

EXTERNALIST JUSTIFICATION WITHOUT RELIABILITY

Michael Bergmann
Purdue University

Externalist analyses of epistemic properties typically include some sort of reliability requirement. Such analyses applied to *justification* fall prey to a simple yet persuasive counterexample: the beliefs of a victim of a deceptive demon can be justified despite their being formed in completely unreliable ways.¹ In this paper I will propose a way of analyzing justification that enables us to hang on to the externalism without the reliability requirement. As an added bonus, the proposed analysis of justification yields an account of the elusive connection between justification and truth.²

I begin with the assumption, congenial to many internalists, that the following *evidentialist* proposal is *prima facie* plausible and at least *close* to being a correct account of at least *one* important sort of epistemic justification:

E_F : S's belief B is justified iff B is a *fitting* doxastic response to S's evidence.³

In the first part of the paper I argue that the notion of fittingness employed in E_F is best understood in terms of the notion of proper function. This conclusion lends support to the nonreliabilist externalist analysis of justification I propose and defend in the second part of the paper. For since that analysis is a refinement of E_F understood in terms of proper function, I expect it to share to a large extent in E_F 's *prima facie* plausibility. The final section deals with objections.

I. Fittingness and Proper Function

My path to defending a nonreliabilist externalism about justification begins with the evidentialist thesis, E_F , and makes its way first to the

conclusion of section I: that E_F is best understood in terms of proper function. Let's begin, then, by trying to understand evidentialism.

A. Evidentialism

What exactly is a subject's evidence? I take it that it is limited to what the subject is aware of.⁴ Thus, in order for a belief to be a fitting doxastic response to a subject's evidence, it must be a fitting doxastic response to what that subject is aware of (or to some relevant part thereof). I will assume that we have some grasp of what it is for a belief to *fit* one's evidence and that, with respect to certain clear cases, we have some fairly firm intuitions concerning whether the belief in question does or doesn't fit the subject's evidence. The very ease with which we make sense of the examples used by evidentialists to illustrate their position testifies to the plausibility of these last two assumptions. When, for example, Feldman and Conee say that the belief that *there is something green before one* fits the evidence that a normal person has in ordinary circumstances when looking at a plush green lawn in broad daylight or that the belief that *sugar is sour* does not fit our gustatory experience, we have no trouble understanding or accepting such claims.⁵

A standard line taken by evidentialists is endorsement of the following three claims:⁶

Nonreliability: the fittingness of doxastic response B to evidence E is not contingent upon E's being a reliable indicator of B's truth.

Objectivity: the fittingness of doxastic response B to evidence E is objective fittingness (in the sense that fittingness from the subject's perspective isn't sufficient for it).

Necessity: the fittingness of doxastic response B to evidence E is an essential property of that response to that evidence.⁷

Nonreliability is extremely plausible. Just as it seems possible for a demon victim to have justified beliefs, so also it seems possible for her beliefs to fit her evidence. Her only problem is that her evidence isn't rightly connected to the world. So I join the evidentialist in accepting Nonreliability. I also accept Objectivity (which will be explained in some detail below). Consequently I follow the evidentialist in denying that objective fittingness depends on reliability. But then what *does* it depend on? The evidentialist's answer to this question is given by Necessity: the objective fittingness of doxastic response B to some evidence E is an *essential* property of that response to that evidence; it isn't a contingent property that depends on some other condition being satisfied. Here is where the evidentialist and I part company.

Notice first that the fact that the fittingness of doxastic response **B** to evidence **E** isn't contingent upon **E**'s being a reliable indicator of **B**'s truth doesn't prove that it isn't contingent upon anything. So Nonreliability doesn't entail Necessity. Thus, we aren't *forced* to give the Necessity-inspired evidentialist answer mentioned above to the question: "If objective fittingness doesn't depend on reliability, what does it depend on?" My answer to that question is that, contrary to what Necessity suggests, objective fittingness depends on proper function. In section I.D I will draw upon some suggestions of Thomas Reid's in order to develop an objection to Necessity⁸—an objection that points to proper function as that on which objective fittingness depends. Before getting to that objection, however, I'll need to say a little more, in the next two subsections, about Objectivity. For unless we first get clear on certain implications of Objectivity, I won't be able to explain the counterexample to Necessity that I have in mind.

B. Objectivity

According to Objectivity, the fittingness *from the subject's perspective* of doxastic response **B** to her evidence isn't sufficient for **B**'s *actually* fitting her evidence. In order to explain this claim, it will be helpful to distinguish between the main evidence for a belief, on the one hand, and connectors and disconnectors on the other.

Consider a person, Jane, who has managed somehow to become an adult without any exposure (directly or via testimony) to the phenomenon of a straight object looking bent when immersed in water. And suppose that Jane for the first time comes upon a straight stick immersed in water and, upon seeing it, forms the false belief that it is bent. The natural thing to say is that Jane's evidence for this belief is her visual experience (which, we may assume, is like the visual experience you and I have when we see such a thing). Consider next Tim, a logically perceptive man who learns from an informer who delights in obfuscation that

- (1) If John goes to the party then {if Judy goes to the party then {if Jan goes to the party then the party will be a lot of fun}}.

Upon learning this, Tim immediately forms the belief that

- (2) If John and Judy and Jan go to the party, then the party will be a lot of fun.

In this case, the natural thing to say is that Tim's evidence for his belief that (2) is his belief that (1). I'll refer to evidence of this sort—the sort we are naturally inclined to point to when identifying a person's evidence—as 'the main evidence' for a belief.

Now, some people think that the main evidence for our beliefs doesn't always constitute our total relevant evidence for them. For example, in the

Jane case, there may be, in addition to the visual experience on which she bases her belief that the stick is bent, the strong *felt* inclination to take her visual sensations as indicative of the truth of the belief in question.⁹ Those of us who are completely familiar with the phenomenon of water distorting the appearance of a straight stick lack this felt inclination (at least it isn't as strong in us). As a result, some think our total evidence is somewhat different from Jane's. We can call Jane's strong felt inclination a 'connector' between her main evidence and her belief based on that evidence. It's a connector we lack. Another (related) difference is that we are inclined to think that the main evidence on which Jane relies is *not* indicative of the truth of the belief she bases on it. This provides us with a *disconnecter* between the main evidence and the belief in question—a disconnecter Jane lacks.¹⁰ So if we think of these connectors and disconnecters as parts of one's total evidence, then *our* total evidence when we view a stick looking bent in water is different from *Jane's* when she considers such a sight (even if the main evidence—the way it visually appears—is the same).

In the Tim case there could be a connector too. Let us suppose it is the strong felt inclination to take the truth of (1) to entail the truth of (2). It is easy to imagine those less logically perceptive than Tim failing to see this sort of connection. We could describe this difference by saying that although these others have the same main evidence Tim has, they don't have the same total relevant evidence since they lack a connector he possesses.¹¹

With this terminology at our disposal, we can return to our discussion of Objectivity. The view with which proponents of Objectivity disagree is the subjectivist view according to which a belief fits the subject's evidence if (though perhaps not only if) the subject's evidence consists of both her main evidence and a connector that connects her main evidence to her belief. In opposition to this suggestion, the supporter of Objectivity claims that if the subject's belief fails to fit the subject's main evidence, it won't help merely to add to her evidence a connector connecting that main evidence with the belief. Consider, for example, the following belief and pieces of evidence:

Belief	Main Evidence	Connector
B1. The first person belief: "There is smallish hard round object in my hand".	ME1. Tactile sensations of the type you experience when you grab a billiard ball.	C1. The strong felt inclination to take ME1 to be indicative of the truth of B1.
	ME2. Olfactory sensations of the type you experience when you smell a meadow full of flowers. ¹²	C2. The strong felt inclination to take ME2 to be indicative of the truth of B1.

