over a particularly deep and swift-moving section of stream. One day, however, Frau Schmidt stumbles and loses her grip on the carriage, sending it and its occupant careening over the side. Horrified, she scrambles to her feet and over the side herself, rescuing a now very testy baby Hitler.

According to TUU, because the eventual result of Frau Schmidt's selfless rescue is, in fact, that the Holocaust takes place (a state of affairs with overwhelmingly negative total utility), it is simply wrong. Hitler's nurse should have left the baby to drown—or perhaps dropped him in herself—regardless of the fact that she had no way (apart from an attack of sudden omniscience) of knowing what he would eventually become.

Such a result seems plainly ridiculous and has led some to reject the TUU account of moral rightness in favor of correlating value with expected total utility rather than with actual total utility. The expected total utility of act A is the probability that a certain state of affairs S will occur (should A be performed) multiplied by the total utility of S. This new interpretation of moral rightness takes into account the intuition most of us have that Frau Schmidt did the right thing in saving the baby Adolf, since it can be assumed that the probability of his becoming the powerful madman he in fact did become was minute.

As inviting as such a move from actual to expected utility may seem, however, its resources for dispelling such as the baby Hitler problem are at best illusory. The probability of S on A¹ can be understood to mean one of two things: (1) the actual probability that S will occur, given A, or (2) the subjective probability that S will occur, given A.

First, it might be assumed that generalizations from all past instances of acts relevantly similar to A will produce something like an actual probability of S on A. But although it may be possible in principle for an agent to know this actual probability, it is all but impossible in practice. To know the actual probability of S on A, the agent would have to be in possession of a general law derived (not necessarily by the agent him/herself) with respect to all past relevantly similar instances of A-like acts and S-like states of affairs. Thus this first account of expected utility utilitarianism (EUU) falls victim to a version of the same objection its supporters leveled at TUU: agents are obliged to act with respect to information they may well not be (and probably aren't) in possession of.

One can see this objection clearly with respect to our above example. A right act, for Frau Schmidt, is that act among her alternatives which is such that its expected total utility is as great as or greater than that of any of its alternatives. One might assume that the actual probability of the baby's growing into a tyrant

1. Here and hereafter, by the probability of S on A I mean the probability that S will occur, given A, and holding, as much as possible, all other things constant.

should the nurse rescue him is extremely low, but that is no more than an assumption. It may in fact be the case that the past acts relevantly similar to Frau Schmidt's hauling little Adolf from the drink are not all those acts of saving babies from certain death, nor even all those acts of saving babies from death by drowning, but all those acts of saving babies with relevantly similar parents, in particular socio-economic spheres, in particular sociological and political settings . . . (etc!) from death by drowning. And if Frau Schmidt did have some way of determining which past acts were relevantly similar to A, it may well be the case that the actual probability of S on A is quite high.

On the other hand, if the probability relevant to determining expected total utility is what the agent thinks is the expected probability of S on A (i.e., the subjective probability of S on A), then the reasonableness of an expected utility interpretation of moral rightness becomes even more suspect. Suppose, for instance, that our friend Frau Schmidt were a follower of a particular theological circle which taught that all babies are born sinless and that, the longer they live, the more sin they accumulate and the closer to eternal damnation they come. Assuming also that she saw eternal damnation as the worst possible state of affairs for any individual to be a part of and that only belief in Jesus (which babies aren't thought to have) can prevent such from happening, she might well—according to this interpretation of expected utility—be right not only in allowing baby Hitler to drown but also in drowning every baby she comes across. This result seems at least as intuitively problematic as was that of TUU.

It might be objected that this latter attack on subjective probabilities is unconvincing, that the example of Frau Schmidt's unusual religious beliefs counters not what is generally meant by subjective probability but a straw version held by no one. Properly, the subjective probability of S on A is the probability of S on A, given some **limited** (as opposed to bizarre) information possessed by the subject.

While this interpretation of subjective probability seems reasonable and supports Frau Schmidt's decision to save little Adolf, it is able to pass a negative judgement on the mass destruction of infants (in the case above) only at the cost of becoming prey to the very objection leveled at both TUU and the first account of EEU. In order to distinguish a set of bizarre beliefs about the utility of an act from some reasonable but limited set of beliefs, some reference to actual total or actual expected utility must be made. But this, again, is to relate rightness or wrongness to objective information possibly inaccessible to the subject.

