16

Respondeo

Alvin Plantinga

My thanks to Jon Kvanvig and the other authors for these penetrating and illuminating essays; I have learned much from them. Replying to them is a privilege—a *scary* privilege. It is also a bit chastening. According to the Heidelberg Catechism, to live and die properly one must know the extent of one's sins and miseries: I find these essays very useful along those lines. I only wish I had read them *before* I wrote the Warrant volumes. But mostly I am delighted with new vistas and new ideas. Reading these essays has convinced me once again of the real importance of epistemology, and of its enormous *interest*. Those who proclaim the death of epistemology ought to read these essays—if this is what death is like, then Socrates was right: we (or epistemology) should be eager to die. It has also convinced me (once more) of the extreme difficulty of epistemology; with further insight, not only the answers but the very shape of the questions keep changing.

From a respondent's point of view the essays present an embarrassment of riches: it's hard to know where to start, which topics to address, how far to go with any topic. By way of compromise, my reply will have two parts. In the first I'll take the topic on which these essays have taught me the most, and try to say something of what I have learned. In the second I'll make specific replies to individual authors. I am not first of all interested in doggedly replying to every objection, even where I think I have a satisfactory reply; and (given the cornucopia of objections, ideas, suggestions) I won't be able to comment on nearly everything of importance. I've replied elsewhere

to some of the objections raised here; where it still seems to me that my answer is apt, I'll simply refer the reader to that answer. I have also been obliged to shape my reply in such a way as to emphasize what I have to say that may be of interest: in many cases I unduly neglect contributions simply because I can't say much more than I already have. There is of course no proportion between the length of my reply to an essay and my judgment of its merit.

I: Rethinking Gettier

One of the really seminal developments in twentieth-century epistemology (its second half, anyway) was Edmund Gettier's three-page paper[1] presenting a couple of counterexamples to the Justified True Belief theory of knowledge. Of course the JTB account was so-called only *after* Gettier, and in fact it isn't really clear that it *was* the received view prior to Gettier.[2] But if the JTB theory had not existed, it would have been necessary to invent it, if only to provide occasion for Gettier and responses to him. I'm grateful to Peter Klein and Marshall Swain for calling to my attention a response to Gettier that I unconscionably neglected in the Warrant volumes: defeasibility theories. And I am grateful to Klein and Richard Feldman for making it crystal clear to me that my own treatment of what I called "quasi-Gettier" problems is defective. Others have expressed skepticism about my treatment, but (like the farmer who had to hit the bull over the head with a two by four to get his attention) Klein finally managed to get me to pay proper attention. I'll say something about defeasibility theories below (concluding that Klein's version, sophisticated and refined as it is, isn't successful as an account of warrant); but I'll begin by looking into Klein's complaints about my account of quasi-Gettier problems.

But first, what *are* Gettier problems? Suppose we quickly review the usual list.[3] First, there is Gettier's original Smith-has-a-Ford-or-Brown-is-in-Barcelona example; here you come to hold a true, justified but unwarranted belief (one that doesn't constitute knowledge) by inferring it from a false but justified belief. Second, there is the kind of example represented by Carl Ginet's[4] proud but impecunious Wisconsinites who erect all those barn facades to make themselves look more prosperous than they are; traveling through the area I see what is in fact a barn, forming the justified, true, but unwarranted belief *Now* that's *a fine barn*. Here there is deception on the part of those

Wisconsinites. Third, there is Bertrand Russell's pre-Gettier Gettier example of the clock that stops at midnight: you see it at noon and form the belief that it is noon; your belief is justified and true, but doesn't constitute knowledge. In this case there is neither deception nor inference from a false belief. In none of these three kinds is there failure of proper function on the part of the believer; but fourth, there are also examples where a cognitive glitch is involved. In Alexius Meinong's example as reported by Roderick Chisholm, an aging forest ranger suffers hearing loss and can no longer hear his wind chimes; he also develops tinnitus, sometimes hearing chime-like sounds due to no external stimulus; this sometimes occurs just when the wind chimes are in fact sounding, thus producing in him the justified, true, but unwarranted belief that the wind is up. Still another kind of example, one involving no malfunction, deception, or inference from a false belief: you are a desert tyro and seem to see an oasis about three-quarters of a mile away, forming the belief that there is an oasis near by. As it happens, you are the victim of a mirage, but in fact there is an oasis near by; your belief is justified and true, but does not constitute knowledge.

One feature these examples have in common, of course, is that they are cases of justified true belief that do not constitute cases of knowledge; hence they refute JTB theories of knowledge. More broadly, Gettier cases also show (and it is clear from some of the essays in this volume that internalists would not disagree) that no strictly internalist account of warrant can be successful. But I think their significance extends further. It is not uncommon to dismiss concern with Gettier problems as so much scholastic niggling; and perhaps the long line of circumventions and countercircumventions, the complicated "fourth conditions" together with their more complicated refutations and the counterrefutations, the cycles and epicycles—perhaps these do give that impression. But I believe we can still learn something interesting about the structure of knowledge by thinking about Gettier problems. In particular, we can learn something by thinking further about another and oft-noted feature of these Gettier situations: in these Gettier cases, *S*'s justified true belief does not constitute knowledge because, given the circumstances, it is only accidentally, or by serendipity, or by just dumb luck that *S* forms a true belief in the situation.

Mea Culpa

Now my own response to Gettier is off the mark, as Klein and Feldman point out. The chief problem is that (given my understanding

of the environmental condition) what I said handles only *some* Gettier problems, leaving others untouched. Let me begin by recalling what I said.

By way of general comment, I noted that Gettier examples involve something like mild cognitive environmental pollution; in each of these cases the cognitive environment, I said, diverges in some small or subtle way from the paradigm or standard sort of environment for which our faculties are designed.[5] Clocks seldom stop precisely twelve hours before you look at them, people don't ordinarily take the trouble to erect phony barns, it's most unusual to develop just this kind of tinnitus after a lifetime of monitoring wind speed by listening to wind chimes, and so on. This still seems to me correct. Gettier examples involve relatively minor failures of fit between cognitive capacities and cognitive environment; as a consequence it is only by happy cognitive accident that a true belief is formed in the circumstance in question. Of course if just *any* example that shows justification and truth insufficient for knowledge is a Gettier example, then there will be Gettier examples that involve not just a bit of retail lack of fit between cognitive capacities and environment, but wholesale lack of fit. A madman, for example, or the victim of a Cartesian evil genius, or someone who has been envatted by Alpha Centaurian cognitive scientists might well be completely justified (in any of the ordinary ways of understanding justification), and nevertheless display not minor but monumental failure of fit between cognitive capacities and cognitive environment. These examples are indeed counterexamples to a JTB theory, but they seem out of the spirit of Gettier examples—perhaps because (as it is plausible to think) JTB theories implicitly presuppose something like a proper function condition.[6]

Well, this seems right, but how is it supposed to handle Gettier problems? I went on to call attention to misleading cognitive responses of a certain sort: Müller-Lyer examples, straight sticks that look bent in water, mirages, dry North Dakota roads that look wet on a hot summer's day, airplanes high in the sky that look small, and so on. In these cases there is no cognitive malfunction (no failure to conform to the design plan); still the cognitive responses are misleading. (I can remember the astonishment with which I learned, as a child, that airplanes are actually large enough to contain people.) What is the explanation of these responses? Why are they in our cognitive repertoire? If the perceptual system is designed to produce true beliefs, why are these misleading responses part of it? The answer, I said, lay in

the area of tradeoffs and compromises. We can imagine that the designer (God or evolution) aims to have a creature of our general style: made of flesh and blood and bone rather than plastic and steel, of roughly our size, in a world with our regularities, with a cognitive system mediated by brain and nervous system. Now of course the designer also aims at producing a cognitive system that furnishes us with true beliefs, and the more accurate the better. But perhaps a maximally excellent cognitive system of this sort would require too large a brain, so large we would have to hold up our heads with both hands, thus being unable to play the piano or go rock climbing (or fight off predators). So tradeoffs and compromises are necessary: trade a bit of cognitive excellence for reasonable brain size and mobility, for example.

Müller-Lyer *et al.*, therefore, are present, not because they serve the general function of providing true beliefs (they don't), but because they are a locus of a best or satisfactory compromise between these competing desiderata. Of course in a way they *do* serve that general function, by way of being part of a best compromise. But this, I said, is to serve it *indirectly* rather than *directly*: "the thing to say is that *R* [the cognitive response member of the relevant triple of the design plan] is joined with *M* [the circumstance member] not in order to *directly* serve the main purpose of providing true beliefs (it doesn't do that) but to do so *indirectly* . . ." (WPF p. 40). But then consider any of these misleading responses (*there is an oasis near by* or *it must have rained on the road recently*): a belief produced in this way can happen, by virtue of cognitive good luck, to be true. If so, the belief is *accidentally* true, and we have a Gettier example. By way of stating a general clause or condition to deal with Gettier problems, I put it like this:

> Take a perceptual illusion or false testimony case and add that the belief produced is true (but by accident): then what you have is a quasi-Gettier case. The belief in question has little warrant and, though true, does not constitute knowledge; *for a belief has warrant for you only when it is produced by a segment of the design plan directly aimed at truth.* (WPF p. 40, italics added)

Now Klein has more than one objection. He suggests that my view is committed to the possibility that a false belief *f* should have warrant sufficient for knowledge; if so (as in Gettier's original example) *S* could disjoin a true belief with *f*, thus deducing from *f* a true belief. This

belief should have as much warrant as *f*, so that (according to my view) the belief should constitute knowledge: but of course in fact it wouldn't.[7] This objection isn't conclusive. There is more than one reason, but chief among them is just the fact that my account isn't committed to the possibility that a false belief should have warrant sufficient for knowledge. (I'll argue below, p. 329, that a false belief *can't* have that degree of warrant.)

He has another objection, however, that reveals a defect in the above clause.[8] Here is his example. I own a Chevrolet van, drive to school one afternoon, park the van, and go to my office. As it turns out, my van is demolished by a runaway gravel truck or (Shope's example) a wayward meteorite. By an astonishing stroke of good fortune, however, I had entered and just won the win-a-van (a Chevrolet) contest sponsored by the local Varsity Club, although I haven't yet heard the good news. You ask me what sort of automobile I own, and I reply cheerfully (and truthfully), "a Chevy van."

This is a classic Gettier case: my belief is true and justified but doesn't constitute knowledge (it is true 'just by accident'; had I not won the win-a-van contest it would have been false). But it also eludes my proposed condition for fending off quasi-Gettier cases. For consider a situation just like the one in the example up to the time where my van is destroyed—perhaps the gravel truck takes an alternate route or the meteorite is deflected by a collision with another and disintegrates before it hits the ground. Had this happened, my van would have suffered no damage, my belief that I own a Chevrolet van would have been formed in just the way it is in the example, and the belief *would* have constituted knowledge. The belief is produced in just the same way (by the same processes governed by the same parts of the design plan) in the two situations; but then the fact that the belief constitutes knowledge in the one situation but not the other cannot be attributed to the belief-producing processes being aimed "directly" at the production of true belief in the one case but "indirectly" in the other. The same bit of the cognitive system, governed by the same bit of the design plan produces the same belief in the two situations; if the process in question is governed by a bit of the design plan aimed directly at the truth in the one case, the same goes in the other. Hence the proposed clause designed to cover Gettier cases doesn't distinguish the two cases.

We can see the same thing even more simply. Consider the Russell-Gettier example: you form the true belief that it is now noon by looking at a clock that happened to stop last night at midnight. The processes

that produce this belief in you are the very ones that would have produced the same belief had the clock not stopped at midnight but continued to keep proper time. If in the one case the processes involved are only indirectly aimed at producing true belief, the same is true in the other case. But in the first case you lack knowledge and in the second you have it. So the proposed clause doesn't do the job.

The Resolution Problem

Consider the Russell-Gettier case, the van case, and the other cases my clause doesn't properly deal with, and consider the cognitive processes that produce the relevant beliefs in those cases: these processes don't do well in those cases. The process of telling time by glancing at the clock won't ordinarily lead to false belief: but in situations where the clock has stopped, it ordinarily *will* lead to false beliefs. (In *those* situations, if the belief produced is true, it will be only by accident, by pure dumb luck.) If your van has been destroyed by a meteorite but you haven't yet heard the bad news, you will ordinarily hold a false belief; your belief will be true only if something very unusual happens—you win the win-a-van lottery, or your grandmother unexpectedly dies and leaves you her Chevy van.

In WPF I spoke of our *cognitive environment*, and I added that a belief has warrant only if it is formed in an *appropriate* cognitive environment. For the most part I was thinking of our cognitive environment as the one we enjoy right here on earth,[9] the one for which we were designed by God or evolution. This environment would include such features as the presence and properties of light and air, the presence of visible objects, of other objects detectable by our kind of cognitive system, of some objects not so detectable, of the regularities of nature, the existence of other people, and so on. Call this our "maxi-environment"; in stating the environmental condition, what I (mostly) had in mind was a maxi-environment. Our cognitive faculties are designed to function in *this* maxi-environment, the one in which we find ourselves, or one like it. They are not designed for a maxi-environment in which, e.g., the only food available contains a substance that seriously inhibits memory, or where there is constant darkness, or where there aren't any distinguishable objects, or where there is little or no regularity, at least of a kind we can detect, or where everything is in a state of constant random flux.

But there is also a much less global cognitive environment. For any belief B and (more relevantly) the exercise E of my cognitive powers

issuing in *B*, there is also a much more specific and detailed state of affairs we might call its "cognitive mini-environment."[10] For example, there is the cognitive mini-environment of the van case, the Gettier-Russell mini-environment in which I happen to look at a clock that has stopped, and the fake barn environment. On the other hand, there are also the cognitive mini-environments in which the clock I glance at is keeping proper time, the ones in which my van remains unmolested in the place I parked it, and the ones where there are only real barns. We can think of a cognitive mini-environment of a given exercise of cognitive powers *E* as a *state of affairs* (or proposition)—one that includes all the relevant epistemic circumstances obtaining when that belief is formed. Consider any current belief *B* I hold and the exercise *E* of cognitive powers producing it: the mini-environment *M* for *E* (call it "MBE") includes the state of affairs specified by my cognitive maxi-environment, but also much more specific features of my epistemic situation. It will include, for example, the presence or absence of fake barns, of my van's being destroyed in unforeseen ways (if it is), of Paul's brother Peter being in the neighborhood, and any other relevant epistemic circumstance. To be on the safe side, let *MBE* be as full as you please, as large a fragment of the actual world as you like. (No doubt elegance would counsel cutting *MBE* back to epistemically relevant features; that, however, would set us the laborious task of giving an account of relevance, thus perversely sacrificing ease to elegance.)

In any event, we must note that a given exercise of (properly functioning) cognitive faculties can be counted on to produce a true belief in *some* mini-environments but not in others. In the ones where my van remains unharmed where I parked it, the processes that ordinarily produce in me the belief that I now own a van will produce a true belief; not so if the van is destroyed in some quirky fashion I could not have anticipated. You form an opinion as to what time it is by glancing at a clock; this exercise of cognitive powers can be counted on to produce true belief in the usual mini-environments (in which the clock is keeping reasonable time); not so for ones in which it has stopped. I form a belief as to the identity of the person standing in Paul's doorway by taking a quick look from across the street; this exercise of cognitive faculties can be counted on to produce a true belief in a mini-environment where Paul is the only person in the neighborhood that looks at all like Paul; not so for one in which Paul's look-alike twin brother Peter is (unbeknownst to me) staying in Paul's house.

Note the relativity to the specific belief-producing exercise of cognitive powers. I take a quick look from across the road: then I am as likely as not to form a false belief about the identity of that person of Pauline appearance standing in the doorway. But if I go right up to him and look closely I could tell that it wasn't Paul. I form a false belief (that my van remained undamaged right where I parked it) in the van example, but I wouldn't have, had I called on other cognitive processes; if I had hired a detective to watch my van, I would have received the bad news. I glance at a stopped clock and form a false belief as to the time; not so if I watch it for a period of ten minutes and see that the minute hand doesn't move. We are therefore thinking here of a specific exercise of properly functioning cognitive faculties: such an exercise can be counted on to produce a true belief with respect to some cognitive mini-environments but not with respect to others.

A possible hitch: *MBE*, of course, is a state of affairs including the circumstances in which a belief *B* is formed by *E*. But does the fact, if it is a fact, that *E produces a true belief* get included in *MBE*? Does the mini-environment of the van case include the proposition that (owing to my luck with the lottery) I *do* in fact own a van and also believe that I do, so that *MBE* includes my forming a true belief with respect to whether I own a van? If so, then of course with respect to *that* situation the exercise of cognitive powers in question can be counted on to produce a true belief. But we want to *ask* whether a certain exercise of cognitive powers is or isn't to be counted on to provide true belief in its cognitive mini-environment; *MBE* must therefore be specified in such a way that it doesn't include *E*'s producing a true belief and also doesn't include *E*'s producing a false belief. The proposition that *S* forms a true belief will be neither true nor false[11] in *MBE*. But of course an actual situation in which someone forms a belief will be maximally specific, and *will* include that *S* forms a true belief, if in fact *S* does so. So let's say that *MBE is* maximally specific except for the truth or falsehood of the proposition that *S* forms a true belief by way of *E*. A cognitive mini-environment will be a state of affairs *diminished with respect to* that proposition—a state of affairs as much as possible like the actual maximally specific situation, given that *MBE* includes neither the proposition that *S* forms a true belief nor its denial.

An exercise of my cognitive powers, therefore, even when those powers are functioning properly (perfectly in accord with my design plan) in the maxi-environment for which they are designed, can be counted on to produce a true belief with respect to *some* cognitive

mini-environments but not with respect to others. Some mini-environments are *favorable* for a given exercise of cognitive powers; others are *misleading*, even when my faculties are functioning properly. These mini-environments, we might say, are such that my faculties are not designed to produce a true belief in or with respect to them—even though they include the maxi-environment for which my faculties have indeed been designed (by God or evolution). The cognitive processes involved in the van case and Russell-Gettier examples are unreliable in the van and Russell-Gettier mini-environments; in *those* circumstances the processes in question, even when functioning properly, are more likely to lead to false beliefs than to true.[12] These mini-environments are misleading with respect to those cognitive exercises; we might also say that in these misleading mini-environments our faculties (more exactly, specific exercises of them) display a certain deplorable *lack of resolution*. Even when my cognitive powers are subject to no dysfunction, I am unable (by a quick glance) to distinguish Peter from Paul, or the case where the clock is keeping good time and shows 12:00 noon from the case where it stopped at midnight. In phony barn country my cognitive faculties don't (by way of a casual look from the road) provide the resolution needed to enable me to distinguish the genuine article from the *papier mâché* mockups. And when the cognitive mini-environment is misleading, when it is one within which the relevant exercise of cognitive powers suffers from this lack of resolution, then if I form a *true* belief, it will be just by accident, just by way of dumb luck. It will not be because the situation is one for which my faculties (or this kind of exercise of them) are designed to produce true belief.