The subjectivist opponent of Objectivity might acknowledge that B1 fits ME1 and that it doesn't fit ME2. But she will add that B1 *does* fit the combined evidence of ME2 together with C2 because by adding C2 to ME2 (and adding no disconnectors¹³) we get subjective fittingness. The proponent of Objectivity rejects this suggestion. She says that since B1 doesn't fit ME2, it won't help merely to add C2 even if we stipulate that no disconnectors are present. The objective failure of B1 to fit the subject's evidence (ME2) isn't altered by adding C2 to the subject's evidence base; if B1 fails to fit ME2, it also fails to fit ME2 + C2 (the combination of sensation ME2 and connector C2).

Paradigm evidentialist Richard Feldman appears to be endorsing Objectivity so construed when he points out that one can't get justification for one's beliefs merely by thinking one has good reasons for them.¹⁴ Other evidentialists such as Bruce Russell, Paul Moser and Richard Fumerton also seem to agree with Objectivity. For although they say that reliability isn't sufficient for E's being good evidence for B, they pay the same compliment to *thinking that E is good evidence for B*.¹⁵ In each case, the idea seems to be that if a belief doesn't fit one's main evidence, then merely adding a connector joining that belief with that main evidence won't help.

C. *Learned and Unlearned Doxastic Responses*

A possible concern about Objectivity arises in connection with its suggestion that if a belief doesn't fit the subject's main evidence then merely adding a connector won't help. Consider B1 and ME2. Why couldn't a person learn, in the way one learns to associate certain smells with gasoline or paint (i.e., by experiencing their constant conjunction), to associate olfactory experience ME2 with the truth of B1? Then one would have learned to have C2 whenever one has ME2, in which case B1 could be a fitting response to ME2 + C2 after all.

That concern about Objectivity strikes me as a sensible one. To handle it, we need to distinguish learned from unlearned doxastic responses. The distinction isn't easy to draw, but it is something like this. *Learned* doxastic responses, such as an experienced birdwatcher's immediate bird identifications after a quick look (or listen), are ones a person comes to have only after first finding out independently (i.e., without relying in any essential way on other instances of that same type of doxastic response) that there is a correlation between the truth of such beliefs and the experiences to which they eventually become immediate responses. By contrast, an *unlearned* doxastic response to experience is a hardwired or automatic response that occurs (perhaps only after a certain level of cognitive development) without the subject first independently finding out that there is a correlation between the truth of the belief in question and the experience to which it is a response.¹⁶

We can also distinguish learned from unlearned *connectors*. Suppose C is a connector that connects main evidence ME with a belief B. If C is acquired by first independently finding out that there is a correlation between the truth of B and the occurrence of ME, then C is a *learned* connector. However, if C isn't acquired by first learning independently of a correlation between the truth of B and the occurrence of ME—if, instead, a person simply has C without learning independently that there is such a correlation—then C is an *unlearned* connector. With these distinctions in mind, we can state the claim of Objectivity more carefully as follows: some unlearned doxastic responses to one's evidence are unfitting and merely adding an *unlearned* connector to one's evidence won't change that.

D. Necessity

So far I'm in agreement with the evidentialist. I accept Objectivity. And, as I noted earlier, because I also accept Nonreliability, I follow the evidentialist in denying that objective fittingness depends on reliability. This brings us to the all-important question: "What *does* it depend on?" You will recall that the evidentialist's Necessity-inspired answer was that the objective fittingness of doxastic response B to some evidence E is an *essential* property of that response to that evidence, not a contingent property that depends on some other condition being satisfied. In this subsection, I will be arguing that Necessity is false.

To do this, I will need to employ some assumptions implicit in Objectivity. In particular, I will be assuming that there can be unlearned doxastic responses that are objectively fitting as well as ones that are objectively *unfitting* (in the ways described in the previous two subsections).¹⁷ What follows is a description of a possible case in which the billiard ball belief B1¹⁸ is a fitting unlearned doxastic response to ME2 and an unfitting unlearned doxastic response to ME1. Such a case is a counterexample to Necessity because in the actual world, B1 is a fitting unlearned doxastic response to ME1¹⁹ and an unfitting unlearned doxastic response to ME2.

Thomas Reid emphasized that there doesn't seem to be any logical connection between our sense experiences and the content of the beliefs based on them. For example, the tactile sensations we experience when touching a hard surface seem to have no logical relation to (nor do they resemble) the content of the hardness beliefs they prompt. In light of this he said that "no man can give a reason why the sensations of smell, or taste, or sound, might not have indicated hardness".²⁰ Considering the matter in the abstract, tactile sensations do not seem to be any more suited than olfactory sensations to being indicators of hardness. Thus, it seems there could have been cognizers like us in outward appearance who experience, upon grabbing a billiard ball, a sensation that is qualitatively of the same type as one of our actual world sensations of smell.²¹ And it seems possible that the natural *unlearned* doxastic response of

such a cognizer to that “olfactory”²² sensation is the first-person belief “There is a smallish hard round object in my hand” (i.e., B1). There is nothing about the process of *grabbing a billiard ball then experiencing ME2 then holding B1* that makes it intrinsically less suitable (as a natural unlearned process for a cognizer to undergo) than the process of *grabbing a billiard ball then experiencing ME1 then holding B1*. In each case, there is a causal chain from an external stimulus to an experience to a belief. And in each case, the experience has the same functional role of connecting that stimulus with B1. The only difference is that in the one case, the experience playing this functional role is ME1 whereas in the other it is ME2.²³

This Reidian example suggests *two* things: first, that it is possible for billiard ball belief B1 to be an *unlearned* doxastic response to “olfactory” sensation ME2; second, that it is possible for such an unlearned response to be natural for a cognizer, even an entire species of cognizers. The first claim seems relatively uncontroversial—the possibility of a certain kind of cognitive malfunction in humans entails it. But the second seems plausible too.²⁴ It is no less plausible than the possibility of a species of cognizers whose experienced color spectrum is inverted with respect to ours.

It might be helpful here to say something in support of a crucial feature of this objection to Necessity, namely, its suggestion that brain-damaged humans could have the *same* evidence base for B1 (i.e., ME2) as do the possible cognizers in the Reidian example. Consider first normal humans. It seems that, for them, B1 is an unlearned doxastic response to ME1—something that occurs without their first learning *independently* that hardness is correlated with such tactile experiences.²⁵ Similar remarks apply to the possible cognizers in the Reidian example: B1 is, for them, an unlearned doxastic response to ME2. And the same thing can be said of the brain-damaged humans: due to some sort of injury, B1 is, for them, an *unlearned* doxastic response to ME2. You might think that normal humans also have, as a part of their evidence base for B1, an unlearned connector C1 (where what this amounts to is that they have a sense that B1 is the appropriate belief to hold given the circumstances, i.e., their experiencing ME1). But we could simply stipulate that the possible cognizers in the Reidian example have in their evidence base for B1 a similar sort of *unlearned* connector (only it is C2 rather than C1). And we could add that, due to the same injury that causes them to form B1 in response to ME2, the brain-damaged humans also have an *unlearned* connector C2—a sense that B1 is an entirely appropriate belief to hold in their circumstances, i.e., their experiencing ME2. Furthermore, we could stipulate that in neither the Reidian case nor the brain-damaged human case is there anything else in the subjects’ evidence base that is relevant to their holding B1.²⁶

What should we say of a species of cognizers for whom the natural unlearned response to grabbing a billiard ball is to experience ME2 and then form B1? It seems we should say that for such cognizers, B1 is a fitting

unlearned response to ME2 and an unfitting unlearned response to ME1. This shows that the fittingness of an unlearned doxastic response is a contingent feature of it, a feature that depends in some cases on the species of the cognizer who has the response. Necessity is, therefore, false.²⁷

