

Reliabilism, Analyses and Defeaters¹

ALVIN PLANTINGA

University of Notre Dame

I'm delighted to reply to William Alston, Carl Ginet, Matthias Steup, Richard Swinburne, and James Taylor; their challenging comments on *Warrant: the Current Debate* (hereafter WCD) and *Warrant and Proper Function* (WPF) raise issues of great interest and push the discussion well beyond where I left it. My aim is not to defend myself in a doggedly self-exculpatory fashion, but to try to make some further progress—sometimes by amending what I said, and sometimes by carrying the discussion further, into areas I didn't deal with properly. (There are plenty of areas of that sort.) There is no proportion between the length of my reply and my gauge of the quality of the comment replied to; all depends on whether I have something reasonably interesting and worthwhile to say in response.

I Alston: Generic Reliabilism?

According to the basic paradigm (so I say), a belief has warrant for me if and only if (1) it is produced in me by cognitive faculties functioning properly in an epistemic environment sufficiently similar to the one for which my faculties were designed, (2) the segment of the design plan governing the production of that belief is aimed at the production of true beliefs (rather than at the production of beliefs that contribute, for example, to survival, or psychological comfort) and (3) the design plan is successful: there is a high objective probability that a belief produced under those conditions will be true. To put it briefly, a belief has warrant if it is produced by cognitive faculties functioning properly in a congenial epistemic environment according to a design plan successfully aimed at truth. Now William Alston makes substantially two comments. First, he notes that my arguments against reliabilism were all arguments against *specific* formulations of it: Dretske's, Goldman's, his own, and so forth; that is indeed true. He then claims that there is a kind of *generic* reliabilism that emerges unscathed from the arguments and counterexamples I proposed against those specific reliabilisms; that is perhaps also true, but I

¹ I am very much indebted for incisive criticism and wise counsel to Michael Bergmann, Andrew Koehl, Trenton Merricks, Michael Rea, and Patrick Kain.

shall argue that this generic reliabilism is open to other objections, and is no more successful than its more specific colleagues. Second, Alston argues that proper function isn't necessary for warrant; here I shall argue that he is mistaken.

In the interest of doing things decently and in order, suppose we take the first comment first. Alston notes that "The third condition quoted above is close to reliabilism"² and suggests that my account "can be viewed as 'reliabilism plus'." "It is not a completely different kind of approach from reliabilism, but one that adds extra requirements." And the question, says Alston, is "Are these additions necessary?"—a question to which his answer is "Probably not."

Now I am inclined to think that they are indeed necessary. Perhaps, by way of facilitating comparison of my account with reliabilism, we could put the former as follows:

- (1) A belief *B* has warrant for me iff it is produced by reliable faculties (processes, mechanisms) working (a) properly, (b) in the sort of environment for which they are designed, and (c) according to a design plan aimed at the production of true beliefs.

(For the reason given in footnote 2, this isn't quite equivalent to my account, but it will do for present purposes.) (1) is stated in such a way that the account is clearly 'reliabilism plus'. But as far as I can see, the three additional conditions are indeed necessary—at any rate the reliabilist component is insufficient, as is its combination with any one or two of the other components of the right-hand side of (1). The first additional component requires that the faculties in question be working properly. And isn't this necessary? Clearly my vision might be a reliable faculty, but on a given occasion produce beliefs with no warrant. I may suffer from fever or delirium tremens, for example; visual beliefs produced under those conditions may have little warrant, even if my vision is a reliable faculty. The second condition also seems necessary—at any rate the first two alone won't do the job. For even if the faculties in question are reliable in the sort of environment for which they are de-

² "Epistemic Warrant as Proper Function," *Philosophy and Phenomenological Research*, 55 (1995), p. 398. As Alston notes, the third condition is *close* to reliabilism; but it doesn't entail it. For on a rough and ready version of reliabilism, a belief has warrant only if it is produced by reliable cognitive faculties or processes; on my view, however, it is possible that a belief have warrant even if it isn't produced by a reliable faculty: if, for example, it is produced by a faculty that is unreliable overall (overall produces too few true beliefs), but has an area of operation in which it functions reliably and in which it is aimed at the production of true beliefs. What really counts is not the reliability of the faculty or process involved, but the specific module of the design plan governing the production of the belief in question.

signed (by God or evolution) they might fail to confer warrant in some alien cognitive environment.

And isn't the third condition also necessary? Perhaps I have a faculty that is in fact reliable, and is ordinarily aimed at the production of true belief, but under certain narrowly defined and relatively rare conditions functions in such a way that its purpose, in so functioning, is not that of producing true belief, but beliefs conducive to some other end, such as survival or comfort. Consider an example from William James. Suppose there is a faculty or process that issues beliefs about how far you can jump: for example, whether you can jump the gully or creek or crevasse in front of you. Suppose this faculty or process is in fact reliable: in the vast majority of cases it issues true beliefs; and suppose further that its purpose is ordinarily that of furnishing true beliefs on the topics in question, thus preventing you from attempting creeks too wide and regularly getting soaked. Under certain unusual conditions, however, it functions in such a way that its purpose in producing beliefs in that situation is not that of producing true beliefs, but rather that of producing beliefs that contribute to survival. For example, you are descending Mt. Blanc; darkness is closing in (you foolishly started much too late); the wind is rising and the temperature is plummeting; if you are benighted, you will die (to compensate for the late start, you decided to go light and left all of your warm clothes in camp); and the only way of getting off the mountain before dark is to leap the crevasse before you. The crevasse is pretty formidable, but you form the belief that you can do it. That is because the design plan for the faculty or process in question includes an 'optimistic override' that operates in just such situations. The purpose of its operation is not that of producing true beliefs (it's a bit unduly optimistic for that) but of beliefs that conduce to survival (the fact is you are much more likely to be able to jump a crevasse if you think you can than if you think you can't). I take it that under those conditions your belief that you can jump the crevasse has little warrant, even though it is produced by a process or faculty that is reliable, that is functioning properly in the right kind of environment, and is furthermore such that its (ordinary) purpose is that of producing true belief.

Now here Alston or some other reliabilist might say that his view requires, for warrant, not just that the belief in question be produced by a *reliable faculty*, but that it be produced by a faculty or process that *is functioning reliably* on the occasion in question—or that the belief in question be produced *in a reliable way*. This way of putting the matter, he might think, obviates the necessity for the three conditions I add. I don't think it does.

Here I can kill two birds with one stone or at least answer two connected questions: (a) Is there a satisfactory generic reliabilism? and (b) What is it for a faculty to be functioning reliably on a given occasion? In WCD I argued against reliabilism by cases; I argued that Dretske's reliabilism, the old

Goldmanian reliabilism, the new Goldmanian reliabilism, Alstonian justification taken as necessary and sufficient for warrant—I argued that each of these fails: none of the conditions suggested is both necessary and sufficient. (And the failures aren't such that a bit of chisholming will mend matters.) Now Alston doesn't suggest that my arguments against these specific versions of reliabilism are unsuccessful. What he does suggest is that in addition to all these *specific* versions, there is a *generic* version of reliabilism:

Plantinga brings different counterexamples against different forms of reliabilism, but in this short discussion I will treat them all as directed against "generic reliabilism"—the view that a belief is warranted iff it is produced (or sustained) in a reliable way, one that is such that in a sufficiently wide and varied set of ordinary cases it would usually produce true beliefs....

And he goes on to argue that my counterexamples do not succeed if taken as aimed at that *generic* version. (This is not surprising, since they weren't aimed at that version.) So the idea is that in addition to the various specific versions, there is a generic version; and while my arguments against the former might be successful, they don't controvert the latter.

Well, suppose we take a closer look at this alleged generic reliabilism. Perhaps, ironically, it will succeed where the more specific versions fail. In the passage just quoted Alston states generic reliabilism as follows: "a belief has warrant for me if and only if it is produced (or sustained) in a reliable way, one that is such that in a sufficiently wide and varied set of ordinary cases it would usually produce true beliefs." A belief has warrant iff it is reliably produced; and a belief is reliably produced iff it is produced in a way such that in a wide and varied set of ordinary cases, a belief produced in that way would usually be true. To spell out more fully the implied counterfactual element, perhaps we could put it like this:

- (2) A belief *B* is reliably produced iff it is produced in a way *W* for which there is a wide and varied set of ordinary circumstances *S* such that if any member *C* of *S* obtained and *B* were produced in *C* in way *W*, then *B* would probably be true

or

- (3) A belief *B* is reliably produced iff it is produced in a way *W* for which there is a wide and varied set of ordinary circumstances *S* such that the conditional (objective) probability of a belief's being true, given that it is produced in way *W* in any of the circumstances in *S*, is high.

But here there is a crucial question: what are these 'ways *W*' in which a belief could be produced? Taking things at face value and as naturally as pos-

sible, one would think that *being produced by perceptual faculties* or *being produced by my perceptual faculties* or *being produced by my vision* (using 'vision' to denote the faculty or module that produces visual beliefs, beliefs like the belief that there is now a computer before me) would be examples of ways in which beliefs could be produced or sustained. Furthermore, if we consider the record of vision or perception generally over large and varied sets of ordinary cases, we find it to be very good; for each of vision and perception there is certainly a wide and varied set of ordinary circumstances *S* such that if a belief were produced by vision or perception in some member of that set, then the belief in question would probably be true. So each of these ways meets the conditions specified in the right-hand sides of (2) and (3). A visual or perceptual belief, therefore, meets the proposed condition of warrant stated by (2) and (3): it is produced in a way that, in a sufficiently wide and varied set of ordinary cases, would usually produce true beliefs.

But does it follow that just any of my visual (or perceptual) beliefs, has warrant? Obviously not. On a given occasion I may be sick or drunk; I may be suffering from delirium tremens again, or suffer from visual hallucinations due to drugs or a high fever. Even though the visual beliefs produced under those conditions are produced in a way that meets (1) and (2), these beliefs have no warrant for me. If the hallucinatory beliefs happen to be true, they do not constitute knowledge. And the reason, I would say, is failure to meet the condition of proper function. So far generic reliabilism seems not to work.

Well, perhaps Alston would respond that the ways I picked—*being produced by vision* or *perception* or *being produced by my vision* or *my perception*—aren't the *right* or *relevant* ways: they aren't what the generic reliabilist has in mind when he states his generic reliabilism. But then what *is* the right way? What way *does* he have in mind? What are the relevant ways? (And of course the problem, here, for one of Alstonian and reliabilist proclivities, is to specify the right way or ways without adverting to proper function or any other condition in addition to reliability.) But isn't this question precisely what the various *specific* versions of reliabilism are answers to? Each of them specifies a way of a belief's being produced and/or sustained (a way that in one way or another involves reliability) that is, according to the author, necessary and sufficient for warrant. There are, for example, Dretske's reliabilism in *Knowledge and the Flow of Information*, the old Goldmanian reliabilism, and the new Goldmanian reliabilism; in WCD I argued that in each case the condition suggested fails to be necessary and sufficient for warrant. As far as I can see, Alston doesn't disagree.