This is the resolution problem. It isn't, of course, a problem for a *knower*; it's a problem for an epistemologist. It's a problem for my account, at least if we neglect cognitive mini-environments in stating the environmental condition. It isn't enough for warrant that the *maxi*-environment be one for which the faculties in question are designed to produce true belief; that is the lesson of the Notre Dame van case and others like it—indeed, of Gettier cases more generally.

The resolution problem is a problem for other accounts as well. Consider justificationism: since the usual Gettier cases fall within the area of insufficient resolution, a satisfactory statement of the elusive "fourth condition" will have to include a solution to the resolution problem. It also afflicts reliabilisms. William Alston thinks of knowledge (minus a bell or whistle or two) as true belief produced by a reliable belief-producing process or mechanism—i.e., a process that

(in our maxi-environment) ordinarily produces true beliefs. But clearly a true belief can meet this condition and fail to have warrant, and that in two quite different ways. First, it can fail to do so by way of failure of proper function. My vision is reliable; I've passed all the tests with flying colors. But if I get drunk and see snakes in my bedroom, my belief has little or no warrant; and even if there happens to be a snake or two lurking in the corner, I certainly don't know that there is. But second, it isn't implausible to suppose that Alston was implicitly presupposing a proper function condition. If so, his view is affected by the resolution problem: the clock stopped at midnight, the van case, Paul's look-alike twin Peter—they and their ilk all lurk in the wings. In these cases the belief in question is produced by reliable processes functioning properly, all right, but in a situation where those processes display insufficient resolution. Things are less clear with coherentism (partly because coherentism itself is less clear), but given any plausible conception of coherence, it certainly seems that a belief could be appropriately coherent with its significant others even when its production falls into the area of irresolution. But then there will be situations where a belief is true and meets the coherence condition, but is nonetheless unwarranted.

Defeasibility to the Rescue?

In WCD I said that internalists take it that some quality—justification of one sort or another, perhaps—gives us the basic structure and form of warrant, with a fillip of some sort needed to evade Gettier. Several contributors—BonJour, Feldman, Swain—quite properly took me to task for this unduly flip way of stating the matter; there is no call for speaking thus disrespectfully of that fourth condition, calling it a mere fillip. Let me put it instead like this: justificationists think of justification as giving the basic structure of warrant; it takes up most but not all of the conceptual space between true belief and knowledge, so that a fourth condition is needed to occupy the rest of the space. But of course an internalist might reply that the fourth condition is every bit as important as the third, and takes up just as much conceptual space. Well, given the lack of a metric for conceptual space, there's not a lot of sense arguing about how much of it is taken up by what here; let's suppose that the internalists are the experts on what it is they think. In any event this fourth condition, whatever exactly the internalist proposes to propose, will have to solve the resolution problem.

Enter defeasibility. A defeasibility clause is not of itself an entire

theory of warrant, of course; it requires a base clause. Defeasibility clauses are essentially ways of grappling with the resolution problem, but they don't themselves provide a base or operative step. Thus they have often (early Klein, Swain, Pollock, others) been grafted onto a *justificationist* account of one sort or another: a true belief constitutes knowledge if it is justified and undefeated. But (as Swain in effect points out) a defeasibility account can in principle be added to any of several *different* base accounts of warrant. Swain says he accepts a defeasibility account in which it is justification[13] that is the base property; but it could also be employed where the base step is *being reliably produced* or *being produced by properly functioning cognitive faculties . . .* , or *being (appropriately) coherent*. In accounts of this sort a defeasibility clause rides on the back of some specification of the base clause for warrant.

A really good first question, clearly enough, is whether some defeasibility account—some defeasibility clause grafted onto a base clause—actually succeeds in resolving the resolution problem. There are several such proposals, but Peter Klein's is as good as any: I shall take a careful look at it. Now Klein cunningly sneaks up on what he takes to be the right account; I'll refer you to the text to see just how he does it, but the final product is something like

(K) *S* knows *p* if and only if

(1) *p* is properly grounded: either justified or reliably produced,

and

(2) there is no defeater *d* such that either (a) if *p* is justified for *S*, adding *d* to *S*'s beliefs is such that *p* is not justified, or (b) if *p* is not justified for *S* and the belief state containing *p* is reliably produced, then adding *d* to *S*'s beliefs moves *p*'s justificatory status too far from *p*'s being justified.

First, what does Klein mean by 'justification'? As he is thinking of it (I take it), a belief is justified for me only if I have *evidence* for it; and I think by 'evidence' he means *propositional* evidence, evidence from other propositions I believe. So a belief *p* is justified for me only if I accept it on the evidential basis of other propositions, and only if those others do in fact evidentially support it. A belief that I accept in the basic way—for example, the belief that I am being appeared to redly, or that 2 + 1 = 3, or that I had an orange for breakfast—is not justified.

Second, note that clause (1) means that the base or operative step in

Klein's account of warrant is a sort of hybrid (which is of course nothing against it): a belief meets that base condition just if it is either justified or reliably produced. This is the "general" theory; as it stands, Klein thinks, it won't quite do the job because the second clause is too strong. It is too strong, because there are certain kinds of defeaters that shouldn't count. For example, you see Tom in Tom's doorway, forming the belief that it is Tom you see; unbeknownst to you, Tom's grandmother is telling a neighbor that Tom has a look-alike brother, Don, indistinguishable from Tom from more than 10 feet away. This is a defeater for your belief that you see Tom. As it happens, however, Tom's grandmother is senile, and in fact Tom has no siblings at all. The defeater, therefore, is a *misleading* defeater, and doesn't relieve your belief of warrant. By way of repair, he proposes that the defeaters whose absence (2) requires be *genuine* ("a defeater is genuine just in case it defeats without depending on a false proposition," p. 117). This rough and ready characterization of genuineness gives way to a more official characterization (p. 117):

> . . . a proposition, d, is an *initiating defeater of S's belief that p* and d* is an *effective defeater of S's belief that p* iff d is the first proposition in an inference chain C, such that the last member of C, call it d*, is such that when d* is added to S's beliefs, p is not justified. . . . So, we can say that an initiating defeater, d, is *genuine* just in case it is true and there is some inference chain, C, to an effective defeater, d*, and there are no false propositions in C.

He then adds that a belief *p* is warranted for *S* just if *p* is properly grounded for *S* and there is no genuine defeater of it.

Will (K) do the trick? We might think, first, that this account suffers from undue vagueness: "*too far* from being justified"? Well, this vagueness does make it harder to assess the theory (to come up with counterexamples, e.g.), but I doubt the complaint is justified. Any satisfactory account will have to be vague. Suppose phony barn country starts a quarter-mile away from where you're looking at barns? Do you know then? Or a half-mile, or a mile, or 30 miles? Suppose those fake barns were there yesterday (or ten minutes ago) but have since been destroyed, or suppose they will be there within ten minutes? Suppose there's only one (or two, or *n*) phony barn(s) in the county, or within a mile, or a half-mile, or within eyeshot of where you are? Suppose they decided at the town meeting not to erect any fake barns, but it was a really close decision, and some people planned

to erect some anyway? There is certainly a substantial and ineliminable area of vagueness here, and it's not clear that (K) displays more of it than the notion of warrant itself.

1. False Beliefs in My Noetic Structure

So the vagueness isn't a problem; in fact it's a virtue. Nevertheless this account does suffer from grave difficulties. First, it seems to have the implication that if I hold a false belief, then any belief I have that is properly grounded but not justified is (or easily can be) defeated. Before I give the argument, however, we must be clear about just how an effective defeater works. According to Klein, a defeater d^*, for p, "is such that when d* is added to S's beliefs, p is not justified." Here I gather that Klein is thinking as follows: consider the set A of S's beliefs, and add the defeater d^*, thus moving to a new set A^*; then check to see whether p is justified with respect to A^*. I gather he's also thinking that when we add d^* to A, we delete p; otherwise, obviously, p will automatically be justified with respect to A^*, being a member of A^*. And of course we must delete not only p, but also any conjunctions of which p is a conjunct and more generally, any proposition that entails p. Still further, for any pair of propositions that together entail p, we must delete one or the other; the same for triples, quadruples, etc. From a rough and ready point of view, however, perhaps we could put it like this: to reach A^*, move first from A to A *diminished with respect to* p (a set of propositions maximally similar to A that does not entail p); then add d^*. This is *very* rough and ready (for one thing it won't really work when p is noncontingent— it's not really *entailment* we want, but a more epistemic notion of some kind), but stating this condition exactly is a project for some other occasion.

Now for the argument that if I hold a false belief, then any belief I have that is properly grounded but not justified is (or easily can be) defeated. Suppose I (mistakenly) believe

(A) Glasgow is larger than Chicago,

and suppose I also hold the properly grounded (by way of memory, say) belief

(B) I had an orange for breakfast.

From (A) I deduce (and add to my beliefs)

(C) If $-A$, then $-B$ (material implication).

Then $-A$, the denial of (A), is both a genuine initiating and an effective defeater for (B), the inference chain in question being just $-A$ itself (or the unit set of $-A$). The only member of this chain is true; $-A$ is an *effective* defeater because the result of adding it to my beliefs (diminished with respect to (B)) entails the denial of (B); hence (B) is "too far" from being justified with respect to that set of beliefs. So, on this account, if I believe any false proposition *A*, I won't know any proposition *B* that isn't justified for me; I won't know any propositions in the basic way. And the fact is we don't really need that basicality condition—won't the argument work just as well for a proposition *B* that is *justified* for me?[14]

The same problem, so it seems to me, afflicts the more elaborate account of defeater-genuineness in Klein's carefully crafted book *Certainty: A Refutation of Scepticism*.[15] The idea is that a genuine defeater is an initiating defeater that isn't misleading; and a *misleading* initiating defeater (again) is one such that the inference chain to the effective defeater depends upon some false proposition. The problem is to explain this lack of dependence. Klein puts it like this:

> Let us suppose, then, that an initiating defeater, d_1, is *misleading* if and only if there is some false proposition, f, in *every* D-chain between d_1 and an effective defeater, d_n, and f occurs in a link in the d-chain prior to every link in which a false proposition in E occurs. (*Certainty* p. 148)

I refer the reader to the text for details; but a *D*-chain is an evidential chain going from d_1 to the effective defeater; this chain can contain propositions from *E*, which is the set of *S*'s beliefs (actually, the subset she can properly take as evidence, but this complication won't be relevant here). And the account stipulates that a defeater is misleading only if there is no *d*-chain in which the first false proposition is a member of *E*. The idea (see note 16) is that a defeater is really misleading only if its defeating power, so to speak, does not depend upon any falsehoods in *S*'s beliefs, but does depend upon falsehood in the initiating defeater itself. Then we can say that you know *p* only if there are no genuine initiating defeaters for *p* for you, i.e., only if all the defeaters of *p* for you are misleading.

But I'm sorry to say this fails to do the job in the same way as the

condition in Klein's paper. Suppose I am a mathematical neophyte; on David Hilbert's authority I mistakenly believe the denial of Gödel's first theorem (GT). Where the arrow expresses material implication,

(H) $(-(GT) \rightarrow ((GT) \rightarrow -(2 + 1 = 3))) \& (GT)$

is now a genuine defeater for the proposition $2 + 1 = 3$ for me. For a defeater is genuine if there is at least one D-chain such that the first false proposition in it is a member of my set of beliefs. Here (H) is the initiating defeater and $-(2 + 1 = 3)$ the effective defeater; one D-chain will begin with the two conjuncts of (H), the third item being $-(GT)$. In this case, of course, there won't be any false proposition prior to $-(GT)$, which is one of my beliefs. If so, however, then on Klein's account of knowledge I have a genuine defeater for my belief that $2 + 1 = 3$ and thus do not know this proposition. Clearly the example can be generalized: if there is any false proposition I believe, I won't know anything at all. But that's unduly demanding.

The problem here is the highly articulate and complex nature of our design plan. False beliefs can indeed *sometimes* prevent a proposition from achieving the status of knowledge, thus depriving it of a status it would otherwise have;[16] but it is also possible to have knowledge that *p* even if (because of false beliefs I hold) there are true propositions not themselves rendering plausible any falsehood which, if added to my evidence, would bring it about that *p* is not justified for me. One sort of condition where this can happen is if *p* has a great deal of warrant in the basic way.

2. *Insufficient Generic Flexibility*

Second, Klein's account suffers from two kinds of unhappy inflexibility. First, an account of knowledge ought not to be tied to the *human* design plan: if it is really an account of *knowledge* (as opposed to *human* knowledge) it ought to apply to all the kinds of knowledge there are or could be. But obviously there could be creatures (God could create creatures) with cognitive styles very different from ours. For example, there could be creatures who had knowledge, but whose design plan did not involve justification in Klein's sense at all; these creatures' cognitive architecture doesn't allow for accepting a belief on the evidential basis of other beliefs. For such creatures, all beliefs would be held in the basic way; none would be accepted on the evidential basis of others, and none would be justified. Such creatures

could be either more impressive than we are, from a cognitive point of view (presumably God's knowledge is of this sort), or less impressive. Suppose we think about such a hypothetical creature: call him Sam. Suppose further that Sam's cognitive faculties sometimes display insufficient resolution; it is therefore possible that a belief *p* of Sam's should be true by accident. According to (K), *p* constitutes knowledge only if it satisfies the second clause of (K): there are no genuine defeaters for it. But how can this clause apply to Sam, given that his design plan is such that none of his beliefs is ever justified? Wouldn't they all always be "too far" from being justified?

Perhaps Klein could think of it like this: there is an objective *is-a-good-reason-for* relation (see Klein p. 117). This is a relation that holds among *propositions*; it does not involve reference to a noetic structure or system of cognitive powers; it makes no reference to knowers and their quirks, idiosyncrasies, and limitations.[17] (Perhaps it involves entailment and objective probability.) Then we can still raise the question whether there is a defeater for Sam's belief *p*: the question is whether there is some true proposition in the neighborhood such that the result of adding it to the propositions Sam does in fact believe is "too far" from standing in that objective relation to *p*.

Suppose there is that relationship. It is still possible that God (or evolution) has designed Sam in such a way that it *doesn't matter* whether there are true propositions which, when added to what Sam believes, yield a set of propositions with respect to which Sam's belief *p* is far from being justified. We can see how this could be the case with respect to us. Suppose I am in phony barn country, but the fact is I'm not able to look at any of the phony barns. Maybe I have a guardian angel who prevents me from ever doing so, or maybe those phony barns are made from a certain substance that causes my eyes to water when I look in the direction of one, thus preventing me from seeing it. Then when I look at a barn and judge that it is a barn, I will be right; furthermore, I will know that I see a barn, even though there is a (Kleinian) defeater for this belief. In the same way, perhaps Sam's cognitive system is designed in such a way that he always or nearly always forms true beliefs, despite the fact that these defeaters are present. And if so, then even though there are genuine defeaters lurking in the neighborhood, Sam doesn't believe the truth just by accident; he believes the truth because of the way he's designed. To put it another way, Klein's implicit characterization of accidentality is too parochial; it depends upon universalizing what are in fact specific

features of the human cognitive design plan, not features that every cognitive design plan will have to display.

3. *Insufficient Specific Flexibility*

But even if we confine ourselves to our own design plan, (K) is still too inflexible. (K) confers the title of knowledge upon some beliefs that are accepted in the basic way: for example, those that Descartes and Locke were especially enthusiastic about, such as self-evident propositions and beliefs appropriately about my own mental life (*I'm being appeared to redly*), as well as other beliefs that are reliably produced. But (K) fails to take account of the fact that warrant comes in degrees, and thus fails to accommodate the fact that some basic beliefs get an enormous amount of warrant in the basic way, enough to overwhelm propositional evidence against them. They can constitute knowledge even if, *so far as available propositional evidence goes*, there is a balance of evidence against them. For example, suppose I am appeared to redly, take note of my phenomenal field, and form the belief

(R) I am being appeared to redly.

(R) is not justified for me in Klein's sense, because I don't believe it on the basis of any other propositions at all. We may assume (at least for purposes of argument) that (R) is reliably produced; it therefore meets the first of (K)'s conditions. But it could run afoul of the second and still constitute knowledge. For the fact is

(B) I display a certain brain structure such that 6 out of 10 people who display that structure are never appeared to redly.

If (B) were added to my beliefs (of course deleting (R), any proposition that entails it, etc.), then, with respect to that set of beliefs, it would be more likely than not that I am not being appeared to redly. (R), then, would be a long way from being justified—adding this proposition would move its justificatory status "too far" (I'd guess) "from *p*'s being justified." (B) is therefore a defeater for (R). But of course the fact is I would know (R); in fact I would know it even if I also knew or believed (B).

Clearly there are many different sorts of examples of this general sort. The problem with (K), here, is that it fails to match the complex,

highly articulated architecture of our design plan. It can happen that a given belief *p* gets a great deal of warrant in the basic way: beliefs about my own mental life, to be sure, but also perceptual beliefs, memory beliefs, simple *a priori* beliefs, perhaps religious beliefs or beliefs about God, and the like. It could also happen that there is a genuine (in Klein's sense) defeater *d* for *p*; that fact needn't be sufficient for denying that *p* constitutes knowledge. Indeed, something stronger is true: it could be both that I know *p* and that I *believe d*. It is possible, that is, that I believe *p* in the basic way, that *p* be very far from being justified for me by virtue of the fact that I believe *d*, and that *p* nonetheless constitutes knowledge.[18]

4. Problems Inherited from Reliabilism

In his version of defeasibility theory (as opposed, e.g., to Swain's), Klein helps himself to a reliabilist notion: *p* is properly grounded just if either it is justified or it is *reliably produced*. But beware of reliabilists bringing gifts: in helping himself to this reliabilist theme, Klein inherits reliabilist problems. I don't have the space to go into this properly,[19] but I think we can see briefly that there will be real problems. For how is Klein understanding 'reliably produced'? Suppose he takes it as in Alstonian generic reliabilism:[20] 'produced by a reliable belief-producing process or faculty'—a *concrete* process or faculty (so far no truck with Goldmanian types). Well, suppose my (reliable) perceptual faculties malfunction (due to drink or drugs) and I think I see pink rats, forming the belief that indeed there are pink rats in the room; and suppose as it happens there are some (in the closet, perhaps). Then this belief is properly grounded (the belief is reliably produced). Furthermore, there need be no defeater—no true proposition which, when added to my beliefs, yields a set with respect to which the belief in question is "too far" from being justified.[21] But then, on this account, I would know that there are pink rats in the room—which I don't.