E. Proper Function

Having rejected Necessity, we must return to the question that prompted our discussion of it, namely: “If objective fittingness doesn’t depend on reliability, what does it depend on?” The counterexample to Necessity considered in section I.D suggests that the fittingness of a doxastic response depends, in some cases at least, on the species of the cognizer who has it. What is it about the species of a cognizer that determines such fittingness in those cases? An answer that immediately suggests itself is that what makes a belief a fitting unlearned doxastic response to an experience has to do with the way the cognitive faculties of the cognizer in question are *supposed to* function. For clearly that is something that can vary from species to species. *Our* cognitive faculties are supposed to function so that when we experience tactile sensation ME1, our unlearned doxastic response is B1. Not so the cognizers described in the previous subsection. *Their* faculties are supposed to function so that when those cognizers experience “olfactory” sensation ME2, their unlearned doxastic response is B1. The sense in which our faculties and theirs are *supposed to* function in the ways just specified is the same as the sense in which our hearts are supposed to function so that they beat less than 200 times a minute when we are at rest. And the ‘supposed to’ of heart function is clearly connected with the notion of proper function. This suggests that the fittingness of a doxastic response to evidence is contingent upon the *proper function* of the cognitive faculties of the person in question. And this, in turn, suggests that the evidentialist claim E_F is equivalent to something like the following (where a *PF-induced* response is one resulting from the proper functioning of the subject’s cognitive faculties):

E_{PF} : S’s belief B is justified iff B is a *PF-induced* doxastic response to S’s evidence.

II. Justification and Proper Function

A. Refining E_{PF}

E_{PF} isn’t intended as a careful analysis. It is meant to be a rough statement (in need of refinement) of the position suggested by combining

evidentialism with my objections to Necessity. The first refinement I want to recommend is that we expand our focus from evidence (thought of as something the subject is *aware of*) to something more general:

I_{PF} : S's belief B is justified iff B is a PF-induced doxastic response to the *input* to S's belief-forming systems.

What besides evidence counts as input to one's belief-forming systems? My proposal is that anything that is what John Pollock calls an 'internal state' of a believer counts as such input.²⁸ According to Pollock, a state or circumstance-type is internal if it is accessible (in a *nonepistemic* sense) to our "automatic processing systems", to the "cognitive mechanisms that direct our epistemic cognition"; it needn't be accessible to *us*.²⁹ The sense in which our cognition has this sort of access—the sense in which it "notes" that we are in these states—is "metaphorical. It is the same as the sense in which a computer program accessing a database might be described as noting that some particular item is contained in it."³⁰ Since the access to these internal states is neither *epistemic* nor access by *us*, we can require input in the form of such states without committing ourselves to internalism.

The reason I prefer I_{PF} to E_{PF} is that E_{PF} suggests that a necessary condition of justification is having evidence. And if having evidence is thought of as involving the subject's being aware of something, then E_{PF} seems to be imposing an awareness requirement on justification. Since I think there are severe problems associated with imposing awareness requirements on justification, I'd like to avoid any suggestion that I'm imposing one.³¹ Furthermore, since evidence is perhaps the most prominent sort of input to *our* belief-forming systems, there will be a wide range of standard cases in which the different implications of I_{PF} and E_{PF} are not noticeable. The result is that the *prima facie* plausibility of E_{PF} (which arises from a consideration of these standard cases together with the section I argument for the rough equivalence of E_F and E_{PF}) will carry over in large part to I_{PF} .

In the remainder of the paper I will propose and defend an analysis of justification that can be viewed as a more careful statement of the idea behind I_{PF} .

B. An Analysis of Justification

Plantinga has proposed the following analysis of *warrant* (i.e., that which makes the difference between knowledge and mere true belief):

W_{PF} : S's belief B is warranted iff each of the following conditions is satisfied:

- (i) the cognitive faculties producing B are *functioning properly*
- (ii) the cognitive environment in which B is produced is sufficiently similar to the one for which S's faculties were "designed"³²

- (iii) the modules of the “design” plan governing the production of B are directly aimed at the production of *true* beliefs
- (iv) there is a high objective probability that a belief formed in accord with those modules in that sort of cognitive environment is true.³³

He adds condition (iv) because he recognizes that the first three conditions don’t entail reliability, and he thinks warrant should entail reliability.

A similar line of reasoning leads me to think he should add a fifth condition we can call a ‘no-defeater condition’:

- (v) S does not take B to be defeated.

For warrant should entail a no-defeater condition, but Plantinga’s four conditions don’t.³⁴ There could be cognizers designed by a literal creator not only to form beliefs in a reliable way but also to take each of their reliably formed beliefs to be defeated. Furthermore, they could have been designed so that although they took their beliefs to be defeated, they ignored this fact and continued to hold them. Such beliefs would *not* be warranted even if they satisfied Plantinga’s four conditions.³⁵

My proposed account of justification is just my modified version of Plantinga’s account of warrant (i.e., his account together with the fifth condition I just mentioned) *without* Plantinga’s condition (ii)—the environmental condition. Once that condition is dropped, his account ceases to be reliability-entailing. Thus, to a first approximation, we may say that:

- J_{PF} : S’s belief B is justified iff (i) S does not take B to be defeated³⁶ and (ii) B is produced by cognitive faculties that are (a) functioning properly, (b) truth-aimed and (c) reliable in the environments for which they were “designed”.

To see how this analysis is a refinement of I_{PF} , it’s important to notice that clause (ii) is to be understood so that it entails that S’s doxastic response to the input to her belief-forming systems is due to the proper functioning of S’s faculties. As Plantinga points out, when we say cognitive faculties are functioning properly, the basic idea is that their functioning results in the appropriate doxastic response to the circumstances in which they are operating (which will include, rather prominently, the input to the subject’s belief-forming systems).³⁷

C. Motivating the Last Two Sub-Clauses of J_{PF}

Why does J_{PF} require that the belief be produced by *truth-aimed* cognitive faculties that are *reliable* in the environments for which they were

“designed”? Why not require, in addition to the no-defeater clause, merely that they be produced by *properly functioning* cognitive faculties? The answer here is very much like the explanation Plantinga gives for including clauses like (ii)(b) and (ii)(c) in his account of *warrant*.³⁸ The main difference is that when Plantinga adds such clauses to his account of warrant, they entail that the belief is produced by a cognitive faculty that produces mostly true beliefs, whereas they do not have that consequence when added to my account of justification. The reason that the addition of such clauses doesn't make my account reliability-entailing is that I've dropped the “right environment” clause included in Plantinga's account. J_{PF}, therefore, is not a *reliabilist* account of justification despite the fact that it includes a clause having to do with reliability. It is because it isn't a reliabilist account that it is able to avoid the standard evil demon objection to reliabilist accounts of justification mentioned in the opening paragraph of this paper.

To return, then, to the question at hand, let me explain why I've added clauses (ii)(b) and (ii)(c) having to do with the cognitive faculty in question being truth-aimed and reliable in the right environment. My inclusion of clause (ii)(b), requiring that the faculty be truth-aimed, is due to the following sort of counterexample to the suggestion that clauses (i) and (ii)(a) are sufficient for justification. Suppose that the cognitive faculties of the cognizer in question were created by some benevolent being interested in the well-being of the cognizer. However, suppose also that the particular faculty producing the belief we are considering is intended by its creator not to produce *true* beliefs but, rather, to produce beliefs that will minimize psychological trauma (even if that involves regularly producing false beliefs). Then it seems that beliefs being produced by such a cognitive faculty won't be *epistemically* fitting responses to the input to the subject's belief forming system though they may be appropriate in some other sense. This would be a case of a belief that isn't justified (since it isn't an epistemically fitting doxastic response) even though it is produced by a properly functioning cognitive faculty.