Clearly, EEU fares no better than does TUU when it comes to delivering the utilitarian from counter-intuitive attributions of rightness to particular acts. But TUU must—if it is to be taken to be a reasonable ethical theory—provide more than an argument by elimination of alternative utilitarianisms in response to the impracticability objection. It must somehow take into account our intuition that it is absurd to hold agents responsible for acting according to information they have no access to. The expected utilitarian response to this was to interpret rightness in terms of the best reasonable choice among possible alternative acts, rather than as choosing among the best of possible alternative acts. But this is

neither the only nor the best response to the argument from impracticability. Another is to accept the account of rightness presented in TUU and introduce a further ethical category—praise- and blameworthiness—to take into account that important intuition.

First, apart from the impracticability objection, it is not at all unreasonable to understand rightness and wrongness in terms of the actual total utility produced by our actions. It is neither unintelligible nor uncommon to judge acts retrospectively according to their consequences, quite apart from what good or bad intentions were instrumental in determining those acts at the time they were performed. Even in our Hitler case, there is certainly a sense in which we would assent that saving baby Adolf turned out to be the wrong thing to do and that Frau Schmidt was in one sense (as contributory cause) responsible for the existence of the future Adolf Hitler.

This need not, however, mean that Frau Schmidt is in any way culpable for her selfless, dutiful act. Her responsibility does not extend to blameworthiness. In fact, the nurse can be said to be praiseworthy for her action, in spite of the fact that it is retrospectively seen to have produced horrendous consequences. Rightness or wrongness is determined by the total utility of the act; praise- or blameworthiness depends on the motives and circumstances of the agent. It is in this way that TUU can take into account the impracticability objection. The force of the objection is the result of our reluctance to attribute blameworthiness to agents acting responsibly and with good intentions. Once it is clear that wrongness and blameworthiness (or rightness and praiseworthiness) are not necessarily conjoined, the objection is no longer compelling.

The Epistemic Line

ACCORDING to one version of epistemic externalism (a slightly altered version of that offered by Robert Nozick), an agent S's claim about the truth of some proposition p counts as an expression of knowledge if and only if it satisfies the following four conditions:

- CKa: (1) p is true.
 (2) S believes that p.
 (3) If p weren't true, S wouldn't believe that p.
 (4) There is no false express or tacit belief (q) central to S's basis for believing that p such that, were p true and S did not believe q, S would not believe that p.

These conditions are intended to be individually necessary and jointly sufficient for knowledge, such that any truth claim satisfying them and no truth claim failing to do so counts as an expression of knowledge. But, while conditions 1 and 2 need no additional explication (being conditions heralded in a long epistemic tradition), that is clearly not the case with conditions 3 and 4.

Until the 1963 paper by Edmund Gettier "Is Justified True Belief Knowledge?" epistemologists commonly held that the following three conditions were necessary and sufficient for S's knowledge of some proposition p:

- CKb: (1) *p* is true.
 (2) *S* believes that *p*.
 (3) *S* is justified in believing that *p*.

Gettier's paper, however, definitively showed that these conditions, even if necessary, are certainly not sufficient. Take, for example, the following Gettier-type case offered by Gail Stine:

if on the basis of visual appearances obtained under optimum conditions while driving through the countryside, Henry identifies an object as a barn, normally we say that Henry knows that it is a barn. Let us suppose, however, that unknown to Henry, the region is full of expertly made papier-mâché facsimiles of barns. In that case, we would not say that Henry knows that the object is a barn, unless he has evidence against it being a papier-mâché facsimile, which is now a relevant alternative. (Nozick, pp. 174–75)

Imagine this case in which Henry stops his car in order to admire a particularly handsome New England barn. There is in fact a barn before him (b1). He believes the object before him to be a barn (b2). (If asked what he was looking at he would not hesitate to tell an inquirer that he was admiring a barn.) And he appears to be perfectly justified in believing that what he sees is a barn (b3); inference from visual appearances obtained under optimum conditions would seem strongly to justify his conclusion. But it seems clear here that we would not want to call Henry's belief knowledge—in spite of the fact that it is a justified true belief—for Henry's belief is not sensitive to the possible falsity of *p*. Because of the widespread presence of barn facsimiles in the area, Henry would believe this object to be a barn whether it were a real barn or not.

