



## II Critiques of Epistemic Externalism





# The Chimerical Appeal of Epistemic Externalism

JOE CRUZ AND JOHN POLLOCK

## 1. Introduction

Internalism in epistemology is the view that all the factors relevant to the justification of a belief are importantly internal to the believer, while externalism is the view that at least some of those factors are external. This extremely modest first approximation cries out for refinement (which we undertake below), but is enough to orient us in the right direction, namely that the debate between internalism and externalism is bound up with the controversy over the correct account of the distinction between justified beliefs and unjustified beliefs.<sup>1</sup> Understanding that distinction has occasionally been obscured by attention to the analysis of knowledge and to the Gettier problem, but our view is that these problems, while interesting, should not completely seduce philosophers away from central questions about epistemic justification.

A plausible starting point in the discussion of justification is that the distinction between justified beliefs and unjustified beliefs is not the same as the distinction between true beliefs and false beliefs. This follows from the mundane observation that it is possible to rationally believe claims that are false. Still, most epistemologists have claimed that truth and falsity play a crucial role in distinguishing between justified and unjustified belief because they maintain that believing truths is the ultimate objective of human rational cognition. Much of the explanation for the continued attraction of externalist theories of epistemic justification derives from a specific perspective on the conviction that truth is the sole or primary evaluative standard against which rationality is judged. By this view, externalism has a strong *prima facie* attraction, as properties of epistemic agents that are truth-aimed—like the actual reliability of their cognitive processes—are (in at least an intuitive sense) external to the cognitive agent. Thus, externalism appears to hold out the most promise of properly respecting the connection between belief formation and truth as the evaluative standard of rationality.

The appeal of externalism, we maintain, is illusory. Our strategy is to consider the motivations behind externalism and to show that each is misguided or can be accommodated in an acceptable internalist theory. This, combined with the difficulties faced by externalist theories, seem to us to constitute a decisive case in favor of internalism.

Contemporary externalist accounts have germinated in light of three sources that constitute refinements of the generic truth motivation, above, even if these sources

---

<sup>1</sup> In this paper we will treat ‘justified’ and its cognates as synonymous with ‘rational’ and its cognates. Thus, we ignore the fact that they are sometimes distinguished in the epistemological literature in light of what some take to be the overly deontological connotation of justification and the greater theoretical neutrality of rationality. See, e.g., Plantinga (1993a).

have not always been articulated explicitly or particularly clearly. Firstly, since truth has been viewed as the ultimate epistemic aim, externalists have attempted to answer to that aim directly by making it a non-negotiable constraint on a theory of justification that whatever justification is will co-vary in some way with truth. Secondly, analyses of justification have often emerged in close affiliation with analyses of knowledge. The connection to knowledge recommends a true-belief oriented approach to justification, and thus externalism has been tempting because it appears uniquely able to build in this orientation. This second source of attraction to externalism is closely related to the first, but is worth considering separately because—in our view—it suffers from an additional burden in that knowledge is a poorly behaved concept. Finally, it has appeared to many epistemologists that the only possibility for a naturalistically acceptable account of rationality is some version of externalism.

We reject all three of these avenues to externalism. Against the first, we maintain that the ways externalists attempt to incorporate truth as the evaluative standard of rationality either conflates rational beliefs with true beliefs or avoids the conflation in a wholly *ad hoc* way. Neither of these is acceptable for a theory of rational belief. Our concern with the second motivation is that there is no stable concept of knowledge that constitutes the target of epistemic inquiry in the way that the externalist envisions. If this is right, strong scruples with respect to forging a connection between justification and knowledge are misguided. In response to the last motivation, we sketch an internalist theory that is recognizably naturalistic, and thereby show that externalism is not uniquely amenable to naturalism.

We turn now to laying out the relevant terrain in order to clarify just what the externalist proposes.

## 2. *Externalism contra Internalism*

We can begin to understand epistemic externalism by trying to make sense of what it denies, namely, epistemic internalism. One species of internalism claims that reflective, careful agents are able to make assessments of their reasons for a belief in order to determine whether that belief is justified. The view is that we can determine, for any given belief, whether it is sustained by reasons that are good enough to make it justified. Chisholm offers a representative statement of this sort of internalism:

...Epistemologists presuppose that they are rational beings. This means, in part, that they have certain properties which are such that, if they ask themselves, with respect to any one of these properties, whether or not they have that property, then it will be evident to them that they have it. It means further that they are able to know what they are thinking and believe and that they can recognize inconsistencies (1989, p. 5).

This and related Chisholmian formulations are notoriously vague, as they do not distinguish between the claims that epistemic agents have reflective access to the content of their beliefs, to the reasons for their beliefs, to the reasonableness of the connection between their reasons and their beliefs, all of these together, or yet some other quantity. The one feature of Chisholm's internalism that is clear is that it proposes a fairly dramatic ability on the part of the epistemic agent to find out through reflection what

the epistemic credentials of her belief are. The sense of ‘internal’ operating here, then, is that the crucial elements of justification are such that an epistemic agent can locate them by reflection alone. This can only be true of elements that are inside the agent, and, more specifically, within her reflective purview.

Many epistemologists have followed Chisholm’s lead in understanding internalism as committed to a strong and encompassing version of internal reflective access to properties that make a belief rational<sup>2</sup>, but this is somewhat misleading. There are different kinds of access to reasons and beliefs that the internalist may propose, and not all kinds demand the thoroughgoing introspective insight that Chisholm does. Insight that is more limited in scope seems to also allow for an internalist view. For instance, one of us has defended a view where epistemic agents are able to determine through introspection whether they are conforming to their rational norms in individual cases, even if the nature of the rational connection between their internal states or the content of the relevant states is not always transparent (Pollock 1974; 1979; 1986; for the latest statement of this view, see Pollock and Cruz 1999). This is nonetheless an internalist theory in the sense that epistemic agents have introspective access to whether their cognitions are rational.