Suppose instead we take 'reliably produced' in terms of Goldmanian types. Out of the frying pan into the fire. For, for any necessary proposition I believe—that there is such a person as God, if there is, that there is not, if there is not—there will be a reliable type of which that belief will be a product, no matter *how* I actually came to the belief (see WPF pp. 205–208). And again, there need be no true proposition which, when added to my beliefs, makes *p* "too far" from being justified. So on this way of taking 'reliably produced', any

necessary proposition I believe will automatically constitute knowledge for me—even if I can't see that the proposition is necessary, know of no arguments for it, and come to the belief in a totally outrageous way. And while I can't take the space to show this here,[22] beliefs of necessary truths will present the same problem if we take 'reliably produced' in such a way as to satisfy the reliabilist conditions laid down by the early Dretske,[23] the later Dretske,[24] Robert Nozick,[25] or Fred Suppe.[26]

5. An Unnecessary Difficulty

I am halfhearted about this final difficulty, because I think it is a problem Klein unnecessarily creates for himself. As Klein thinks of justification (personal communication), it isn't necessary, for me to be justified in believing p on the evidential basis of q, that I see the connection between q and p, see that the former really does offer evidential support for the latter. All that is required is that I believe the former, accept the former on the basis of the latter, and the former actually support the latter. But this leads to real difficulties. Suppose I believe, say, Gödel's second theorem, not because I see that it is true, or have followed a proof, or have been told by someone in the know, but because of some cognitive glitch: my favorite comic book character asserts it, and I always believe anything my favorite comic book character asserts. But suppose also the fact is I believe propositions from which Gödel's theorem deductively follows, by an argument that is easy for a normal human being to follow. Furthermore, there is no true proposition p that can be added to my beliefs, such that Gödel's theorem isn't justified with respect to the resulting set of propositions. Then my belief in Gödel's theorem satisfies the conditions for knowledge laid down in (K); but surely it doesn't constitute knowledge. I say I am halfhearted about this objection, because I can't see why Klein doesn't simply require that I am justified in believing p on the basis of q only if I (at least implicitly and dimly) see that p does in fact support q. He fears an infinite regress here, but I fail to discern it.

The Environmental Condition Revisited

But of course none of this helps me with the resolution problem. All well enough to complain that *Klein* hasn't really solved it: do I have anything better to offer?

Our problem is that a true belief B can fail to be knowledge by way

of being 'accidentally true'—i.e., it can be just by accident that I form a true belief on this occasion—even though *B* is produced in an appropriate maxi-environment and the other conditions of warrant are met. The problem is that there is a certain retail lack of fit between a particular exercise of cognitive powers and a particular cognitive mini-environment. I can't visually distinguish a real barn from a barn facade—that is, I can't do so from the highway and given a certain angle of vision. (Of course I can tell it's a real barn if I walk around behind it, or if it is a barn we are building.) In the van case, the processes that lead me to and sustain the belief that I own a Chevy van don't enable me to distinguish between the situation where I do in fact own a van, and the sort of situation where my van is destroyed by that meteorite or errant gravel truck. As I pointed out above (p. 314), some mini-environments are favorable for a given exercise of cognitive capacities, and others are unfavorable, misleading, for a given exercise of cognitive capacities.

But then the solution to the resolution problem (and hence also the Gettier problem) is simple enough: the conditions for warrant as I stated them include an environmental condition; and that condition must be understood in such a way that it specifies an appropriate cognitive mini-environment (as well as an appropriate maxi-environment). A belief *B* has warrant only if *MBE* is favorable for the exercise of cognitive powers *E* by which *B* was produced, is the sort of mini-environment for which the powers of which *E* is an exercise are designed (by God or evolution), the sort of mini-environment in which *E* can be counted on to produce a true belief. Together with the other conditions of warrant, this clarification or amplification of the environmental condition will enable us to solve the resolution problem and elude Gettier without resort to the distinction between processes directly aimed at producing true belief and those that are not directly so aimed.[27]

But can we say anything more precise than 'can be counted on'? Perhaps not; perhaps that sort of precision isn't attainable or necessary here. Still, it won't hurt to give it a try. Shall we construe 'can be counted on' probabilistically, for example? There may be possibilities along this line; however, I'd like to explore a slightly different direction, a direction which will also display connections with some other views.

In some of his very early work Fred Dretske suggested that we think about *counterfactuals* in this context, a thought later taken up and amplified by Robert Nozick. According to Nozick, *S* knows *p* if and only if (1) *S* believes *p*, (2) *p* is true, (3) if *p* were false, *S* would not

believe p, and (4) if p were true S would believe p. (Nozick adds an epicycle involving methods of knowledge that won't concern us here.) Taken as an attempt to explicate warrant, these conditions have foundered. First, as Richard Fumerton shows, they encounter crippling difficulty when it comes to knowledge of necessary truths.[28] They also come to grief because of various possibilities of cognitive malfunction, for knowledge of contingent as well as of necessary truths.[29] Still further, that third clause seems to have the consequence that we can't know (for example) that our cognitive faculties are functioning properly (if they weren't, we might still believe that they are); but surely we (sometimes) can. Taken as a total account of warrant, then, these counterfactuals won't do the job.

In "Postscript to 'Proper Functionalism and Virtue Epistemology' " (pp. 275–77), Ernest Sosa offers an alternative to Nozickian tracking—"Cartesian tracking," as he calls it—as a naturalistic account of proper function. So taken, Cartesian tracking too, I believe, has serious problems (see pp. 369–70). Still, something in this general neighborhood could perhaps serve to explain what it is for an exercise of cognitive powers to be such that it can be counted on to deliver a true belief. So consider a given belief B, the exercise E of cognitive powers that produces B, and the cognitive mini-environment MBE in which it is formed. Then

(F) MBE is *favorable* for E, if and only if, if S were to form a belief by way of E, S would form a true belief.

Then we can say that

(Resolution Condition) A belief B produced by an exercise E of cognitive powers has warrant only if MBE is favorable for E.

We may add that a belief has warrant if and only if it meets the conditions of warrant as I stated them in WPF, the environmental condition understood to include favorability on the part of the relevant mini-environment as well as suitability of the maxi-environment.

A couple of comments. First, of course, we aren't to suppose that my forming a true belief B by way of E is sufficient for the favorability of MBE; the truth of p and q is not sufficient for the truth of the counterfactual *if* p *then* q. This is a point on which the usual (Lewisian, Stalnakerian) semantics for counterfactuals is inadequate. Consider quantum effects: perhaps in fact the photon went through the right slit

rather than the left; that is not enough to entail that if it had gone through either slit, it would have gone through the right. I toss the die; it comes up 5. That is not sufficient to entail that if I had tossed the die, it would have come up 5. The truth of a counterfactual requires not just that *p and* $-q$ be false in fact; it is also necessary that even if things had been moderately different, it still wouldn't have been the case that p and $-q$. To put it in familiar semantical terms, the counterfactual is true only if there is no *sufficiently close* possible world in which p is true but q is not. How close is sufficiently close? That is of course a question without an answer; counterfactuals are in this way quite properly vague.

Second: note that (RC) guarantees that no false belief has warrant: if in fact my belief that p is false, then the counterfactual specified in (RC) has a true antecedent and a false consequent.

(RC) rules out many cases where the other warrant conditions are met but the belief doesn't constitute knowledge—many, but perhaps not all. For suppose you have a box in which there is a vase. If I open the box, however, I don't directly see the vase—what I see is its reflection in a cleverly placed mirror. I open the box, take a quick look, and form the firm belief that there is a vase inside. The warrant conditions are met, and so is (RC): if I were to form a belief on the topic of the focal proposition of this mini-environment by way of an exercise of those cognitive powers, I would form a true belief. The situation is one in which that exercise of cognitive powers can be counted on to produce a true belief. But does this true belief really constitute knowledge? If I *know* about the arrangement (maybe I'm inspecting boxes of this sort to find the few in which the vase has been left out), then perhaps I do—just as I know that Mount Hood is conical by seeing a photograph of it or by seeing it on TV. But suppose I *don't* know about the arrangement, have no idea about any mirrors in that box. *Then* do I know? The 'causal chain' by way of which I come to form the belief, some might say, is somewhat 'deviant'. Perhaps that's true, but does it destroy knowledge? Not all deviant causal chains do; does this one? The answer isn't clear. It isn't *obvious* that I know; in fact one is a bit less strongly inclined to think so than to think not. But it also isn't obvious that I *don't* know; furthermore, it isn't obvious that this case falls into one of those areas of vagueness in the penumbra of the concept of knowledge. So perhaps it isn't really clear whether or not we have a counterexample here. Even if we do, however, we may still rest reasonably satisfied, for the moment anyway, with the vaguer formulation of the environmental condition.

II: Replies to Individual Authors

Ad Lycan

William Lycan chides me for failing to address the sort of explanationist coherentism made famous by Nelson Goodman, W. V. Quine, Wilfrid Sellars—and Lycan himself.[30] Of course he's right; this version of coherentism does indeed deserve close attention. (I can only plead human frailty and insufficient space-time.) He puts his "typical" explanationism as follows:

> (EXP) . . . s/he holds that a hypothesis is warranted by its ability to explain a particular set of data better than any other available hypothesis. . . . (p. 4)

This leaves us wondering about those data: how do *they* get warrant? Here Lycan appeals to what he calls "the Principle of Credulity," which is "Accept at the outset each of those things that seems to be true" (p. 5). He claims further that this principle is a consequence of the "more general characteristically explanationist claim that *conservativeness* . . . is a theoretical virtue" (p. 6); and he argues that "whatever epistemic or justifying status inheres in the other standard pragmatic theoretical virtues (simplicity, testability, fruitfulness, power, and the like), conservativeness shares that same status" (p. 6). His suggestion seems to be that one of my beliefs can acquire warrant just by virtue of seeming to me to be true.[31] A belief doesn't get *much* warrant in this way ("A belief is justified by the bare fact of our holding it, I maintain, but only to the smallest degree," p. 7), but it does get a smidgen.

Here I have a question and a problem. First, the question: theoretical virtues are ordinarily thought of as characteristic of *hypotheses*, or *explanations*, or *scientific theories*; how then do other sorts of beliefs—for example, "spontaneous beliefs" (p. 5) such as those produced by memory and perception—get in on the largess? That's the question; the problem is this. Although in (EXP) Lycan uses the term 'warrant', he soon starts focusing his attention on *justification*. It sounds as if he thinks my belief or I myself might be *justified* just by virtue of the belief's seeming to me to be true. Agreed: at least for deontological and responsibility-related construals of justification. If a belief simply seems right to me, perhaps *obviously* right, perhaps as obviously right as any belief I have, how can I be irresponsible or go contrary to duty in holding it? So perhaps I'm *justified* by a belief's

simply seeming to me to be true; but of course the belief in question might have no warrant. Lycan mentions (pp. 14–15) the Cartesian madmen who think they are gourds—summer squash, perhaps, or pumpkins. Such beliefs might be *justified* (better, the believer might be justified in holding them) just by virtue of the fact that they seem true to him, but of course in this way it doesn't follow that they have or thus acquire any degree of *warrant*.

This problem with ambivalence (I hesitate to say "equivocation") between justification and warrant persists and reappears at other points in Lycan's paper. Some of my objections to coherentism, as he points out, involve the possibility of a person's forming a mad belief, his other beliefs then settling into a coherent pattern around it. Lycan's comment:

> Plantinga's examples are . . . cases of someone's starting with one or more weird ideas that have no foundation in the real world, and then constructing a coherent system of beliefs that surround and support those ideas. . . . I say the weirdo *is* justified in accepting the wild belief and the system that goes with it. (pp. 18–19)

Elsewhere (note 20):

> I take the position that a smoothly and globally deceived victim [e.g., a brain in a vat] is not only conceptually possible, but is exactly as well justified as is a comparable, counterpart subject whose experience is similar but veridical.

Well, I think so too. If we take justification in anything like its original deontological sense, it is hard to disagree: the weirdo is indeed justified, as is the smoothly and globally deceived victim. But their beliefs might nonetheless have no *warrant* at all: mad beliefs don't acquire warrant just by way of seeming true, or even by way of being integrated into a coherent system of some sort.

So perhaps Lycan is really thinking throughout of *justification*, not warrant; perhaps he thinks a certain kind of coherence is necessary and sufficient for justification, with some further fourth condition (a condition designed to deal with Gettier or the resolution problem) required for warrant. If so, he and I, I think, still have a disagreement, in fact a pair of disagreements. First, it seems quite clear that justification—at least if taken as responsibility or in some other way fairly close to its original deontological sense—does not require explanatory coherence. Jimmie G., Oliver Sacks's Lost Mariner (Lycan pp. 14–16),

had a noetic structure that displayed minimal explanatory coherence; nevertheless he might have been as responsible as you please; he might have been doing his level best; he might have been wholly justified.[32] Whether my beliefs display coherence might not be up to me; if it isn't, my failure to be coherent will not reflect on my responsibility.

But what if Lycan isn't thinking of justification in this way—what if he just *defines* 'justification' in terms of coherence? Then of course coherence will be necessary for justification; and Lycan's overall view would simply be that explanatory coherence is necessary for warrant, along with another and so far unspecified condition to deal with Gettier or the resolution problem. This view also seems to me mistaken. Just as coherence is clearly *insufficient* for warrant, so it is also *unnecessary*. Or rather (as it seems to me), no more coherence is necessary than is required for proper function. Noetic structures such as ours must display a certain degree of coherence in order to satisfy that condition; although I can't go into the matter here, it is clear that various kinds and degrees of incoherence are incompatible with a belief's being produced under the conditions of warrant. But why think warrant requires more coherence than is required by proper function? In particular, why think that much by way of *explanatory* coherence is required? I believe the corresponding conditional of *modus ponens*; I also believe that it is wrong to lie about one's colleagues in order to advance one's career. I believe both with great, near maximal firmness. But neither of these propositions, as far as I can see, is in any straightforward sense an *explanation* of other propositions I believe; nor is either *explained by* any other propositions I believe. And yet they have, I should judge, maximal or near maximal warrant.

Following Sosa, BonJour (p. 60) distinguishes reflective knowledge from *animal* knowledge, the sort of knowledge an animal or small child can have; and he claims that a certain kind and degree of coherence is required for the former, even if not for the latter. Perhaps he is right; and perhaps there is also an important variety of knowledge for which explanatory coherence is crucial. We can be grateful to Lycan and explanatory coherentists for calling our attention to this variety of knowledge,[33] even if their comments do not apply to knowledge as such.

Ad Lehrer

Keith Lehrer's characteristically acute and penetrating essay contains much by way of sound philosophy; I'd like to begin by noting a

couple of important points where we agree. First, each of us hopes to play Elisha to Thomas Reid's Elijah, thus inheriting Reid's mantle (II Kings 2:11–15). In particular, each of us sees crucial significance in Reid's First Principle 7:

> Another first principle is—*That the natural faculties, by which we distinguish truth from error, are not fallacious*. If any man should demand a proof of this, it is impossible to satisfy him . . . because, to judge of a demonstration, a man must trust his faculties, and take for granted the very thing in question. If a man's honesty were called in question, it would be ridiculous to refer it to the man's own work, whether he be honest or not.[34]

Lehrer's esteem for Principle 7 is unveiled in his (T) (p. 41); mine is displayed in the contention (WPF chap. 12) that acquiring an undefeated defeater for this principle wreaks unholy havoc with one's entire noetic system, perhaps even bringing it about that whatever else you believe, you believe irrationally. Lehrer also seems to me approximately right in his diagnosis of Mr. Truetemp's predicament (p. 32–33). As I see it, Truetemp has a defeater for his belief in the fact that (as he no doubt thinks) he is constructed like other human beings and none of *them* has this ability; furthermore, everyone he meets scoffs or smiles at his claim that he does have it. Truetemp's defeater means that his belief does not meet the conditions for warrant; hence (contra Lehrer) he doesn't constitute a counterexample to my analysis of warrant. (If Truetemp doesn't have a defeater here, he also lacks warrant, since proper function, in his situation, requires that he *have* a defeating belief.)

Still further, I agree with Lehrer in holding that at least certain kinds of knowledge demand the sort of coherence of which he speaks. And even further yet, I believe we agree that Gettier problems depend essentially on the resolution problem, i.e., on the fallibility of our cognitive faculties even when they are functioning properly (and even if we disagree about how to resolve the resolution problem).

Still, there remain several interesting points of disagreement. Although I very much admire Lehrer's *Theory of Knowledge* (as other of his works) I am disinclined to believe that his account of knowledge there is successful; but I can't take the space here to explain why.[35] Instead, I'd like to respond to his comments on my evolutionary argument against naturalism (WPF chap. 12), and on his suggestion that my account of warrant is incompatible with the idea that God

could cause me to know a proposition by specially causing me to believe it.

Turning to the first topic, Lehrer represents me as arguing that naturalism and evolution taken together cannot supply us with any reason for thinking that our cognitive faculties function reliably (in such a way as to provide us with mostly true beliefs), agrees that this is in fact so, but then taxes me with inferring from this "the much stronger conclusion that our beliefs are likely to be false if evolutionary theory is correct" (p. 27). That would indeed be a miserable inference, but I plead not guilty. What I did argue is as follows. Consider philosophical naturalism (N)—as van Fraassen says (this volume p. 172), a view not easy to characterize, but at any rate including the idea that neither theism nor any view similar to it is true; consider the view (E) that our cognitive faculties have come to be by way of the sorts of mechanisms to which contemporary evolutionary theory directs our attention; and consider

(R) My cognitive faculties are reliable.

P(R/N&E), the objective conditional probability of (R) on N&E, I argued, is either low or inscrutable—such that we can't even make a decent estimate of it. Now (R) is really a version of Reid's first principle; I agree with Reid and, I hope, Lehrer in holding that (R) has a great deal of warrant in itself, in the basic way. Like other beliefs that get warrant in the basic way, however, (R) can be defeated, and my claim is that the low or inscrutable probability of (R) on N&E furnishes one who accepts N&E with a defeater for (R)—a defeater that can't be itself defeated. One who has a defeater for (R), furthermore, has a reason for rejecting, withholding, being agnostic with respect to any belief she holds, including, of course, N&E itself. The latter, therefore, is in a certain crucial sense self-defeating and hence (for someone aware of the argument) can't rationally be held.