Nozick's condition a3 (if not-*p*, then *S* wouldn't believe that *p*) is intended to deal with just such "lucky evidence" cases—cases in which a true belief is justified but remains insensitive to the possible falsity of *p*. Were Henry not looking at a real barn—were *p* false, he would still believe that he were. This violates the condition and invalidates Henry's truth claim as an expression of knowledge.

Two terms central to a4 must be explained before the importance of the fourth condition can be illustrated. First, a belief *q* is central to *S*'s basis for believing that *p*, if and only if *S*'s degree of confidence in *p* would be greatly reduced in the absence of *q*. For example, in the above case, Henry's belief (*q*) that his sense organs are generally reliable is central to his belief (*p*) that the object he sees is in fact a barn, while his belief (*q'*) that he packed his lunch in the trunk this morning before leaving his driveway is not. Were Henry seriously to doubt or cease to believe in the reliability of his sense organs, his confidence in the belief that there is a barn in front of him would be significantly reduced. Were, on the other hand, Henry to doubt or cease to believe that his lunch is in the trunk, his confidence in himself or his memory may be reduced, but his barn belief would likely remain unaffected.

Second, among the beliefs important in a4 are tacit beliefs. For our purposes, a tacit belief (*q*) is such that *S* would believe *q* expressly if questioned about it. In the above paragraph, Henry's belief (*q*) that his sense organs are generally reliable is a tacit belief. That Henry believes *q* is indicated by his confidence in

beliefs formed as the result of the operations of his sense organs; but until asked whether he believes that *q*, Henry may not even know that he holds such a belief. Roughly, *S* is said to have a tacit belief that *q* just in case (1) *S* would expressly believe that *q* if questioned,² and (2) *q* is not primarily the result of the questioning itself.³

Condition *a4* is directed at what may be called “false evidence” cases like the following advanced by Saul Kripke:

Driving through the countryside, Henry is again confronted with a real barn, which he correctly identifies as such. But although he is again in a region plagued with papier-mâché facsimiles, the ground on which this particular barn is built is such that it is chemically impossible to erect phoney barns there; on this particular plot of ground, phoney barns vanish.

Here Henry satisfies conditions *a1* through *a3*. There is a barn (*a1*) which he is seeing; he believes that what he sees is a real barn (*a2*); and, were there no real barn there, he would not believe there to be (*a3*), since no phoney barn could exist there to cause him to be deceived. However, Henry possesses a false tacit belief *q*—that there are not such things as phoney barns in the area—such that, did Henry not believe *q*, he would certainly come to doubt that what he is seeing is in fact a real barn (thereby failing condition *a2*), and thus no longer have knowledge.

It should be mentioned that although *a4* takes care of false evidence cases, it is not the condition as originally suggested by Nozick. Nozick gives the following fourth condition:

(*a4'*) If *p* were true, *S* would believe that *p*.

A4' was originally directed at a particular species of the brain-in-a-vat cases, with which Nozick deals in *Philosophical Explanations*: “. . . someone floating in a tank oblivious to everything around him is given (by direct electrical and chemical stimulation of the brain) the belief that he is floating in a tank with his brain being stimulated” (Nozick, p. 172). The truth-claim of this individual satisfies all three of the first three conditions. He is a brain in a vat (*a1*). Because his brain is being manipulated in a particular way, he believes himself to be a brain in a vat (*a2*). And were he not at the mercy of clever experimenters, he would not believe that he is a brain in a vat (*a3*). According to Nozick, this should not be considered an instance of knowledge for the reason that it might have been the case that the superscientists had stimulated his brain so that he had the belief that he was something other than a brain in a vat: an embodied Italian opera singer, for example.

2. This is not to say that *S* would admit to a belief that *q*, only that *S* would explicitly believe that *q*. Nor does an expressed (verbally) belief that *q* constitute a belief that *q*. That *S* verbally expresses a belief that *q* may imply either that *S* did tacitly or expressly believe that *q* or that *S* is being deceptive. *S* may in fact hold contradictory tacit beliefs, express beliefs which contradict *S*'s tacit beliefs, or express contradictory beliefs.