If we like, we could add Robert Nozick's suggestion³ as to what this way might be: according to him (with a qualification or two) a belief that *p* has

³ In *Philosophical Explanations* (Cambridge, Massachusetts: Harvard University Press, 1981), pp. 172–78. Substantially this suggestion was made by Dretske a decade earlier

warrant if and only if, if p were true, S would believe it, and if it were false, S would not believe it. In “Positive Epistemic Status and Proper Function”⁴ I argued that this won’t work either. First, *a priori* knowledge, which is usually knowledge of necessary truths, is one of our most important kinds of knowledge: but the Nozickian suggestion won’t work for necessary truths, due to the vagaries of counterfactuals with necessarily false antecedents. But I argued also that it won’t work for a large variety of contingent propositions. Further, it has wholly bizarre consequences: for example, while I do not know that

- (4) I am not a brain in an Alpha-Centaurian vat,

I *do* know that

- (5) I am at home in Indiana and am not a brain in an Alpha-Centaurian vat.⁵

I know the conjunction but do not know the second conjunct. Nozick represents himself as a friend of the skeptic; but with friends like that...

Shall we say, following Fred Suppe, that the relevant way is this: my belief that P has warrant for me just if I am in an experiential state R that is a cause of my belief that P and the circumstances are such that R is a conclusive reason for P (where “ R is a conclusive reason for P under the circumstances if and only if there is a C descriptive of the actual circumstances (including both the presence and absence of various factors) such that $\neg(C \& \neg P \& R) \& \Diamond(C \& P) \& \Diamond R$ ”⁶)? There are problems here having to do with “the actual circumstances”: what precisely are they? But whatever the relevant circumstances might be, once more this account won’t work for knowledge of necessary truths: no necessary truth P is such that it is causally possible that $(C \& P)$.

Taken neat, therefore, generic reliabilism construed as in (2) and (3) above is clearly inadequate; on the other hand, if we try to specify the way in question—the way of being produced or sustained such that being produced or sustained in *that* way is necessary and sufficient for warrant—then we are ringing the changes on the various specific reliabilisms, which are also wholly inadequate. But isn’t there a quite different version of generic reliabilism? Perhaps, we could say: a belief has warrant for me at a time just if it is produced by

in “Epistemic Operators,” *Journal of Philosophy*, vol. 67, 1970, pp. 1007 ff., and “Conclusive Reason,” *Australasian Journal of Philosophy*, vol. 49, 1971, pp. 1 ff.

⁴ *Philosophical Perspectives*; 2, *Epistemology*, 1988, ed. James Tomberlin (Atascadero: Ridgeview Publishing Co., 1988), pp. 15–18.

⁵ P. 16.

⁶ *The Semantic Conception of Theories and Scientific Realism* (Chicago: University of Illinois Press, 1989), p. 368. (‘ \Diamond ’ is to be understood as ‘It is causally possible that...’)

my faculties functioning reliably *at that very time*, on that very occasion. It's not sufficient that the belief be produced by a faculty that is reliable, i.e., in many or most or nearly all circumstances produces true belief: in addition, the faculty must be working reliably on the occasion in question, the occasion on which it produces *B*. But then won't it turn out that *every* true belief has warrant? If my cognitive faculties (or some module of them) produce a true belief, then aren't they on that occasion functioning reliably? Well, perhaps not. My thermometer is stuck on 48 degrees, and it is now 48 degrees; it doesn't follow that my thermometer is now functioning reliably. Presumably thermometers and cognitive faculties function reliably on a given occasion only if they *would* produce truth in a variety of situations somewhat different from the actual situation. The thermometer would not give us the truth if the temperature were anything other than 48 degrees; therefore it is not now functioning reliably, even though the temperature just now is 48 degrees, just as it says.

Well then, what is it for your vision, say, to be functioning reliably *now*, or *on this very occasion*? Merely producing a true belief is not sufficient: what more is required? Presumably your vision would have to be such that even if things were a little different, it would still produce true beliefs. More generally, it would have to be such that even if things were a little different, a belief produced in the way in which *this* one is presently produced would probably be true. Equivalently (or nearly so) we could say that the probability that a belief is true, given that it is produced the way this one is produced, is high. But now we are back at those 'ways'; now we are back again at that litany of specific reliabilisms. For example, in the case of the thermometer, one thinks that it is functioning reliably only if, if the temperature were *not* 48 degrees, the thermometer wouldn't say that it is. Transposed to the case of belief, the suggestion would be that a cognitive faculty *F* is functioning reliably in a given case in producing a belief *B* only if, if *B* were false, *F* would not have produced *B*—i.e., if *B* were false, *S* would not have believed *B*. This of course is the primrose path that leads to Nozick and the Dretske of "Conclusive Reason." Alternatively, we might say: *F* is functioning reliably only if there is a high objective probability that a belief is true, given that it is produced by *F* in the circumstances; this is the dead end that leads to the Dretske of *Knowledge and the Flow of Information*. Or perhaps we will say: *F* is functioning reliably at *t* just if it produces beliefs that are permitted by a right system of J-rules; this is the wide road that leads to the new Goldman, and, like some other wide roads, it leads to destruction.

The general point, in brief, is this. Generic reliabilism taken neat clearly won't do the job: if we leave the "ways of production" unspecified as in (2) and (3) (above, p. 430), then clearly a belief (a visual belief, for example) can be produced in the approved way but nonetheless (due perhaps to malfunction)

have little or no warrant. But if on the other hand we try to specify more precisely the relevant ways of production, then we are back at that litany of failed specific reliabilisms. My suggestion is that we should just break down and accept proper function and the teleology that goes with it.

So much for generic reliabilism: by way of concluding this reply to Alston, I turn briefly to his argument that proper function isn't necessary for warrant. This is an argument by counterexample: Alston proposes to present a possible case in which warrant is present, but proper function is not. The essentials of the example are that we were in fact designed with a faculty of innate knowledge that was destroyed by sin; as a result we now use complex (inductive?) inferences from what we know immediately. (Since you can just turn back a few pages to see Alston's case, I won't repeat the details.) In this situation, he says, proper function and reliability "fall apart. If we were to operate according to our design plan (indeed, those segments of it that are aimed at truth), we would fail to satisfy the reliability requirement. And if we use such resources as will give us a reasonable degree of reliability, we are not proceeding in accordance with the design plan" (p. 401). But this example doesn't seem to me to be a case in which, in the relevant way, there isn't proper function. There isn't proper function of that original faculty of innate implicit knowledge, all right, but in the example we aren't forming belief by way of *that* faculty, but by way of some other processes, perhaps the processes of inductive inference we now use. In the example, these processes are functioning properly; but then (if they also meet the other conditions for warrant) that is all that is necessary for them to produce beliefs with warrant. My conditions for warrant, for a given belief *B*, don't require that *all* my cognitive faculties be functioning properly, but only those producing the belief in question. As far as I can see, this example is entirely compatible with that condition's being met; thus it isn't a counterexample.

II Ginet: Defeaters

I turn now to Carl Ginet's comments (and I don't have the space to reply to all of his penetrating observations). I will reply briefly to three of Ginet's critical comments, and then reply at slightly greater length to a fourth.

1. Ginet argues that while I claim that (epistemic) justification is not necessary for knowledge, my argument for this claim is absurdly weak: it consists in no more than an argument for the claim that justification as construed by *Chisholm* isn't necessary for warrant (along with the sly suggestion that justification *tout court* isn't necessary). He's right: that's shamelessly inadequate. There are of course many conceptions of epistemic justification, and no space to go into this matter properly here. All I can do is ask the reader to consider the 5 or 6 conceptions of justification mentioned in the first chapter

of WCD, and ask herself whether justification taken *any* of those ways is necessary for warrant. I think the answer is clear: they are not.

Furthermore, Ginet's own conception of epistemic justification, so it seems to me, is also such that justification so thought of is not necessary for knowledge. According to Ginet, a person's beliefs are justified, at a time, only if at that time the person has something of a grasp of the concept of justification:

The problem is that the application of either of the pair "justified"/"unjustified" to a particular person's beliefs presupposes that the person satisfies a certain condition that need not be satisfied by every person who has beliefs. The beliefs of a small child, for instance, cannot be said to be either justified or unjustified because the child altogether lacks any concept of the justification of belief, any sense of that sort of consideration in the formation and alteration of belief.... Unless a person has at least a modest hold on the notion of justification of belief, it would be absurd to reproach her for failing to measure up to standards of rational justification in some belief she has.⁷

Here the suggestion is that a person can't be said to be justified (or unjustified) unless she has a concept of justification; this implies, I take it, that a person can't *be* justified unless she has a concept of justification. But isn't it clear that a person could know much even if she had no such concept? Ginet seems to think of justification in a broadly deontological way: justification is a matter of being within one's rights, of flouting no duties or obligations, and the principles of justification are principles one has an obligation not to flout. A person who had no conception of duty, obligation or rights, therefore, would have no conception of justification. But couldn't someone who had no such conception, perhaps due to cognitive malfunction of some sort, nevertheless know a good deal? According to my dictionary, "psychopathic personality" is "a disorder of behavior toward other individuals or toward society in which reality is usually perceived except for an individual's social responsibilities or moral obligations...." Some people claim there are no psychopaths, but even if they are right, there certainly *could* be some; and perhaps the reason for their failure to perceive their moral obligations would be that they lack the very idea of moral obligation. Such persons would then lack the idea of Ginetian justification as well. But might they not nonetheless know a good deal, for example that the First National Bank is located on the corner of 4th and Main and that its security system will fail if the power to the bank is cut off? But then on Ginet's account, justification is not necessary for knowledge. Of course this is a bare beginning of a decent response to Ginet on this point; Steup brings up a related consideration, however, so I will defer the rest of what I have to say to my response to Steup.

⁷ "Justification of Belief: a Primer," in *Knowledge and Mind*, ed. Carl Ginet and Sydney Shoemaker (New York: Oxford University Press, 1983), p. 28.

2. Ginet proposes a counterexample to my account: his famous example of a countryside (no doubt in Wisconsin) strewn with fake barns (barn facades) where I believe (looking at one of the few real barns) that I am seeing a barn. I discuss a version of this example in WPF,⁸ adding there that the impecunious Wisconsinites had erected the barn facades in an effort to make themselves look more prosperous. My reply there was that man is a social animal; more inclusively, human beings are social creatures. Our faculty of credulity (to use Reid's name for it), by virtue of which we learn from the testimony of others, isn't designed to work, of course, in human beings who live in solitary splendor; it is designed for *social groups* of some sort, for function in families and larger groupings. And it works properly only when our fellows don't lie to us or try in some other way to deceive us; hence, so I said, the warrant a belief has for me doesn't depend merely upon myself and how things stand with me, but also upon the others around me. (In this way *warrant* is much less an *individual* matter than justification; the latter inherits its individualistic character from the individualistic character of deontology.) In my example, however, the Wisconsinites *were* trying to deceive us, so that the condition of proper function isn't met; hence my belief that I am seeing a barn does not have warrant. (If the liar tells me a truth *t* in the midst of all those lies, my belief that *t*, if it is based on nothing but the liar's say-so, does not have warrant.)