A view like this is also internalist in a weaker but—in our view—ultimately preferable sense. We maintain that the fundamental idea behind internalism is that the justifiability of a belief is exclusively a function of internal states, where internal means simply inside an agent’s cognitive apparatus. What that means is that we can vary everything about a situation other than an epistemic agent’s internal states without affecting which of her beliefs are justifiable. In particular, varying contingent properties such as the reliability of the cognitive processes that produce or sustain a belief will not affect whether that belief is justified. We have called this cognitive essentialism (Pollock 1986; Pollock and Cruz 1999). According to cognitive essentialism, the epistemic correctness of a cognitive process is an essential and inherent feature of that process. It is not necessary to reference some external quantity such as reliability in order to certify the correctness of the cognitive process. We aim to make this plausible later in this essay, but at this point it should be clear that cognitive essentialism does not implicate or emphasize access. We think that this is a virtue, in part because we think that the genuine details of cognitive access are to be revealed by cognitive psychology, not by the *a priori* fiat of a theory of rationality. Be that as it may, we will continue to talk about internalism in terms of access issues. This is useful because it accords so well with the way most internalists themselves conceive of internalism.

It cannot be denied that 20th century epistemology has been primarily internalist (Plantinga 1990). Internalist considerations have been accorded a special place in epistemology because of several related concerns: First, one of the projects that sometimes rides piggyback on an assessment of rationality has been to illuminate how one might improve the quality of one’s beliefs. If improvement is to be possible, however, it needs to be possible to determine which belief among many candidate beliefs is most justified. Internalists have been driven, then, by what we might think of as rational solip-

---

2 See (Moser 1985) or (Foley 1987; 1993).

sism since, in their view, we are alone in the task of refining our beliefs and the only resources we have to work with are the ones that can be brought to bear internally. Second, epistemic concepts have been viewed as deontological in character. That is, it has seemed to many epistemologists that intellectual achievement is at least partly a matter of duty-fulfillment. Fulfilling a duty, however, seems to require that one be able to do the things that duty requires and to determine what one's duties are. In order to secure the means to an intellectual duty, an epistemic agent will need to be able to reflect on her condition and on the resources she has available. Thus, a reflective requirement has been built into epistemic theorizing. Third, intellectual achievements have had some historical association with defeating the skeptic, and it has sometimes been suggested that the ability to answer the skeptic is a crucial component of epistemology. The only answers that the epistemic agent can give in reply to a skeptic's claims, though, are the ones that are available to her reflective capacities.

In principal, externalism seems to gain a foothold when at least some of these motivations for internalism are denied. That they can be denied should not be a complete surprise. Introspective access to the properties of a belief that make it rational or even a more weakly internalist cognitive essentialism do not seem demanded by the goals of epistemology as such. There is the ubiquitous goal of epistemology where the epistemologist takes it upon herself to say what justification or rationality is. There is a second current where the epistemologist maintains that the believer must be able, in general, to determine whether or not her beliefs are rational. In principle, an epistemologist may pursue the first kind of theory without pursuing the second. It is open to the epistemologist to reject the claim that the right philosophical account of belief quality must enable actual regulative, duty-oriented, or skeptic-answering evaluations to take place. This is just the opening that the externalist needs.

What makes things unfortunately murky here is that both the deontological heritage of epistemology and the need to answer the skeptic have been denied by some internalist epistemologists. It is our view that these sources of internalism are best thought of as closer to optional than the first, regulative source. It is the rejection of the regulative character of justification that provide a more powerful point from which to reject internalism, and the development of the literature on externalism bears this out. For example, Goldman (1980) specifically repudiates the regulative character of epistemology, where the regulation must take place in a first-person way. In his positive view, he claims—to a first approximation—that what makes a belief justified is that it is produced by a psychological process that usually produces true belief (Goldman 1979; Goldman 1986). The process must in some sense instantiate a set of reasons, but they need not be reasons that are accessible to the rational agent. Plantinga (1993a; 1993b) is also associated with rejecting the first personal regulative conception of rationality. His theory of proper functions claims that a belief is justified in case it is probably true and is the product of a properly functioning cognitive process working in an environment for which it is appropriate (Plantinga 1993b, pp. 4-20).

### 3. *Externalism as a Proposal about Justified Belief Formation*

Understanding externalism as the denial of the regulative strand of cognitive access is a plausible way of generating a taxonomy of theories of justification, but it should be emphasized that this is only a heuristic device. The actual motivations behind externalist theories are not, of course, principally or solely concerned with rejecting internalism.

At the beginning of this essay we identified the putative truth-aimedness of justification as a source of motivation for externalism. Since truth is thought to be the ultimate normative aim of belief formation, the externalist has found it inviting to simply build into her theory some covariation between justified belief and true belief. This way of putting the externalist's ambition highlights an important distinction that must be made when considering externalism. Externalism can be understood as providing answers to two quite different questions. On the one hand, the externalist might describe the correct epistemic norms—where norms dictate how to come to a belief<sup>3</sup>—with an eye toward emphasizing some external feature of those norms. An example of such a proposed norm might be “It is permissible to hold a belief if it is generated by a reliable cognitive process.” On this view, the norms for belief themselves appeal to reliability. On the other hand, the externalist might aim solely to tell us what makes epistemic norms correct and appeal to external properties only with respect to that meta-justificational question. So, it might be urged that from the set of possible epistemic norms a cognizer can have, the norms that are aimed at the truth are the correct ones. We have labeled these two kinds of externalism *belief externalism* and *norm externalism* (for further discussion, see Pollock and Cruz 1999, pp. 130-42).