Or suppose some evil demon intentionally designs creatures to form mostly false beliefs. For example, suppose that this malevolent creator decides to make Vic in such a way that his natural unlearned doxastic response to ME2 (the “olfactory” experience) is to form the billiard ball belief B1. Vic's cognitive faculties are, therefore, designed to function in much the same way as are the faculties of the possible cognizers described in my Reidian counterexample to Necessity in section I.D. The difference is that the beliefs of those cognizers are mostly true if their cognitive faculties are functioning properly in the environment for which they were “designed” whereas Vic's beliefs (including B1) are not. Now we noted earlier that the beliefs of those possible cognizers mentioned in I.D—beliefs like B1—are (epistemically) fitting responses to their evidence. But consider Vic who has been designed (by the demon who is intent on having Vic form mostly false

beliefs) to form the false billiard ball belief B1 in response to the “olfactory” experience ME2. Is Vic’s belief B1 an *epistemically* fitting response to ME2? It seems not. It’s true that we can’t blame Vic for holding the belief; so his epistemic blamelessness might give his belief a *subjective* sort of epistemic justification. But that won’t be enough to make his belief *objectively* epistemically fitting. Nor will the fact that he was created by a designer who *intended* that he respond in this non-truth-conducive way to his experiences.

For the same reason, if the evil demon wanted a creature to have mostly false beliefs but decided to achieve this, in part, by having that creature believe B1 in response to the “tactile” experience ME1 (in a context where experiencing ME1 rarely occurs when B1 is true), we should conclude that *that* creature’s belief B1 is not an epistemically fitting response to ME1. Hence, that creature’s belief B1 isn’t objectively justified (since it isn’t an epistemically fitting response to his evidence) despite the fact that it is produced by cognitive faculties functioning properly in the sense that they are functioning so as to achieve their intended purpose. To avoid these sorts of counterexamples (ones like those given in this paragraph as well as ones like those mentioned in the previous two paragraphs), it is important to add clause (ii)(b) to J_{PF} requiring that the belief be formed by a properly functioning faculty that is *aimed at truth*.

To see the importance of clause (ii)(c), consider a creature designed by one of Hume’s infant deities. And suppose that, although this incompetent creator was trying to make a believer with reliable faculties, it instead created one whose faculties produce mostly false beliefs when placed in the environment in which it was intended by its creator to produce true beliefs. For example, suppose this infant deity intentionally created Ric in such a way that, like Vic described above, his natural unlearned doxastic response to ME2 is to form B1. But, contrary to what this bumbling creator hoped, this belief (like most of the other beliefs Ric forms) is false when produced in the intended environment by faculties functioning as they were designed to function. Is Ric’s belief B1 an epistemically fitting response to ME2? It seems not. And the same would be true if the infant deity had designed a creature to form B1 in response to ME1 (rather than ME2), intending (but utterly failing to bring it about) that such a belief is true when formed in that way in the intended environment. Given the failed design attempt, the belief in question may be, in some sense, the output of properly functioning cognitive faculties (since those faculties are, in some sense, operating in the way they were designed to operate). But it doesn’t seem to be an epistemically fitting response to the subject’s evidence.

What the considerations in the previous paragraph suggest is that beliefs produced by properly functioning truth-aimed cognitive faculties do not result in epistemically fitting doxastic responses to one’s evidence *if* the faculties in question aren’t *successfully* aimed at truth—i.e., if they aren’t

likely to produce true beliefs when operating in the environment for which they were “designed”. Hence, clause (ii)(c) of J_{PF} .

I should comment on the fact that what I’ve been saying in this subsection about demon cases might seem to conflict with my suggestion, in the first paragraph of the paper, that the evil demon counterexamples to reliabilism are plausible. According to those counterexamples, a belief can be justified, even objectively justified, in cases where the belief isn’t reliably formed but is instead the result of misleading experiences produced by a deceptive demon. Yet in this subsection I point out the implausibility of the suggestion that the beliefs of a creature created by an evil demon to have unreliable beliefs are objectively justified. The explanation for this apparent conflict is that although I agree that in *some* evil demon cases (in which the subject’s beliefs aren’t reliably formed) the subject’s beliefs *are* objectively justified, I don’t think this is true in *all* such cases. What is true in most demon cases is that the subject’s beliefs are subjectively justified, in the sense that the subject is epistemically blameless in holding them (since it is beyond her control to do otherwise). Furthermore, it is also true that we are sometimes quite sure that the demon victim is *supposed to* hold a certain perceptual belief in response to some sensory experience (as, for example, when an ordinary *human* comes, part way through her life, under the influence of a deceptive demon). If we *are* sure about such “supposed to” claims, then, with respect to cases in which such a demon victim holds the belief we think she is supposed to hold (given the experiences she is having), we may conclude that that belief is objectively justified, despite the fact that the experiences in question are illusory ones produced by the demon. However, when the demon victim is designed and created by the demon to hold unreliably formed beliefs, things are different.

Now that we have J_{PF} before us and understand some of the motivation for it, let’s consider two of its virtues.

D. Virtue One of J_{PF} : Handling of Cases

One virtue of J_{PF} is that it handles certain examples better than reliabilism or Necessity-endorsing internalism handles them. Some of these examples involve the reliable cognizers described in section I.D, whose natural unlearned response to ME2 is B1 and whose natural reaction to grabbing a billiard ball is to experience ME2. I’ll call the complete description of the natural way of functioning for those cognizers ‘design plan B’. And I’ll call the complete description of our natural way of functioning ‘the human design plan’.³⁹

Here are the six cases (in each case, the doxastic response is unlearned):

Case I: A human forms B1 in response to the tactile sensation ME1. B1 is a normal reliably formed belief.

Case II: A human forms B1 in response to ME1. However, ME1 is produced by a deceptive demon, not by actual contact with a billiard ball (nothing like such contact occurs on the occasion in question). The result is that B1 is not reliably formed.

Case III: Due to cognitive malfunction caused by a radiation overdose, a human forms B1 in response to ME2. The same overdose also prevents her from noticing anything wrong with forming this belief in this way. Since ME2 isn't a reliable indicator of the truth of B1, the result is that B1 is not reliably formed.

Case IV: A nonhuman cognizer with "design" plan B forms B1 in response to "olfactory" sensation ME2. This belief is a normal reliably formed belief of such a cognizer.

Case V: A nonhuman cognizer with "design" plan B forms B1 in response to ME2. This belief is a normal unlearned doxastic response for such a creature to ME2. However, ME2 is produced by a deceptive demon, not by actual contact with a billiard ball (nothing like such contact occurs on the occasion in question). The result is that B1 is not reliably formed.

Case VI: Due to cognitive malfunction caused by a radiation overdose, a nonhuman cognizer with "design" plan B forms B1 in response to "tactile" sensation ME1. The same overdose also prevents her from noticing anything wrong with forming this belief in this way. Since ME1 isn't (in this situation) a reliable indicator of the truth of B1, the result is that B1 is not reliably formed.

In both case IV and case V, B1 seems to be justified. For in those cases, just as in cases I and II, B1 is a natural and fitting unlearned response to the main evidence despite the fact that, in case V, the belief is formed in an unreliable way. And since case VI is like case III in that B1 is a malfunctioning and *unfitting* unlearned response to the subject's main evidence, B1 seems to be unjustified in case VI.

We can summarize all six cases as follows:⁴⁰

Case	Which "Design" Plan?	Functioning Properly?	Which Environment?	Which Move From Ground to Belief?	Reliably Formed?	Justified?
I	human	yes	appropriate	ME1 → B1	yes	yes
II	human	yes	demon	ME1 → B1	no	yes
III	human	no	appropriate	ME2 → B1	no	no
IV	plan B	yes	appropriate	ME2 → B1	yes	yes
V	plan B	yes	demon	ME2 → B1	no	yes
VI	plan B	no	appropriate	ME1 → B1	no	no

Reliabilist accounts get cases II and V wrong. Internalist accounts endorsing Necessity get cases IV, V and VI wrong. For the belief ground in cases IV and V is the same as the belief ground in case III and yet, contrary to what such internalist accounts entail, B1 is *not* unjustified in cases IV and V even though it *is* unjustified in case III. Likewise, despite the fact that the belief ground in case VI is the same as the belief ground in cases I and II, B1 is justified in cases I and II but not in case VI. By contrast, J_{PF} gets all six cases right.