Obviously many questions can be (and have been[36]) raised about this argument, but it wasn't at all a matter of making the inference Lehrer mentions. Still, what Lehrer says at a neighboring juncture raises an extremely interesting point about the argument. He suggests that

> it may be that, though our justification or warrant for the things we believe does not depend on an argument concerning God, the existence of God does supply a better philosophical explanation for why the proper

> functioning of our faculties yields true beliefs or has a trustworthy if fallible tendency to do so. (p. 28)

He then argues that in fact this isn't so: the existence of God does not supply a better explanation. Why not? Here he neatly turns the tables. I've argued in various places that the existence of evil doesn't provide a strong (probabilistic) argument for the nonexistence of God; Lehrer argues, correspondingly, that the existence of God doesn't provide the materials for a strong (probabilistic) argument that our cognitive faculties are reliable. Now in one way this is obvious; as Reid notes in Principle 7 (quoted above), one can't sensibly give any argument at all for the reliability of our cognitive faculties. But Lehrer really has something different up his sleeve:

> Compare, finally,
>
> S. Satan and his cohorts produce incredible deceps of error
>
> with
>
> E. Evolutionary processes produce incredible deceps of error.
>
> I find little to choose between them. A naturalist wishing to assign a high probability to the conclusion that the proper functioning of our faculties yields truth because they are the result of evolution must assign a low probability to E, while a supernaturalist wishing to assign a high probability to the conclusion that the proper functioning of our faculties yields truth because they are designed by God must assign a low probability to S. (pp. 29–30)

This is relevant to the evolutionary argument against naturalism as follows. I argued that P(R/N&E) is low or inscrutable, and I claimed that this gives the devotee of N&E a defeater for R. But Lehrer gives us a reason for supposing that something similar goes for the theist. P(R/Theism), we can take him as suggesting, is also low or inscrutable: for even if we have been created by God, it is possible that God allows Satan to deceive us massively (just as it is possible that God allows both us and Satan to create other kinds of havoc); furthermore, we can't make a decent estimate of the probability that God *would* allow us to be thus deceived. But then P(R/T) is inscrutable for the theist, like P(R/N&E) for the evolutionary naturalist; the former, therefore, is in the same boat as the latter, having the same sort of defeater for (R).

Very clever! This is a *tu quoque* (you're another) objection; it resembles an objection brought by Carl Ginet in the piece mentioned in note 36. Consider austere theism, the view that we have been

created by a very powerful and knowledgeable being. Austere theism (call it "A") is obviously and immediately entailed by theism *simpliciter*; it differs from the latter in not including the proposition that we human beings have been created in the image of God, part of which involves our resembling God by way of our ability to form true beliefs and have knowledge. Ginet pointed out that P(R/A) is low or inscrutable; but any theist who isn't deductively challenged believes A; so any such theist has a defeater for R—just like the evolutionary naturalist! My reply was that the low or inscrutable value of P(R/A) doesn't give the theist a defeater for R. The reason is that he knows or believes that the warrant A has for him is derivative from the warrant T has for him; it is T, not A, that he knows by way of the *Sensus Divinitatis* or by way of faith and the "Internal Testimony of the Holy Spirit." But if I rationally believe that the warrant, for me, of a proposition Q is derivative from that of a proposition P, for me, then for any proposition R, if I don't believe that P is a defeater, for me, of R, then Q is not a defeater, for me, of R.[37] So, for example, I know that

N You have an old Nissan;

acting on the principle that it is always nice to believe an extra truth or two, I infer:

J You have a Japanese car

and

O You own an old car.

I then note that P(J/O) is low (most people who own an old car do not own a Japanese car), and conclude (in considerable puzzlement) that I have a defeater for J. But of course I am mistaken: P(J/O) may be low, but O is not a defeater for J. The reason is just that the warrant of both O and J is derivative, for me, from that of N; since N is not a defeater, for me, of J, neither is O. Lehrer's objection is similar to Ginet's, but stronger; its beauty lies in the fact that it concedes creation in the image of God but still produces a defeater. Granted, that's how we were created, but there is also (at least on Christian theism) the fall into sin and the possibility of deceit by Satan (that father of lies) and his cohorts. According to Christian doctrine, furthermore, the fall into sin has indeed damaged the image of God in us, and has damaged our cognitive faculties: were it not for sin and the fall, we human beings would find the existence of God as obvious and uncontroversial as that of trees and horses. To put it in terms of John Calvin's way of looking at the matter, sin has damaged and

compromised the *Sensus Divinitatis*, and also to some degree corrupted our apprehension of moral truth. So the probability of R on T plus this Christian teaching is either low or inscrutable. And doesn't this provide the Christian with a defeater for R, and hence for everything else she believes, including Christianity?

But there is a reply, and substantially the same reply as to Ginet. The Christian believes she knows these central Christian truths—creation and fall into sin—by way of divine revelation. This can be construed in more than one way; for simplicity, take it Calvin's way. Very much oversimplified, the idea is that the Christian knows these truths by way of the Internal Testimony of the Holy Spirit, which prompts acceptance of what the Bible teaches; more exactly, what God intends to teach in the Bible. Part of what the Christian thus learns is that divine grace *restores* the image of God in the believer; part of the effect of the work of the Holy Spirit is for the doleful effects of sin to be increasingly mitigated. (In particular, this restoration cures, repairs the damage to the *Sensus Divinitatis*; it removes our blindness to the existence of God and enables us to see, once more, some of his glory and majestic beauty.) As a Christian sees it, then, she is a person in whom the image of God has been partly restored, so that once more she resembles God with respect to the ability to form true beliefs and have knowledge.

If so, however, she needn't after all have a defeater for (R). For what she thinks has warrant for her is the *whole* message of Christianity, the *whole* of what the Holy Spirit testifies to. What she knows by way of Revelation and the Internal Testimony of the Holy Spirit isn't just T (theism), but the whole Christian story, including fall, redemption, and restoration of the image of God. But then (by the principle enunciated three paragraphs back) the fact that the probability of R with respect to some *part* of that whole is low or inscrutable doesn't give her a defeater for R, unless R is also unlikely with respect to that whole; and of course it isn't. More generally, if *S* rationally believes that the warrant, for him, of a belief *B* is derivative from that of a belief *A*, then *B* won't be a defeater, for him, for any belief *C* if he doesn't believe that *A* is a defeater for *C*.

Finally and briefly, I'd like to comment on Lehrer's argument that if God specially *causes* me to believe something or other, then, on my view of warrant, that belief does not constitute knowledge. I hope this isn't correct, because in *Warranted Christian Belief*, the third and (I devoutly hope) last in the Warrant series, I follow John Calvin and much of the rest of the Christian tradition in arguing that Christian faith is produced by way of the Internal Testimony of the Holy Spirit;

I also mean to hold, of course, that Christian faith so produced can constitute knowledge. And in fact it *isn't* correct. What happens in this case, we can suppose, is that there is a smallish revision in the design plan governing the relevant parts of my cognitive architecture, as well as the architecture of will and what Jonathan Edwards calls "affections," in particular the *religious* affections.[38] The relevant cognitive powers are, so to say, mildly redesigned, and they work properly according to this new design plan.

Lehrer's case is a little different, however: there isn't any alteration in the design plan, but God simply, directly, and specially causes a belief to well up within you. But can't I take this to be a special limiting case of cognitive faculties or belief-producing processes functioning properly? God instills a true belief in you, intending in so doing to instill in you a true belief. Why can't I think of his doing that as itself a belief-producing process, especially since that is precisely what it is? True, this belief isn't exactly produced by a cognitive *faculty*, or at least by one of *my* cognitive faculties; but it is produced by a properly functioning cognitive *process*, and I think that's sufficient.

Ad BonJour

Laurence BonJour's paper is lucid and thoughtful; it also enters deep waters. I don't have the space to pursue nearly all the important topics he broaches; I shall have to content myself with commenting on just two.

1. Justification Not Even Necessary *for Warrant?*

BonJour points out that my way with justificationists was a bit quick. I argued that justification, taken either deontologically or in terms of having a reason or evidence, isn't *nearly* sufficient for warrant.[39] He replies, substantially, that there are no near misses in logic; an analysis is either correct or it isn't (there's no such thing as being *nearly* correct); we've known since Gettier that we need a fourth condition; fourth conditions can't be said to vary in size or significance; and there isn't any metric for conceptual space. Of course I realize that distance is a metaphor here, and that it isn't possible to quantify distance, so taken. Nonetheless, I think it is a metaphor we understand and can usefully employ, and that in fact justification *isn't* anywhere nearly sufficient for warrant—contrary to what I took BonJour to mean when he said that the justified true belief account is substantially correct.

But suppose the internalists (BonJour, Feldman) insist that they mean no such thing: justification, they say, is *one* condition needed for warrant, and a fourth condition is *also* needed; neither is more sizable or important than the other, and they never intended to suggest anything to the contrary. Well, the internalists are presumably the authorities on what they do and did mean. What I shall argue instead, then, is that justification isn't even a *necessary* condition for warrant. (If that's right, then internalists don't have so much as a *third* condition, let alone a fourth.)

First, take justification in the sense of evidence, or reason, or ground—the way BonJour seems to take it.[40] Clearly one can know many propositions without having evidence or a reason for them: obvious truths of arithmetic and logic, for example, such as $2 + 1 = 3$ are not accepted on the basis of reasons or grounds (see WPF p. 188), but can nonetheless constitute knowledge.[41] The same goes for memory beliefs (see WPF pp. 188–89).

As for justification taken deontologically (in terms of duty or obligation), things are a bit more complicated. Here the idea would be that you don't really know some proposition p unless you have been *responsible* in forming and sustaining the belief that p, and you have satisfied the relevant duties or obligations. Now clearly there are *some* duties such that one can't know p without satisfying them: for example, the duty not to destroy your cognitive capacities by taking mind-altering drugs. But presumably that isn't the sort of duty the deontological internalist has in mind. If he were satisfied to stop here, he could argue in the same way that justification is a necessary condition of, say, *digestion*. You won't digest properly if you don't conform to the duty not to commit suicide, as well as the duty not to destroy your digestive system in thoughtless and ill-conceived experimentation with ground glass or sword swallowing. So what duties or obligations, with respect to a belief, *are* the ones satisfaction of which is relevantly required by warrant? In WCD (p. 45) I argued that one could know much even if violating the epistemic duties Roderick Chisholm suggests. Of course other deontological internalists might suggest other duties; perhaps there are other good candidates for the relevant epistemic duty. If there are, however, it is exceedingly hard to see what they might be. True, there might be specific beliefs B such that in fact B won't constitute knowledge unless some duty has been satisfied; but of course that isn't sufficient. What is needed is some duty such that satisfaction of *that* duty is required for any belief to enjoy warrant—or perhaps the idea would have to be (less demandingly) that for each

belief *B*, there is some duty such that satisfaction of that duty is necessary for *B* to have warrant. But what could such duties be? No matter how undutiful I am, no matter what relevant duties and obligations I flagrantly flout, won't I still be able to know that 3 + 1 = 4 and that I am being appeared to redly? It is up to the deontological internalist to specify the duty or duties she has in mind; but it is at the least exceedingly difficult to see what they might be. The fact is, I don't believe there are any such duties.

But if this is true, then justification is neither anywhere nearly sufficient for knowledge, nor even necessary for it; it isn't even a third condition.[42]

2. *Skepticism?*

BonJour asks (p. 63), "Why isn't Plantinga's view itself a deep and troubling version of skepticism?" By way of reply: first, the claim that we *do* have knowledge is no part of my official account of warrant; that account says only what conditions are necessary and sufficient[43] for a belief's having it, not that any of our beliefs meet those conditions. But of course I do think we have a good deal of knowledge, and depend on that opinion in presenting counterexamples to various alleged necessary conditions for knowledge; I should therefore be distressed if it turned out that my conditions for warrant guarantee skepticism. Why does BonJour think they do so? His basic reason, I think, is contained in the following:

> On his [i.e., my] view, while we may have "warrant" and even knowledge *if* the right conditions are satisfied, we apparently have no way to tell from the inside whether those conditions are ever satisfied, nor any reason at all to think that they are (though we might still, of course, be "warranted" in believing and even know that they are satisfied, if the belief that they are satisfied should itself happen to satisfy Plantinga's conditions). If a particular belief is called into question, either theoretically or as a possible basis for action, we are apparently helpless to resolve the resulting issue in any way that we can understand to be adequate, as helpless as the dumbest, least reflective animal. (p. 63)

And of course BonJour thinks that if we *are* thus helpless, in these ways, then we don't in fact have knowledge.

Now, first, it seems to me wholly mistaken to say that if a belief is called into question, then, on my view, we are helpless to resolve the resulting issue: if you claim that Newark is in New Jersey and I that it

is in New York, a quick look at a map ought to settle the issue, and there is nothing in my account of warrant to suggest otherwise. But of course this isn't the sort of "settling" BonJour has in mind; what he means is that on my view we can't provide a certain kind of ultimate answer to the skeptic. We can't satisfy the skeptical demand that we be able to tell "from the inside" that the conditions of warrant are in fact satisfied. But what is it to tell "from the inside that *p*"? And wouldn't the same skeptical questions arise about whatever it was that one appealed to in settling this from the inside? You will believe such things as that you are appeared to thus and so: but don't you then, on BonJour's showing, need some reason to think that you really *are* appeared to in the way you think you are? And what about the answers to *those* questions, and so on, world without end? I *think* BonJour believes that *a priori* knowledge will fend off the looming regress; and I think he means to say that we can tell from the inside that the conditions of warrant are satisfied only if we can tell *a priori* that they are. (As he says in a slightly different connection (p. 54), "An empirical argument would seem to be obviously question-begging in relation to the larger problem of whether all of our perceptual experiences might be caused in some aberrant way.")

But again: don't the same questions arise about *a priori* belief and alleged knowledge? Can't we raise the same skeptical problems? I believe the corresponding conditional of *modus ponens* and that 2 + 1 = 3; indeed, I believe each necessarily true. These beliefs seem wholly obvious; I find myself utterly convinced. They have about them, furthermore, the peculiar feel that *a priori* beliefs have—that feel that somehow they just couldn't possibly be false. But of course such a feel could be misleading. A *false* belief, obviously enough, could have that sort of feel for me: I could be mad, or a victim of an Alpha Centaurian cognitive scientist, or a brain in a vat, or a victim of a Cartesian evil demon. Indeed, it was in this very context of *a priori* knowledge that Descartes turned to that evil demon scenario. So obviously I can be wrong, even when it seems *a priori* for all the world that the belief is true. As a matter of fact, this isn't merely an abstract possibility: some propositions that have that *a priori* feel about them *are* false, as is shown by certain versions of the Russell paradox (see WCD pp. 104–5). So do I have any reason to think that 2 + 1 really does equal 3? Of course it *seems* to me that it does and indeed necessarily does; but how much is that worth? BonJour complains that on my accounts of induction and our knowledge of other minds, we *just find* ourselves believing, under certain circumstances, that the sun will rise tomorrow

or that Sally is angry, without any real insight into how it is that the sun's having risen lo! these many days makes it likely that it will rise tomorrow. But are we any better off in the *a priori* case? When we contemplate the corresponding conditional of *modus ponens*, we just find ourselves with this powerful inclination to believe that this proposition is true, and indeed couldn't be false. But (as we also know) such inclinations are by no means infallible. We really don't have any reasons or grounds for this belief; we simply, so to say, *start* with it.

But then the question for BonJour is this: why are simple *a priori* beliefs exempt from this demand that we must have a good reason for thinking a belief true, if it is to constitute knowledge? He seems to me to face a dilemma at this point. If he treats *a priori* belief differently from perceptual belief, belief about other minds, inductive beliefs, and the like, then he's guilty of a sort of arbitrariness. As Reid says,

> Reason, says the skeptic, is the only judge of truth, and you ought to throw off every opinion and every belief that is not grounded on reason. Why, sir, should I believe the faculty of reason more than that of perception? They came both out of the same shop, and were made by the same artist; and if he puts one piece of false ware into my hands, what should hinder him from putting another?[44]

On the other hand, if BonJour insists that *all* beliefs, *a priori* beliefs as well as others, must meet this condition, then he insists on a condition for knowledge that cannot possibly be met, at least by finite beings like us, beings with only finitely many beliefs. Indeed, can it be met by God himself, essentially, yea, necessarily omniscient as he is? Would God have an answer to the question 'What is your reason for believing that 2 + 1 = 3?' I doubt it. He knows that 2 + 1 = 3 and that's that; he doesn't know it by way of some infinite chain of propositions, each his reason for the succeeding one. But however things stand with God, it is clear that you and I do not have a reason for believing such elementary truths of logic and arithmetic; and yet we do know them, and know that we know them.

So does my account of warrant entail skepticism? Only if we add this further condition BonJour insists on—and given that condition, *every* account of warrant will entail skepticism. For let the conditions of warrant be whatever you like: if we insist that knowledge of p requires knowledge that the conditions of warrant with respect to p be satisfied, we immediately start something we can't finish, embark upon a regress we can't complete. In other words, it isn't my account of warrant that guarantees skepticism: it is this condition BonJour adds.

Ad Foley

In an uncharacteristic display of intolerance, Richard Foley declares a pox on all our houses. Justification, reliability, proper function, causal connection between belief and its object, fit between cognitive powers and cognitive environment—none is either necessary or sufficient for warrant. What counts, rather, is *accuracy* and *comprehensiveness*: knowledge is sufficiently accurate, comprehensive belief. It doesn't matter how you acquire that splendid set of beliefs. It could be by the sheerest chance; perhaps, by lucky happenstance, it happens when your brain is scrambled by a stroke or an errant bolt of lightning. It doesn't matter whether your beliefs are justified in the deontological sense: you could be a real epistemic villain and nonetheless know an enormous amount. It doesn't matter whether you have evidence for any or all of these beliefs. It doesn't matter whether your cognitive faculties are reliable: they can be as unreliable as you please, but if by dumb luck they happen on some occasion to produce such a fine upstanding set of beliefs, you have knowledge.