What would make someone adopt belief externalism? In “What is Justified Belief?” Goldman writes:

Consider some faulty process of belief-formation, i.e., processes whose belief-outputs would be classed as unjustified. Here are some examples: confused reasoning, wishful thinking, reliance on emotional attachment, mere hunch or guesswork, and hasty generalization. What do these faulty processes have in common? They share the feature of unreliability: They tend to produce error a large proportion of the time. By contrast, which species of belief-forming (or belief-sustaining) processes are intuitively justification-conferring? They include standard perceptual processes, remembering, good reasoning, and introspection. What these

3 In a reasonably non-committal sense we take epistemic norms to be recipes that capture a peculiarly epistemic propriety between cognitive states. Some epistemologists, for instance, have proposed that there is an epistemic norm regarding the good inferential relationship between perceptual states and beliefs about those perceptual states. One way of putting this norm is, “If a person has a clear sensory impression that  $x$  is  $F$  (or of  $x$ 's being  $F$ ) and on that basis believes that  $x$  is  $F$ , then this belief is *prima facie* justified” (Audi 2001, p. 43). Another way is: “If  $S$  believes that he perceives something to have a certain property  $F$ , then the proposition that he does perceive something to be  $F$ , as well as the proposition that there is something that is  $F$ , is one that is reasonable for  $S$ ” (Chisholm 1966, p. 47). A third is: “Having a percept at time  $t$  with the content  $P$  is a defeasible reason for the cognizer to believe  $P$ -at- $t$ ” (Pollock and Cruz 1999, p. 201). Other formulations of this norm vary (cf., Bonjour 2000, p. 30; Field 2000, p. 134; Plantinga 1993, p. 99; Sosa 1980, p. 22), but they all seem to be after the same thing, namely the claim that sensing that  $p$  is a fallible but good basis for an inference to the belief that  $p$ . For a longer treatment of epistemic norms, see (Pollock and Cruz 1999, chapter 7) and (Cruz, unpublished manuscript).

processes seem to have in common is reliability: the beliefs they produce are generally reliable (1979, pp. 9-10).

Goldman is claiming that rational beliefs have something in common and that that common thing, reliability, is constitutive of justified belief. Goldman thus proposes an early version of his process reliabilist theory. The basic idea behind process reliabilism is that a belief is justified if and only if a reliable cognitive process produces it. For example, the reliabilist explains why perceptual beliefs are justified by pointing to the fact that, in the actual world, perception is a reliable cognitive process. Similarly, deduction is a reliable cognitive process, so beliefs deduced from other justified beliefs will be justified. On the other hand, wishful thinking is not a reliable cognitive process, so beliefs produced by wishful thinking are not justified. The reliability of a cognitive process is the indefinite probability of beliefs produced by it being true. It makes the justifiedness of a belief depend not only on the processes that produced it, but also on whether those processes happen to be reliable in the actual world.

Reliability is neither something that the epistemic agent will have access to nor is it cognitively essential in our sense (i.e., justification may vary even if properties internal to an agent are kept constant). This proposal is belief externalist because Goldman is claiming that epistemic norms must themselves reference reliability. He signals as much when he writes, early in the piece, that he assumes that “a justified belief gets its status of being justified from some processes or properties that make it justified” (p. 2). This seems to be saying that there is some property occurring at the level of the epistemic norms that confers justification on beliefs. Goldman is driven to belief externalism by his sense that he has identified reliability as the unifying property of epistemic norms.<sup>4</sup>

Belief externalism of this sort has been thoroughly criticized in the literature. The simplest objection is that, intuitively, it is not the extrinsic properties of the processes that generate a belief that is responsible for its justifiedness. This intuition has been the source of repeated attempts by internalists to generate cases that show that, for instance, reliability is neither necessary nor sufficient for justification. The best known of these is Bonjour’s case of Norman who has reliable clairvoyance and who believes that he has clairvoyance, although he lacks any justification for that belief (1980, p. 62; 1985, p. 41). Suppose Norman comes to believe that the President is in New York City on the basis of his clairvoyance. Bonjour claims that our intuitive response is that, in the absence of any justification for his belief in his own clairvoyance, Norman is irrational to accept his belief that the President is in New York. Since reliabilism claims that Norman is justified in that belief, it gives the intuitively wrong results

In response, reliabilists either deny the intuition or generate ancillary principles to address it. For instance, Goldman claims that the Norman case cannot be coherently described in such a way that it does not violate a plausible no-undermining—or, as we prefer to put it, a no defeater—clause of reliabilism. This no defeater clause maintains that a belief is justified if it is the product of a reliable cognitive process and if there

<sup>4</sup> Indeed, the simplicity and unity of the underlying justification-conferring property is a feature of externalist theories that externalists have sometimes been keen to emphasize, since it seems to contrast so starkly with the theoretically heterogeneous collection of norms proposed by internalists.

is no more reliable cognitive process that, if available and employed, would defeat the belief. Goldman says that Norman ought to reason in such a way as to verify his clairvoyant powers. If he does not, he is failing to use a process available to him that would render his belief unjustified (1986, p. 112).

We do not need to rely solely on intuitions about cases in order to criticize reliabilism. Belief externalism has been criticized in light of a very fundamental difficulty it faces. There seems to be no non-ad hoc way to circumscribe a cognitive process in order to assess its truth-aimedness. In the specific case of reliability, this has come to be known as the generality problem (Pollock 1984; Feldman 1986; Conee and Feldman 1998).<sup>5</sup> There are really two problems that fall under the generality heading. The first has to do with individuating process types. There seems to be no non ad hoc way to isolate a cognitive process in order to assess its reliability. Suppose someone forms a belief about an object by using her normal vision in good light. Which process should be assessed for its reliability? Should we individuate the process as vision right at this moment, vision over the last ten years, the whole vision/cognitive system, or by some other individuating scheme? The difficulty with specifying a principled answer to this challenge has dogged this brand of reliabilism nearly from the start.

The other problem has to do with the environment. Which elements of the environment are relevant to assessing reliability? It might first be proposed that in judging the reliability of a cognitive process in a particular instance we should take account of everything about the circumstances in which it is used. If it makes any sense at all to talk about the reliability of the cognitive process “under the present circumstances” (in all their specificity), it seems that it must be the (indefinite) probability of producing a true belief, conditional on everything true of the present circumstances. But the present circumstances are infinitely specific and include, among other things, the truth-value of the belief being produced by the cognitive process and the fact that that is the belief being produced. Consequently, this indefinite probability must go the same way as objective definite probabilities and be either 1 or 0 depending upon whether the belief in question is true or false. Thus this reliabilist criterion entails the consequence that in order for a belief to be justified it must be true, and this is a denial of the starting point of a theory of rationality. Goldman has tried to evade these challenges by claiming that the processes and environmental conditions are individuated intuitively (1979, p. 12). He claims that, if there is vagueness involved this only reflects the vagueness of our concept of justification. This response, however, is not particularly persuasive (cf. Alston 1995).