E. Virtue Two of J_{PF}: Justification and Truth

Another virtue of J_{PF} is that it provides an account of the justification-truth connection that doesn't lead to the usual troubles. Stewart Cohen has argued that although there seems to be a nontrivial connection between justification and truth, there seems to be no good account of that connection.⁴¹ He considers and, for reasons I find persuasive, rejects two types of accounts: (i) reliabilist accounts, according to which there is an objective connection between justification and truth, and (ii) subjectivist accounts that say that what matters is that there is a justification-truth connection *from the subject's perspective*.⁴²

But there are two other accounts he doesn't consider—understandably since they hadn't been proposed at the time. One is J_{PF}. According to it, the connection between justification and truth comes through the notion of an appropriate environment: there is a high objective probability that a justified belief will be a true belief *if* the properly functioning faculties that produce it are operating in the environment *for which they were "designed"*. This explains the connection between justification and truth while avoiding the evil demon counterexamples to reliabilism. The other account is Sosa's virtue reliabilism. According to it, "justification is relative to environment".⁴³ Thus, the demon victim's beliefs are justified relative to *our* environment because the virtuous dispositions producing her belief are reliable in our environment; but those same beliefs aren't justified relative to the victim's environment because the dispositions in question aren't reliable in the victim's environment.⁴⁴

The advantage of J_{PF} over Sosa's virtue reliabilism is that the proponent of J_{PF} can recognize a nonrelativized concept of justification whereas Sosa can't. But it seems that we do have the notion of justification *simpliciter*. This is important because although it is natural to think that the beliefs of demon victims aren't reliably formed in their environments, it isn't very natural at all to think of their beliefs as *unjustified relative to their environments*. That relativized concept of justification doesn't seem to be one we have in our pretheoretical repertoire of concepts. What we think of the demon victims is just that their beliefs exemplify justification *simpliciter*.⁴⁵ The problem is that Sosa needs, but doesn't seem to have, some

principled way of selecting *the* environment in which reliability matters for nonrelativized justification. J_{PF} has a principled way of doing that: the environment that matters is the one for which the cognizer's faculties were "designed".

III. Objections

A. Objection One: Proper Function and Naturalism⁴⁶

J_{PF} employs several notions in a way that might seem uncongenial to a naturalistic frame of mind. The three "offensive" notions are the *proper function* of a cognitive faculty, its *aim* and the environment for which it was *designed*. The proper function of a thing is the way it is supposed to function. The aim of a thing is the goal to which it is supposed to contribute. The environment for which a thing was "designed" is the environment in which its functioning the way it is supposed to function *is supposed to* result in that thing's contributing to the goal to which it is supposed to contribute. What makes J_{PF} 's employment of these notions potentially unacceptable to the naturalist is that, although they may not be problematic when applied to artifacts, they do seem problematic when applied to living organisms or their parts, which, according to naturalists, aren't artifacts. Some people express this sort of worry by saying that proper function analyses commit their proponents to theism or some other view according to which humans have an intelligent designer.

There is a large and growing literature on functions and goals.⁴⁷ Many naturalists are quite confident that the notion of a goal and of a function can be understood naturalistically and applied to living organisms (or their parts) even if these organisms are not the product of intelligent design. But when it comes to the idea of a function a thing is *supposed to* have or a goal to which it is *supposed to* contribute, some naturalists are more skeptical. Let's examine some of the possible views one could hold concerning the normativity implied by the "supposed to" talk.

First, consider the following three claims: the human heart is supposed to function so that it beats less than 200 times a minute when the person is at rest; the human heart is supposed to contribute to the goal of circulating blood and, ultimately, to the survival of the human in question; the type of environment in which a heart's functioning that way is supposed to result in the survival of the human of which it is a part is an environment relevantly similar to the one in which we find ourselves. These "supposed to" claims seem very sensible and natural. But can they survive as more than just ordinary ways of speaking? Is it literally true that there is a function a human heart is *supposed to* have, that there is goal to which it is *supposed to* contribute and that there is such a thing as the type of environment in which

its functioning in this way is *supposed to* contribute to that goal? And are such claims, taken literally and seriously, compatible with naturalism?

There are (at least) three positions one could take in response to these questions. Let's say that "supposed to" talk is *reducible* to naturalistically acceptable talk just in case it is possible to give truth conditions for "supposed to" claims (applied to parts of living organisms) using only naturalistically acceptable concepts. The *reductivist* position is that "supposed to" talk is reducible to naturalistically acceptable talk. According to this view, the serious "supposed to" claims mentioned in the previous paragraph are true and compatible with naturalism.⁴⁸ The other two positions—nonreductive realism and eliminativism—are nonreductivist views insofar as they agree that "supposed to" talk is not reducible to naturalistically acceptable talk. According to *nonreductive realism*, serious "supposed to" claims are true and incompatible with naturalism (so naturalism is false).⁴⁹ According to *eliminativism*, serious "supposed to" claims are neither true nor compatible with naturalism.⁵⁰ The eliminativist with respect to the serious "supposed to" claims says that, strictly speaking, there is no such thing as the *proper* function of organisms (or their parts) or the *right* environment for them; these normative notions need to be eliminated from our careful conversation which will instead employ notions that are more friendly to naturalism— notions such as *statistically normal* or *evolutionarily selected* functions or environments (or even functions *valued by us*).

Each of these three positions has its proponents; each has a legitimate claim to be taken seriously even by those who reject it. Which position is the correct one? Obviously, I can't resolve this issue here. But I will say this. In sections I.D and I.E I have defended the view that the notion of fittingness is conceptually tied to the idea of there being a *right* way for a cognizer's cognitive faculties to function—a way its faculties are *supposed to* function. If I'm right about *that*, then whatever position one takes with respect to the normativity of "supposed to" claims about an organism's (or its parts') functions and goals—whether reductivism, nonreductive realism or eliminativism—one ought to take that same position with respect to fittingness. But Objectivity is opposed to eliminativism with respect to fittingness. And Nonreliability is opposed to at least one strand of reductivism with respect to fittingness. So those who find Objectivity and Nonreliability plausible are forced to be nonreductive realists with respect to fittingness or to come up with some nonreliabilist way of being a reductivist about fittingness. Such people should, therefore, be either nonreductive realists or a certain kind of reductivist about the commonsense "supposed to" claims concerning functions and goals; eliminativism is not an option for them.

In sum, my response to those naturalists who find the notion of proper function naturalistically unacceptable (i.e., my response to eliminativists about this sort of "supposed to" talk) is twofold. First, my claim that the notion of fittingness is conceptually tied to the notion of proper function is

compatible with eliminativism about the normativity of proper function. So at least that part of my argument avoids this naturalistic challenge. Second, upon accepting the conclusion that there is that conceptual tie between fittingness and proper function (and, therefore, between justification and proper function), one can take the plausibility of Objectivity and Nonreliability to provide us with a reason for rejecting eliminativism and certain versions of reductivism about proper function.

B. Objection Two: The Supervenience Thesis

The supervenience thesis (as it is sometimes called⁵¹) says that normative properties supervene on nonnormative ones. Applied to epistemology, and, in particular, to justification, it says that justification supervenes on nonnormative properties—that there couldn't be a difference in justification without a difference in nonnormative properties. The plausibility of the supervenience thesis might give rise to the following complaint about proper function accounts of epistemic justification: “The goal in understanding epistemic concepts such as justification is to give analyses of them in terms of nonnormative properties; but since *proper function* is a normative concept, analyses in terms of it are unsatisfying or uninteresting or in some other way unacceptable.”