Now I don't believe that accuracy and comprehensiveness by themselves are sufficient for knowledge. (Foley and I first had this conversation some ten years ago; apparently neither has succeeded in convincing the other.) Let me explain why I'm doubtful. Note first that Foley's suggestion flies in the face of one of the strongest intuitions (as it seems to me) that we have in this area: the intuition that if I know p, then it can't be just by accident that I form the true belief that p.[45] On Foley's account, this can be as accidental as you please. Second, knowledge requires a certain degree of stability. Suppose I am captured by Alpha Centaurian super scientists, or am a victim of a Cartesian evil demon: my tormentors run a nasty experiment in which my beliefs alternate between two-second periods where I have very accurate and comprehensive true belief, and two-second periods in which my beliefs are pitiably confused and mistaken. Do I have knowledge during the two-second periods when I have accurate and comprehensive true belief? I doubt it, just as I would my mechanic's claim that my van works just fine seven-eighths of the time (and who can sensibly ask for more?), since it is only one of the eight cylinders that refuses to fire. Foley tries to justify himself by claiming that "it is not impossible to have fleeting moments of insight" (p. 92); that's true, but it doesn't follow that just any amount of instability is compatible with knowledge.

There is another problem here, one of which Foley is doubtless

aware. Clearly the account needs an addition: that the belief system in question is accurate and comprehensive *in the neighborhood*, so to speak, of *p*, the belief that putatively constitutes knowledge. For suppose that I am magnificently knowledgeable, knowing as much science, even, as Carl Sagan. I drive through that phony barn country in Wisconsin, having no idea about what those sly Wisconsinites have been up to. Then couldn't my (true) belief that I see a barn (it stands in the middle of a clump of four fake barns) be an element of an accurate and comprehensive system of beliefs about barns (their history, characteristic employment, construction, etc.) but still fail to constitute knowledge? We'd have to construe 'accurate and comprehensive' in such a way that it somehow guarantees that I wouldn't or couldn't drive through phony barn country without believing that it was indeed phony barn country. Similarly for the van case (pp. 312 ff.) 'accurate and comprehensive' would have to be construed, somehow, in such a way that it precludes either my failing to learn that my van has been destroyed, or my failing to learn that I've won the Varsity Club win-a-van lottery. We'd also have to construe this condition in such a way that it eludes the original Gettier case. But how can we do that? Couldn't I have a very accurate and comprehensive system of beliefs that nonetheless contained a false belief *p*? Maybe I know an enormous amount about Napoleon, but am mistaken about the birthplace of his maternal grandfather, thinking he was born in X, a village three miles from his actual birthplace. If so, couldn't I disjoin some belief *q* with *p*, which was such that while it was true, I have no reason to think it is? Maybe I have no idea what his maternal grandfather's name was (that fact has been obscured by the mists of time); but I idly deduce *Napoleon was born in X or his maternal grandfather's name was Pierre* (and as it happens, his name was indeed Pierre) from my belief about his birthplace. We'd have to construe 'accurate and comprehensive' in such a way that it precludes this possibility too. In general, we'd have to construe this condition in such a way that its satisfaction doesn't permit an area of lack of resolution, or else doesn't permit my having true beliefs in that area. But how can we do this, short of insisting on omniscience or near omniscience in the vicinity of *p*?

This condition of accuracy and comprehensiveness, therefore, is going to have to be *very* strong. (When Foley is arguing that nothing beyond accuracy and completeness is necessary, he speaks of a set of beliefs much more accurate and comprehensive than yours or mine.) But won't that lead to an opposite problem? Can't I know something

even if I don't enjoy much by way of accurate and comprehensive belief? As my grandmother aged, she knew less and less. Eventually she didn't know much at all—but she still knew her name, that the man living with her was her husband, that she had four children, and that she didn't have a dog. My problem with Foley's account is that I don't see how he can sail between Scylla and Charybdis. I don't see how his account can simultaneously meet two conditions: first that 'accurate and comprehensive' be so construed that Gettier problems with *sufficiency* can't arise, and second, that it be so construed that I can know *p* even if I don't know (or believe) a great deal more.

Ad Klein

I've already had my say about much of Peter Klein's paper; here I want to add just one brief point. Klein proposes the Garden of Eden Case (pp. 109–110) to show that even if I am in fact designed and created by God, it could be that I should come to know something God didn't intend me to know: "My story shows that we can pull apart our having knowledge from the intentions of our maker—if there were one. Were our maker to intend that we remain ignorant and design us with that plan in mind, we could rebel and gain knowledge. In other words, I think the root notion of proper function is such that we *can* gain knowledge, even though doing so would not be to function properly" (p. 110).[46]

But here we need a distinction. God didn't intend that Cain kill Abel; when Cain did so, however, it wasn't that his limbs (or other systems or organs) were malfunctioning. It was rather that he was using the powers God gave him for a wicked purpose, a purpose contrary to what God intended. In fact, if his arms (e.g.) *had* malfunctioned at the crucial moment, *then* he would not have been able to kill Abel. We must distinguish proper function of system or organ (function in accord with design plan) from our using the power in question in a way God intends us to. The former is required for warrant, not the latter.

Ad Swain

Marshall Swain argues that "the theory that knowledge is ultimately undefeated justified true belief (the defeasibility theory, hereafter) is superior to Plantinga's version of the warranted true belief view . . . as an account of knowledge." He adds that the former is also superior to

the latter in that "the defeasibility theory does not commit us, or even threaten to commit us, to the idea that our faculties must function in accordance with some design plan, which is the feature of Plantinga's view of warrant that leads to supernaturalism" (p. 132). Swain has two main claims: first, that the defeasibility theory is superior to my account of warrant, just as an account of warrant (its proposal as to the necessary and sufficient conditions of warrant is closer to the truth than mine); and second, that it is also superior in that it doesn't have the supernaturalistic implications my theory is alleged to have. I have two brief comments.

First, Swain says he prefers the view according to which warrant is "ultimately undefeated justified true belief." But what does he mean by 'justified'? He doesn't tell us; but if we take it in either of the most common ways—either in terms of epistemic responsibility or in terms of evidence—it doesn't seem to be necessary (see pp. 317–18). Taken the second way, justification seems clearly unnecessary; as I argued in WPF (pp. 186 ff.) many beliefs constitute knowledge without having either propositional evidence or phenomenological evidence. Perhaps they must have what I called impulsional evidence, this feeling that the belief in question is indeed true (WPF pp. 190 ff.), but this is really an inevitable accompaniment of belief and isn't a further condition in addition to believing p.[47] On the other hand, I've already argued (p. 330–31) that there doesn't seem to be any epistemic duty the satisfaction of which is a necessary condition of knowledge; I can be as undutiful as you please and nevertheless know, e.g., that I am being appeared to redly or that $2 + 1 = 3$. (That is not to say that there are no specific kinds of knowledge such that having knowledge of that kind requires conformity to some duty or other.) As I see it, then, Swain's account doesn't succeed in providing a necessary condition of warrant (unless he has some quite unusual account of justification up his sleeve).

As to sufficiency, that would depend upon just how Swain proposes to gloss 'undefeated'. Peter Klein's proposed condition doesn't seem sufficient, and it's about as good as defeasibility accounts get. I therefore persist in believing that my account (with the amplified environmental condition) does better than any defeasibility account currently on offer.[48] Indeed, a defeasibility theory can't be successful without adding a clause involving proper function. For clearly there are beliefs that achieve warrant in the basic way; as far as I can see, however, any successful attempt to specify *how* they receive warrant in the basic way will require the notion of proper function. The only

plausible alternative would be to follow Klein and appeal to reliability; as we saw in thinking about Klein's proposal, however, appeal to reliability without proper function is bootless.

This leads directly to Swain's second contention; a defeasibility account of warrant, he says, is superior to mine in that it doesn't threaten to commit us to supernaturalism. But that is at best a dubious advantage. In WPF, I do indeed argue that there is no way to give a naturalistic account of the notion of proper function. More exactly, what I argue is that there is no way to give an account of proper function that doesn't involve the activity of conscious, intelligent agents; anything that functions properly or improperly is in this respect like an artifact. That means that on my account, if we human beings have knowledge, then we have been designed by one or more conscious intelligent agents. In my own view this agent would of course be God; but it is also at least abstractly possible that it be an agent or agents of some other sort.

But first, my claim that proper function can't be analyzed or explained naturalistically isn't really part of the account of warrant; it isn't itself an *epistemological* claim at all. Someone could share my epistemological views even if he added that he was agnostic as to whether there is a decent naturalistic analysis of proper function, or even if he went on to endorse some naturalistic analysis of it. And second, if I am right about proper function, it will be a lot more than warrant or knowledge that carries an apparent commitment to supernaturalism. The same will go, for example, for the notions of sanity and health. Medical science is full of descriptions of how organs or systems work when they function properly, as well as descriptions of what happens when they fail to function properly. A naturalist *already* has the problem of deciding what to do with this absolutely pervasive concept; he'll need some *general* way of accommodating or dealing with proper function. Perhaps he will deny my suggestion that it implies design by a conscious and intelligent agent; or perhaps he will take refuge in some form of fictionalism, suggesting that we can adopt the "functional stance" even if in reality there is no such thing as proper (or improper) function; or perhaps he will try something else. The point is only this: whatever problem he has with my account of warrant is a problem he *already* has; hence he need not reject this account because of that problem.

Ad van Fraassen

Bas van Fraassen's subtle picture of science is powerfully drawn, presented with irony, force, passion. To echo Jonathan Edwards in

another connection, "I could not but stand and admire"—as indeed with all van Fraassen's work. (As with some of his work, however, I also find my admiration deeper than my understanding.) Much in his paper cries out for comment. His topics differ considerably from everyone else's, however; a proper comment would require much by way of stage setting and considerable shifting of gears; furthermore I hope to write elsewhere on the topics he takes up. I shall therefore restrict myself to a brief comment on just one of the many hares he starts.

As arresting as any theme in the paper is the claim that materialism, that venerable and widely revered piece of metaphysics, really isn't a *claim* or a *view* at all, but instead an *attitude*:

> I propose the following diagnosis of materialism: it is not identifiable with a theory about what there is, but only with an attitude or cluster of attitudes. These attitudes include strong deference to science in matters of opinion about what there is, and the inclination to accept (approximative) completeness claims for science as actually constituted at any given time. Given this diagnosis, the apparent knowledge of what is and what is not material among newly hypothesized entities is mere appearance. The ability to adjust the content of the thesis that all is matter again and again is then explained instead by a *knowing-how* to retrench which derives from invariant attitudes. (p. 170)

Here van Fraassen harks back to an old empiricist or positivist theme: much of what you are initially inclined to take to be a claim about the world, a position, an opinion as to what there is, is really no such thing. (Carnap suggested that metaphysics is perhaps a form of *music*: *bad* music no doubt—nothing to rival Mozart or maybe even Madonna—but music nonetheless.) But while the positivists were crude, wholesale, disdainful, and dismissive in their assaults on metaphysics, theology, ethical views, and the like, van Fraassen is subtle, piecemeal, and non-dismissive—though a certain disdain remains. In claiming that materialism is an attitude rather than a view, he is only, he says, giving the straight story about materialism, not attacking or dismissing it: "This does not reflect badly on materialism; on the contrary, it gives materialism its due" (p. 170).[49] Indeed, van Fraassen regards his own *empiricism* as an attitude rather than a claim. But while there is no condemnation as such in calling materialism an attitude rather than a thesis, the materialist still has a problem: false consciousness. She often confuses "theses held with attitudes expressed" (p. 170). In declaring "matter is all," says van Fraassen, she

is really expressing an attitude rather than (as she confusedly sees herself) stating a view as to what there is.

This is a fascinating thesis, presenting a powerful challenge to *soi-disant* materialists. Of course questions arise about attitudes: what sort of things *are* they? That's not wholly clear, but at any rate they can include "strong deferences," e.g., to science in matters of opinion, and "inclinations to accept" certain factual claims; they can also include commitments—e.g., to form one's opinions in terms of contemporary science—as well as judgments of worth or value. The heart of van Fraassen's claim is that there is more to materialism than factual judgment: there is also this cluster of attitudes. In fact materialism, as he thinks of it, doesn't include, just as such, any factual judgment at all. At any given time it is, so to say, embodied in opinion of one sort or another ("Its incarnation at any moment will be some position distinguished by certain empirical consequences, and these will either stand or fall as science evolves," p. 170). But the opinions associated with materialism at a moment are not essential to it; one can be a materialist without accepting those opinions; at bottom materialism is a cluster of attitudes rather than a factual opinion or judgment. Materialism may include an "inclination to accept" certain factual judgments, but there is no factual judgment such that accepting that judgment is necessary to being a materialist. For first, the inclination may be to accept different opinions in response to different circumstances: if science says at t that matter is particulate, then at t the materialist will be inclined to believe that all is particulate or made of particles; but if at t^* science says matter is *not* particulate, then at t^* the proper materialist will be inclined to reject the previous opinion. And second, of course, the attitude includes a *tendency* to accept certain factual opinions; but (I should think) the tendency isn't necessarily realized within the breast of any given materialist.

This suggestion is at the least insouciant; the materialist may have stronger terms for it. But what leads van Fraassen to make the suggestion? That the materialist always knows how to retrench in the face of new scientific developments and hypotheses. Perhaps first she thinks that matter consists just in elementary particles and things made of them; but then science posits forces, which are not particles; her reaction, of course, is not to *give up* materialism, but instead *amend* it. As science marches (or sashays) along, positing its enormous variety of entities of ever-stranger sorts, she always knows how to retrench. And that is because, says van Fraassen, her materialism really consists, not in a view as to what there is, but in a cluster of

attitudes including a commitment to endorse whatever science comes up with.

This is indeed a striking feature of materialism. But here I find in myself a wholly unfamiliar inclination to defend the materialist, or at any rate her self-understanding; this knowing how to retrench is compatible with a view of materialism that departs less from her own idea as to what she is up to. It is indeed hard to say precisely what matter is; but perhaps this is due, not to the term's having no cognitive content at all, but to more than usual vagueness. Perhaps the materialist, at any given time, says something like this: all is matter, and matter is what current science says there is, *together with anything sufficiently similar* to what current science says there is. What constitutes *sufficient* similarity is of course unstated; that is the source of the vagueness. The crucial respects of similarity are also unstated; perhaps there is in fact no way to make them wholly explicit, and perhaps they change with changes in science. But materialism could then still be an opinion, a claim, even if a bit nebulous and inspecific, a little like the claim that the truth is whatever is sufficiently similar to what the pope or the Synod of the Christian Reformed Church says. So taken, materialism would still have the reflexive character van Fraassen mentions: what is sufficiently similar to what science *now* says there is may not be sufficiently similar to what it will say one hundred years from now—more likely, what *isn't* sufficiently similar to what it now says there is may be sufficiently similar to what it will then say there is. (So materialism may be less a specific opinion than an evolving group of opinions.) But (provided science doesn't change too rapidly) at any given time the materialist may know how to retrench: perhaps current science says there are things of kind k but none of kind k^*; the materialist then denies that there are things of kind k^*; but then science moves on and says there are indeed things of kind k^*; provided k^* isn't too different from k, the materialist can easily retrench. She thought there were no things of that kind, but now thinks there are—while remaining a materialist.

Materialism thus thought of has or may have "cognitive content"; it needn't be so vague as not to eliminate *anything*. Perhaps there are no explicitly stateable and reasonably precise necessary and sufficient conditions for something's being material, but there may be necessary conditions. If materialism is true, for example, there won't be any such person as the God of traditional theism: no all-knowing, everlasting, all-powerful, and wholly good being who has created the world. (If the materialist could retrench as far as all *that*, then materialism

would indeed be contentless.) Still further, perhaps (but this is only a guess) an entity could be a material entity only if it had a spatial location of some sort. Perhaps it need not be any distance from *us*; but perhaps it will have to be at some distance from *something*, if only itself. This condition isn't wholly vacuous, at least for those of us who adopt more or less traditional views as to the nature of numbers, propositions, states of affairs, possible worlds, and the like; these things exist but have no spatial location. Furthermore *God* isn't at any distance from anything, not even himself. I'm not sure the materialist ought to prefer this diagnosis of his condition to van Fraassen's but it is at any rate an additional option.

Now van Fraassen adopts the same line with respect to *naturalism*. ("Yet I venture to assert: we see here too a position that only purports to be a factual thesis," p. 172). It too, he suggests, is an attitude or a congeries of attitudes—perhaps overlapping or even identical with materialism ("Perhaps it is just the materialism I have been discussing here, under a different name," p. 172). And this means that the evolutionary argument I gave against naturalism in chapter 12 of WPF can't possibly be successful: that argument is stated as turning on the probability that our cognitive faculties are reliable, given a certain proposition that includes naturalism as a conjunct. But if naturalism is an *attitude* rather than an opinion, then, of course, there *is* no proposition that includes naturalism as a conjunct. The best my argument could do would be to refute "some temporary tradeable asset of the position" (p. 172), leaving the central core, the constituting attitude, untouched.

Well, perhaps a modest philosopher ought to be content with refuting the current assets of the position, even if those assets are tradeable; its subsequent assets can be left to the attention of subsequent modest philosophers. More fundamentally, however, I propose the same claim here as with materialism (which won't be hard if indeed naturalism, as van Fraassen suggests or at any rate entertains, is indeed the very same thing as materialism). Perhaps naturalism too is a view, all right, but also an unusually vague view. Perhaps it is the view that there is no such person as God, or anything sufficiently similar to God. As van Fraassen suggests, it is exceedingly hard to see precisely what naturalism—or worse, naturalistic epistemology—is supposed to be (see pp. 172 ff.). But surely you can't consistently be both a naturalist and a traditional theist, i.e., someone who believes there is an all-powerful, all-knowing, wholly good person who has created the heavens and the earth. Perhaps it is also not possible to be (consistently)

both a naturalist and, say, an absolute idealist in the style of Hegel or Bradley (as the saying goes, absolute Idealism is really a continuation of Protestantism by other means), or a believer in Plato's Idea of the Good, or a Neo-Platonic emanationist, or a believer in the Stoic's *Nous*, or in Spinoza's *Deus sive Natura*, or Tillich's Ground of Being, or angels. Some of these may not be coherently characterizable; assuming that they are, however, some seem to fall clearly outside what a naturalist can admit, and others (Spinoza, Tillich) seem to fall into that area of vagueness. If this is right, then naturalism is a view after all, even if only a hazy view; naturalism is a saying, even if a dark saying. If so, my argument against naturalism doesn't fail for this reason, at any rate, whatever its other blemishes.