We think that the failure of this prominent kind of externalism is instructive, but we prefer not to build our case against externalism around the missteps of a particular view. Notice, though, that there is a much more general point to be made here about belief externalism. It appears that the strategy of attempting to build a truth-aimed component into norms of belief formation will always founder on a tension between belief formation processes that fall short of truth and ones that if employed, would guarantee

---

5 The label for the problem comes from Goldman's own recognition of it in his (1979).

truth. The lesson appears to be that if truth-aimedness is treated as the critical value that an epistemic norm must have, then there is no principled way to resist the slide to epistemic norms that yield only true beliefs. The alternative strategy is to attempt to articulate epistemic norms that can fall short of truth, but the degree of reliability demanded will always seem ad hoc.<sup>6</sup> This is just the generality problem writ large for belief externalist proposals that respond to the truth-aim motivation by trying to build the external component into the belief formation processes themselves.

#### 4. *Externalism about the Correctness of Epistemic Norms*

We take the above considerations as showing that belief externalism is ill-advised, and we take it as some confirmation of this that Goldman, for one, has since admitted that he can and does accept internalist rules for belief formation (1994, p. 304). But this is to say nothing about another sense of externalism, norm externalism. Norm externalism acknowledges that we must employ wholly internalist norms in our reasoning, but proposes that alternative sets of internalist norms should be evaluated in terms of external considerations. The premiere external consideration is, again, truth-aimedness. The truth motivation appears again in a slightly different form and at a higher level. For the norm externalist, it is the entire framework of belief formation processes that must answer to truth as the fundamental normative aim of epistemology.

Why would someone be a norm externalist? As a first pass, norm externalism is attractive because it seems as if norms should be criticizable on some grounds apart from the norms themselves. Apparently, the internalist must assume that our actual ultimate epistemic norms are immune from criticism because whatever norms she expects to employ in her self-criticism will be norms that are internal. This is surely odd, and the norm externalist's thought is that there must be a higher standard against which the norms are evaluated. Consider two people who employ different epistemic norms. They could then hold the same belief under the same circumstances and on the basis of the same evidence and yet the first could be conforming to her norms and the second not conforming to hers. If a person's epistemic norms were always beyond criticism, it would follow that the first person is justified in her beliefs and the second is not, despite the fact that their identical beliefs are based upon the same evidence. Because this case is possible, it seems as if we ought to be able to determine which of those norms are correct. This makes a strong *prima facie* case for norm externalism.

These are just the sorts of motivations that lie behind current versions of norm externalism. For example, in "An Internalist Externalism," William Alston takes the simple course of proposing that a belief is justified if it enjoys the sorts of access properties that

<sup>6</sup> Another way to put this point would be in terms of defeat. When the belief externalist attempts to build in a no defeater condition, she is moved toward perfect reliability. This conflates rationality and truth. There is pressure, however, to include a no defeater condition because any epistemic norm will be truth-aimed some of the time and will fail to be truth-aimed at other times, depending on how circumstances are described. No defeater conditions are introduced to handle this possibility.

are typical of an internalist view and is also probably true. Both features are necessary for a belief to be justified, and this goes to show that Alston is concerned to maintain an external evaluative standard for beliefs, i.e., truth.<sup>7</sup> The drive to norm externalism is operative in Goldman's later work on reliabilism as well. In *Epistemology and Cognition*, Goldman first defends a rule framework for epistemology, where the conditions for having a justified belief can be captured as a set of justificatory rules (J-rules). He acknowledges that the content of the J-rules might turn out to appeal only to internal considerations, while the criterion of rightness for the J-rules might be objective (p. 72). This is Goldman's way of making the distinction between belief externalism and norm externalism. At the level of questions of the correctness of the norms, Goldman favors a truth-aimed, or *verific*, theory. He gets there by defending a consequentialist rather than deontological account of the criterion of correctness. The argument is that no purely deontological account satisfies the aims of epistemology. Once Goldman has concluded that a consequentialist view of the criterion of correctness is necessary, he assesses various possible consequentialist conceptions including non-*verific* consequentialist accounts.

According to Goldman, the failure of non-*verific* consequentialist theories is that they would allow beliefs to be justified even if they were not true. This is the point of what he calls 'fantasizer counterexamples.' These are alleged counterexamples to theories of justification where the epistemic agent is guilty of fantasizing to such a degree that wildly false beliefs turn out justified because they are sanctioned by his norms. He writes, "...if these fanciful constructions were consistently false, I would feel little inclination to say that this sort of process confers justifiedness. Certainly it would not yield the kind of justifiedness that brings us close to 'knowledge'" (p. 101). The external property of truth serves as insurance against this kind of massive fantasy.

We think that Goldman is fundamentally confused on this point. We insist that a theory of justification needs to respect the constraint that justified beliefs are not the same as true beliefs (though, of course, it is possible for them to be coextensive). This applies to belief sets as well as belief tokens. A set of beliefs can be justified while being false. Goldman's complaint about theories of rationality that are predicated on non-*verific* consequentialist theories of the criterion of correctness is that they can yield beliefs that are not true. But being not true is not in any way a criticism of a belief that is otherwise justified. Are we begging the question against *verific* consequentialist theories here? We do not think so. The kind of judgments we would make of intellectual achievement if we insisted that justified beliefs must be mostly true beliefs would be completely bizarre. We would end up insisting that most of what was believed about the natural world for the last 6000 years was not only false but also irrational. Worse, we would be in the position of thinking that our own scientific beliefs about the world are irrational, since we have some inductive reason to think that they are probably false. That seems completely wrong (see also Cohen 1984).