Although this objection seems to be taken seriously by some people,⁵² I fail to see what the trouble is. The proponent of a proper function account of epistemic justification is like the Chisholmian who thinks epistemic concepts are reducible to ethical concepts.⁵³ Each thinks that epistemic evaluation is reducible (at least in part) to another sort of evaluation. But that in no way commits them to a rejection of the supervenience thesis. If epistemic properties are reducible to proper function (or ethical) properties and proper function (or ethical) properties are reducible to nonnormative properties, then the supervenience thesis holds. So I see no reason to think the supervenience thesis conflicts in any way with the analysis of justification given in this paper. It's true that it would be nice to conjoin my proper function analysis of justification with an analysis of proper function in terms of nonnormative properties and that I haven't done that here.⁵⁴ But that doesn't make it uninteresting to explain one normative notion in terms of another in the way Chisholm does—not unless the *only* sort of analysis of interest in epistemology is an analysis of epistemic properties in terms of nonnormative (rather than merely nonepistemic) ones.

C. Objection Three: Swampman

Against proper function analyses of epistemic properties, Ernest Sosa (1993) uses Donald Davidson's example of Swampman who comes into

existence by accident as a result of a lightning strike. Swampman is supposed to be a molecule for molecule replica of Donald Davidson, complete with Davidson's beliefs, memories, experiences and dispositions. Given his origin, it seems that Swampman has no design plan, not even in the sense in which faculties "designed" by evolution have a design plan. Thus, according to this objection, we can't say that his beliefs satisfy the proper function requirement. But since Davidson's beliefs are justified why not say the same of Swampman's?

By way of response, consider what we'd say of Swampman's heart and lungs. Assuming Davidson's heart and lungs were in fine shape at the time of the accidental duplication, the thing to say about Swampman's heart and lungs is that they are healthy. Furthermore, it seems plausible to say that Swampman is a human (molecular replication implies DNA replication after all). We might feel *some* hesitation about saying Swampman's heart is healthy and that he's a human. But saying so is far *more* plausible than saying that Swampman's heart and lungs are not healthy (either because they are unhealthy or because the concept of health doesn't apply) or that Swampman has Davidson's DNA but is not a human.

Now, if we find it plausible to say that Swampman's heart and lungs are healthy, it is reasonable also to conclude that his heart and lungs are functioning the way they are *supposed to* function. And if we can say that about his heart and lungs, we can say the same about his cognitive faculties. It's true there is no literal designer and no evolutionary origin. But because of the physical similarities (down to the DNA level), we are inclined to think of Swampman as being human. And that leads us to think that Swampman is supposed to function in the way humans are supposed to function.⁵⁵

In this paper I've argued that J_{PF} is superior to both reliabilist and Necessity-endorsing internalist accounts of justification. In doing so, I've defended the following two claims in a way that recognizes a nonrelativized concept of justification: (i) there is a satisfying externalist account of the justification-truth connection; (ii) there is a plausible externalist account of justification that avoids the evil demon counterexamples to reliabilism. The truth of these claims shows that epistemologists need to be careful *not* to take the failures of reliabilism to be failures of externalism.⁵⁶

Notes

1. This objection to reliabilism is proposed by Cohen (1984, 280–82), Foley (1985, section I), Lehrer (1990, 166) and Moser (1985, 240–241).
2. The larger project here is that of defending externalism and objecting to internalism. By presenting a nonreliabilist but still externalist account of justification, I take the sting out of the now standard evil demon objection (just mentioned in the text) to externalism about justification. See my 2000a, my 2000b and my forthcoming paper "A Dilemma for Internalism" for other contributions to this

larger project, which includes both attacks on internalism as well as defenses of externalism against other objections.

3. The more standard way of stating the evidentialist's thesis is to have the right-hand side of the 'iff' say something like "B fits S's evidence" (see Feldman and Conee 1985). I'm assuming that doxastic response B is a fitting response to S's evidence only if (i) B fits S's evidence and (ii) B is formed in response to S's evidence. Clause (ii) is a way of adding a sort of basing requirement. Since justification, as I'll be thinking of it in this paper, is *doxastic* justification rather than merely *propositional* justification, the addition of such a requirement is quite unremarkable. Doxastic justification is justification a person's *belief* has only if that belief is properly based. Propositional justification is justification enjoyed by a *proposition* for a person if that person's evidence is such that a belief in that proposition would be doxastically justified if she were to hold it on the basis of that evidence. (See Firth 1978, 218 for more on this distinction.) The notion of doxastic justification corresponds (roughly) to Feldman and Conee's notion of well-foundedness (see their 1985, 24). And the notion of propositional justification seems to correspond to what Feldman and Conee think of as justification *simpliciter*.
4. See Feldman 1988a.
5. Feldman and Conee 1985, 15.
6. Of these three claims—Nonreliability, Objectivity and Necessity—the first is the one most obviously endorsed by evidentialists. I'll explain later (in section I.B) why I think Feldman and other evidentialists hold Objectivity. As for Necessity, I'll just point out that Chisholm seems to hold it when he asserts (1986, 53) that epistemic properties supervene on the psychological properties that constitute our evidence base. And Feldman and Conee seem to hold it when they combine their claim (1985, 16) that "the epistemic justification of an attitude depends only on evidence" with their claim (1985, 15) that an attitude is epistemically justified if and only if that attitude *fits* the evidence. Who counts as an evidentialist? Perhaps the clearest examples are Feldman and Conee (1985). But evidentialist sympathies can also be discerned in the writings of Chisholm (1977), Moser (1985, 1989), Fumerton (1995), Haack (1993, 1997), and Russell (2001).
7. The idea is that if B is a fitting response to E, then, even if B could be an unfitting response to evidence that includes E *and more besides*, it couldn't be an unfitting response to E by itself.
8. Notice that in rejecting Necessity I am not saying that the relation of supervenience holding between justification and its supervenience base is contingent. For even if it were true that belief B's *justification* supervened (of necessity) on the fittingness of doxastic response B to the subject's evidence (which happened to be E), that wouldn't entail that the *fittingness* of doxastic response B to evidence E was an essential property of it.
9. A felt inclination to do X is an experience of feeling inclined to do X. It isn't merely a disposition.
10. Thus, disconnectors will be a sort of undercutting defeater (see Plantinga 2000, 359 and Pollock 1986, 39 for a discussion of undercutting defeaters). Notice that *lacking a connector* between a belief and one's main evidence isn't by itself sufficient for *having a disconnector*.

11. I am *not* suggesting that justification *requires* a connector between one's beliefs and one's main evidence. That smells of regress problems. I'm just explaining what connectors are and noting that they may be present in some cases.
12. Since a meadow full of flowers usually has a number of different kinds of flower (as well as a variety of nonflowering plants) each of which has a distinctive odor, the type of olfactory experience one has upon smelling a meadow full of flowers has many components. In this way it resembles the tactile experience one has upon grabbing a billiard ball, an experience that also has many components.
13. How could a person's total relevant evidence for B1 consist of only ME2 and C2 without any disconnectors? Wouldn't she glance at her hand or attempt to use the hand in which she thinks there is a small hard ball? Wouldn't others inform her that it is crazy for her to think, on the basis of ME2, that she is holding a small hard ball? These are certainly possibilities, likely ones even. But all we need is a possible example where no such disconnectors are present.
14. See his 1988b, 411. See also sections II and III of Feldman and Conee 1985 where they say that if B fails to fit one's evidence, it won't help to add subjective factors such as *trying one's best to hold only beliefs that fit one's evidence* or *being blameless in holding B*. They insist that even with those factors present, B still fails to fit the subject's evidence.
15. See Russell's 2001, 37–38, especially his comments on the believers raised in the benighted religious community whom he thinks of as subjectively justified but not objectively justified. See also Moser's 1989 (38–42, 47–52 and 202–3) and Fumerton's 1995 (8–20, 113–16 and chapter 7), where they emphasize that merely believing that one's evidence E is good evidence for B isn't sufficient for B's justification since E might not *in fact* make B objectively probable.
16. This distinction is similar to and inspired by Reid's distinction between original and acquired perception. See Reid 1969, 300–305 and Reid 1970, 210–11 (parts of the latter are in Reid 1983, 87–88). It is basically the same as Goldman's distinction between beliefs produced by native processes and beliefs produced by acquired methods (see his 1986, 93–95).
17. It is because my objection to Necessity relies on a counterexample employing these assumptions that I had to discuss Objectivity and explain its implications *before* giving the objection.
18. I call B1 a 'billiard ball belief' for convenience even though the concept of a billiard ball is not included in its content.
19. Objection: Is B1 really an *unlearned* doxastic response to ME1 for normal humans? Perhaps psychological research will show (or has shown) that it is *learned*.
Reply: Maybe so. But suppose psychological research had shown that B1 is (as Reid believed) an *unlearned* doxastic response to ME1. Would that have given us a reason to take such beliefs to be unjustified? No. So it seems possible for there to be creatures for whom B1 is justified in virtue of being believed on the basis of ME1 even though B1 is (for them) an *unlearned* doxastic response to ME1. The possibility of such creatures is all I need to make my case.
20. Reid 1970, 63–64/Reid 1983, 39–40.
21. Perhaps sensations supervene on brain states. Then we have to imagine, first, that sensations of the same qualitative type as our olfactory sensations could supervene on physical states (call them X-states) other than those on which ours