But of course it isn't necessary that those barn facades get there by way of deception. We can imagine instead that they are just being stored there, awaiting shipment to France, or that they are part of an elaborate movie set, or are entries in a local barn-look-alike contest. Then my response would be in terms of tradeoffs and compromises (WPF pp. 38–40). I see a straight stick partly in and partly out of water; unwise in the ways of sticks in water, I form the belief that the stick is bent. I am a desert tyro, see a mirage, and form the belief that there is an oasis just over *there*. In these cases there isn't malfunction, but there also isn't warrant. In WPF (pp. 38–40) I said that, due to other constraints, the design plan governing these bits of our cognitive establishment isn't aimed directly at the production of true beliefs, and the cognitive response doesn't directly serve the purpose of providing us with true beliefs; rather, in these cases that response is there because it is the locus of a best compromise among competing desiderata. The same would go for many other sorts of mistakes we make, including inductions that go wrong as in the present barn case. I added that where a belief of this kind happens accidentally to be true (I form the belief that there is an oasis just 3/4 of a mile away; as it happens, there is one 3/4 of a mile away in a different direction and just over the horizon) then what we have is a quasi-Gettier case, since the bit of the design plan governing the formation of the belief in question is not

⁸ P. 37.

aimed directly at the truth, but is instead the locus of a compromise, and the belief in question has little warrant.

But here Ginet gives the example a special twist.⁹ He adds that our cognitive faculties have been designed by a very knowledgeable being with a blind spot or two: this being has “overlooked the possibility that something might look very like a barn and not be one and thought that by constituting me so that I always respond to experience of a sufficiently barn-looking visual appearance with the belief that I see a barn he was giving me a disposition that would always produce true beliefs” (p. 404). Here my response in terms of tradeoffs and compromises doesn’t seem satisfactory. It isn’t that such a compromise necessarily requires an intelligent agent who consciously and deliberately settles on it in full knowledge of the relevant factors. For purposes of argument, at any rate, I was conceding that we might get our design plan by way of unguided evolution rather than intelligent design—so, at any rate, a nontheist will have to hold.¹⁰ It is rather that in this case we have a conscious and intelligent agent who designs cognitive faculties, but whose choice of one bit of the design plan is to be attributed to a false belief on his part; there it hardly seems we can attribute that bit of the design plan to tradeoffs and compromises.

This is an interesting emendation, and I am not sure of the right response: here I shall just mention some possibilities. First, my account as it stands doesn’t imply or entail that, in the conditions in question, I *would* know that it is a barn I am looking at. The account entails that the belief in question has warrant for me; it does not entail that it has sufficient warrant for knowledge. We can consistently add to my account that no false belief has a degree of warrant sufficient for knowledge;¹¹ and we can also add that if a true belief *B* is such that some false belief *B** (e.g., *all the things around here that look like that are barns*) has as much warrant for me as *B* does, then *B* does not have, for me, warrant sufficient for knowledge. This is one attractive possibility. Here is another. One of the conditions of warrant is that the bit of the design plan governing the formation of the belief in question meet a certain condition of goodness. In WPF I took this condition to be a matter of its being objectively probable that a belief formed according to that bit (given the satisfaction of the proper function and environmental conditions) would be true. But perhaps that is only necessary and not sufficient for goodness; perhaps it is also necessary that if the design plan is the result of conscious and intelligent design, then the bit of the design plan in question must not be due to an erroneous belief on the part of the designer. There are other possibilities

⁹ I am indebted to Andrew Koehl for calling this forcibly to my attention.

¹⁰ Although I argue in chapter 11 of WPF that there don’t seem to be any satisfactory naturalistic accounts of the notion of proper function.

¹¹ See Trenton Merricks’ “Warrant Entails Truth” forthcoming in *Philosophy and Phenomenological Research*.

as well, but the matter requires more thought. What we have here is, so to speak, a Gettier problem distributed over designer and designee; it isn't wholly clear to me how to think about design flaws due to mistake or error on the part of the designer. Of course if it is part of our concept of proper function that we have been designed by God if by anyone, then the problem, as Ginet notes, disappears; God doesn't fall into error or make mistakes.

3. I analyze our concept of knowledge in terms of proper function in an appropriate environment according to a design plan successfully aimed at truth. Ginet makes the following point: on some occasions *the basis on which we judge* that a given belief has warrant, or constitutes knowledge, is not the satisfaction of the conditions I offer as necessary and sufficient for knowledge. He apparently concludes that the proposed analysis can't be correct. But this seems to me mistaken: why suppose that an analysis is successful only if the conditions stated in the *analysans* are that by which we determine that the *analysandum* applies? The medievals made a distinction between the *ratio cognoscendi* and the *ratio essendi*; the first has to do with *how we know* that a certain condition obtains, and the second with *what it is that makes it the case that* it obtains. Even if we aren't wholly clear about this 'making it the case that' something obtains, it is clear that there is something to this distinction. A rockclimbing duffer sometimes identifies the holds on a climb by the chalk previous climbers have left on them. That's how he tells there is a hold there; but that isn't what makes the bit of rock a hold. Similarly here: while there may be a problem stating precisely what an analysis is supposed to do, clearly it isn't supposed to provide conditions which are such that it is by our grasp of their obtaining that we find out or know that the *analysandum* obtains. Here, of course, there are deep waters; in the interest of saving space, however, I'll defer the rest of what I have to say to my reply to Swinburne.

4. I turn finally and at a bit greater length to Ginet's comments on my argument (WPF chap. 12) that a certain very widely held combination of views can't be rationally accepted: the conjunction of metaphysical naturalism with contemporary evolutionary accounts of our origin. This belief, I said, is self-referentially incoherent and thus can't rationally be held. I can't recapitulate the whole argument, but its basic form is as follows. Let R be the proposition that my cognitive faculties are reliable, let N be metaphysical naturalism (which entails at the least that there is no such person as God), let E be currently accepted views about the evolutionary origin of humanity, and let $P(R/N\&E)$ be the (objective) probability of R on the conjunction of N with E. I argued in WPF chapter 12 that the proper epistemic attitude to take towards this probability¹² is that it is either low or inscrutable, i.e., such that

¹² Actually it was the slightly more complex probability ($P(R/N\&E\&C)$) (where 'C' refers to certain conditions we may here ignore), but I now think the extra complexity unne-

one just can't tell what it is. One who accepts or believes N&E, I said, thus has a defeater for R, and also a defeater for any other belief she has, including N&E itself. Hence N&E is self-defeating. Further, given our current epistemic condition, the rational thing to do is to accept E if you accept N (and perhaps even if you *don't*). Naturalism, therefore, is self-referentially incoherent, and (at least in that regard) is inferior to theism.

Now Ginet counters as follows. Contrary to what I said, he believes, the theist is in the very same leaky epistemic boat as the naturalist. For

...if we delete this component [that human beings are created in God's image] from theism and consider just the hypothesis (T-) that there is a perfect being who creates everything else, then it looks as if we could argue in just the same way Plantinga argues concerning P(R/N&E&C) to the dismal conclusion that P(R/T-&E&C) is low or unknown. Now how is it that the theist is allowed to build into her metaphysical hypothesis something that entails R or a high probability for R but the naturalist isn't? Why isn't it just as reasonable for the naturalist to take it as one of the tenets of naturalism that our cognitive systems are on the whole reliable (especially since it seems to be in our nature to have it as a basic belief)?¹³

This sounds like an eminently fair question; is there an equally fair reply? I think Ginet is raising at least two questions here. First, says Ginet, the theist believes that she has been created by a perfect being, one who is omnipotent, omniscient and wholly good, and who, furthermore, has created her in his image, so that in significant respects she resembles him: call this 'T'. T does not, of course, give her a defeater for her belief or assumption that R, that her cognitive faculties are reliable. But, says Ginet, she also believes T-, the weaker proposition that she has been created by a very powerful being.¹⁴ And consider P(R/T-&E): isn't this low or inscrutable, just as is P(R/N&E)? But if so, then I am hoist with my own petard: like the naturalist, the theist too has a defeater for R, and therefore a defeater for everything else she believes, including her theism. Second, suppose this doesn't really pose a problem for the theist after all: she doesn't really have a defeater for R, even though she believes T-, a proposition such that the probability of R on it is improbable or inscrutable. What preserves her from defeat, presumably, is the

essary. It is important to note that what is at issue here is not subjective probability taken as a measure of degree of belief, but an *objective* probability. (If it were the former, then, since the naturalist presumably believes R, N&E, and the conjunction of R with N&E, P(R/N&E), i.e., P(R&N&E)/P(N&E), would be very high.)

¹³ P. 407. Stephen Wykstra, William Alston, Richard Otte, Earl Conee and John Perry have urged this or a similar objection in discussion.

¹⁴ Here I ask Ginet's indulgence for amending his example. *His* T- is the proposition that we human beings have been created by a perfect being; but it isn't at all clear that the probability of my cognitive faculties' being reliable is low or inscrutable on the proposition that I have been created by a perfect being. If we have been created by a perfect being, then, as Descartes pointed out, we have been created by a being who is utterly trustworthy, in no way a deceiver. But the likelihood, just on that proposition, that our cognitive faculties are basically reliable, will be fairly high.

fact that she believes T as well as T-. T can be described as T- plus a little something extra, T+, we might say. So the theist has a potential defeater T-; she staves off defeat by adding a little something extra, thus moving to T+. But then why can't the naturalist do the same thing? Why can't he too add a little something extra, perhaps the proposition that "our cognitive systems are on the whole reliable," thus moving to N+? P(R/N+&E) is not low or improbable; N+ entails R. If the theist can stave off defeat by adding something to T-, why can't the naturalist do the same by adding something to N?

Take the second question first. Can the naturalist safeguard R from defeat just by adding a little something to N&E, just by embracing (for example) N+, the conjunction of N with the proposition that, probable or not, as a matter of fact we have won the lottery and our cognitive faculties are indeed reliable? "Why isn't it just as reasonable for the naturalist to take it as one of the tenets of naturalism that our cognitive systems are on the whole reliable (especially since it seems to be in our nature to have it as a basic belief)?" Now it's pretty clear that this sort of thing can't be right *in general*; if it were, no defeater could succeed. For example, suppose I believe the Bible is a special revelation from God and is therefore infallible: everything affirmed therein is true. I injudiciously read Mark Twain and contemporary Scripture scholarship, however, and come to believe that the Bible is chock-full of errors; as I see it, the likelihood that a proposition is true, given that it is affirmed in the Bible, is no greater than that for any other ancient book, and perhaps about 1/2. I then have a defeater *D* for anything I believe just on the basis of Biblical teaching. Now consider any belief *B* I do hold just on that basis—for example, the belief that God was in Christ, reconciling the world to himself. *D* is a potential defeater for *B*; could I protect the latter from defeat by adding a little something to *D*, maybe *B* itself, thus moving to *D*+? *B* is not very unlikely on *D*+, right? But of course this is not the method of true philosophy. If you discover a defeater *D* for one of your beliefs *B*, you can't in general deliver *B* from defeat just by conjoining it or something that entails it with *D*, noting triumphantly that it is not improbable with respect to that conjunction. In particular, if your belief that N&E is a potential defeater for R, you can't preserve R from defeat just by conjoining it with N and pointing out that R is not improbable with respect to N&E&R. So the naturalist cannot preserve N&E from defeat in this way.

But isn't this addition procedure just what the theist does? And if the naturalist can't properly do it, why think the theist can? This takes us to the first question: the theist believes T-; R is unlikely or inscrutable with respect to T-; the theist apparently preserves R from defeat just by adding a little something to T-, thus arriving at T+, i.e., T. How can the theist get away with this, if the corresponding course for the naturalist is disreputable? R is im-

probable or inscrutable on T-; the theist believes T-; doesn't she therefore have a defeater for R?