Ad Conee

Earl Conee is scandalized by my claim that proper functionalism, with what he sees as its ultimate reference to God, is a brand of *naturalistic* epistemology.[50]

> Thus Plantinga's theory of warrant—"proper functionalism" for short—invokes God, the epitome of a supernatural entity. Amazingly, Plantinga nonetheless regards his theory of warrant as a naturalistic proposal. He is quite open about this. He explicitly advocates a supernatural ontology for a naturalistic epistemology. (p. 183)

He goes on to make an impassioned argument that my epistemological views are about as unnaturalistic as you can get. But here what we really face is an illustration of van Fraassen's point about the elusiveness of naturalism, or more exactly, of the project of naturalizing epistemology. Conee himself seems to recognize this:

> Perhaps in the end Plantinga does not intend to be taken completely seriously in arguing in the ways that we have reviewed that his theory of warrant is a naturalistic theory. After counting his theory as a radical naturalism, he makes the following lighthearted comment: "[S]triking the naturalistic pose is all the rage these days, and it's a great pleasure to be able to join the fun." (p. 195)

"I take no pleasure in spoiling anyone's fun," he says, but then he goes on to scold me for undue frivolity:

> But it would be too costly to allow this recreational endeavor to go unchallenged. Stretching our interpretation to accommodate proper func-

> tionalism in the category of "naturalistic epistemology" would so distort the meaning of the phrase as to ruin it for those who are earnestly trying to mean something natural by it. (p. 195)

Earnestly trying, no doubt, but so far without much success, which was really my point in claiming that my account of warrant is naturalistic. The advocates of naturalism and more particularly naturalistic epistemology clearly suppose the latter a very good thing; beyond that things are wide open. To illustrate this fact, I argued that my account of warrant, despite its supernaturalistic setting, is really naturalistic according to at least some current accounts. Jaegwon Kim and Hilary Kornblith, for example, see the essence of naturalistic epistemology as a matter of rejecting normativity in epistemology. Kornblith asks us to consider the questions

> (1) How ought we to arrive at our beliefs?

and

> (2) How do we arrive at our beliefs?

and then characterizes naturalism in epistemology as follows:

> I take the naturalistic approach in epistemology to consist in this: question 1 cannot be answered independently of question 2. Questions about how we actually arrive at our beliefs are thus relevant to questions about how we ought to arrive at our beliefs. Descriptive questions about belief acquisition have an important bearing on normative questions about belief acquisition.[51]

Well, this leaves quite a lot of leeway. What is meant by the first question—what kind of normativity is it that is at issue, what is the force of the 'ought'? Is Kornblith thinking of duty, obligation, requirement, deontology? What he says elsewhere makes it clear that he isn't. There is a use of 'ought', however, that might be relevant: its use when we say such things as that your blood pressure ought to be fairly close to 120/80, that when you press the starter button, the engine ought to turn over, and that a 60-year-old man ought to be able to run a mile in 10 minutes. This is the 'ought' of proper function. And of course (as I argued in WPF, p. 46) how we ought to arrive at our beliefs in this sense of 'ought' is indeed relevant to how we actually arrive at our beliefs, at least if we construe the latter as how we arrive

at them when there is no cognitive malfunction; indeed, so taken, questions (1) and (2) are identical. Furthermore, it seems likely (though he doesn't say) that Kornblith *is* construing 'actually arrive at' as 'actually arrive at when there is no dysfunction'; how we arrive at our beliefs when there is cognitive malfunction, or madness, or insanity, I take it, will not be of much relevance to the question how we *ought* to arrive at them, no matter how we construe the latter. So on this way of understanding naturalistic epistemology, my account of warrant qualifies. Conee demurs (p. 187), but his demurral depends, as far as I can see, upon confusing the normativity that goes with warrant with warrant itself. It is indeed true, as he claims, that the latter involves more than proper function; it doesn't follow, of course, that the same goes for the former. I therefore persist in thinking that my account of warrant passes Kornblith's test for being naturalistic. If this be heresy, I invite Kornblith to shoulder some of the blame.

I distinguished three grades of normative involvement, suggesting that there are three corresponding grades of naturalism in epistemology: the first renounces deontology in epistemology; the second is the one outlined by Kornblith; and the third, inspired by Quine, eschews any kind of normativity not invoked in science. I've just argued that my account is a naturalism of the second grade; it is obviously one of the first grade as well. But it is also a naturalism of the third grade, in that medical science, psychology, biology, and other sciences employ, explicitly or implicitly, the notion of proper function, and thus employ the only sort of normativity involved in my account of warrant. True, the teleology involved in proper function is a source of a certain embarrassment for those who think science shouldn't or doesn't involve any such thing, perhaps taking refuge in the pious if unfounded[52] hope that teleology can be reduced, somehow, to notions more naturalistically antiseptic. Whether science should or shouldn't involve teleology, however, it does.[53]

Again, Conee demurs, claiming that what I see as the notion of proper function in the sciences I mention is really no such thing. He points out first that this notion isn't involved in physics and chemistry, apparently inferring (p. 191) that this means it isn't really involved in other sciences (such as medicine, biology, and psychology) either. I fail to follow the inference. *Pliotropy* and *gross national product* aren't to be found in physics and chemistry; does it follow that they aren't to be found in biology or economics either? (Or that the latter aren't sciences?) He next suggests that what is really involved in science

is something quite different from the proper function I profess to find there:

> . . . accurate psychological generalizations say no more than what must happen in sentient organisms when the distinct kind ["level," "scale"] of causal process that is governed by psychological forces is operative. . . . The intended qualification of psychological laws just sets aside other potential influences, nonpsychological factors that must be at least controllable in their impact if theoretical psychology is to be practicable. (p. 191)

I'm a bit in the dark about this claim: what, for example, are "psychological forces"? Be that as it may (and no doubt it will), this seems to me to be inaccurate. What is to be ruled out, in the functional generalizations in question, is more than just potential influences from outside the 'level' under consideration. A hungry cat confronted by a plump mouse will give chase—unless she is suffering from congestive heart failure, or is paralyzed, or . . . —or is *insane*. Psychological malfunction need not involve invasion from some other level.

Finally, Conee's own characterization of naturalism in epistemology (a little softly focused, he admits) seems to me to let in my account of warrant. "Using broad and blurry terms," he says, "the theme can be said to be that of locating the epistemic within nature, and finding epistemic topics to be suitable for investigation by scientific methods" (pp. 183–84). Once more, I'm not sure how to take this: what it is to locate the epistemic within nature? Part of epistemology, presumably, would be to say what knowledge or warrant is: but what would it be for such an account to locate them within nature? Further, why can't my account of knowledge meet these conditions? On my view, human beings are to be found, naturally enough, within nature: is this sufficient for locating the epistemic within nature? I should also think, furthermore, that at least some epistemic topics are "suitable for investigation by scientific methods." One such epistemic topic would involve description of how human cognitive faculties function when they function properly—or improperly. Under the latter heading would come such topics as the cognitive disturbances involved in schizophrenia, the way in which cognitive faculties are impaired by dementia praecox, and many more; these topics not only can be but have been investigated scientifically.

But I am afraid this will not satisfy Conee. What he really means, I think, is given in the following:

> He [i.e., I] begins to support it by making the stunning assertion that naturalistic epistemology is "quite compatible with supernaturalistic theism" (p. 46). This seems quite doubtful. If naturalistic epistemology is what it appears to be—an application of a naturalistic world view and methodology to epistemic topics—then it is incompatible with any sort of supernaturalism. (p. 184)

So it looks as if a naturalistic epistemology, by Conee's lights, must as such entail, e.g., that there is no such person as God, whom Conee rightly sees as paradigmatically supernatural. (It might have to entail more, at any rate if atheism and naturalism do not coincide; perhaps it would have to entail that Bradley's metaphysics is false (if coherent) and that there is no such thing as Plato's idea of the Good.) Alvin Goldman's reliabilism, then, would not be a naturalistic epistemology, since it is clearly compatible with the existence of God (assuming the latter self-consistent). The same would go for Fred Dretske's epistemology. The same would also go for Conee's own evidentialism: it too is clearly compatible with the existence of God. (Not that Conee claims that his own epistemology *is* naturalistic—so far as I know, he doesn't.) Quine's suggestions are ordinarily taken as definitive or at least paradigmatic of naturalistic epistemology; but clearly a theist could consistently share Quine's opinion that there is nothing to epistemology except empirical psychology, at least if neither theism nor this opinion is itself necessarily false. So by this standard, no one we know, not even Quine, the old master himself, is a naturalistic epistemologist. Indeed, is naturalistic epistemology thus conceived even possible? How could you have an epistemology that entails the nonexistence of God? Of course you could say: "Here is my epistemology: first there is no such person as God, second, . . ."; but would the first conjunct really be *epistemology* as opposed to an irrelevant theological addition? Could you have an evolutionary epistemology just by proclaiming, as its first tenet, that Darwinism is indeed true? (If so, could you also have, e.g., a *vegetarian* epistemology?) This standard for naturalistic epistemology seems a bit high. I don't doubt there can be an epistemology that fits naturally (to use some more softly focused terminology) with philosophical naturalism or atheism, but the relation between it and naturalism will be subtler than entailment.

A more moderate suggestion is embodied in the passage I quoted at the beginning of this section. Perhaps what Conee really means is something like this: "Whatever exactly naturalistic epistemology is

(and it may be hard to say just what), at any rate it can't entail the existence of God." That seems initially sensible. Of course if traditional theism is true, God is a necessary being, so that every proposition entails the existence of God; but perhaps we can bracket that difficulty, or at any rate ignore it and hope for the best. But *does* my account entail (in the relevant sense, whatever precisely that is) the existence of God? I don't think so. The account does, of course, invoke the notion of proper function; and as a matter of fact I also argue (see WPF chap. 11) that there is no correct naturalistic analysis of proper function. But the latter claim, as I said in response to Swain, is not really part of the epistemology. Someone else could have the same epistemological views as mine even if he thought there *is* an accurate naturalistic analysis of proper function, or was agnostic about whether there is. I also believe that the existence of matter entails the existence of God (and not just by way of the latter's being a necessary being); but I wouldn't think of this claim as part of physics. Suppose I agree with Quine: the content of epistemology is really empirical psychology; but suppose I also believe, e.g., that there couldn't be any living beings unless there were such a person as God: would that mean that my epistemology wasn't naturalistic after all?

There is much more to be said here, and the topic, despite the lightheartedness Conee notes at the end of his paper, seems to me really important. It is of a piece with the question whether there can be such a thing as Christian or theistic philosophy, science, literary criticism, art, and the like. Newton thought God had created the world and instituted the scientific laws and regularities; he also thought gravity was, or was a manifestation of, divine activity; does that make Newtonian physics theistic? If not, what more would be required? Suppose you accept one of the inflationary cosmologies presently on offer because you think they avoid singularities that suggest divine creation, and you think no such suggestions should emanate from science: does that make your physics naturalistic? Van Fraassen addresses some of these questions in his paper. These are excellent questions, but will have to await another occasion.

Ad Feldman

Richard Feldman's penetrating comment makes three claims: (1) that I erred in thinking those whom I called internalists are internalists with respect to *warrant* (they are rather internalists with respect to *justification*, taken one way or another), (2) that my objections to

internalistic views are wanting, and (3) that my own view is deficient in that it doesn't properly handle certain Gettier cases. He's right on (3); the deficiency, I hope, is remedied by the amplification of the environmental clause in Part I.

As to (1), I took internalists to be saying that the "basic shape" of warrant is given by an internalist notion (ordinarily justification of one or another variety), together with what I called a codicil or fillip to appease Gettier. Feldman demurs, claiming that internalists don't think of the additional condition needed to evade Gettier as a fillip, or a codicil, or a smallish addition, or anything else of the sort; they don't think of any internalist notion as nearly sufficient for warrant; and while they are internalists about justification, they are not internalists about the fourth condition. Well, as I said above, there is nothing to be gained by arguing about how close to sufficiency internalists think justification is; we must give them the last word about that. So let's agree that internalists think justification is necessary for warrant, but are not prepared to claim that it is anywhere nearly sufficient for it. (Of course, insofar as they are internalists about justification and also think it necessary for warrant, they are internalists about warrant in the sense that they are internalists about (what they see as) a necessary condition of warrant.)

But *is* justification necessary for warrant? The real question, as it seems to me, is whether internalists really succeed in stating even a *third* condition of knowledge, a condition that is both necessary for warrant or knowledge and independent of belief. I argued this above with respect to justification taken deontologically, as a matter of doxastic or epistemic responsibility. But the same goes, I think, for the other main branch of the justification tradition: the idea that justification is a matter of forming and sustaining belief only on the basis of evidence. I should like to illustrate this problem for internalists with reference to Feldman's own internalist view, *evidentialism*.[54] This is an important view, and one with a long and distinguished history. On Feldman's version, a belief is justified for *S* if and only if it fits *S*'s evidence. I criticized evidentialism in WCD; Feldman responded in "Proper Functionalism," *Nous* XXVII, 1 (March 1993); my riposte, involving a further criticism of evidentialism, was in "Why We Need Proper Function" (*loc. cit.*); and here we have Feldman's reply to *that*. Feldman's present response, however, involves misunderstanding, no doubt due as much to expository ineptness on my part as hermeneutical inadequacy on his. Let me try again.

My criticism went as follows. First, concede that the evidentialist

internalist is making only the limited claim that justification is necessary for warrant and that *S*'s belief *B* is justified just if it fits *S*'s evidence. To evaluate this claim, of course, we must know what is to count as evidence. My claim was that the evidentialist faces a dilemma (or possibly a trilemma); no matter how he construes *evidence*, he winds up in hot water. More specifically, either he winds up with a condition that isn't necessary, or he winds up with one that isn't independent of belief (and thus isn't even a third condition). We can see this as follows. What are the possibilities for construing this notion of evidence? First, there is *propositional* evidence: my propositional evidence for one of my beliefs is one or more other beliefs on the evidential basis of which I hold the belief in question.[55] Although most twentieth-century discussions have tended to limit evidence to propositional evidence, the Feldmanian evidentialist can't sensibly do so. The reason is familiar: we know (for example) simple arithmetical truths as well as certain propositions about our own mental life, neither of which are accepted on the evidential basis of other propositions; propositional evidence is not necessary for knowledge. Second, then, there is *sensuous* evidence, for example, sensuous imagery of the sort involved in visual (and other sorts of) perception. What we have in these cases is 'the evidence of the senses'—the phenomenological sensuous evidence furnished by way of being appeared to in those characteristic ways involved in perception. Can the evidentialist rest here, i.e., can he sensibly hold that evidence is exhausted in propositional and sensuous evidence? I don't think so; fitting the evidence thus construed isn't necessary for warrant. Consider my knowledge that 2 + 1 = 3 or any other obvious mathematical or logical truth. I don't believe this proposition on the evidential basis of any other proposition, but I also don't believe it on the basis of sensuous imagery.[56] So it doesn't fit my evidence;[57] but it is obviously warranted for me. If evidence is construed in either of these two ways, or as their disjunction, justification (fitting the evidence) isn't necessary for warrant or knowledge.

There is one further possibility: what in WPF and "Why We Need Proper Function" I called "impulsional" evidence. That sensuous imagery I mentioned isn't the only kind of phenomenology that can be thought of as evidence. Consider again the belief that 2 + 1 = 3; in addition to the sensuous imagery there is also something like a certain *felt attractiveness* of the belief; it *feels* right, somehow, and other beliefs you might consider in its place (*2 + 1 =* **5**?) feel wrong, weird, absurd, eminently rejectable. *2 + 1 = 3* has about it a sense of

rightness, or fittingness, or appropriateness. Perhaps the thing to say is that there is a sort of felt inclination, or impulse, to accept this proposition as opposed to others; indeed, perhaps impulsional evidence is no more than the phenomenal reflection of the fact that you do in fact believe the proposition in question. In any event, this kind of phenomenology is involved in *a priori* belief, but in all other sorts of belief as well. In the *a priori* case, then, there are really two quite different sorts of experience or phenomenology: the fleeting, fragmentary, indistinct, unstable, sometimes random sensuous imagery, on the one hand, and on the other, the felt inclination or impulse to believe *that* proposition, as opposed to others that might suggest themselves. Many other kinds of beliefs involve both these kinds of phenomenal experience and, as far as I can see, all beliefs involve impulsional evidence.

Now return to the Feldmanian evidentialist, who holds that a necessary condition of *p*'s having warrant for *S* is that *p* fits *S*'s evidence. As we saw, if he means by 'evidence' just propositional evidence, or propositional evidence together with sensuous evidence, then he is mistaken: a belief can be warranted without enjoying the benefit of either of those kinds of evidence. But perhaps he can construe 'evidence' very broadly, including under that rubric this felt impulse to believe, this sense of the proposition's being true or right. If he takes 'evidence' in *that* way, then, I think, he's right in claiming that evidence is a necessary condition of warrant (and knowledge). But then he has *another* problem. For the evidentialist, clearly enough, means to assert that justification is a necessary condition of warrant, and a necessary condition *in addition* to the truth and belief conditions. But taken this way the evidential condition adds nothing to the belief condition. You have impulsional evidence for *p* just by virtue of *believing p*; *p* fits *this* kind of evidence just if you believe it. It isn't even possible that you believe *p* but lack impulsional evidence for it: how could it be that you believe *p* although *p* doesn't seem to you to be true?[58] If so, however, what the evidentialist really needs is not a *fourth* condition, but a *third*. For evidential justification (construed thus broadly) guarantees only belief and a necessary accompaniment of belief: the nonsensuous phenomenal counterpart of the belief, the sense that the proposition believed is indeed true. So the proposition that *S*'s belief that *p* fits *S*'s evidence, on this construal of 'evidence', is satisfied by *S*'s merely believing *p*.

My point, then, was (and is) that the evidentialist faces a dilemma. If he takes the notion of evidence *narrowly*, so that it includes proposi-

tional evidence together with sensuous phenomenal evidence, but does not include impulsional evidence, then he is mistaken in claiming that justification is necessary for warrant (knowledge): in many cases of warranted belief—many cases of memory and *a priori* belief, for example—the belief need fit neither phenomenal nor propositional evidence. On the other hand, if he takes 'evidence' broadly (so that it includes impulsional evidence) then justification is indeed necessary for warrant—but only because it is necessary for belief itself. If so, however, the evidentialist isn't proposing a further condition of warrant independent of truth and belief; what he proposes is only a necessary condition of belief. Given that the evidentialist intends one or the other of these construals of evidence (and he certainly hasn't suggested any others), his evidential condition is either not necessary or not independent of the belief condition. In either case he fails to state a further necessary condition of warrant; so the fact is he hasn't so far stated a necessary condition of knowledge independent of the belief condition.