supervene (cf. the Martians in section I of Lewis 1980). Second, we imagine a creature who is like us in outward appearance and who, upon grabbing a billiard ball in its hand, causes itself to go into an X-state.

22. I use quotation marks to indicate that although the sensation in question is qualitatively of the same type as *our* olfactory sensations, it may not, in the circumstances in question, be properly thought of as an olfactory sensation in the ordinary sense (since that sense may have functional role implications).
23. I should note that the cognizers in this Reidian example are such that *all* of the experiences produced in them by the sorts of activities that produce tactile experiences in us are what we would call ‘olfactory experiences’. So it isn’t just *BI*’s ground that is different in this way.
24. Cf. Greco 1999, 277 and 2000, 173–74.
25. See note 19.
26. See note 13.
27. Plantinga’s argument (1993a, 54–63) that the *warrant*-conferring power of a belief ground is a contingent feature of it is similar to my objection to Necessity. The main difference (aside from the fact that he focuses on warrant and I focus on justification) is that his examples are less convincing than the Reidian one I employed in this subsection. One problem with Plantinga’s examples is that many of them involve *learned* doxastic responses. This leads to the concern discussed in section I.C—a concern that the Reidian example, which focuses on *unlearned* doxastic responses, sidesteps. Another problem is that one of Plantinga’s examples (1993a, 62–63) is met with understandable resistance by Bruce Russell (see Russell’s 2001, 45–46) who seems to think that Plantinga’s example involves treating a doxastic response that must be *learned* as if it can be an appropriate *unlearned* doxastic response. I think Plantinga’s example can survive Russell’s criticisms. But the Reidian example is better because it involves a belief that clearly *can* be an appropriate unlearned doxastic response.
28. See Pollock 1986, 133–34. I should note that I *reject* the view that the term ‘internal’, as Pollock uses it, is helpful in making sense of the internalism-externalism debate in epistemology. See Plantinga 1993a, 180–81 for more on this.
29. See Pollock 1986, 133 and Pollock and Cruz 1999, 133 & 136.
30. Pollock and Cruz 1999, 136.
31. For a discussion of these problems as well as a defense of the view that imposing such requirements is at the heart of internalism, see my forthcoming paper “A Dilemma for Internalism”.
32. The quotation marks are to indicate that the design in question needn’t involve a literal designer. For discussion of how to make sense of design talk without postulating a literal designer, see section III.A below.
33. This is *not* a direct quotation (for his exact wording, see chapters 1 and 2 of Plantinga 1993b, especially p. 19 and pp. 46–47). Plantinga insists that this is still only a first approximation. For further details, qualifications and emendations see his 1993b as well as his 1996, his 1997 and his 2000 (156–61).
34. It is widely believed, even by externalists, that a belief isn’t warranted if the person holding it takes it to be defeated (see for example Goldman 1986, 62–63 & 112–12 and Nozick 1981, 196). The same, I assume, is true of objective justification. See my 1997 for a discussion of this condition and its connection with internalism and externalism.

35. Plantinga says that warrant depends on the proper functioning of our defeater systems (1993b, 41). So he is aware of the concern I'm raising. However, what he doesn't acknowledge is that conditions (i)-(iv) don't entail the satisfaction of condition (v).
36. This clause requires that the subject not *take* her belief to be defeated. But shouldn't there also be a clause requiring that the belief not *in fact* be defeated, whether the subject takes it to be or not? Or, if there is such a requirement on justification (i.e., that the belief not in fact be defeated) but its satisfaction is guaranteed by the satisfaction of the proper function requirement, then why not think the satisfaction of the "no believed defeater" requirement is also guaranteed by the satisfaction of the proper function requirement? (My thanks to an anonymous referee for drawing these questions to my attention.)

There is much to be said here. But the short response is this. Yes, there is a requirement that the belief (we'll call it 'B') not *in fact* be defeated. What this amounts to is the requirement that the subject have no belief or experience that, when added to her previous beliefs and experiences, has the result that B becomes an epistemically inappropriate response to the input to her belief-forming systems. (For more on this way of thinking about defeaters, see Plantinga 2000, 359–66 and my forthcoming paper "Defeaters and Higher-Level Requirements".) Does the fact that the subject's cognitive faculties are functioning properly guarantee that this "no actual defeater" requirement will be satisfied? No. For the faculties might not be truth-aimed; and they might not be reliable in the environment for which they were "designed". (See section II.C for further discussion of these possibilities.) Would the satisfaction of the proper function requirement, together with the satisfaction of clauses (ii)(b) and (ii)(c) from J_{PF} , guarantee that the "no actual defeater" requirement is satisfied? No. For it may be that those conditions are satisfied and yet the subject takes her belief B to be defeated (see the example just given in the text to motivate the addition of clause (v) to Plantinga's account). That would be enough to make B an epistemically inappropriate response to the input to her belief-forming systems. In short, the "no actual defeater" requirement is satisfied only if the "no believed defeater" requirement is satisfied. Furthermore, if each of the conditions mentioned in J_{PF} is satisfied, the "no actual defeater" requirement is satisfied too.

37. 1993b, 22–24.
38. See Plantinga's 1993b, 11–20 and 26–28.
39. As Plantinga points out (1993b, 22–24), we can think of "design" plans as sets of circumstance-response pairs (or circumstance-response-goal triples).
40. The six cases can also be described so that the belief ground in each case includes an *unlearned* connector. Just replace all appearances of 'ME1' (in the above descriptions of the six cases as well as in the table below) with 'ME1 + C1'; likewise, replace all appearances of 'ME2' with 'ME2 + C2'. In the demon and malfunction cases, this connector in question is caused by the demon or by the radiation overdose.
41. See Cohen 1984.
42. In section II of his 1984, Cohen objects to reliabilist accounts of justification on the grounds that they fail to attribute justification to demon victims. And in

section III he objects to subjectivist accounts on the grounds that they implausibly require for justification what we rarely have, namely, further beliefs about the reliability of our beliefs.

43. Sosa 1991,144.
44. This view is developed in Sosa 1991,140–44 and 284–90.
45. For a similar concern about views like Sosa's, see Goldman 1993, 101–3.
46. For other ways to handle a variety of objections to talk about proper function, see Plantinga's 1993b, 1993c, 1995 and 1996.
47. See Nissen 1997, Boorse (unpublished manuscript) and the introduction to Allen, Bekoff and Lauder 1998 for summaries of the relevant literature.
48. Karen Neander (1991a and 1991b) holds this view as do Fred Dretske (1994 and 2000) and Robert Koons (2000).
49. This view is endorsed by Alvin Plantinga (1993b) and perhaps also by Bedau (1993).
50. This view seems to be endorsed by John Searle (1992, 238) and by Christopher Boorse (unpublished manuscript).
51. Steup 1996, 30–36.
52. For example, by one of the referees for this paper.
53. See Chisholm 1986 and 1991. For objections to Chisholm's position, see Firth 1978.
54. But see the references in notes 48 and 49 for attempts at and discussions of such analyses.
55. Compare with Plantinga's response to the Swampman objection in his 1993c, 76–78.
56. Thanks to John Greco, Kevin Meeker, Trenton Merricks, Alvin Plantinga, three anonymous referees and especially Michael Rea and Matthias Steup for reading and providing insightful comments on earlier drafts. Thanks also to Christopher Boorse and Richard Feldman for helpful discussions of some of the ideas presented here. Finally, I would like to acknowledge the support of the Pew Evangelical Scholars Program while I worked on this paper.