3. The purpose of requirement 2 is to discount as tacit a certain class of express beliefs caused directly by the questioning itself. For example, “that I am being questioned” or “that someone is speaking to me” are beliefs such that *S* would believe them expressly were *S* so questioned. Without requirement 2, such beliefs would by definition be tacit beliefs universally held. These beliefs are to be contrasted with beliefs such as “that blue whales are bigger than guppies,” which is made express by a question to that effect but is primarily the result not of the question itself but of beliefs about physical relationships in the external world.

According to Nozick, a4' is necessary so that knowledge will be sensitive not only to the possible falsity of the truth claim (a3) but also to its truth. "To be sure, conditions 1 and 2 tell us that p is true and he does believe it, but it does not follow that his believing q is sensitive to p's being true" (Nozick, p. 176).

The restated fourth condition (a4) does not, however, rule out the above brain-in-a-vat case as an instance of knowledge. There is in that case no false belief (q) (tacit or otherwise) central to the hapless floater's basis for believing himself to be a brain in a vat (p) such that were p true and he not believe q he would not believe that p. While the belief that his brain is being directly stimulated by superscientists (q') might be expected to result in S's not believing that p (because of the consequent awareness that ". . . [t]he operators of the tank could have produced any belief, including the false belief that he wasn't in the tank; if they had, he would have believed that . . ." [Nozick, pp. 175–76]), this has, in fact, not occurred—the case stipulates that he does believe that p. Moreover, this belief q' is a true belief; a4 applies only to knowledge claims involving false central beliefs.

There are two closely related objections to the inclusion of a4' in a list of necessary and sufficient conditions for knowledge. First, *prima facie*, a4' seems far too strong a requirement for knowledge. But certainly by a4' we cannot intend to require, nor does Nozick intend to require, that in all possible worlds where p is true S believes that p. Rather we must be making a requirement concerning states of affairs in possible worlds in which p is true and which are relevantly similar to the actual world. But, second, if this is enough to dispel the objection that a4' is too strict a condition for knowledge to be necessary, it achieves that end only at the expense of its usefulness. Take, for example, a case in which Henry correctly identifies a real barn and is justified in that truth claim (assuming there are no such things as barn facsimiles and Henry is justified in his claim on the basis of visual appearances obtained under optimum conditions), but makes this claim at all only because he has happened to glance attentively over his right shoulder while driving in the country. But all other things—including the truth of p—held constant, had he continued to look straight ahead, or glanced over his left shoulder instead, or glanced over his right shoulder while daydreaming (three possible worlds almost inarguably close to the actual one), he would not believe that p were it true. A4' gives no hints when it comes to its interpretation relative to problem cases—a severe fault in a necessary condition for knowledge directed at problem cases.

There are two important sorts of cases which have led many epistemologists to reject externalist accounts of knowledge in favor of some internalist variety: (1) cases which focus on the reliability of the source of the belief, and (2) cases which deal with the irrationality or irresponsibility of the believer. The brain-in-a-vat case above is a case of the first sort. According to CKa, the case illustrates an instance of knowledge, but, for reasons mentioned above, such a conclusion seems suspect. Cases of the second sort have been advanced by Bonjour. Bonjour remarks that it is consistent with externalism that ". . . a person may be highly irrational and irresponsible in accepting a belief, when judged in the light of his

own subjective conception of the situation, and may still turn out to be epistemically justified . . . ” (Bonjour, p. 38).

Consider the following case:

Samantha believes herself to have the power of clairvoyance, though she has no reasons for or against this belief. One day she comes to believe, for no apparent reason, that the President is in New York City. She maintains this belief, appealing to her alleged clairvoyant power, even though she is at the same time aware of a massive amount of apparently cogent evidence, consisting of news reports, press releases, allegedly live television pictures and so on, indication that the President is at that time in Washington, D.C. Now the President is in fact in New York City, the evidence to the contrary being part of a massive official hoax mounted in the face of an assassination threat. Moreover, Samantha does in fact have completely reliable clairvoyant power under the conditions which were then satisfied, and her belief about the President did result from the operation of that power. (Bonjour, p. 38)

In this case, Samantha seems to be behaving both irrationally and epistemically irresponsibly in holding on to her belief in the face of contrary evidence; it seems absurd therefore to count such a case as an instance of knowledge.