<sup>7</sup> One problem with Alston's view that we do not pursue here is that it relies on a theory of objective probability without actually offering one. We have persistent doubts that a theory of objective probability that is appropriate for epistemic justification can be constructed. See (Pollock 1989).

The same problem is in stark evidence in Plantinga's work. Plantinga is also a norm externalist. He proposes that:

A belief has warrant for me if (1) it has been produced in me by cognitive faculties that are working properly...in a cognitive environment that is appropriate for my kinds of cognitive faculties, (2) the segment of the design plan governing the production of that belief is aimed at the production of true beliefs, and (3) there is a high statistical probability that a belief produced under those conditions will be true (1993b, pp. 46-7).

(1) is a distinctive feature of Plantinga's view that has attracted a great deal of attention. Plantinga himself heroically attempts to fend off some of the most pressing difficulties faced by the notion of proper function (1993b, chapter 2). (3) is an attempt to address generality problems faced by reliabilism. Plantinga aims to meet these challenges by proposing an account of objective probability (1993b, chapters 8 and 9). Though we have flagged our concern about objective probability and while Plantinga's defense of proper functions is ripe for continued careful scrutiny, we do not wish to challenge (1) or (3) here. (2) is the most relevant for our project.

Apparently, Plantinga thinks that an orientation toward truth is simply demanded by having warrant as one's subject matter. Unfortunately, there is a perennial problem in dealing with Plantinga's theory in a discussion of internalism and externalism. He is at pains to emphasize that he seeks a theory of warrant and not a theory of justification, where warrant is whatever yields knowledge when added to true belief. That makes it so that warrant has to itself incorporate truth in order to allow knowledge to meet Gettier counterexamples in a way that is reminiscent of some now quite old proposals on knowledge (e.g., Lehrer and Paxson 1969). More importantly, this project makes 'warranted' not at all the same thing as 'rational.' Rationality, in our view, is not about helping knowledge to meet the Gettier problem. Perhaps all of this is just as Plantinga intends. That is, perhaps Plantinga does not mean to be talking about rationality at all. One might think that this is completely evident on the basis of his rejection of justification, but Plantinga rejects justification because he thinks that it has an ineliminable deontological cast. We do not think that that is the crucial thing about justification, even if talking about justification in terms of duty is a useful way to make connections to traditional epistemology, and a useful way to capture the notion of epistemic permissibility (as we do in Pollock and Cruz, p. 123). Our view is that justification is consequentialist while at the same time being distinct from warrant.<sup>8</sup>

Insofar as Plantinga might be thought to be discussing the topic we are concerned about, he sometimes marshals counterexamples to non-verify accounts of the correctness of epistemic norms. These counterexamples are really no different than Goldman's fantasizer counterexamples. For example, Plantinga offers Paul, who "suffers from a brain lesion induced by radioactive fallout from a Soviet missile test. He now reasons differently from the rest of us; when appeared to in the church-bell fashion, he forms the belief that something is appearing to him in that fashion, and that it is orange"

8 Furthermore, if Plantinga's warrant has nothing at all to do with rationality, then he is not entitled to criticize Pollock's view as not providing a theory of warrant (as he does in 1993a, chapter 8). After all, Pollock never meant to be talking about warrant.

(p. 170). Paul is reasoning according to his norms, but Plantinga is worried that Paul's norms are incorrect. Plantinga, also following Goldman, sees fit to point out here that Paul's reasonings according to his pathological norms would not, when added to true belief, yield knowledge. This leads Plantinga to say that a non-*verific* account of the criterion of correctness will not do. If Plantinga means to be giving an account of rationality, it appears that Plantinga, like Goldman, is conflating rational beliefs with beliefs that are true or are somehow very closely connected with truth.

Goldman's and Plantinga's appeal to knowledge is spurious, and it is precisely here where the truth-aimed motivation for externalism shades into a mistaken reverence for the analysis of knowledge. Traditionally, knowledge requires justified true belief plus some measure to guard against Gettier counterexamples.<sup>9</sup> Thus, on the face of it, truth and justification are independent elements of knowledge. It is inappropriate to demand that a theory of justification needs to bring a belief closer to knowledge by incorporating truth. The role of a theory of justification in knowledge is to distinguish between beliefs that are true but only accidentally so, versus beliefs that are true because the believer has some positive reason to think so (Unger 1975). Goldman's defense of *verific* accounts of the criterion of correctness therefore seems in grave danger of conflating rational beliefs with true beliefs. This is just the charge we made against belief externalism above. Though it is not always easy to see, externalists appear to be repeatedly making the same error.

Norm externalists think that being aimed at the truth is the criterion of correctness for epistemic norms. It appears that the reason that they think this is that they are driven by a concept of knowledge that requires something more than true belief, but where what must be added must still be intrinsically truth-aimed. We doubt that any such concept of knowledge can be made sense of. There are at least two concepts of knowledge that have drawn the attention of epistemologists. One is a kind of connection between the knower and the fact known where truth is the sole important element. Thus, true belief is sufficient for knowledge of this sort.<sup>10</sup> The second concept of knowledge relevant to epistemology—and the concept that has received more attention—brings together true belief with justifying reasons. It is important to see how this second concept is supposed to operate. The justifying reasons are seen as the mechanism by which the genuine knower can be distinguished from the person who believes the truth, but does so accidentally or spuriously. Having reasons will ensure that the knower is brought along to the belief according to standards that are rationally defensible as procedures for belief formation. Unfortunately, coming to a true conclusion through rationally defensible procedures for belief formation can sometimes still fail to yield knowledge. This is the force of Gettier counterexamples (1963). The problem

9 Here we are restricting our attention to 'justificationist' strategies for analyzing knowledge. There are at least three other prominent ways with knowledge, however, namely causal theories, tracking theories and contextualist theories. See, e.g., Goldman (1970), Nozick (1981) and DeRose (1991) and Lewis (1996) respectively.