I hope this clarifies the point.

Ad Markie

Peter Markie concludes (1) that my account of degrees is not well developed, (2) that a certain conditional which on my account would have to be necessarily true is in fact contingent if true at all, and (3) that my account, as it stands, has no place for epistemically inappropriate degrees of warranted belief. As to (1), he is of course right, and I thank him for helping me think about the matter to a bit better effect.

Things are less clear with respect to (2) and (3), however. Begin with (3):

> Plantinga's theory does not allow for cases of warranted belief in which the degree of belief is epistemically inappropriate. Consider first the cases in which our confidence is too low, those in which we are warranted in believing p but believe p with less intensity than is called for by its degree of warrant. According to Plantinga's theory, since our degree of belief is inappropriate, our belief is not the result of a properly functioning cognitive faculty aimed at the truth, in which case our belief is unwarranted and our degree of belief is not epistemically insufficient after all. . . . Now consider the cases in which our confidence is inappropriately high, those in which we are warranted in believing p but believe p with more intensity than is called for by its degree of warrant. Once again, according to Plantinga's theory, since our level of belief is inappropriate,

> our belief is . . . unwarranted and our degree of belief is excessive, not because it is more than the belief's positive degree of warrant deserves, but because the belief has no warrant at all. Plantinga's theory thus precludes the existence of warranted beliefs in which the belief level is too high. (p. 231)

As it stands, this criticism is question begging. I proposed that degree of warrant (given the satisfaction of the conditions of warrant) is proportional to degree of belief (or is at least a monotonic, increasing function of the latter); but this quotation from Markie presupposes that degree of warrant is *not* such a function of degree of belief. So suppose we restate the objection: the claim is that there are some cases where someone's belief is warranted for her, but she holds the belief in question with too little firmness, and others where her belief is warranted for her, but she holds it with too *much* firmness. An example of the first would be the sort of case in which I have a great deal of evidence for some proposition *p*—that my mother's name is 'Lettie' or that it is likely that I will die within the next couple of years—but for one reason or another do not believe *p* with the firmness it deserves. I do indeed *believe p* (I don't merely think it probable) but I don't believe it firmly. In that case, says Markie, the fact is the belief obviously has a great deal of warrant for me. This is a problem for me, in that the firmness of the belief diverges from that dictated by the design plan: hence there is cognitive malfunction, in which case on my view the belief has little or no warrant.

But *does* this belief have much by way of warrant for me? If I don't believe it firmly, then surely I *don't know* it, no matter how much evidence I have for it. That seems to me to be a pretty obvious feature of our concept of knowledge. Someone might say that a diffident student really knows the answers to the exam questions, even if (due to lack of self-confidence) he is quite unconfident of them. But this (I think) is an analogical extension of the central concept: he *should* know (he has evidence sufficient for knowing) and he *would* know if it weren't for that lack of self-confidence; but the fact is (in the literal and central use) he *doesn't* know; he isn't convinced of the answer, or he isn't firmly enough convinced. So I doubt that there is a real problem here.

The other sort of case—the case where someone has enough evidence for knowledge, but believes too firmly—worries me more. To think about these cases to good effect, we must think about the degrees of belief compatible with knowledge. First, to know one must of

course believe *simpliciter*; this is distinct from being willing to bestow some degree of probability—even a probability of 1—on the proposition in question (see WPF p. 8). Second, there are degrees of belief, and also degrees of belief compatible with knowledge; that is, there are propositions p and q such that I know both, but believe p more firmly than q. So suppose I know (and hence believe) both that I went to the beach yesterday (I clearly remember doing so) and that $2 + 1 = 3$. The design plan dictates (so I say) that someone in my position should believe these two to different degrees: the second should be believed with maximum firmness, but the first to some degree a bit short of that. But suppose (contrary to the design plan) I believe both first and second to the max. Then don't I still know the first, even though I believe it to a degree different from that specified by the design plan, so that my faculties are malfunctioning (to at least a moderate degree) in producing in me a belief as firm as all that? Well, it is clear, first, that this belief doesn't have as much warrant for me as *2+1 = 3*, even though I believe it as firmly as the latter. But doesn't it have *some* warrant, and indeed enough warrant for knowledge? You and my other friends will indulgently concede that I do know that I went to the beach yesterday, even if my belief on that head is a little too enthusiastic.

The answer, I think, is that while there is a *bit* of malfunction here, it isn't significant enough to destroy warrant.[59] For my faculties to be functioning properly, in the relevant sense, it isn't necessary that they be functioning *perfectly* (see WPF pp. 10–11). I can know a lot by virtue of vision, even if my vision isn't 20–20; I can know much by way of memory even if (as with other people my age) it doesn't work quite up to the specifications of the design plan. I'd say the same here: warrant is determined by the degree of belief specified in the design plan, given that the conditions of warrant are met. But the proper function condition can be met even if it isn't met perfectly, and I think that is what is going on in this case.

Finally, and perhaps most worrisome, Markie argues that what he calls the Degrees of Warrant Principle,

> (DW) Under the conditions of warrant, degree of warrant is an increasing function of degree of belief,

contrary to what my view requires, is a contingent truth if true at all. He suggests that the crucial premise in the most plausible argument I

could give for the necessity of (DW) would be the necessity of his (3) (p. 227):

> (3) The degrees of belief dictated by the design plan for any faculty, insofar as the design plan is aimed at producing true beliefs, are proportional to the degrees of warrant the beliefs formed by the faculty will have if warranted. (p. 227)

Couldn't we (as Markie argues) have been designed by an incompetent designer,[60] one who assigns the wrong degrees of belief to a given belief in a given situation? (3) does seem contingent, at least within the parameters of this inquiry.

But my account does not require that (3) *be* necessary. According to (DW), *under the conditions of warrant*, degree of warrant is a monotonically increasing function of degree of belief. (That is compatible with different degrees of belief being mapped to the same degree of warrant and indeed, strictly speaking, with all degrees of belief being mapped to the same degree of warrant; it is not compatible, of course, with the same degree of belief being mapped to different degrees of warrant.) But among the conditions of warrant is that of the design plan's being a *good* one; so what I am committed to is the claim that if the design plan in question is a *good* one, then it will mandate that degree of warrant be an increasing function of degree of belief.

Of course that doesn't automatically get me out of trouble; it all depends upon what goodness for a design plan requires. I explained goodness in terms of reliability: if a belief has warrant, then "the module of the design plan governing its production must be such that it is objectively highly probable that a belief produced by cognitive faculties functioning properly according to that module (in a congenial environment) will be true or verisimilitudinous" (WPF p. 18). I was thinking of this reliability condition as an *explication* of goodness for a design plan; but Markie's points incline me to think that it should instead be thought of as just one of the conditions entailed by goodness for a design plan. Perhaps another condition is that the design plan be such that degree of warrant be a monotonically increasing function of degree of belief. This condition is (trivially) satisfied by such a being as God, all of whose beliefs (presumably) are held to the same degree and enjoy the same (maximal) degree of warrant; it is also met by any other knowers all of whose beliefs have the same degree of warrant. But for creatures like us, creatures with beliefs of differing degrees of warrant, it requires differing degrees of belief, and requires that where

belief *A* has more warrant than *B*, then *A* will be believed more firmly than *B*. It isn't hard to see how this would go. Indeed, perhaps degree of warrant is correlated both with degree of firmness of belief and also with degree of reliability. For example, consider *a priori* belief. This faculty or power produces belief of maximal firmness with respect to simple logical and arithmetical truths. These beliefs are also the ones with maximal warrant; and no doubt they are also the ones (so we think, anyway) where the faculty in question is most reliable, most likely to produce true belief.

Ad Pappas

"Ad Pappas" is a misnomer: since all of what George Pappas says seems eminently sensible, I have nothing to say in reply. I have perhaps one question. Pappas suggests that people who have learned how to reason in accord with a rule like *modus ponens* do not now but at one time had to reason by way of paying explicit attention to the rule:

> . . . they have reached a stage at which they achieve new knowledge in this domain—for example, that this sequence admits of modus ponens detachment—without the need to rely on rules which describe permissible moves when formulas have specific structures. No doubt when they first learned elementary logic they needed to fall back on the rule, perhaps recalling to mind what the rule was and then carefully noting that a given sequence falls under the rule. But now. . . . (p. 245)

He goes on to compare such a reasoner with a person whose job it is to classify buttons according to a rule: at first she must constantly consult the rule, but later on the rule becomes internalized and she follows it automatically.

Perhaps this is right if we are thinking of exercises in formal logic. There we have sentential or propositional variables (or schematic letters) and formalized patterns of "inference"; and we have to check carefully to see whether a pair of sentences do in fact meet the conditions for being the premises of *modus ponens*. But if we are thinking of actual reasoning (and perhaps this is not what Pappas is thinking of), the sort we all do unselfconsciously in ordinary life and more self-consciously and explicitly in philosophy, then, so it seems to me, this description wouldn't be accurate. I doubt that there is ever a stage in our learning to reason in which we explicitly reason by way

of applying the rule *modus ponens*. Rather, even as very young children we see, with respect to a specific conditional proposition, that it must be the case that if its antecedent is true, then so is its consequent; if we also believe or learn the antecedent, we automatically draw the conclusion. Following Aristotle, I conjecture we first see these relations in specific cases and then later on derive or (more likely) hear the rule from others. But of course Pappas may not have intended anything to the contrary.

Ad Sosa

Ernest Sosa is widely and rightly known as a fine epistemologist, but also as a person of generous spirit. This generosity is made manifest in his writing *two* essays for this volume! I have a brief comment on each. In the first, "Plantinga on Epistemic Internalism," Sosa reflects on my suggestion that we can understand current internalism about justification (and warrant) if we note that justification is conceived *deontologically*, in terms of epistemic duty, obligation, requirement.[61] My idea was that there is a certain internal connection (or congeries of connections) between internalism and justificationism taken deontologically.

Now Sosa distinguishes three grades of justificatory involvement and suggests that the one relevant to my discussion ("Let us focus on epistemic justification of the third degree—the sort of epistemic justification used by Plantinga to defend deontologism as a route to epistemic internalism," p. 81) goes like this:

> (Sosa Justification) S is *justified*$_3$ (justified to the third degree) in ø'ing iff in ø'ing S abides by an objective duty through a knowledgeable choice or at least through a choice based on a correct belief as to what one ought then to do. . . . (p. 77)

Sosa then argues that deontologism falls into an infinite regress here (p. 8): you are justified in believing a proposition P_1 only if you justifiably believe

> P2 Believing P1 is right;

but of course you are justified in believing P2 only if you justifiably believe

> P3 Believing P2 is right;

and so on. So you would be justified in believing any proposition P1 only if you believe infinitely many propositions, a feat most of us can't manage.

But is the requisite sense of justification really as demanding as (Sosa Justification)? The analogy, of course, is with justified actions, permissible actions, actions with respect to which I am within my rights. But can't I be justified, in taking such an action, even if I do not explicitly hold the belief that the action in question *is* right? I am within my rights in crossing my right leg over my left; no doubt I was within my rights on every occasion when I performed that maneuver; but I doubt that I ever formed the belief that this was indeed permissible. The question never really came up. Is it really required, for me to be within my rights, that I raise the question whether this *is* the right thing to do? I'm inclined to think not. So I'm inclined to think the regress never begins: what objective justification requires is that the action in question be in fact in accord with the right rule; what subjective justification requires is that I be nonculpable in not believing that the action in question is wrong; neither requires that I have formed the belief that it is in fact right.

But if the regress *does* begin, it ends almost as quickly as it begins. For suppose I am wrong: suppose it *is* required, somehow, that I believe, perhaps in some implicit and *sub tabula* way, that the action in question is justified. And suppose furthermore I do indeed believe that it is. I don't, it seems to me, have to raise the *further* question whether I am (deontologically) justified in believing that the action is justified, in order to be within my rights in believing that it *is* justified. For in many or most cases it simply isn't up to me whether or not I believe this. It's as obvious as $7 + 5 = 12$ that there is nothing wrong with drawing a breath, or crossing my right leg over my left (rather than *vice versa*). I am certainly within my rights in believing the former, if only because I can't believe anything else; the same goes, I think, for the latter. But then I will be within my rights in believing that I may cross my right leg over my left, no matter what else I believe; in particular, it isn't required that I raise the question whether I *am* within my rights in believing that.

As it looks to me, then, the regress doesn't get started at all, or else it ends after the first step.

Sosa's other contribution, "Postscript to 'Proper Functionalism and Virtue Epistemology,' " is a continuation of a discussion which began with his "Proper Functionalism and Virtue Epistemology" in a *Nous* symposium on my Warrant books and continued with my reply,

"Why We Need Proper Function."[62] Sosa argues that his own "virtue epistemology" is simpler and more plausible than mine, inasmuch as mine involves the notions of proper function and design plan. He and I agree that a belief is warranted only if it is produced by an ability, or faculty, or power; we also agree that faculties are the sorts of things that have functions, and that they can function properly or improperly. Where we disagree, though, is over my suggestion that the notion of proper function involves that of design: "And the notion of 'functioning *properly*' is not far to seek. In none of that, however, do I see a need to import any notion of design, either theological or merely teleological" (p. 273).

Well, at the level of *epistemology*, I'm not sure we disagree. I say the notion of warrant involves the notion of proper function; Sosa apparently agrees. I go on to say that it isn't possible to give a 'naturalistic' account or explanation or analysis on the notion of proper function; here apparently Sosa disagrees. But the latter claim isn't really part of the epistemology, as I see it; I should have thought it's a *metaphysical* claim, not an epistemological claim.[63] Sosa's *epistemology*, as far as I can see, involves the notions of faculties having functions (including that of furnishing us with true beliefs), of their functioning properly (and surely here he'd want to note the relativity of the latter to the right kind of environment), and of their being 'well designed'—i.e., such that when they do function properly, they do furnish us with true beliefs. But these are also the basic elements of my account of warrant, which can be expressed in just those terms: a belief has warrant for *S*, so I say, if and only if it is produced by cognitive faculties that (a) are functioning properly in the right kind of (maxi- and mini-) environment, (b) have the production of true beliefs as their function, and (c) are 'well designed' in the sense that there is a high objective probability that a belief meeting the two preceding conditions will be true. So I can't see that Sosa's *epistemology* is simpler and more streamlined than mine, although his *metaphysics* might be. Of course if I am right and the notion of proper function *entails* that of design, then his metaphysics will be simpler only at the cost of failing to include an important truth, or, worse, embracing necessary falsehood.

In "Proper Functionalism and Virtue Epistemology" Sosa argues that proper function (cognitive proper function, anyway) doesn't entail design by offering an account of what it is for a cognitive faculty to function properly (an account that doesn't involve intelligent design): it does so if it "tracks the truth." I mistakenly took him to intend this

in *Nozick's* sense and offered objections; in "Postscript to 'Proper Functionalism and Virtue Epistemology' " he now points out that he *wasn't* following Nozick here. What he did and does intend is

> *S*'s cognitive faculty *F* tracks the truth (and functions properly) if and only if, (1) if *P* were true *F* would produce (in *S*) the belief *P*, and (2) if *F* were to produce (in *S*) the belief that *P*, *P* would be true. (p. 276)

He then goes on to point out that this account of tracking has several advantages over Nozick's.

Now this isn't a *general* ('naturalistic') account of proper function; it applies only to belief-producing faculties or processes. Of course that's no problem in the present context. If (contrary to expectations) Sosa can give a 'naturalistic' account of proper function for cognitive faculties, more power to him—whether or not it is possible to give such an account of proper function generally. And his account *does* seem superior to Nozick's. But it still won't do the trick. A cognitive faculty can meet this condition without functioning properly; this account (like the more general naturalistic accounts of proper function[64]) isn't sufficient. Consider, first, necessary truths, e.g., Gödel's First Theorem, or the Fundamental Theorem of the Calculus. Note that the instantiations of the first of the two tracking conditions with such a necessary truth yield a conditional that has a necessary antecedent and a contingent consequent; instantiations of the second condition with a necessary truth yield a conditional with a necessary consequent and a contingent antecedent. Both are problematic. Take the second:

> If faculty *F* were to produce belief in Gödel's theorem (i.e., produce a belief whose content was Gödel's theorem) Gödel's theorem would be true.

Well, no doubt; and also, if F were to produce belief in the *denial* of the theorem, it would be true. I suppose the plausible thing to think here is that a faculty automatically meets this condition for any necessary truth; this conditional is a necessary truth.

Not so, of course, for the first condition:

> If Gödel's theorem were true, *F* would produce belief in it.

What does this *mean*, and when or how would it be true? It's not wholly clear. Suppose we think about it from the perspective of

possible worlds: this conditional will be true if and only if, in the nearby worlds in which Gödel's theorem is true, F produces belief in it. Of course the theorem is true in all the nearby possible worlds (as well as those afar off); this condition, therefore, reduces to the claim that in the nearby possible worlds, *F* produces belief (in *S*) in Gödel's theorem. But that is quite compatible with *F*'s malfunctioning. Perhaps *S* believes the theorem, not because he can see or prove that it is true, but because he suffers from a malady which causes him to believe any mathematical proposition whose name in English begins with 'G', or perhaps any mathematical statement of a certain level of complexity; or because he is obsessed with German mathematicians and automatically believes any mathematical claim he thinks put forward by any German mathematician. (Sadly enough, he also believes that arithmetic is complete, thinking that this claim was put forward by Hilbert.) Then his cognitive faculties, at least the ones involved in producing the belief in question, are not functioning properly, despite their meeting this tracking condition. And of course if I'm right about this case, the same will go for nearly any other necessary truth. Thus (as I see it) the existence of God is a necessary truth. Now while of course I don't for a moment agree with those sociologists of religion who think any belief in God a matter of cognitive malfunction, I should think it would be *possible* that belief in the existence of God be produced by malfunctioning cognitive faculties. If so, however, those beliefs would meet Sosa's tracking conditions but result from cognitive malfunction (and would therefore lack warrant).

The same problem will arise for very many contingent truths—they too will be such that a malfunctioning faculty can meet Sosa's tracking conditions in producing belief in them. Any proposition true in all the nearby possible worlds will be of that sort. This would include propositions stating physical regularities (the velocity of light, the values for the force of gravity and for the weak and strong nuclear forces, the fact that the earth is very old, the distance from the earth to the sun, and the like). Each automatically is such that if I were to believe it, it would be true; each also meets the converse condition if in fact I believe it in all the nearby possible worlds—no matter what the source of the belief, and even if the source is in some kind of cognitive dysfunction. So I doubt that Sosa's naturalistic account of cognitive proper function is successful.