An obvious response to such counterexamples is to reject externalist accounts of knowledge in favor of an internalist view. Bonjour makes just such a move. “[E]xternal or objective reliability is not enough to offset subjective irrationality. If the acceptance of a belief is seriously unreasonable or unwarranted from the believer’s own standpoint, then the mere fact that unbeknownst to him its existence in those circumstances lawfully guarantees its truth will not suffice to render the belief epistemically justified and thereby an instance of knowledge” (Bonjour, p. 41). Instead, knowledge is understood not in terms of an external relation between the objects of knowledge and the true beliefs of some subject (a relation to which the subject ultimately has no access), but in terms of certain characteristic (for knowledge) conditions of the subject’s belief set.

In Bonjour’s case, these characteristic conditions are (1) consistency, (2) the presence of a significant proportion of inferential interconnections among component beliefs, (3) the lack of a significant number of unexplained anomalies in the believed content, (4) requisite putative input, and (5) persistent coherence in the long run. Again, some explanation is in order.

Bonjour’s view is a coherence view of justification, and he explains coherence as follows:

Intuitively, coherence is a matter of how well a body of beliefs “hangs together”: how well its component beliefs fit together, agree or dovetail with each other, so as to produce an organized, tightly structured system of beliefs, rather than either a helter-skelter collection or a set of conflicting sub-systems. (Bonjour, p. 93)

The necessity of the first condition (consistency) is clear from this passage: beliefs which more or less explicitly contradict one another cannot well “hang together” in a belief system. Both belief sets A and B below contain logically contradictory beliefs (A more explicitly so than B); neither set hangs together in the requisite way.

- A {dogwoods are deciduous, dogwoods are not deciduous}
- B {chimps are omnivorous, chimps are herbivorous}

But Bonjour has more than mere logical consistency in mind by condition 1. Logical consistency is necessary but not sufficient for what he terms probabilistic consistency. For example, a belief system containing both the belief that *p* and the belief that it is extremely improbable that *p* is intuitively less coherent than that system would be without these beliefs (Bonjour, p. 95); while it is logically consistent, it is not probabilistically so. A system of beliefs is coherent in proportion to its degree of probabilistic consistency.

Consistency alone is not, however, sufficient for coherence. Consider the following belief sets C and D.

- C {bats have wings, there is no "q" in "meringue," Gladys has hidden the car keys}
 D {all ravens are black, this bird is a raven, this bird is black}

Although the beliefs in C do not conflict with one another (any more than do the beliefs in D), C does not hang together as well as D. This is because the beliefs in C avoid conflict by having no connections whatsoever to one another, while the beliefs in D are inferentially connected (Bonjour, p. 96). The coherence of a system of beliefs is increased in proportion to the number and strength of inferential interconnections among its component beliefs (Bonjour, p. 98).

One important consequence of this condition is that the coherence of a belief system is decreased in proportion to the number of belief subsystems relatively unconnected to other beliefs in the system as a whole (Bonjour, p. 98). Take, for example, the case of Edmund the engineer, whose belief system is characterized, for the most part, by beliefs inferentially interconnected to a belief in mechanistic (causal) determinism. Edmund's belief system, however, also contains a subsystem of beliefs inferentially interconnected with the belief in his own free will, but relatively unconnected to beliefs in the rest of the belief system. Even assuming that this belief in his own freedom of will does not involve explicit contradiction with Edmund's belief in determinism, it still seems clear that Edmund's belief system is less coherent than it would be without the free will subsystem, and this in spite of the fact that that belief system is both relatively consistent and contains a great many inferential interconnections.

Condition 3 is actually a special case of the requirement that there be a significant number of inferential interconnections. One type of inferential connection important to the coherence of a belief system is explanatory connection. A purpose of scientific explanation, for example, is to ". . . exhibit events of widely differing kinds as manifestations of a relatively small number of basic explanatory principles . . ." (Bonjour, p. 99), to unify the system according to these principles. The greater the extent to which this can be done, the more coherence the belief system exhibits. Anomalies (events unexplained within the system) cut against that system's unity, and therefore reduce its coherence; a system's coherence is decreased in proportion to the presence of anomalies in its believed content (Bonjour, p. 99).