10 Lehrer (1990) discusses this concept of knowledge, and it can be found in Goldman's relatively recent work on social epistemology (1999).

is that the truth believed and the rational procedures used may well be disconnected in the sense that the rational procedure accidentally arrives at truth. For example, in the classic Brown in Barcelona case, Smith reasons impeccably from his justified belief that Jones owns a Ford to arrive at the belief that either Jones owns a Ford or Brown is in Barcelona. Gettier points out that it can turn out that it is the second disjunct that makes the belief true, while the first (plus the innocent disjunction introduction) is what makes the belief justified. That the two quantities are present in the same belief is an accident, and so the specter of spurious epistemic achievement which initially motivated the second concept of knowledge looms again. Solving the Gettier problem amounts to forging a connection between the justifier and whatever makes for the truth of the belief.

In our view, the Gettier problem is a curiosity, and the vast literature that it has inspired has shown that the second concept of knowledge is not at all well behaved. The point here, however, is that the externalist treats knowledge as being composed of true belief plus a notion of justification that also has truth-aimedness built in. It is far from obvious how building in truth aimedness is going to help, though, since it looks as if Gettier problems are going to be replicated at the level of justification itself unless justification is identical to the truth. This is simply and inevitably a consequence of justification and truth being different. Thanks to Gettier, the difficulty is now obvious in the case of the second concept of knowledge but it should be clear in the case of externalism's concept of knowledge, too. Externalism grasps for a concept of knowledge where the justification component is sufficiently separate from truth to maintain a distinction between mere true belief and knowledge, but where justification is at the same time sufficiently attached to truth so that having it will make knowledge likely and attainable. We are skeptical that there is any sense to be made of this blended sense of justification and the concept of knowledge that drives the externalist to hold out for it. Success in articulating this sense of justification appears just as likely as success in solving the Gettier problem.

In summary, one can be an externalist by being either a belief externalist or a norm externalist. The belief externalist tries to formulate epistemic norms directly in terms of externalist considerations, but it is impossible to construct non ad-hoc procedural norms in this way. The norm externalist proposes instead to recommend evaluations of procedural norms on the basis of considerations of reliability. Norm externalism initially appeared compelling because it provided a way to preserve the internal and procedural nature of epistemic norms while still allowing for external assessment of those norms. Combined with internalism's apparent inability to make sense of the comparative evaluation of norms, it seemed that norm externalism was very promising. Unfortunately, norm externalism conflates justification and truth or relies on a concept of knowledge that is suspect. So, norm externalism must be rejected.

As far as we can see, externalism has nothing to contribute to the solution to traditional epistemological problems.<sup>11</sup>

<sup>11</sup> Might the externalist be stalking a completely different epistemic achievement that has accidentally been called 'justification'? This is one way of reading the claim in, e.g., Goldman (1988) where he distin-

### 5. *A Naturalistic Internalism*

We have seen that externalism is motivated by the view that rationality must have a close association with truth. The externalist has tried to forge that close bond by attempting—either at the level of the norms or at the level of the criterion of correctness for the norms—to make belief-forming processes answer to truth ‘directly’. That is, externalists about epistemic justification have tried to build into their theories a metaphysical connection to the truth that is much stronger than an epistemic agent’s internal norms with regard to what it takes to be reliable or true. In rejecting this on behalf of internalism, it may seem that we are rejecting truth as the normative aim of epistemology. In a way, this is correct. Our conclusion is that truth is not a very interesting part of our best understanding of rationality. This is out of step with a very common view on the part of epistemologists, both internalists and externalists alike.

The most common response that we have encountered to the rejection of a close tie between rationality and truth is the question, “what else is there for rationality to answer to?” The worry is that, if we do not think that the goal of rationality is to secure truths, then we have somehow missed the only imaginable target of rationality or that we have opened the door to a completely subjective or arbitrary conception of rationality. Moreover, the *prima facie* case for truth-aimed epistemology is clear. When we visit the doctor, we hope that she will form beliefs that are true. Even if her diagnoses sometimes fall short of truth, truth is what we expect her goal to be. In more mundane cases of belief formation we may find ourselves with extremely little evidence on which to base our beliefs. The evidence that we finally go with, though, appears to have a little something extra in its favor, and it seems plausible to think that that extra something is higher likelihood of pointing to the truth.

Describing belief formation in this way is intelligible if the unit of analysis is the *person*. In a fairly straightforward sense, people have aims, and sometimes, perhaps often, people have the aim of reaching the truth. Epistemologists, however, typically do not focus on the personal level in their discussions of belief formation.<sup>12</sup> If the commitment is to be taken seriously with respect to belief formation, the ‘aims’ talk will ultimately need to be replaced by a specification of *how* belief formation aims at the truth when the relevant level is the subpersonal. Our understanding of the way that epistemologists have talked about subpersonal belief formation is through epistemic norms. When we take account of the role of epistemic norms in belief formation, it becomes much more difficult to sustain the claim that the aim of belief formation is truth because the norms themselves identify a proximal endpoint for the formation of beliefs. Aims, goals, and targets are appropriate for agent-level discussions of explicit decisions, as in

---

guishes weak and strong justification. There was a time when we were inclined to accept that internalist and externalist epistemologists were simply after two different things. It is now not so clear what the point of externalist theories of justification might be, except to accommodate the externalist’s problematic intuitions about the concept of knowledge. At any rate, we would be happier with externalism if externalists themselves regarded their enterprise as simply different from traditional theorizing about rationality. Our experience is that they rarely embrace this view.

12 It seems that some participants in the debate have been misled into thinking otherwise again by concentrating on the case of knowledge, where the level of analysis *is* often enough the personal level.

the case of the doctor. We hope that *she* has the goal of achieving truth, as opposed to merely operating in a way that avoids malpractice suits. The metaphorical extension of this goal to her epistemic norms is strained, however. Beliefs formed in accordance with epistemic norms obey the constraints and demands of the norms themselves.