This suggestion seems initially to presuppose or endorse

- (1) for any propositions *A* and *B* I believe, if *B* is improbable or inscrutable with respect to *A*, then *A* is a defeater for *B*.

But this 'principle' is false. For example, I believe

- (2) all dogs live at or near the surface of the earth,

and I also believe

- (3) either all dogs live at or near the surface of the earth or some dogs don't.

(2) is either objectively improbable or inscrutable with respect to (3); but (3) is not a defeater, for me, of (2). I also believe both that I own an old Nissan Sentra, and that the population of China vastly exceeds that of Scotland; each of these is objectively improbable or inscrutable with respect to the other, but neither is a defeater for the other. Still further, I believe that I own an old Nissan Sentra, and also that I own an old automobile; the former, although entailing the latter, is objectively improbable with respect to it (the vast majority of those who own old cars, do not own old Nissan Sentras). But the latter isn't a defeater for the former. Therefore (1) is false.

These three cases are significantly different from each other, and here I shall have to be brief and enigmatic. (I hope to address these and some other questions about defeaters later.¹⁵) In the first, the prospective defeater is entailed by the prospective defeatee, but is an obvious necessary truth; therefore the epistemic probability¹⁶ of the prospective defeatee on the prospective defeater and the relevant rest of my noetic structure¹⁷ is equal to its epistemic probability on my noetic structure minus the prospective defeater. But a defeater *D* for a proposition (for a person *S*) must be such that it *lowers* the epistemic probability of the prospective defeatee: it must be the case that the epistemic probability of the proposed defeatee on the conjunction of *D* with the relevant rest of *S*'s noetic structure is lower than on that relevant rest alone. In the second case, the prospective defeatee is not entailed by the

¹⁵ In "Naturalism Defeated," not yet committed for publication.

¹⁶ For the notion of epistemic probability, and the distinction between the latter and objective probability, see WPF chaps. 8 and 9.

¹⁷ For present purposes, my noetic structure is what I remember and believe together with my present experience. I speak of the *relevant* rest of my noetic structure: that relevant rest cannot, of course, include the prospective defeatee itself, or the conjunction of that proposition with something else, or any proposition whose warrant is derivative from that of the prospective defeatee, and so on. More about this in "Naturalism Defeated."

prospective defeater, but the latter is nevertheless probabilistically irrelevant to the former, in that (as in the last case) the epistemic probability of the prospective defeatee on the prospective defeater together with the rest of my noetic structure equals the probability of the prospective defeatee on the rest of my noetic structure alone. As in the first case, therefore, the proposed defeater is unsuccessful, since it does not lower the epistemic probability of the proposed defeatee. In the third case, the prospective defeatee entails the prospective defeater, but my warrant for the latter comes via my warrant for the former. I know that I have an old car, all right, and the probability that I have an old Sentra is low on that proposition, all right, but I know that I have an old car only because I know that I have an old Nissan Sentra. The warrant enjoyed by the first belief, we might say, is derivative from or parasitic upon that enjoyed by the second. But in such a case the first belief does not function as a defeater of the second.¹⁸

Well then, we can agree that (1) is false; still, why isn't T- a defeater for R, for the theist? Isn't she just adding something to T- in order to stave off defeat? And how can she get away with that, if the corresponding course is villainous for the naturalist? The answer is intimately related to the reason that my knowledge that I own an old automobile is not a defeater for my belief that I own an old Nissan Sentra. We saw above that, where *A* is improbable on *B* (together with the relevant rest of *S*'s noetic structure), *B* is not a defeater of *A* if the warrant *B* enjoys for *S* is derivative from the warrant *A* enjoys for *S*. But there is a stronger principle lurking in the neighborhood. Consider again my belief that I own an old car: this does not constitute a defeater for my belief that I own an old Nissan Sentra, even though the latter is improbable on the former; the reason is that the warrant the former has for me is derivative from the warrant the latter has for me. But note that the former is also not a defeater for my beliefs that I own a Japanese car, that I own

¹⁸ But if (1) is false, why think that N&E (together with the proposition that $P(R/N\&E)$ is low or inscrutable) offers a defeater for R? This raises more general questions about the nature of defeaters and how they work (questions I address in "Naturalism Defeated"); here I shall have to leave matters at the more or less intuitive level I left them in WPF. There I argued by analogy with other cases that N&E (together with the proposition that $P(R/N\&E)$ is low or inscrutable) is indeed a defeater for R. There were the two widest cases. There was also the case of the theist who unwisely read too much Freud, coming to the conclusion that her theistic belief was really produced by wish fulfillment: she has a defeater for T if she thinks that the probability of a belief's being true, given that it is produced by wish fulfillment, is either low or inscrutable. We might add another case: suppose I foolishly read too much brain-in-vat literature; as a result I find myself with the belief that I am myself a brain in an Alpha-Centaurian vat, being given the beliefs and experiences I have by Alpha-Centaurian cognitive super-scientists who are running an experiment of some kind, but aren't at all interested in my having true beliefs. Call this belief 'BV': BV and $P(R/BV)$ is low or inscrutable together give me a defeater for R. And, so I claim, the case of R and N&E relevantly resembles these cases.

an old Nissan, that either I own an old Nissan Sentra or Brown is in Barcelona, etc.—even though each of these is improbable on the former, and even though the warrant of the former, for me, isn't derivative from that of any of these. The reason, of course, is that each of these gets the warrant it possesses from the proposition that I own an old Nissan Sentra, the very proposition from which *I own an old car* gets its warrant. So we can add another principle here: If propositions *B* and *C* are both such that the warrant they have, for *S*, is derivative from the warrant *A* has for *S*, then *B* is not a defeater for *C* (even if *C* is objectively improbable on *B*).

More generally,

- (4) If the warrant *A* has for *S* is derivative from the warrant *B* has for *S*, then *A* is a defeater for a belief *C* for *S*, only if *B* is a defeater for *C* for *S*.¹⁹

(We can conveniently take it that if *A* has no warrant for *S*, then the warrant *A* has for *S* is derivative from the warrant *B* has for *S*, for any *B*.) But now return to T, T- and R. It is clear that the warrant T- has for the theist is derivative from the warrant T has for her; she knows or believes that she is created by a very powerful being only because she knows or believes that she is created by God, an omnipotent, omniscient and wholly good being. But then T- isn't a defeater, for her, for any belief for which T isn't a defeater. Clearly T is not a defeater, for her, for R. Neither, therefore, is T-.

III Steup: Use vs. Function

Matthias Steup makes essentially three points. First, he proposes an important distinction: that between the proper *function* of our cognitive capacities, and their proper *use*. "We can distinguish two kinds of drivers: those who perform poorly even with a properly functioning Mercedes Benz, and those who do well even with a beat up and improperly functioning Yugo" (p. 409). There *is* such a distinction to be sure, but it's a bit delicate and must be employed with caution. My relationship to my cognitive faculties is very different from my relation to my automobile. To use my Yugo, or even my Mercedes, I must of course use my cognitive faculties, which are not part of the automobile in question. To drive my Yugo to Cleveland, for example, I must know how to start it, where Cleveland is from here, how to get it pointed towards Cleveland, that it is probably better to follow roads than take a direct cross country compass course, and a good bit more. To use my Yugo well, I have to use it intelligently; to use it intelligently I must employ my cognitive faculties. My use of my own cognitive faculties, on the other hand, is quite different; I don't have *another* set of cognitive faculties which I employ

¹⁹ Here I am heavily indebted to correspondence with Stephen Wykstra.

to use *them*. Further, I can evaluate any aspect of my automobile's functioning in order to determine whether it is working properly or working the way I want it to. To put it metaphorically, I can stand outside my automobile and measure its performance against my goals; and in so doing, I am not obliged to rely on its proper functioning.

This is just what I can't do with my cognitive faculties. Obviously (and trivially) I can't make use of any of my cognitive faculties without using some of my cognitive faculties: more to the point, I can't use any of my cognitive faculties without *relying on* some part of my cognitive faculties. In any use of those faculties, I must rely upon them, trusting that they are indeed functioning properly. Clearly I can't determine that they are functioning properly or are reliable before I rely on them. As Reid understood Descartes, the latter refused to trust his faculties until he had established that they are indeed worthy of trust; and as Reid points out, this is a silly and self-deluded dream. I could establish the trustworthiness of my faculties only by assuming that at least some of them are indeed trustworthy: I am, so to speak, at the mercy of my cognitive faculties in a way in which (happily) I am not at the mercy of my Yugo. This makes my relation to my cognitive faculties vastly different from my relation to my automobile. In particular, at a certain basic level I must simply trust my cognitive faculties; I have no alternative. At that basic level, the issue of using them well or ill doesn't arise: how they function and whether they function reliably (a) isn't up to me, and (b) is such that I can't determine it without assuming it.

Steup goes on to point out that one of my beliefs may fail to have warrant because of impatience, overeagerness, or other psychological conditions on my part; these conditions may interfere with warrant, he says, even though they do not involve anything quite as portentous as cognitive malfunction. (Hence my *analysans* is not sufficient for my *analysandum*: proper function (together with the other requirements) isn't enough; there must also be no interference from these possibly impeding psychological factors.)

Now I certainly agree that proper function can be impeded by pride, ambition, lust, anger, patriotism, fear, greed, impatience, buck fever, mother love, avarice, hate, undue sensitivity, excessive pessimism (or optimism) and the like; and when this happens warrant is often excluded. (We speak quite properly of someone as *blinded* by ambition, or pride, or lust, or mother love.) I was thinking of such lack of impedance or interference as included in proper function; cognitive faculties function properly, in my sense, only where their functioning isn't impeded by these conditions. Although this is how I was thinking of the matter, however, I note to my chagrin that I don't say so in the appropriate place in *Warrant and Proper Function*.²⁰ So the proper func-

²⁰ I did say so elsewhere, in "Positive Epistemic Status and Proper Function": "It may be (and in fact is) the case that it is not at all unusual or abnormal for a person to form a

tion clause should be understood as including that the function of the faculties or powers in question is not impeded by conditions of that sort. Such impedance, however, isn't clearly a matter of cognitive malfunction; to that extent, therefore, my use of the term in the *analysans* is technical. Of course a really proper job here would explore these conditions and more besides: just how do or can they impede proper function? Here we need psychology as well as (or more than) philosophy; in any event this is far beyond our present scope. But Steup is certainly right: these conditions can interfere with warrant by impeding proper function.

Now I propose to describe these phenomena in terms of 'impeding proper function'; Steup prefers to put it in terms of 'failing to use cognitive faculties properly'. I believe this difference between us reflects a deeper difference: Steup is thinking of failure to use cognitive faculties properly, I suspect, as a matter of using them in a way which contravenes epistemic duty, so that one is blameworthy for so using them. On the other hand, I am thinking of impedance of proper function as possibly but not necessarily due to one's own blameworthy behavior. I get drunk and can't see properly; here the impedance is due to my own misbehavior. On the other hand, loyalty and love for your friend prevents you from seeing what you otherwise might have, namely that she has been lying to you; here you need not be at fault or properly subject to reproach. Steup is really seeing the problem, in these cases, as lack of justification; I am not. The impedance in question *can* be a result of unjustified behavior, but it need not be so. And even when it is, it isn't the lack of justification (the dereliction of duty) that is responsible for lack of warrant: it is the impedance.