Ad Kvanvig

Jonathan Kvanvig pleads for more breadth, more scope. We should see warrant (the property or quantity enough of which distinguishes

knowledge from mere true belief) as a special case of something broader—indeed, of *two* things broader. On the one hand, he proposes that we think of warrant not just as a property of *beliefs*, but also as a property of (or relation among) propositions just as such, regardless of believers or knowers and their idiosyncrasies. On the other hand, there are many propositional attitudes in addition to belief: doubt, withholding, assuming, hope, fear, desire, disgust, and the like. Kvanvig thinks that warrant (as a property of beliefs) is just a special case of a property that can attach to these other propositional attitudes as well: he laments the fact that I don't provide an account of this more general property.[65] He also claims that I am an Aristotelian in epistemology. In his use, this is not a compliment: "Aristotelianism in epistemology offers a distinctive answer to the question of where to begin, an answer I will argue is incorrect" (p. 282).

As to the suggestion that there is a great deal more to epistemology than an exploration of warrant, I enthusiastically agree. First, warrant is only one kind among others of positive epistemic status that can be enjoyed by beliefs: there is an additional whole litany of epistemic values that can be enjoyed by beliefs (WPF p. 2). And second, Kvanvig is quite right in pointing to all those other propositional attitudes and their importance.[66] By way of self-exculpation, I suppose I might offer a couple of bromides: we have to start somewhere, and warrant is as good a place as any; we can't do everything, and hence can't be faulted for picking one worthwhile project as opposed to all the others. At bottom, I think this has to be my defense; but I am prepared to concede that possibility that one can't do a really proper job on warrant without exploring the analogically related properties of these other propositional attitudes.

Indeed, the fact is work of this kind is very much needed. Since Descartes, Locke, and the Enlightenment, the standard view (among intellectuals at any rate) has been that one can never go wrong by withholding belief.[67] The mark of the wise person is a certain chariness about belief, a certain doxastic standoffishness, reluctance, restraint. What is required is temperance. You can never go wrong by withholding, but you can easily go wrong by believing. You can easily make a fool of yourself by believing unwisely, but not by abstaining from belief. Belief is a little like alcohol, and the standard view concurs with the attitude of the WCTU towards the latter: possibly it's acceptable in moderation, but it can be very dangerous, and all things considered, perhaps abstinence is the only completely safe stance. But of course this can't really be correct for belief; the standard view is mistaken

here. A person who fails to believe that there are other people or that there has been a past is not an epistemic saint whose virtues far exceed that of the run of humanity; she is more like an epistemic invalid suffering from a cognitive disorder. Obviously there is such a thing as undue fastidiousness with respect to belief. One can go wrong (epistemically speaking) by failing to believe just as well as by believing; there is no safe haven. This has important application in philosophy of religion: we aren't just given in advance that agnosticism is a safe and secure epistemic haven, with belief in God a somewhat risky and speculative epistemic venture.

So I applaud Kvanvig's desire for a broader look; certainly there is at least an analogue of warrant for other propositional attitudes, and I have no doubt that a careful look at these analogues would throw light on warrant (for beliefs). But this is a task for someone else: I commend it to Kvanvig himself.

On the other hand, I am less happy with his (anti-Aristotelian, he says) suggestions about an analogue of warrant for *propositions* (as opposed to beliefs). In part this is because I don't really understand him. Sufficient for being an Aristotelian, he seems to say, is thinking that (according to the central and paradigmatic core of our notion of warrant) it is fundamentally *beliefs* that have warrant (though by extension we can speak of a proposition *p*'s having warrant for a person when her belief that *p* does). But is there an alternative? Warrant is (roughly) whatever must be added to true belief to get knowledge; that is, it is a property or quantity had by *beliefs*. So how could the basic notion of warrant fail to attach initially to beliefs? How could it be that the basic notion here is warrant as had by *propositions*?

Kvanvig also says that "for the Aristotelian epistemologist, doxastic warrant has primacy over propositional support" (p. 283). This and other things he says suggest that what he is really interested in is a *supports* relation that holds between (among) propositions—an *evidential* relation that holds between *p* and *q* whether or not anyone recognizes or responds to its holding. Perhaps he is thinking of either logical or objective conditional probability (including entailment as a special case) (see WPF pp. 140–42, 144–51); or perhaps he is thinking of conditional epistemic probability (WPF chap. 8 and 9). (The difference would be that conditional epistemic probability is independent of *individual* believers but does involve a reference to the human (or other) design plan, while objective conditional probability does not.) What leads me to doubt that these are what Kvanvig has in mind is

that in fact I did say a fair amount about these notions, and Kvanvig says I didn't deal with the property or relation *he's* interested in.

Kvanvig's main concern, however, is in the question whether 'propositional support' can be defined in terms of what he calls 'propositional warrant'—the sort of warrant a belief *p* can receive by being accepted on the evidential basis of another belief *q* that evidentially supports it. Put in his terms, this is the question whether 'propositional support' can be understood in terms of what he calls 'psychological support'. The former is (apparently) just a relation among propositions; the latter is a relation among propositions and people. (*That the butler did it* has psychological support for Holmes, who sees how the evidence bears on the question; it has only propositional support for Watson, who has the same evidence (believes those propositions that in fact support the proposition that the butler did it) but doesn't see the evidential connection and believes just on a hunch.) Kvanvig argues—convincingly, I think—that it is at any rate extremely difficult to see how propositional support could be defined or explained in terms of psychological support.

As I say, I am not entirely confident of my understanding of Kvanvig here. Insofar as I do understand him, I take him to be claiming that (to put things in my terminology) neither objective nor logical (conditional) probability can be defined or explained in terms of epistemic probability (WPF chap. 8 and 9). This seems right, and the account of epistemic conditional probability I gave makes an essential reference to logical probability. (On the other hand, epistemic probability is not *just* a special case of logical or objective probability. The logical probability of *p* on *q* can be very high when the epistemic probability of *p* on *q* is not—if, for example, no human being could grasp or understand the propositions or see the relevant relationships.) So neither is explicitly definable just in terms of the other, but epistemic probability is to be explained partly in terms of objective probability. But doesn't that make me an *anti-Aristotelian*, rather than an Aristotelian, as Kvanvig claims?

There is much more to be said about Kvanvig's contribution, as indeed about all the others. Sufficient unto the day is the discussion therein, however, and I am obliged to stop here—though not without thanking the contributors once more for these fine essays.

Notes

I'm extremely grateful for penetrating criticism and wise counsel to Mike Bergmann, Andrew Koehl, Kevin Meeker, Trenton Merricks, and Mike Rea.

1. "Is Justified True Belief Knowledge?" *Analysis* 23 (1963) pp. 121–23.
2. See WCD pp. 6–8.
3. For a fuller explanation of these examples see WCD pp. 31 ff.
4. Though often attributed to Alvin Goldman, this example was first used by Ginet; see G. C. Stine, "Skepticism and Relevant Alternatives," *Philosophical Studies* 29, 1976, p. 254. Stine also reports that Goldman attributes the example to Ginet.
5. Designed, whether by God or evolution. In what follows, I'll understand this qualification but not constantly repeat it.
6. More exactly (in view of the case of the aging forest ranger), proper function of cognitive powers and processes *internal* to the epistemic agent.
7. See Trenton Merricks, "Warrant Entails Truth," *Philosophy and Phenomenological Research* 55 (1995), 841–55.
8. Feldman proposes a similar example on pp. 217–18, as does Robert Shope in "Gettier Problems" (*Routledge Encyclopedia of Philosophy*), and his forthcoming book *Knowledge as Power*. Like Klein's, Shope's example is automotive.
9. Although of course there could be similar environments elsewhere in the universe, and environments sufficiently similar to permit us to have knowledge, were we somehow transported to them.
10. In WPF I wasn't clear about the distinction between cognitive maxi- and mini-environments; Swain (esp. p. 140) asks penetrating questions about my treatment of cognitive environments.
11. A proposition *p* is true (false) in a situation *S* just if necessarily, if *S* had been actual, *p* would have been true (false). So what I mean here, of course, is not that the proposition *S forms a true belief* has the property of being neither true nor false in the cognitive mini-environment in question, but only that it doesn't have the property of being true in that situation, and also doesn't have the property of being false in it.
12. We must therefore say that a cognitive faculty—vision, say—can be unreliable in a given mini-environment *M* even though it is reliable in the maxi-environment included in *M*.
13. Though I am unable to tell what he means by 'justification': see p. 346.
14. Furthermore, the official account was supposed to explicate the rough and ready genuineness condition according to which "a defeater is genuine just in case it defeats without depending upon a false proposition"; −A, in the example, meets the official condition for genuineness, but it obviously *does* depend upon a false proposition—my mistaken belief (C). What has gone awry? Klein's meaning, I think, is that a *misleading* defeater is one such that the falsehood upon which its defeating force depends, attaches to the defeater *itself*, not to one of my beliefs. (Clearly I could acquire a genuine defeater for one of my beliefs by coming to believe something false.)
15. (Minneapolis: University of Minnesota Press, 1981), pp. 137–50.
16. The issues here are relevant to philosophy of religion. Suppose theism

is true and there is such a thing as the *Sensus Divinitatis* (perhaps restored and extended by what Calvin calls "the internal testimony of the Holy Spirit") and suppose my belief (G) that there is such a person as God is properly based on these sources, thus constituting knowledge. But then I reflect on the evil the world contains and come to believe that the probability of (G) given the existence of (e.g.) suffering on the part of innocent children is low. Even if this probability belief is in fact mistaken, and even if I continue to believe (G), it could be that I no longer *know* (G).

17. For strictures on this suggestion, see WPF p. 164.

18. For example, it could be that the existence of evil together with the rest of what I believe is such that belief in God is not justified for me (in Klein's sense); it doesn't follow that I don't know that there is such a person as God. For perhaps the nonpropositional warrant theistic belief gets (e.g., from Calvin's *Sensus Divinitatis*) is strong enough to counterbalance the weakness of the propositional evidence.

19. But see "Reliabilism, Analyses and Defeaters," pp. 427–34, for hints.

20. *Ibid.*, p. 428.

21. In particular the proposition *I am drunk and my faculties are malfunctioning* won't be such a defeater. Adding it to a set of propositions standing in the is-a-good-reason-for relation to another doesn't yield a set that doesn't stand in that relation to that other.

22. See "Reliabilism, Analyses and Defeaters," pp. 429–34.

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

24. In *Knowledge and the Flow of Information* (Cambridge, Mass.: MIT Press, 1981), pp. 172–178.

25. In *Philosophical Explanations* (Cambridge, Mass.: Harvard University Press, 1981). (Nozick's account is very similar to the early Dretske's.)

26. *The Semantic Conception of Theories and Scientific Realism* (Chicago: University of Illinois Press, 1989), p. 368.

27. There is such a distinction, and it is important; but it isn't needed to evade Gettier.

28. See his *Metaepistemology and Skepticism* (Lanham, Md., and London: 1995), p. 123.

29. See my "Positive Epistemic Status and Proper Function" in *Philosophical Perspectives 2, Epistemology*, 1988, ed. James Tomberlin (Atascadero, Calif.: Ridgeview, 1988), pp. 15–18.

30. *Judgement and Justification* (Cambridge: Cambridge University Press, 1987).

31. Compare my "impulsional evidence" (WPF pp. 192–93). In my view, however, a belief produced by cognitive malfunction—e.g., the belief that I am Napoleon—need have no warrant at all, despite the fact that it enjoys impulsional evidence.

32. And if you think that his beliefs really *were* fairly coherent, think about a person whose beliefs (by virtue of cognitive malfunction, say) are not.

33. But what about beliefs like the two I just mentioned? Do they constitute reflective knowledge or animal knowledge? If the former, then not all reflective knowledge requires explanatory coherence; if the latter, then animal knowledge is necessary for reflective knowledge, so that the latter is not reflective all the way down.

34. *Inquiries and Essays*, ed. by Ronald Beanblossom and Keith Lehrer (Indianapolis: Hackett, 1983), pp. 275–76.

35. Briefly, I think it requires too much; for example, it isn't necessary, for knowledge, to believe or accept that you aren't isolated in his sense; you might know much, even if this thought has never occurred to you.

36. Among published and semipublished objections are William Alston's comments on the paper, presented at a conference at Santa Clara University in the spring of 1992, Carl Ginet's "Comment's on Plantinga's Two Volume Work on Warrant" *Philosophy and Phenomenological Research* (Vol. LV, No. 2), June 1995, pp. 403 ff., Timothy O'Connor's "An Evolutionary Argument Against Naturalism?" *Canadian Journal of Philosophy* (Vol. 24, No. 4) Dec. 1994, pp. 527 ff., Richard Otte's comments on the paper, presented at the same symposium as Alston's, Glenn Ross's "Undefeated Naturalism" and David Hunt's "Is Metaphysical Naturalism Self-Defeating?" presented at the Pacific Division meetings of the APA in March 1994, Leopold Stubenberg's "Is Naturalism Really Defeated?" presented at a colloquium at the University of Notre Dame in 1994, Wesley Robbins's "Is Naturalism Irrational?" *Faith and Philosophy* (Vol. 11, No. 2) April 1994, pp. 255 ff., and a paper by Evan Fales presented at the Central Division meetings of the APA in April 1995. In "Naturalism Defeated" (presently unpublished but copies available) I consider these and other objections to the argument, concluding that they are inconclusive, and that the argument stands.

37. In "Naturalism Defeated," this is the "Second Principle of Defeat" (I owe this principle to Steven Wykstra). For more detail on this principle, and on defeaters generally, see that paper.

38. See his *A Treatise Concerning the Religious Affections*, ed. by John E. Smith (New Haven: Yale University Press, 1959), pp. 95 ff.

39. BonJour thinks my stipulative notion of warrant "unfortunate"; his reason is that what is required in addition to true belief are perhaps *two* conditions (not just one) and furthermore two conditions that are "incommensurable." I agree with the latter (modulo a bit of a puzzlement about incommensurability) but don't see the problem. As a matter of fact I take warrant to be something like the conjunction of *four* conditions: proper function, appropriate (mini- and maxi-) environment, relevant part of the design plan both *aimed* at truth, and *successfully* aimed at truth. I don't know whether or not these conditions are incommensurable in BonJour's sense, but they *are* incommensurable in the sense that no amount or degree of one of them can make up for absence of another.

40. "The *third* condition, for which the term 'justification' is most standardly employed, has to do with the presence of something like a reason or ground for the truth or likely truth of the belief" (p. 49); he's clearly thinking, furthermore, that the ground has to be something of which the epistemic agent is aware.

41. I return to this point on pp. 343 ff. Bonjour says: "I think that his [my] account of *a priori* knowledge as being based merely on a distinctive sort of phenomenology (see *WPF* Ch. 6) is highly misleading at best" (note 16). But here I think he has uncharacteristically erred: my claim was that *a priori* belief is *not* based on evidence, phenomenological or otherwise. See WPF pp. 104, 191–92.

42. See my reply to Feldman, pp. 357–61.

43. BonJour suggests that I was proposing my conditions as an account of *justification* or the third condition. That wasn't what I had in mind: I was intending to give necessary and sufficient conditions for warrant, and that is how I also intend the revised statement in Part I (The Resolution Problem).

44. *An Inquiry into the Human Mind*, in *Inquiries and Essays*, ed. Ronald Beanblossom and Keith Lehrer (Indianapolis: Hackett, 1983), pp. 84–85.

45. See Klein, p. 98.

46. See also Michael Shope's forthcoming *Knowledge as Power*, draft pp. 28–29.

47. For more detail, see pp. 359–60.

48. Swain also refers (pp. 143–45) to some examples of BonJour's involving Norman the clairvoyant; these examples, he says, show that my account isn't sufficient for warrant. I don't find these examples convincing. What makes them initially *seem* convincing, I think, is that we think of Norman as knowing and believing the same sorts of things the rest of us know and believe. If so, he has a defeater for his suddenly and inexplicably acquired belief that the president is in New York. On the other hand, if we *all* had these clairvoyant powers, if there was nothing special about them, then he wouldn't have a defeater and would indeed know.

49. But then what is the force of the 'but only' in the second line of the above quotation?

50. Rather as I am when theologians like Paul Tillich and Gordon Kaufmann advertise their theological views as more sophisticated and satisfactory varieties of theism.

51. *Naturalizing Epistemology* (Cambridge: MIT Press, 1987) p. 3.

52. See WPF chap. 11.

53. As J. B. S. Haldane once quipped, teleology is like a mistress to a biologist: he can't live without her, but he is unwilling to be seen with her in public.

54. Feldman and Conee, "Evidentialism," *Philosophical Studies* 48, pp. 15 ff.

55. Perhaps other conditions should be added; I don't have the space here to go into the matter.

56. There *is* sensuous imagery connected with the belief (I am appeared to as with a fleeting and partial glimpse of a blackboard on which the English sentence 2 + 1 = 3 is written), but I don't believe the proposition on the basis of that imagery (as I argue on pp. 104 ff. of WPF).

57. Unless we stipulate that a proposition for which I have no evidence automatically fits my evidence, in which case, obviously, of course even worse problems rear their ugly heads.

58. It is of course possible that in some sense *p* seems to you to be true, when the fact is you don't believe *p*. I still find in myself an inclination to believe the proposition *every property has a complement and there is such a property as non-self-exemplification*, even though I also believe that proposition false. But here the question is whether one can believe a proposition when it doesn't so much as seem to you to be true; that's what I am inclined to think is impossible.

59. This response also applies to the first case.

60. It might be argued that God is a necessary being, is essentially omniscient, and in every possible world designs the faculties of cognitive beings; if that were so (and something like it has been a part of much traditional thought about God) then it isn't possible that there be an incompetent designer of cognitive faculties.

61. See WCD chap. 1.

62. Both in *Nous* Vol. XXVII, No. 1 (March 1993).

63. See my replies to Swain and Conee.

64. See WPF chap. 11.

65. "Unfortunately, we do not find in Plantinga's two-volume work any discussion of this further property" (p. 285).

66. See van Fraassen's paper for the suggestion that *materialism* and *naturalism* are not really beliefs but another kind of propositional attitude.

67. According to Descartes (at least as he is ordinarily understood: but see note 29 on p. 13 of WCD) the central epistemic duty is to refrain from believing what is not clear and distinct for you. Since it is not possible, according to him, to refrain from believing what *is* clear and distinct, you can never go wrong by abstaining from belief.