Thus, there is an important and compelling alternative for the overall normative aim of belief, namely that the formation of beliefs follows the directives of the epistemic norms. That this is so should not be a surprise. When epistemic norms are introduced as an intermediary between the raw materials for belief formation and the truth, there is inevitably the possibility of treating the norms themselves as the proximal endpoints of justified beliefs. Showing that truth is the distal aim of justified belief—with norms as the vehicle—would then require a further argument on the part of the externalist. Arguments of that sort were shown above to be inadequate.

The framework tells against another kind of attempted vindication that has been advanced for truth-aimed epistemology. It might be thought that a commitment to veritism offers a way to mark off the distinctive nature of *epistemic* reasons (Harman 1997). The thought seems to be that, if belief formation is not aimed at truth, then there is no way to distinguish epistemic reasons from non-epistemic reasons. The reply to this claim is now straightforward, however. The argument only goes through if it can be shown that there is no adequate way of characterizing epistemic reasons apart from their truth-aimedness. But that is patently not the case given the nature of epistemic norms. Epistemic reasons will be just those contentful states that figure in epistemic norms. The norms, in turn, are identified by their output of epistemically laudable belief.

We conclude from this that a truth-aimed externalist epistemology is not entitled to carry the day solely on the allegation that it is intrinsic to beliefs or the processes of belief formation that they must aim at truth. Instead, the right picture in our view involves truth being a person's aim (at least some of the time), while beliefs are formed through epistemic norms that only answer to themselves as evaluative standards.

At the start of this essay, we claimed that a final motivation for externalism is its apparent affinity with naturalistic approaches in epistemology (Goldman 1994; 1999; Kitcher 1992). The reason that externalism has been viewed to have an advantage in this regard is partly an accident of the literature. As a matter of recent historical fact, it has primarily been externalists that have proposed self-consciously naturalized epistemologies. Another reason naturalism and externalism are sometimes associated is that the best know naturalized epistemology—that of Goldman—is also the best-known externalist epistemology. As we indicated above, his view articulates a crucial role for cognitive belief-forming processes in a theory of justification, and he has proposed that it is the job of cognitive science to understand those processes. Since cognitive science is a naturalistic inquiry, Goldman's process reliabilism has been seen to be especially amenable to naturalistic motivations. There is room here to worry about the specific proposals of naturalistically inclined externalists. We propose, however, simply to show that externalism does not have exclusive rights to naturalism. Thus we end with a sketch of an internalist theory whose contours reflect the anti-externalist considerations above and is at the same time plausibly naturalistic. (For the full discussion, see Pollock 1986; 1995; Pollock and Cruz 1999).

We take it that theories of human rationality describe contingent psychological features of the human cognitive architecture, though it is not entirely an accident that humans are built the way they are. Environmental pressures have led to our evolving in particular ways. We represent one solution to various engineering problems that were solved by evolution. There is no reason to think that these problems always have a single, or even a single best, solution. Certain engineering problems call for arbitrary choices between solutions that work equally well. On the other hand, many of the more general features of our cognitive architecture may reflect the only, or one of a small number of, solutions to general problems of cognitive engineering. For example, it can be argued that very general logical and computational constraints on cognition dictate that any sophisticated cognitive agent will engage in defeasible reasoning, will reason defeasibly from perceptual input, will reason inductively, and will engage in certain kinds of planning behavior (Pollock 1995). When that is true, it provides an informative explanation for why human cognition works as it does.

A simplistic view of evolution has it that evolutionary pressures select for traits that enhance survivability of an organism. In describing a naturalistic alternative to externalism, we do not need to go beyond this simplistic view, but it should be kept in mind that a serious appraisal of evolution will reveal a much more subtle theory than we are relying on here. We might regard the design problem for rationality to be that of creating an agent that can survive in a hostile world by virtue of its cognitive capabilities. This, however, presupposes a prior understanding of what cognition is and how it might contribute to survivability. We take it as characteristic of rational cognition that a cognitive agent has doxastic states (“beliefs”, broadly construed) reflecting the state of its environment, and conative states evaluating the environment as represented by the agent’s beliefs. It is also equipped with cognitive mechanisms whereby it uses its beliefs and conations to select actions aimed at making the world more to its liking. A rational agent has beliefs reflecting the state of its environment, and it likes or dislikes its situation. When it finds the world not entirely to its liking, it tries to change that. Its cognitive architecture is the mechanism whereby it chooses courses of action aimed at making the world more to its liking. Within this “doxastic-conative” loop, epistemic cognition is, in an important sense, subservient to practical cognition. The principal function of cognition is to direct activity (practical cognition), and the role of epistemic cognition in rationality is to provide the factual background required for practical cognition.

We can evaluate how well rational agents implement the doxastic-conative loop. A judgment of how well an implementation performs must always be relative to a set of design goals that one implementation may achieve better than another. As remarked above, a natural design goal is to construct an agent capable of using its cognitive capabilities to survive in an uncooperative environment. This is motivated by considerations of evolution and natural selection. There is probably no privileged design goal in terms of which to evaluate cognitive architectures, and truth certainly does not recommend itself as the only consideration.

The key point is that cognition is evaluated relative to practical goals. Thus, our view is a consequentialist view. We can accept the arguments for a consequentialist conception of epistemic norms offered by externalists like Goldman and Plantinga.

Where Goldman and Plantinga go wrong is in their insistence that attaining truths is the only sensible goal of rational cognition. The overall human cognitive architecture is what it is, and human rationality is defined relative to it. We can evaluate cognitive architectures in terms of how well they achieve their design goal, and this is reminiscent of norm externalism's evaluation of epistemic norms in terms of reliability. But one big difference is that practical and epistemic cognition are evaluated as a package. The ultimate objective is not truth, but practical success through the operation of epistemic norms.