This helps me with the next step in Steup's dialectic. He holds that if I include such conditions as the above among the causes of cognitive dysfunction, then I am confronted with a new problem:

While Plantinga's initial account of malfunction is not specific enough to allow him to handle the case of the hasty observer, the expanded account is too specific to handle the following variation of the case....Suppose that when our observer sees the squirrel which he mistakes for a chipmunk, he also sees, just a bit behind the squirrel, a rabbit. Hopping around, the rabbit remains in view just long enough to be clearly discerned. So now our observer forms *two* beliefs at the same time: "There's a rabbit over there: and There's a chipmunk in the bushes." While the first is warranted, the second is not.

And the alleged problem for my account is that "since both beliefs are formed simultaneously and produced by exactly the same faculties,...[i]t follows that

belief out of jealousy, lust, contrariness, desire for fame, or wishful thinking; nevertheless when I form a belief in this way my cognitive equipment is not functioning properly; it is not functioning the way it ought to." *Philosophical Perspectives*, 2 *Epistemology*, 1988, ed. James Tomberlin (Atascadero, California: Ridgeview Publishing Co.), p. 32.

either both of them are produced by properly functioning faculties or neither is” (p. 411)—in which case on my account, says Steup, either they both have warrant or neither does, contrary to the plain facts of the case.

But here I think he’s mistaken, at least if we see the the warrant inhibitor in this case as impedance rather than lack of justification. I may be blinded by pride or lust or ambition; but of course this will interfere only with *some* belief formation, not with all of it. Blinded by avarice, I unwisely invest in Florida real estate, forming the belief that it is prime resort property which is bound to appreciate at an enormous rate, just as the salesman says. That belief does not have warrant for me; but of course other beliefs formed in the course of the same conversation—e.g., that the real estate in question is in Florida, that the salesman’s name is Paul M. Zwier and that he represents Pot O’Gold Ventures—may well have warrant, even if those beliefs are produced by the same faculties or powers. Avarice can affect the formation of different beliefs differently, even if they are produced at the same time and by the same faculties. Avarice interferes with warrant by impeding the proper function of the faculties or powers in question; it can do this with respect to one belief, while leaving wholly unimpeded their function with respect to the production of another belief. And thus, if we include being unimpeded by these conditions as part of proper function, we can say that (at a given time) a faculty or process (or ensemble of the latter) is functioning properly in producing one belief, but not functioning properly in producing *another*.²¹

So suppose I have buck fever and nervously fire at what I mistakenly take to be the flash of a buck’s white tail: my undue eagerness may interfere with the proper function of perception in the production of *that* belief, but it needn’t necessarily interfere with my perception of the clump of trees in which I think I see the buck. And the same thing goes, so it seems to me, for Steup’s example. My inordinate eagerness to see wildlife may interfere with the proper function of my perceptual faculties, thus producing the belief that I see the chipmunk; their proper function in producing the belief that I see a rabbit is not thereby necessarily compromised. (After all, it is much harder to see a chipmunk at 35 yards than a rabbit.) So there is no reason for supposing that, on my account, two beliefs of the same strength formed at the same time by virtue of the operation of the same faculties can’t differ with respect to warrant.²² Here there may be a difference with justification, as Steup con-

²¹ Here a remark (private communication) of Steup’s was helpful.

²² Steup claims that “Epistemologists who analyze warrant in terms of cognitive faculties face the general challenge of specifying which cognitive processes, or intellectual abilities, qualify *as faculties*” (pp. 412–13, his italics). But is this really true? If the notion of faculties or powers were in total disarray, then perhaps this would be true—or rather, then the analysis would be of no use at all. But of course things aren’t like that. As Steup points out, we know of plenty of paradigm cases of faculties, and plenty of paradigmatic cases of things that are not. (We also know, for example, that a faculty is precisely the sort of thing that can function well or ill, can function properly or be

ceives it; perhaps if I am unjustified or irresponsible in forming the one belief, I am also unjustified or irresponsible in forming the other. And perhaps this is why Steup thinks the two beliefs must share the same status with respect to warrant.

Finally, Steup offers me a way out: he suggests that perhaps we should think here in terms of *meta-faculties*, faculties that regulate the output of ordinary first level doxastic faculties; and he adds that if we do, I can perhaps elude the previous objection. I must decline the offer with thanks; I prefer to elude the objection by pointing out that an interfering condition can impede proper function with respect to one belief while not impeding it with respect to another formed at the same time and by the same faculties. But I don't mean to pour cold water on the suggestion that there are meta-faculties of the kind he proposes; perhaps he's right. Steup goes on to say that such faculties, if there are any, would in some ways resemble virtues, such as courage and temperance. As he points out, virtues are dispositions to behavior; but they are also dispositions to belief. Courageous and cowardly persons differ in belief as well as in action; and a generous person is much less likely to put the worst possible construction on your behavior than the departmental misanthrope. This leads to the deeper question here: what is the relation between warrant and justification? (Or intellect and will, or virtue and knowledge?) How does epistemic duty fit in with warrant? According to Steup "Plantinga attempts to capture what warrant is by eliminating epistemic deontology without a trace" (p. 412). True: I don't think (epistemic) justification construed deontologically is either necessary or sufficient for *warrant*; a person can fulfill all relevant intellectual duties but fail to know, and a person can know (so I am inclined to think, anyway) even when flouting intellectual duties. But of course it doesn't follow that satisfying intellectual duty is irrelevant to the well conducted cognitive life. Obviously justification can be of prime importance, even if it is not necessary for warrant. Furthermore, there may be certain *kinds* of knowledge such that satisfaction of intellectual duty is a necessary condition of possessing knowledge of *those* kinds, even if it is neither necessary nor sufficient for knowledge in general.

And here I return to Ginet's concern about justification. As you recall, he complained (quite sensibly) that I said justification wasn't necessary for warrant, but then supported this contentious claim only by arguing that *Chisholm's* version of justification isn't necessary for warrant; aren't there many other versions? Of course there are; so let's look into the matter just a

subject to dysfunction.) Thus memory, vision and other sensory modalities are clear examples of faculties or powers; forming beliefs on Fridays, for example, is not. Further, we make a significant advance if we see the connection between warrant and the proper function of cognitive powers and faculties, even if we don't have a proper theory of the latter. (It would be a fine thing to have a proper theory of the latter; but please, one thing at a time.)

bit further. When I said that justification isn't required by warrant, I was thinking of the root component of the contemporary versions of justification: the idea of justification deontologically considered. The basic question here is whether the fulfillment of duty (presumably *epistemic* duty) is necessary for warrant. But this question itself needs some work. First and trivially, it might be that warrant *as such* doesn't require the fulfillment of epistemic duty, but a degree of warrant sufficient for knowledge does. But second and less trivial, what, exactly, *is* it to say that deontological justification is necessary for warrant? What is the justificationist holding? Is it that for each of my beliefs there is a duty such that my satisfying that duty is necessary for the belief to have warrant? Or perhaps the stronger claim that there is at least one duty such that my satisfying that duty is necessary for any of my beliefs to have warrant? This may be stronger, but it is still too weak to be interesting. One duty I plausibly have is to refrain from knowingly destroying or permanently damaging my cognitive faculties, by, for example, taking mind altering drugs with long-term ill effects. Clearly a belief of mine might have warrant only if I have satisfied a duty of this sort; if the relevant faculties are functioning properly, it follows that I have not permanently damaged them, and hence *a fortiori* have not *knowingly* inflicted permanent damage on them. But this kind of duty is presumably not what the justificationist has in mind.

You might reply that my duty really is to *try* not to abuse my faculties in that way, and that the satisfaction of *that* duty is not entailed by their proper function. Satisfaction of that duty, however, is also unnecessary for warrant. I try to destroy my vision by taking a drug I think will have that effect; the drug doesn't work, and even as I take it I continue to form beliefs due to vision (I see that the bottle is half empty, for example); couldn't those beliefs perfectly well constitute knowledge for me?

But of course this also is not what the justificationist has in mind. She thinks instead of such alleged duties as that of trying to maximize true belief and minimize false belief, or of trying to bring it about that you believe only those propositions for which you have good reasons, or of trying to pursue your doxastic goals in effective fashion.²³ But suppose there are such duties: couldn't I know much even if I didn't satisfy them? I could know what my name is, where I live, and that $7+5 = 12$, for example, even if I pay no attention at all to maximizing true belief, even if I regularly believe lots of things for which I have no good reasons, and even if I am in general thoroughly lax with respect to pursuing my doxastic goals. I therefore can't see that for each belief I have, there is a duty of the relevant sort such that satisfying that duty is a necessary condition of that belief's having warrant for me. As I said above, however, it is compatible with this that there be certain duties and certain kinds of beliefs such that the latter have warrant for me only if I satisfy

²³ See chap. 1 of WCD.

the former. Perhaps duty, obligation, responsibility have a more significant place in more developed, less elementary areas of my noetic establishment, areas where it is plausible to think I have more control over belief formation. Science and scholarship might be such areas; and of course one feature of justificationist accounts is that they fit that kind of intellectual activity better than ground level cognitive activity (as when I remember what I had for breakfast this morning). But it is very hard to think of a duty of this sort such that satisfying it is necessary for *all* kinds of knowledge, and equally hard to see how for each kind of knowledge there is a duty of the relevant sort such that satisfying that duty is necessary for forming warranted beliefs of that kind.

As I am painfully aware, what I have said here is wholly inadequate to this large, difficult and important topic of the relation of warrant to (deontological) justification; I hope to say more about it in *Warranted Christian Belief*.

IV Swinburne, Analyses and Univocal Concepts

I turn now to Richard Swinburne's characteristically provocative comments. His remarks have that good news-bad news structure. His main claim, so far as I can see, is that "Plantinga's account of knowledge is not a useful analysis even of the central and paradigmatic cases..." (p. 419); this is the bad news. The good news, however, is that my failure is excusable or at least explicable: there *aren't* any necessary and sufficient conditions for knowledge or warrant. What we must recognize, he thinks, is "that no necessary-and-sufficient conditions account is going to capture all and only cases of knowledge, or even the central and paradigmatic cases of knowledge, if such there are" (p. 419). This is because knowledge is too *loose* a concept to *have* necessary and sufficient conditions: "Though not as loose a concept as proper functioning, knowledge is still too loose a concept for that" (p. 419). So the idea is that there just isn't any set of conditions (not including knowledge or warrant itself) that are severally necessary and jointly sufficient for knowledge (warrant); if so, then of course my proposed necessary and sufficient conditions are bound to be either unnecessary or insufficient.

On the other hand, my account is also deficient, thinks Swinburne, in that it crucially invokes the notion of proper function, a notion that is much less clear than that of knowledge (or warrant) itself, so that my account provides not clarification but obfuscation: "The trouble with this account is that it is much clearer what counts as a case of knowledge, than what counts as satisfying some of the clauses in this account; the account does not clarify but obfuscates our understanding of knowledge" (p. 415).