References

- Allen, Colin, Mark Bekoff and George Lauder (eds.). 1998. *Nature's Purposes: Analyses of Function and Design in Biology*. Cambridge, MA: MIT Press.
- Bedau, Mark. 1993. "Naturalism and Teleology." In Steven Wagner and Richard Warner (eds.), *Naturalism: A Critical Appraisal*, pp. 23–5. Notre Dame: University of Notre Dame Press.
- Bergmann, Michael. 1997. "Internalism, Externalism and the No-Defeater Condition." *Synthese* 110: 399–417.
- . 2000a. "Deontology and Defeat." *Philosophy and Phenomenological Research* 60: 113–28.
- . 2000b. "Externalism and Skepticism." *The Philosophical Review* 109: 159–94.
- . Forthcoming. "A Dilemma for Internalism." In Thomas Crisp, Matthew Davidson and David Vander Laan (eds.), *Knowledge and Reality: Essays in Honor of Alvin Plantinga*. Dordrecht: Kluwer Academic Publishers.
- . Forthcoming. "Defeaters and Higher-Level Requirements." *The Philosophical Quarterly*.
- Boorse, Christopher. Unpublished. "Functions: A Current Scorecard."

- Chisholm, Roderick. 1977. *Theory of Knowledge* (2nd Edition). Englewood Cliffs, NJ: Prentice Hall.
- . 1986. "Self-Profile." In Radu Bogdan (ed.), *Roderick M. Chisholm*, pp. 3–77. Dordrecht: D. Reidel.
- . 1991. "Firth and the Ethics of Belief." *Philosophy and Phenomenological Research* 51: 119–28.
- Cohen, Stewart. 1984. "Justification and Truth." *Philosophical Studies* 46: 279–95.
- Dretske, Fred. 1994. "If You Can't Make One, You Don't Know How it Works." In Peter French, Theodore Uehling and Howard Wettstein (eds.), *Midwest Studies in Philosophy, 19, Philosophical Naturalism*, pp. 468–82. Notre Dame: University of Notre Dame Press.
- . 2000. "Norms, History, and the Constitution of the Mental." In *Perception, Knowledge and Belief: Selected Essays*, pp. 242–58. New York: Cambridge University Press.
- Feldman, Richard. 1988a. "Having Evidence." In D.F. Austin (ed.), *Philosophical Analysis*, pp. 83–104. Dordrecht: Kluwer Academic Publishers.
- . 1988b. "Subjective and Objective Justification in Ethics and Epistemology." *The Monist* 71: 405–19.
- Feldman, Richard and Earl Conee. 1985. "Evidentialism." *Philosophical Studies* 48: 15–34.
- Firth, Roderick. 1978. "Are Epistemic Concepts Reducible to Ethical Concepts?" In A.I. Goldman and J. Kim (eds.), *Values and Morals*, pp. 215–29. Dordrecht, Holland: D. Reidel.
- Foley, Richard. 1985. "What's Wrong With Reliabilism?" *The Monist* 68: 188–202.
- Fumerton, Richard. 1995. *Metaepistemology and Skepticism*. Lanham: Rowman and Littlefield.
- Goldman, Alvin. 1986. *Epistemology and Cognition*. Cambridge: Harvard University Press.
- . 1993. "Epistemic Folkways and Scientific Epistemology." In Alvin Goldman (ed.), *Readings in Philosophy and Cognitive Science*, pp. 95–116. Cambridge, MA: MIT Press.
- Greco, John. 1999. "Agent Reliabilism." In James Tomberlin (ed.), *Philosophical Perspectives, 13, Epistemology (A Supplement to Noûs)*, pp. 273–96. Cambridge, MA: Blackwell Publishers.
- . 2000. *Putting Skeptics In Their Place*. New York: Cambridge University Press.
- Haack, Susan. 1993. *Evidence and Inquiry: Towards Reconstruction in Epistemology*. Oxford: Blackwell Publishers.
- . 1997. "A Foundherentist Theory of Empirical Justification." In Louis Pojman (ed.), *The Theory of Knowledge: Classical and Contemporary Readings*, Second Edition, pp. 283–93. Belmont, CA: Wadsworth.
- Koons, Robert. 2000. *Realism Regained: An Exact Theory of Causation, Teleology and the Mind*. New York: Oxford University Press.
- Lehrer, Keith. 1990. *Theory of Knowledge*. Boulder: Westview Press
- Lewis, David. 1980. "Mad Pain and Martian Pain." In Ned Block (ed.), *Readings in Philosophy of Psychology, Volume 1*, pp. 216–22. Cambridge, MA: Harvard University Press.
- Moser, Paul. 1985. *Empirical Justification*. Boston: D. Reidel.
- . 1989. *Knowledge and Evidence*. New York: Cambridge University Press.
- Neander, Karen. 1991a. "Functions as Selected Effects: The Conceptual Analyst's Defense." *The Philosophy of Science* 58: 168–84.
- . 1991b. "The Teleological Notion of Function." *Australasian Journal of Philosophy* 69: 454–68.
- Nissen, Lowell. 1997. *Teleological Language in the Life Sciences*. Lanham: Rowman and Littlefield.
- Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge: Belknap of Harvard University Press.
- Plantinga, Alvin. 1993a. *Warrant: The Current Debate*. New York: Oxford University Press.
- . 1993b. *Warrant and Proper Function*. New York: Oxford University Press.

- . 1993c. “Why We Need Proper Function.” *Noûs* 27: 66–82.
- . 1995. “Reliabilism, Analyses and Defeaters.” *Philosophy and Phenomenological Research* 55: 427–64.
- . 1996. “Respondeo.” In Jonathan Kvanvig (ed.), *Warrant in Contemporary Epistemology: Essays in Honor of Plantinga’s Theory of Knowledge*, pp. 307–78. New York: Rowman and Littlefield.
- . 1997. “Warrant and Accidentally True Belief.” *Analysis* 57:140–45.
- . 2000. *Warranted Christian Belief*. New York: Oxford University Press.
- Pollock, John. 1986. *Contemporary Theories of Knowledge*. Savage, MD: Rowman and Littlefield.
- Pollock, John and Joseph Cruz. 1999. *Contemporary Theories of Knowledge* (2nd Edition). Lanham, MD: Rowman and Littlefield.
- Reid, Thomas. 1969. *Essays on the Intellectual Powers*. Cambridge, MA: MIT Press.
- . 1970. *An Inquiry into the Human Mind*, (ed.) Timothy Duggan. Chicago: University of Chicago Press.
- . 1983. *Inquiries and Essays*, (eds.) Ronald Beanblossom and Keith Lehrer. Indianapolis: Hackett Publishing.
- Russell, Bruce. 2001. “Epistemic and Moral Duty.” In Matthias Steup (ed.), *Knowledge, Truth and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, pp. 34–48. New York: Oxford University Press.
- Searle, John. 1992. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- Sosa, Ernest. 1991. *Knowledge in Perspective: Selected Essays in Epistemology*. New York: Cambridge University Press.
- . 1993. “Proper Function and Virtue Epistemology.” *Noûs* 27:51–65.
- Steup, Matthias. 1996. *An Introduction to Contemporary Epistemology*. Upper Saddle River, NJ: Prentice Hall.