Our diagnosis of the attraction of externalism (simpliciter) is that it accommodates the widely held conviction that rationality is extremely closely connected with the attainment of truths. This has led externalists to propose a connection between truth and rationality that threatens to blur the distinction between the two. We think that it is misguided to forge the connection so closely, but without a connection to truth, there is an apparent lacuna in the theory of rationality, namely the puzzle about what rationality answers to. We propose that this lacuna is filled with an account of the relationship between practical and theoretical cognition sketched above. This broadly naturalistic framework seems most faithful to the goal of distinguishing rational beliefs from irrational beliefs within internalist framework.<sup>13</sup>

### References

- Alston, W. (1986). Internalism and Externalism. In *Epistemology. Philosophical Topics*, 14(1).
- Alston, W. (1988). An Internalist Externalism. *Synthese*, 74, pp. 265-283.
- Alston, W. (1995). How to Think About Reliability. *Philosophical Topics*, 23, pp. 1-29.
- Audi, R. (2001) *The Architecture of Reason*. New York: Oxford UP.
- BonJour, L. (1980). Externalist Theories of Empirical Knowledge. In P. A. French, T. E. Uehling, & H. K. Wettstein (eds.), *Midwest Studies of Epistemology V: Studies in Epistemology*. Minneapolis, MN: University of Minnesota Press.
- BonJour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge: Harvard UP.
- BonJour, L. (1999) The Dialectic of Foundationalism and Coherentism. In J. Greco and E. Sosa (eds.), *The Blackwell Guide to Epistemology*, Malden: Blackwell.
- Chisholm, R. (1989). *Theory of Knowledge*. (3<sup>rd</sup> ed.), Englewood Cliffs, NJ: Prentice-Hall.
- Cohen, S. (1984). Justification and Truth. *Philosophical Studies*, 46.
- Conee, E., & Feldman, R. (1998). The Generality Problem for Reliabilism. *Philosophical Studies*, 89(1), pp. 1-29.
- Cruz, J. (unpublished manuscript). *Epistemic Norms and the Sellarsian Dilemma for Foundationalism*.
- DeRose, K. (1995). Solving the Skeptical Problem. *The Philosophical Review*, 104, pp. 1-52.
- Feldman, R. (1985). Reliability and Justification. *The Monist*, 68(2), pp. 159-174.
- Field, H. (2000). Apriority as an Evaluative Notion. In Boghossian, P. and Peacocke, C. (eds.), *New Essays on the A Priori*, New York: Oxford UP.
- Foley, R. (1987). *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard UP.
- Foley, R. (1993). *Working Without a Net: A Study of Egocentric Epistemology*. Oxford: Oxford UP.
- Gettier, E. (1963). Is Justified True Belief Knowledge? *Analysis*, 23, pp. 121-23.

13 Thanks are owed to Jim Pryor, Tom Kelly, and Melissa Barry for conversation and comments on the themes of this essay.

- Goldman, A. (1979). What is Justified Belief? In G. Pappas (ed.), *Justification and Knowledge*. Boston, MA: D. Reidel.
- Goldman, A. (1980). The Internalist Conception of Justification. In P. A. French, T. E. Uehling, & H. K. Wettstein (eds.), *Midwest Studies in Philosophy V*, pp. 27-51, Minneapolis: University of Minnesota Press.
- Goldman, A. (1985). The relation between epistemology and psychology. *Synthese*, 64, pp. 29-68.
- Goldman, A. (1986). *Epistemology and Cognition*. Cambridge: Harvard UP.
- Goldman, A. (1994). Naturalistic Epistemology and Reliabilism. In P. A. French, T. E. Uehling, & H. K. Wettstein (eds.), *Midwest Studies in Philosophy XIX: Philosophical Naturalism*, pp. 301-320, Notre Dame, IN: University of Notre Dame Press.
- Goldman, A. (1999). Internalism Exposed. *The Journal of Philosophy*, 46(6), pp. 271-93
- Goldman, A. (1999b). *Knowledge in a Social World*. Oxford: Oxford UP.
- Harman, G. (1997). Pragmatism and Reasons for Belief. Reprinted in *Reasoning, Meaning, and Mind*, pp. 93-116, Oxford: Clarendon Press.
- Kitcher, P. (1992). The Naturalists Return. *The Philosophical Review*, 101(1), pp. 53-114.
- Lehrer, K. & Paxon, T. (1969). Knowledge: Undefeated justified true belief. *Journal of Philosophy*, 66, pp. 225-37.
- Lewis, D. (1996). Elusive Knowledge. *Australasian Journal of Philosophy*, 74, pp. 549-67.
- Moser, P. (1985). *Empirical Justification*. Dordrecht, Holland: D. Reidel.
- Nozick, R. (1981). *Philosophical Explanations*. Cambridge, MA: Harvard UP.
- Plantinga, A. (1988). Positive Epistemic Status and Proper Function. In J. E. Tomberlin (ed.), *Philosophical Perspectives*, 2, Epistemology, 1988, pp. 1-50, Atascadero, CA: Ridgeview.
- Plantinga, A. (1990). *Justification in the 20th Century. Philosophy and Phenomenological Research 1*(supplement), pp. 45-71.
- Plantinga, A. (1993a). *Warrant: The Current Debate*. New York: Oxford University Press.
- Plantinga, A. (1993b). *Warrant and Proper Function*. New York: Oxford University Press.
- Pollock, J. (1974). *Knowledge and Justification*. Princeton: Princeton University Press.
- Pollock, J. (1984). Reliability and Justified Belief. *Canadian Journal of Philosophy* 14, pp. 103-114.
- Pollock, J. (1986). *Contemporary Theories of Knowledge*. Savage, MD: Rowman and Littlefield Pub.
- Pollock, J. (1990). *Nomic Probability and the Foundations of Induction*. New York, NY: Oxford University Press.
- Pollock, J. (1995). *Cognitive Carpentry*. Cambridge, MA: MIT Press.
- Pollock, J. & Cruz, J. (1999). *Contemporary Theories of Knowledge*, 2<sup>nd</sup> ed., Lanham, MD: Rowman & Littlefield Pub.
- Sosa, E. (1980) The Raft and the Pyramid. In *Midwest Studies in Philosophy Vol. V: Epistemology*
- Unger, P. (1975). *Ignorance: A Case for Skepticism*. Oxford: Clarendon Press.