So the problem is two-fold: on the one hand, warrant doesn't have necessary and sufficient conditions; so (naturally enough) the conditions I propose

aren't necessary and sufficient. On the other hand, the conditions I propose are less *clear* than the concept of knowledge, in the sense that in some cases it is clearer that knowledge is present than that the conditions in question are or are not satisfied. I shall comment on each of these, and conclude with a remark about Swinburne's suggestion that the concept of proper function is not univocal among theists and non-theists.

A. No Necessary and Sufficient Conditions for Warrant?

Swinburne claims the conditions I provide are not in fact necessary and sufficient, because there aren't any necessary and sufficient conditions for warrant; and there aren't any necessary and sufficient conditions of warrant because the concept is too loose. But what is this 'looseness' and what is it for a concept to be loose? We know what it is for a *word* to be *vague*; is looseness for concepts analogous to vagueness for words? We also know what it is for a word to be *ambiguous*; is looseness for concepts analogous to ambiguity for terms? To speak of our concept of so and so is to speak of our grasp of a certain property: can *properties* be loose? Or is the looseness located in something like a weak or partial grasp of the property or properties in question? (Or still another possibility: is looseness really a property, not of concepts or properties or even words, but of someone's *use* of a word?) And suppose a concept were loose: why would it follow that it doesn't have interesting necessary and sufficient conditions? Swinburne doesn't answer any of these questions and doesn't say what he thinks looseness is; extrapolating from his comments, however, I conjecture that a concept is loose, as he sees it, when there are many possible states of affairs such that with respect to each "there seems no clear answer" to the question whether or not the concept applies in it. (We must add, I suppose, that mere ignorance on our part is not the reason there seems no clear answer to this question.) This sounds like *vagueness* to me; at any rate a vague concept (*bald*, e.g.) meets this condition by virtue of its borderline cases. But a concept's being vague is by no means inconsistent with its having interesting necessary and sufficient conditions, although it might be that the necessary and sufficient condition itself involves a vague concept. (A tadpole is a non-adult frog; the vagueness of the concept *tadpole* is matched by that of *adult* and *frog*.)

Whatever precisely looseness is, however, the best way to show that my proposed necessary and sufficient condition for warrant isn't both necessary and sufficient, I should have thought, would be to give a clear counterexample or two: cases where the condition I suggest is met, but there isn't warrant, or cases where there is warrant but the condition I suggest is not met. Swinburne does not follow this course. So far as I can see, his argument for the claim that the condition I specify is not necessary and sufficient, depends wholly upon his claim that there *are* no (nontrivial) necessary and sufficient

conditions for warrant. If this is true, then of course my conditions will not be necessary and sufficient, in which case there will be counterexamples to my proposed analysis: why not exhibit a case or two just to put our minds at rest? Further, why think that there aren't any necessary and sufficient conditions for warrant? One suggestion he makes is that the concept is too loose to have necessary and sufficient conditions; but he neither argues that it *is* loose nor explains why looseness precludes interesting necessary and sufficient conditions. He does suggest, however, that "there is good inductive reason for supposing that no satisfactory general account can be given of that third element involved in knowledge which Plantinga calls 'warrant'" (p. 419). The good inductive reason, I take it, is that none of the other accounts of warrant succeeds; but then wouldn't it be all the more imperative to show that the account I suggest doesn't do the job either?

Perhaps Swinburne would reply that he is not obliged to produce counterexamples, because the conditions I propose are too loose, much looser than knowledge itself; and if proposed conditions are sufficiently loose, counterexamples are not to be expected. Now as I said, I don't know just what Swinburne has in mind by 'looseness', but one can still understand the response. Suppose I propose an analysis of the concept *horse*: *x* is a horse, I say, if and only if *x* is a nonhuman animal and is such that if it were healthy, it would be both alluring and audacious. Well, it isn't so easy to tell when a nonhuman animal is audacious, let alone alluring, and even harder to tell when it meets the condition of being such that, if it were healthy, it would be both audacious and alluring. To come up with a clear counterexample showing that the proposed condition is either unnecessary or insufficient, we would need an example of a horse that wasn't a nonhuman animal or else wasn't healthy only if alluring and audacious; alternatively, we would need an example of a nonhuman animal that was healthy only if alluring and audacious, but not a horse (a peculiarly attractive golden retriever?). It might not be easy to come up with either. Still, that difficulty, one thinks, hardly makes it obligatory to accept the analysis. And perhaps Swinburne thinks something similar holds for the conditions I propose. I have some sympathy with that complaint; some of the conditions involved in my account of warrant are at any rate vague, whatever the case with looseness. But of course so is warrant itself; I will argue below that the conditions in question aren't really any vaguer than what they are the proposed conditions for. Proper function may be a vague and wavering notion but so is warrant; and they waver together, just as they should if my way of linking them is right.

With respect to this first difficulty, therefore, all I can say is that Swinburne's general claim (that there are no necessary and sufficient conditions for warrant) seems to me mistaken. This is because the condition I suggest (along with some qualifications) in WPF seems to me to be necessary and

sufficient. More modestly, it is *nearly* necessary and sufficient (I know people say there are no near misses in logic, but I think they are wrong); or, still more modestly, it is at least on the right track, such that with some further chisholming, it will do the trick. Since Swinburne doesn't argue that they aren't necessary and sufficient, and since his only argument for there being no necessary and sufficient conditions is just that no one, he thinks, has so far succeeded in producing a good analysis, there isn't much for me to say by way of reply; the ball is still on his side of the net.

B. Proper Function Less Clear than Warrant?

Swinburne suggests that the analysis I propose isn't useful because the notion of proper function is *less clear* than that of warrant (knowledge). Now I'll argue below that this isn't true. But suppose it were: would it really follow that the analysis isn't useful? It might follow that there is a particular way in which it wasn't useful: one couldn't use it, presumably, to help decide whether a disputed case was or wasn't a case of knowledge. And this seems to be the problem Swinburne finds here: "The trouble with this account is that it is much clearer what counts as a case of knowledge, than what counts as satisfying some of the clauses in this account; the account does not clarify but obfuscates our understanding of knowledge" (p. 415). But is there any reason to think that the only or main use an analysis can have is that of enabling us to tell, in difficult cases, whether the *analysandum* holds?²⁴ In a successful analysis, the *analysans* must of course be equivalent, in the broadly logical sense, to the *analysandum*. It must meet some further conditions as well: for example, it can't be the very *same* property or concept (or perhaps the very same property or concept presented the same way), or a construction on a set of properties including the property in question; further, perhaps the *analysandum* and *analysans* must be such that it can be known *a priori* (by human beings) that they are equivalent; still further, the *analysans* must perhaps be *simpler*, in some sense, than the *analysandum*. All of this still won't be enough, and it isn't easy to say what more is required. But one thing that is *not* required is that it be easier to tell whether the *analysans* applies in a given case than whether the *analysandum* does.

Thus I am inclined to think that moral obligation is to be analyzed in terms of God's will; but I don't think this will necessarily be of much help when it comes to trying to see what my duty is in a given situation. (The same goes, notoriously, for analyses of obligation in terms of the greatest good of the greatest number.) Analyses of personal identity don't ordinarily help with disputed cases; and the same goes, again notoriously, for functionalist and other 'naturalistic' analyses of mental terms. An analysis can be useful and informative, can contribute to real understanding, even if it is harder

²⁴ See my comments on a similar suggestion by Ginet, above, p. 438.

to determine whether the *analysans* applies than the *analysandum*. So I doubt that this criticism of Swinburne's points to a genuine difficulty—even if in fact it is harder to determine the application of my conditions for warrant than the application of warrant itself.

But *is* it? I doubt it. The alleged problem is with the notion of proper function: “Or to take the example of human cognitive faculties—are the cognitive faculties of the mathematical *idiot savant* producing immediately the answers to very complex arithmetical questions, or those of the man with total recall memory, or those of the tipster who has the knack of guessing the winner of horse races or those of the successful water diviner functioning properly or not? There seems no clear answer” (p. 416). Now it might be that there seems no clear answer because there are important things we don't know here; it could be that the problem here is like when I can't give a clear answer to the question “How far is it from where I am to where you are?” because I don't know where you are. That wouldn't suffice to show that the concept of distance is particularly unclear, and perhaps the same goes here.

But whether or not the same does go here, isn't it also the case that there is no clear answer to the question whether or not the *idiot savant* really *does* know the answer to those questions? He *gives* the right answer, all right, but does he really *know* the answer? The calculating twins Oliver Sacks describes²⁵ were retarded and cognitively damaged in some ways, but could also perform marvelous feats of apparent cognition: they could just see, so it seems, whether a given 6 and even 8 digit number is a prime. But do they really *know* the relevant propositions? According to Sacks, the twins don't seem to have the concept of multiplication at all (p. 197); it is therefore far from clear that they can form the belief *n* is prime, for some six digit *n*, or indeed for any *n* at all. Their cognitive faculties are seriously defective, one doesn't know enough about what they can and cannot believe to be sure precisely what beliefs they do and do not have. One therefore can't really tell whether they know these propositions or not. Similarly for the person who successfully guesses the winner of horse races and the successful water diviner: Do they *know*? The answer isn't clear. I lean towards thinking that they *do* know, at least in some cases, but I'm not sure: there seems to be no sure answer. But then the wavering character of the notion of proper function, here, is matched by that of the notion of warrant, just as it should be if they are linked in the way I suggest.

C. Two Concepts of Proper Function?

Finally, a comment on Swinburne's suggestion that there are really two distinct concepts of proper function, a supernaturalistic one and a naturalistic

²⁵ “The Twins” in *The Man Who Mistook his Wife for a Hat* (New York: Harper and Row, 1970), p. 195 ff.

one. His idea, apparently, is that the naturalist has one concept of proper function and the theist or supernaturalist quite another: "...Plantinga...supposes that there is a universal concept of proper function common to the naturalist and the supernaturalist. That assumption seems to me mistaken" (p. 415). But he also thinks that we all, theist or not, have the naturalistic concept of proper function: "We certainly have a concept of proper function which we can apply in a rough and ready way in our talk about bodily organs....All of this gives us a loose naturalistic concept—naturalistic in that its application does not assume design by a creator..." (p. 416). So the theist has *two* concepts of proper function: the supernaturalistic one and the naturalistic one. Swinburne thinks the two concepts differ in that the supernaturalistic one is much more definite: something is working properly when it is working in accord with the intentions of its designer(s). The naturalistic concept, he says, is much looser; it is one we learn from paradigm cases (a bird's wing working properly, a heart's malfunctioning), but it has little connection with other concepts. It is also "loose" in the sense (I think) that it has a large number of borderline cases, or in any event cases where there seems no clear answer to the question whether or not the concept applies. Further, as a result of this looseness, the naturalistic concept of proper function has no necessary and sufficient conditions.²⁶ Still further, Swinburne apparently believes that neither of these concepts (unhappily for me) will work in my analysis. The naturalistic one is too loose to work. The theistic one, although it does not suffer from *that* debility, suffers from another: if, in my analysis, we understand *proper function* as *supernaturalist* proper function, then the proposition that I know that today is Monday entails that I have been designed by a conscious and intelligent designer. But "...clearly the claim that I know that today is Monday does not *entail* the claim that the cognitive faculties involved in the production of my true belief that today is Monday are functioning in the way their creator intended" (his emphasis).