Condition 4 (requisite putative input) comes about because of the goal of a coherence theory such as Bonjour's: to provide a basis for the justification of individual empirical truth claims. Coherence of a belief system is sought not

primarily because of aesthetic considerations but because it is believed that a coherent system of beliefs can provide knowledge of an extra-systematic reality. As a result, belief systems must be somehow “connected up” to the reality they purport to describe; only so can the consequence, that empirical beliefs having nothing to do with—or false with respect to—the reality they are meant to describe be fully justified by virtue of their place in a coherent belief system, be avoided. It is to avoid such a consequence that Bonjour introduces cognitively spontaneous beliefs.

A cognitively spontaneous belief is a belief which results not from any inference or deliberative process but rather occurs spontaneously in the belief system of the believer in a way that is both involuntary and coercive (Bonjour, p. 117). For example, sitting here at my desk, I hear the sound of birds outside my window; that there is the sound of birds now is a cognitively spontaneous belief.⁴ “It is cognitive spontaneity which marks the belief as putatively observational . . .” (Bonjour, p. 117), as having to do with extra-systematic reality.

That there be such putative input to a coherent belief system (if it is to provide knowledge of an extra-systematic reality) is what Bonjour calls the Observation Requirement. This requirement is in two parts: (1) there *must* be putative input to the system (if it is to be connected up to extra-systematic reality), and (2) the system must also contain “. . . beliefs to the effect that recognizable kinds of [these] cognitively spontaneous beliefs are likely to be true . . .” (Bonjour, pp. 139–40). The support for these latter beliefs consists in the further belief that the best explanation for the existence of a significant body of cognitively spontaneous beliefs largely in agreement with one another is that they are caused by extra-systematic reality (Bonjour, p. 140).

The final condition (persistent coherence in the long run) is where much of the weight of Bonjour’s view rests. To be truth-conducive (to be able to provide adequate justification for individual empirical truth claims), a system of beliefs satisfying the Observation Requirement must remain coherent in the long run. Again, the support for this consists in an argument from the best explanation:

The best explanation, the likeliest to be true, for a system of beliefs remaining coherent (and stable) over the long run while continuing to satisfy the Observation Requirement is that (a) the cognitively spontaneous beliefs which are claimed, within the system, to be reliable are systematically caused by the sorts of situations which are depicted by their content, and (b) the entire system of beliefs corresponds, within a reasonable degree of approximation, to the independent reality which it purports to describe; and the probability of this explanation increases in proportion to the degree of coherence (and stability) and the longness of the run. (Bonjour, p. 171)

A system of beliefs satisfying conditions 1 through 4 is, according to Bonjour, truth-conducive; that is, empirical beliefs justified according to such a system are highly likely to be true, and thus highly likely to be instances of empirical knowledge.

4. Calling such beliefs involuntary and coercive may be somewhat misleading. Certainly, I have a fair amount of control over which cognitively spontaneous beliefs I will have. I might, for example, decide not to have the belief five minutes from now that there is the sound of birds now, and immerse myself in my work or wear ear plugs to ensure that I don’t. Rather, cognitively spontaneous beliefs are involuntary and coercive in the counterfactual sense that if one were to be in a particular state (e.g., sitting attentively in my office not wearing ear plugs at a time when there are indeed birds chirping outside my window), **then** one could not help but have a particular spontaneous belief (that there is the sound of birds now).

One criticism of Bonjour's view directly parallels the criticism of expected utilitarianism discussed above. One of two interpretations of the function of his conditions can be adopted: in instances of knowledge either (1) these conditions are actually satisfied for the subject, or (2) the subject believes these conditions to be satisfied for his/her system of beliefs. The former interpretation appears to be part of Bonjour's understanding of his conditions. In his doxastic presumption, Bonjour makes the requirement that the subject does indeed hold approximately the system of beliefs which s/he believes him/herself to hold. Such a presumption is necessary, he claims, if a subject is to have cognitive access to the fact of the coherence of that system of beliefs (Bonjour, p. 103). This assumption seems itself to be dubitable. If tacit beliefs, for example, are to be included in one's system of beliefs—and it seems that they should, since they are regularly acted upon—it seems clear that they are only “believed to be held” if made express.