But why does Swinburne think there are two different concepts of proper function here? And why does he think one naturalistic and the other supernaturalistic? Perhaps 'supernaturalistic' isn't quite the right term: if there are two concepts here, the important point about the second is not that it involves

²⁶ Swinburne thinks I regard this looseness, or perhaps what he sees as the consequent lack of necessary and sufficient conditions for the naturalistic concept, as a defect, as grounds for moving beyond it to the supernaturalistic concept (pp. 416–17). He also thinks I hold that the fact that the naturalist can't give necessary and sufficient conditions for the naturalistic concept means that the concept in question is 'unintelligible' (p. 416). But here I must demur. Perhaps it is true that our concept of proper function doesn't have nontrivial necessary and sufficient conditions, but if that is true, it is not in my opinion a defect in the concept; in particular, it isn't grounds for supposing the concept unintelligible. (Many magnificently intelligible concepts do not have non-trivial necessary and sufficient conditions—for example, *red*, *horse*, and *belief*.)

something beyond nature, but that it involves the concept of *intelligent design*. True, in the case of natural organisms and their parts, if proper function *does* involve intelligent design, the best candidate for being the intelligent designer would be God. But it isn't analytic that if plants and animals have conscious, intelligent designers, then they are designed by God. Nor is it obviously necessarily true. Is it clearly impossible that God *delegate* the task of designing certain creatures to other beings? Is it clearly impossible that some creatures have been *redesigned* by others? Not obviously.

If there are two concepts of proper function here, therefore, one would involve conscious and intelligent design and the other would not. But *are* there two concepts here? Why think so? In particular, is it true, as Swinburne thinks, that theists have both the naturalistic concept (one not involving conscious and intelligent design) and also another one that does involve conscious and intelligent design? And how could we tell whether this is true? I am certainly not aware that *I* have two concepts of proper function; and when I talk about proper function with my nontheistic friends, there is no indication, so far as I can see, that I (as a theist) have a concept of proper function they don't have. Rather, it seems to me that there is a single concept, one we all (or nearly all) in fact employ. With respect to this single concept, there is a question as to whether or not it does imply conscious and intelligent design. (And of course if it does, then Swinburne's knowledge that today is Monday *does* entail that (some of) his cognitive faculties are functioning in the way their creator intended.)

What is Swinburne's evidence for the claim that there are these two concepts of proper function? And is the claim plausible? It seems to me a little like saying that there are two concepts associated with terms like 'belief'; materialists, like Daniel Dennett, e.g., use the term one way, and dualists, like some of the rest of us, use it the other. But this duality of concepts seems implausible in both cases. There is just one concept, so it seems to me, where Swinburne sees two. It is of course possible that that one concept of proper function doesn't fit well with naturalism, just as our ordinary concept of belief doesn't fit well with materialism. If materialism were true, it would be hard to see how we human beings could be capable of having beliefs and intentions, of thinking about, intending things of any kind at all, and of having beliefs that played a causal role in our behavior. In the face of this difficulty, the materialist philosopher has several options. One is to say that as a matter of fact there just aren't any such things as beliefs and intentions; the categories of a future and adequate science of mind will have no room for those obsolescent categories. Another is to argue that appearances are deceiving: it *looks* as if, if materialism were true, there could be no such thing as belief or intention, but as a matter of fact the two are perfectly consistent. And one way to argue the latter would be to give an analysis, or at any rate

necessary and sufficient conditions, of belief and intention in naturalistic terms.

Precisely the same options, so it seems to me, are open to the naturalist with respect to the notion of proper function. As I argue in WPF, it seems initially that the notion of proper function involves the notion of conscious and intelligent design: something is functioning properly when it is functioning in the way in which it was designed to function. (If this entailment does indeed hold, perhaps it is because for many centuries our forebears thought of themselves and other creatures as designed and created by God.) Presented with this appearance, the naturalist has several options. One is to argue that since we have not in fact been designed, the notion of proper function cannot in fact apply to our cognitive faculties, or to any other of our organs, systems and powers. Indeed one who follows this path will have to argue that we don't, strictly speaking, have faculties at all; for faculties are among the sorts of things that can function well or ill. One who thinks in this way can follow Kant, Vaihinger²⁷ and Dennett in thinking of the notion of purpose as a kind of convenient fiction, a matter of speaking with the vulgar (while of course thinking with the learned).

A second possibility, however, is to hold that appearances are deceiving: as a matter of fact the notion of proper function does *not* involve or entail that of conscious and intelligent design. One way to show that this entailment does not hold would be to propose a satisfactory analysis of proper function in terms that don't involve conscious and intelligent design. Several such analyses have been proposed, but none of them, as I argued (see chapter 11 of WPF), is satisfactory. This lends some credence to the idea, which is in fact my own view, that the one notion of proper function we all employ—theist and nontheist alike—does indeed entail the notion of conscious and intelligent design. If so, however, then the claim that Swinburne knows that today is Monday does indeed entail something like the proposition that Swinburne's cognitive faculties (or at least some of them) are functioning in the way intended by their designer(s).

V Taylor: Dilemmatic Troubles?

I turn finally to James Taylor's intriguing contribution. This is a further installment in a discussion that began with Taylor's "Plantinga's Proper Functioning Analysis of Epistemic Warrant"²⁸ which prompted my "Warrant and Designing Agents: A Reply to James Taylor."²⁹ We now have before us Taylor's riposte, the next installment of the discussion, to which I will reply. Now my account of warrant, of course, crucially involves the notion of

²⁷ For Kant and Vaihinger on this, see WPF, pp. 211–13.

²⁸ *Philosophical Studies* 64 (1991), pp. 185 ff.

²⁹ *Philosophical Studies* 64 (1991), pp. 203 ff.

proper function: a belief has warrant for someone only if the cognitive faculties involved in the production of that belief are functioning properly. In his first paper, Taylor asked whether the notion of proper function is analyzable in terms of the notion of an “actually designing agent or process” such as, for example, God or evolution. The question, I suppose, was whether any correct analysis of proper function will involve the notion of an actually designing agent or process: if so, then, necessarily, a thing functions properly (is working right) only if it functions the way it was designed to function by the being or beings who designed it.

Taylor next pointed out that either the concept of proper function is analyzable in terms of an actually designing agent or it isn't. So far so good. But either way, he said, there is trouble for my account:

- (1) If the concept of proper function is analyzed in terms of an actually designing process or agent [if proper function entails actual design], then Plantinga's account is subject to at least one of two types of counterexamples which show that proper function is not necessary for warrant (p. 185),

and

- (2) ...if the concept of proper function is *not* analyzed in terms of an actually designing process or agent [if proper function does *not* entail actual design], then Plantinga's account either offers no clear advantage over reliabilism, or it is not clearly distinguishable from reliabilism (p. 185).

My response in “Warrant and Designing Agents” was as follows. Consider

- (3) A belief is warranted for a person *S* if and only if it is produced in *S* by cognitive faculties functioning properly according to a design plan successfully aimed at truth; and propositions of the form *X functions properly* entail the corresponding propositions of the form *X was designed by an actually designing process or agent*

and

- (4) A belief is warranted for a person *S* if and only if it is produced in *S* by cognitive faculties functioning properly according to a design plan successfully aimed at truth; and propositions of the form *X functions properly* do *not* entail the corresponding propositions of the form *X was designed by an actually designing process or agent*.

The first conjunct of each of (3) and (4) is, of course, my account of warrant; I am therefore understandably enthusiastic about it. Furthermore, either the second conjunct of (3) is true, or else the second conjunct of (4) is true. Therefore I am committed to thinking that either (3) or (4) is true. And what Taylor proposed is that each of (3) and (4) leads to trouble.

My response, essentially, was to concede that each of (3) and (4) *does* lead to a kind of trouble, all right, but that the trouble is to be assigned to the second conjunct of each. Each of (3) and (4) is doubtful or problematic, and *being doubtful or problematic* is one kind of trouble for a thesis to which one is committed. Either (3) or (4) is true, on my view, but there seem fairly strong arguments against each; that is the source of the trouble (such as it is) with (3) and (4). But the problem with (3) and (4), so I said, is a problem with the second conjunct of each. It is fairly plausible to think that the notion of proper function *does* entail design by an actually designing process or agent (and hence doubtful or problematic that it *doesn't*), and also fairly plausible to think that it *doesn't* (and hence doubtful or problematic that it *does*.) That's what leads to the problematic nature of both (3) and (4). Agreed: I am committed to the view that one or the other is true; agreed: each is problematic; but it doesn't follow that my analysis of knowledge is problematic.³⁰

All of that seems correct; but Taylor points out in *this* piece that the trouble I conceded with (3) and (4) isn't the trouble he had in mind; and the trouble *he* had in mind, he says, *does* make a problem for my analysis of knowledge:

I make the unimpeachable claim that proper function is either analyzable in terms of actual design or it is not. I then argue that each of these disjuncts entails a proposition expressing a difficulty for Plantinga's view of warrant. I grant it is not clear which disjunct is true. However, this is neither here nor there; one of them must be true, Plantinga's account must be understood in light of one of them, and both have consequences unattractive for his position. So his view is problematic even if we don't have a clue how to understand proper functioning (pp. 421–22).

Point taken: Taylor is arguing not just that each of (3) and (4) leads to trouble of *some* kind, but something more specific:

- (5) If necessarily (a thing functions properly or improperly only if it has been designed by a personal agent, and functions properly only if it functions in accord with the intentions of that personal

³⁰ I might have added that in being committed to the truth of either (3) or (4), I am not thereby committed to something problematic. I am committed to their disjunction, but that isn't problematic; what is problematic are the disjuncts. *Being problematic* is like contingency: each of *p* and *q* may be contingent; but it doesn't follow that *p or q* is; and *being committed to* is like necessity: *p or q* may be necessary, but it doesn't follow that either *p* or *q* is.

agent³¹), then there are counterexamples to Plantinga's account of knowledge

and

- (6) If the antecedent of (5) is false, then "Plantinga's account either offers no clear advantage over reliabilism, or it is not clearly distinguishable from reliabilism" (p. 185).

As Taylor points out, to undermine this argument, I must attack (5) or (6).

A. *The First Horn.*

Well then, let's begin with (5). Suppose its antecedent is true: necessarily, a thing functions properly or improperly only if it has been designed by a personal agent, and it functions properly only if it functions in accord with the intentions of that agent. Then, says Taylor, it follows that there are two kinds of counterexamples to my account. The first kind goes as follows:

Theodore is not designed but is an unintended consequence of a clumsy angel's intentional action to create something incapable of warranted belief. Theodore has cognitive faculties that function like those of a normal human. Intuitively, Theodore has warranted beliefs. So warrant does not require actual design (Xref/Taylor p. 2).

But here I believe things have gone awry. First, Taylor says that Theodore has cognitive faculties. But a faculty—something like vision or memory—by its nature functions properly or improperly; as opposed, say, to a tumor or wart, a faculty has a purpose or function, and can work in such a way as to accomplish that purpose (fulfill that function) or fail to do so. But if the antecedent of (5) is true (if proper function entails design by a personal agent), then it isn't possible that Theodore have properly functioning cognitive faculties but not be designed by a personal agent: the case is impossible.³² So again if the antecedent of (5) is true, it follows that the case is impossible. Hence the antecedent of (5) entails that the case *is* possible only if the antecedent of (5) is *itself* impossible. But Taylor doesn't give us any reason to think that it *is* impossible. The first kind of counterexample, therefore, gives us no reason to accept (5).

³¹ I move from 'designed by an actually designing process or agent' to 'designed by a personal agent and functions in accord with the intentions of that personal agent', because in the present paper Taylor apparently means to hold that design entails *personal* design, and *intentional* design: "Arguably, something is designed only if it is intended (I recant my previous openness to impersonal design)" (p. 422).

³² Furthermore, the idea, I take it, is that Theodore is a *person*. It is a necessary truth, however, that a person is (among other things) a being with cognitive faculties; and it is a necessary truth that a cognitive faculty functions properly or improperly. But according to the antecedent of (5), such a being has been designed by an intending and intelligent agent, which, again, is incompatible with the description of the case.

But perhaps this is an uncharitable construction of Taylor's example.³³ Perhaps his thought really goes as follows: "Faculties, schmaceuties; what I mean is that there just pops into existence (unplanned, unintended, undesigned) a being that *functions* just like we do. It has beliefs and experience, and under the circumstances in which we form a given belief, it forms the same beliefs. For example, when it is appeared to in that characteristic way that accompanies speaking with another person, say, it forms the belief that it is speaking with another person, just as we do. We might say that it has the same *max plan* (WPF pp. 22–24) as we do, but doesn't have a *design plan* at all. The question whether it has faculties or not is really irrelevant; what counts is that it behaves just like we do. And if it does, then clearly its beliefs have warrant. But then, given the assumption that proper function requires conscious, intelligent design, it follows that warrant does not require proper function."

So the claim is that if (or perhaps even if) proper function entails conscious and intelligent design, then it would be (is) possible that a being with a max plan just like ours (a being capable of belief, of moral agency, and the like) should just pop into existence, uncreated, unintended, and undesigned. But this is far from obviously so. First, Taylor's scenario entails that it is possible to have a max plan just like ours, but not have a design plan like ours, and in fact not have a design plan at all. But this is far from clearly possible. It is also not clearly *impossible*; but that's not good enough for present purposes. A good counterexample has to be more than not clearly impossible: it must be clearly possible. But further, I am myself strongly inclined to think that if proper function *does* require conscious and intelligent design, then it *isn't* possible that a being capable of belief and moral agency should just pop into existence purely as a result of a clumsy angel's fumbling and incompetent activity. Here is a less than conclusive argument. Suppose proper function requires conscious and intelligent design. Various parts, organs and systems of human beings clearly function properly: human beings, therefore, are consciously designed. The only plausible candidate for the post of being the designer of human beings would be God; human beings, therefore, are designed by God. But if there is such a person as God, it is unlikely that it is possible that a being capable of belief and moral agency should just pop into existence, unintended and undesigned by God. According to the Christian tradition, only God can create beings capable of belief and moral agency; I am inclined to think this is right. But even if it isn't, even if it is possible that God should delegate the task of creating such beings to some of his creatures, it still wouldn't be possible that such a creature pop into existence unintended by God.

³³ As I was compelled to see by the people mentioned in footnote 1.

Thus my response to (5) is that it is false or dubious. I am inclined to think it false, but am satisfied with dubiety; for, as I say, a good counterexample must be clearly possible. The most that Taylor should allege against my account is that there is a state of affairs *S* such that if proper function entails conscious and intelligent design, then my account implies that *S* is impossible, when it is not independently clear that *S* is impossible. But I don't see that this is much by way of an objection to an analysis or account. Indeed, it can be a positive advantage. If you are convinced that the notion of proper function *does* entail conscious and intelligent design, and furthermore you accept my account of warrant, then you have a way of coming to see something you wouldn't have seen otherwise.

Similar remarks apply to Taylor's second alleged counterexample; in the interests of brevity I shall not spell them out, but instead conclude that Taylor has given us no reason to think that (5) is true. Perhaps we also have no reason (independent of my account itself) for supposing (5) false; but that isn't much by way of criticism of my account.

B. The Second Horn

The other horn of Taylor's dilemma is the claim ((6) above) that if proper function does *not* entail a personal designing agent, then "Plantinga's account either offers no clear advantage over reliabilism or it is not clearly distinguishable from reliabilism" (p. 421) (and in the latter case also, of course, offers no clear advantage over reliabilism). In a way I am less interested in this horn, because I am inclined to believe that the central core notion of proper function (as opposed to various analogical extensions) *does* entail intentional design; difficulties that might arise if it *didn't*, therefore, are not so pressing. Nevertheless I'd like to take a brief look at some of the interesting issues arising with respect to (6).

So why does Taylor think that if

- (7) Proper function entails design by an intelligent agent (and also entails that a thing functions properly only if it functions in accord with the designer's intentions)

is false, then my account offers no advantage over reliabilism? His answer seems to be that "The actual design condition provides the only plausible grounds for the claim that a proper functioning view of warrant is superior to reliabilism" (p. 425). What this means, I take it, is that (7) constitutes the only grounds there are or could be for claiming that my account is superior to reliabilism; my account is superior to reliabilism only if (7) is true. Why think that? I claim that reliabilism, in all its protean manifestations, is open to clear counterexample with respect to sufficiency; how is it that if (7) is false, then my account fares no better? Taylor's answer, apparently, is that if

(7) is false, then my account is open to the very same counterexamples on which reliabilism founders. His answer begins with the case of the epistemically serendipitous tumor, a case I used against Dretske's and Goldman's versions of reliabilism: "Suppose Simon acquires a brain tumor which causes him to believe he has a brain tumor (causes him to have T-beliefs). However, though these beliefs are reliably produced, intuitively, they are not instances of knowledge" (p. 424). So far so good. But then Taylor suggests, alarmingly enough, that "...once actual design is no longer required for proper functioning, Plantinga's view of warrant arguably succumbs to the same counterexample" (p. 424).

Why so? How does the argument go? First, Taylor notes that in my previous reply to him (and elsewhere) I said that sometimes the way a thing works can be *adopted* as its design plan. Your radio falls off the roof and now unaccountably gets stations much further away than before (and perhaps even corrects mistakes in the performances of Beethoven sonatas); you might *adopt* its new way of working as its design plan, so that if the radio stops getting those stations from eastern Europe you complain that it isn't working right any longer, and try to fix it by dropping it off the roof again. I try to make a refrigerator; I am miserably inept and my project is a flop; the thing acts more like a warming oven. Making a virtue of necessity, I then decide to use it for a warming oven, claiming that it works right when it serves to keep things warm. I have adopted its way of working, which is not at all what I originally intended, as its design plan.³⁴ (It is important to note that when I complain about the radio's not working properly, and when I claim that my failed refrigerator is working properly in keeping the biscuits warm, I am in each case making an analogical extension from the central core of the concept of proper function, and am using the terms involved in a way that involves an analogical extension of their core meaning.)

But if it is thus possible for us to adopt a thing's way of working as its design plan, no matter how it acquires that way of working, then, says Taylor, why can't the same thing go for Simon? Why can't we adopt *his* new way of functioning (believing that he has a brain tumor, the belief being caused by the tumor itself in an accidental way) as his design plan, claim consequently that the cognitive faculties involved in the production of those T-beliefs are in fact functioning properly, and add that his T-beliefs have warrant? If we can, he says, there are problems for my account:

³⁴ This case is to be contrasted with the case where I make a refrigerator which works at least tolerably well, but later decide that I need a warming oven and redesign it. In this sort of case it is infelicitous to speak of my 'adopting' this new design plan; the latter happens only when a thing works or begins to work in a way I didn't intend and perhaps don't initially understand.

Suppose we do. Then, according to Plantinga, Simon's T-beliefs are a result of proper functioning. However, intuitively, they are not instances of knowledge. The only way to avoid this is to place objective constraints on the ways of working that can appropriately be adopted by us as design plans. However there seem none which would preserve the distinctiveness of proper function besides actual design (p. 424).

Now here the thread of argument is a bit elusive, but perhaps it is to go as follows. Suppose (7) is false, i.e., suppose proper function does not entail design by a personal agent. Then it would be possible that Simon acquire a new design plan, one that involved his acquiring T-beliefs by virtue of their being caused by the tumor. This would be possible, because it would be possible for us to *adopt*, as his design plan, this new way of working, including his forming T-beliefs in that way. But then it would be possible that his circumstances be as described and his T-beliefs have warrant. That isn't possible, however; therefore if (7) is false, my account is open to the Epistemically Serendipitous Tumor counterexample.

The questions Taylor raises are interesting, and worth pursuing in depth. For example: can we really adopt just any way of working we please as the design plan of a thing? I suffer an injury: my knee joint is now frozen, so that it has no mobility at all. Could I 'adopt' this way of working, correctly claiming that my knee now works properly, after all those years of loose and floppy malfunction? Not obviously. I'd be just wrong, so it seems to me, in claiming that my knee worked properly. Could it be, as Taylor suggests, that this is due to an intentional design feature of our concept of proper function; i.e., could it be that (7) is true, and the reason I can't adopt just anything as a design plan is that, really, the design plan of a thing is that given it by its designers (or perhaps *most recent* designer(s))? Perhaps so, but it isn't at all clearly so. It seems that even if (7) were false, it might still be that I couldn't adopt just any old way of working as a design plan. Of course things become difficult here: since I am inclined to think (7) true, and since (7) is necessary if true, it isn't easy to think about what things would be like if (7) were false. (On the standard views, if (7) were false, everything would be false, and true as well. Even if *that's* false, there will still be monumental difficulties in thinking about what things would be like if some proposition which is in fact necessary were false.)

These are fine questions; my reply, however, is already too long (and anyway I don't know the answer to the questions); I shall therefore limit myself to the following reply. First, of course, I needn't argue that (6) is mistaken in order to defend my account of warrant against Taylor's dilemma; it is quite sufficient if (5) is dubious or mistaken. In any event, second, I am not at all sure that (6) is mistaken; I am inclined to think (7) true, and have no opinion or at least no firm opinion on how things would stand if it were false. But third and perhaps most important, consider those cases where we do

adopt a way of working as a design plan—the case of the radio that falls off the roof, for example. In such a case, so it seems to me, when we speak of the radio as no longer working properly when it reverts to its old way of working, we are using the relevant terms in an analogically extended sense—unless, that is, we think that God or some other unknown person revised its design plan. In these cases both the term ‘proper function’ (and its relatives) and the concept *proper function* get extended by analogy (a familiar process or phenomenon). But my account of warrant is of the basic core concept, which involves the basic core concept of proper function. Suppose, therefore, that in some analogically extended concept of *proper function*, Simon’s T-beliefs are produced by cognitive faculties functioning properly. Then perhaps there will be an analogically extended concept of warrant to match this analogically extended concept of proper function, and perhaps in that analogically extended concept of warrant, Simon’s T-beliefs would have warrant. But that is no problem for my account: given the central core concept of warrant, the concept of which my account is an account, Simon’s beliefs don’t have warrant. No doubt my account permits the existence of an analogically extended notion of warrant in which Simon’s T-beliefs *do* have warrant: that fails to show any problem with my account of the central core notion.

There is much more in what Taylor says that deserves comment—as, indeed, in what the other critics say. But my reply is already too long; I shall have to stop here, adding only that I am grateful to my critics for their searching and illuminating comments.