But even if the doxastic presumption is granted, it is obviously not enough that the system of beliefs objectively satisfies Bonjour's conditions for justification. That a truth claim happens to fit in a long-run coherent system of beliefs is no less an external condition than that a belief happens to be lawfully related to an objective fact. According to Bonjour, to be justified in a truth claim the subject must have cognitive access to the relationship between belief and belief set that implies cognitive access to the coherence of the system of beliefs itself.

It is also not enough that the subject believes the conditions to be satisfied (interpretation 2). This would be to make knowledge entirely subjective, a result Bonjour rejects as illegitimate in an account of knowledge. As a result, Bonjour requires that the conditions be both objectively satisfied and that the knower have subjective access to that fact. But this is achieved only by assumption; nowhere does he argue (nor is there an argument available to him) that the objective satisfaction of the conditions is accessible to the knower. One cannot step outside of one's belief set to check the truth of one's beliefs about that belief set, and the attempt to do so involves an infinite regress into beliefs about belief sets.

* * *

MERELY to note that coherentist internalism (or, as I believe, every form of internalism) fails to provide an acceptable means of avoiding certain counterintuitive attributions of knowledge (specifically, those involved in the brain-in-a-vat and clairvoyance cases)⁵ is not, of course, to make externalism a preferable doctrine. Something more—something parallel to the categories of praise- and blameworthiness in ethics—is necessary. Just as the categories of praise- and blameworthiness had to be provided to account for our intuitions that individuals are not culpable for acting according to what they reasonably believe to be right (for example, Frau Schmidt's selfless rescue of the infant Adolf), some

5. Bonjour's internalism, for example, becomes externalist in requiring that his five conditions in fact be satisfied for the believer. It is not enough, nor will it suffice to account for problems presented by such as the brain-in-a-vat and clairvoyance cases, that the believer merely believe his/her belief to have satisfied the conditions.

account must be taken of our dissatisfaction with simply classifying the beliefs in the brain-in-a-vat and clairvoyance cases as instances of knowledge.

One possibility is to associate epistemic justification exclusively with the reasonableness of the believer in asserting a truth claim. However, the term “justification” is not infrequently used with respect to objective conditions for knowledge as well. To stipulate that justification has to do with the conditions for reasonably holding a belief rather than the conditions for knowledge is to invite confusion in a debate already fraught with it.

I would suggest, then, the use of the terms “responsible” and “irresponsible” for referring to the respectability (reasonableness) of particular truth claims and the subjects advancing them. That any given truth claim is in fact an expression of knowledge tells us nothing about whether it is responsible or not, just as the rightness of an act does not tell us whether or not it is praiseworthy.

To illustrate the application of this approach, we look once again at our two problem cases:

(1) Although the belief in the brain-in-a-vat case (pp. 93–94) satisfies each of our external conditions for knowledge, the subject’s belief that he is a brain in a vat is not a responsible one. Adopting Bonjour’s very reasonable five epistemic conditions as conditions characteristic not for knowledge (as was his intent) but for epistemic responsibility, we can see that our believer in a vat’s belief fails the second part of the Observation Requirement (condition 4). The floating believer has no good reason to believe that his spontaneous beliefs about his condition are likely to be true, since their source (the experimenters) is questionably reliable. They might as easily have caused him to have false beliefs.

(2) In the same way Samantha’s belief in the President’s whereabouts (p. 95) satisfies each of the four conditions for knowledge but fails the test for epistemic responsibility. As stated in the case, she has no reason to believe that her clairvoyant (cognitively spontaneous) belief is of a kind likely to be true. Her belief fails part two of the Observation Requirement (condition 4) and thus is irresponsible.

Each of these cases, then, illustrates an instance of knowledge which is irresponsible, while a false truth claim arrived at via reasonable methods such as Bonjour’s conditions would constitute a responsible instance of non-knowledge.

Sources Used

- AUNE, BRUCE. “Epistemically Justified Opinion.” Forthcoming in *The Current State of the Coherence Theory*. Ed. J. Bender.
- BONJOUR, LAWRENCE. *The Structure of Empirical Knowledge*. Harvard Univ. Press, Cambridge, 1985.
- GETTIER, EDMUND. “Is Justified True Belief Knowledge?” *Analysis*, 23 (1963), 121–23.
- NOZICK, ROBERT. *Philosophical Explanations*. Belknap Press, Cambridge, 1981.