

5

Proper Function

Part I of this book was devoted to my attack on internalism. The main argument of Chapter 1 ran as follows:

- I. An essential feature of internalism is that it makes a subject's actual or potential *awareness* of some justification-contributor a necessary condition for the justification of any belief held by that subject.
- II. The awareness required by internalism is either *strong* awareness or *weak* awareness.
- III. If the awareness required by internalism is strong awareness, then internalism has vicious regress problems leading to radical skepticism.
- IV. If the awareness required by internalism is weak awareness, then internalism is vulnerable to the SPO, in which case internalism loses its main motivation for imposing the awareness requirement.
- V. If internalism either leads to radical skepticism or loses its main motivation for imposing the awareness requirement (i.e. avoiding the SPO), then we should not endorse internalism.
- VI. Therefore, we should not endorse internalism.

Since premise (II) is uncontroversial, only four premises needed defending. In the first chapter I gave arguments for premises (III) and (IV) and offered a partial defense of premise (V). Chapter 2 was aimed at showing that not even those internalists sensitive to the similar-sounding *Sellarsian* dilemma for internalist foundationalism are able to escape the dilemma for internalism presented in Chapter 1. The purpose of Chapter 3 was twofold: I defended premise (I) by arguing against mentalist accounts of internalism which conflict with the account of internalism proposed in premise (I); then, having argued that mentalism is different from internalism, I argued against mentalism itself (thereby objecting to the main nonexternalist position besides internalism). Finally, in Chapter 4, I added to my defense of premise (V) by arguing that deontologism was an inadequate replacement for the main motivation for internalism.

Having argued in Part I against the two main competitors to externalism—internalism and mentalism—I will turn in Part II of this book to a defense of externalism. I won't defend externalism by identifying necessary

and sufficient conditions for being an externalist and arguing that in order for a theory to be correct it must satisfy those conditions. (In Chapter 1, I identified a necessary condition for being an internalist, indicating that I didn't think it was advisable to seek a complete list of necessary and sufficient conditions for being an internalist. I have the same reluctance about trying to identify a complete list of necessary and sufficient conditions for being an externalist.) Instead, I will defend externalism by defending a view that is clearly a version of externalism. By defending that view, I will be defending externalism itself, even though I'll be making no attempt to give an analysis of externalism.

According to the version of externalism I'll be defending, the two most important requirements on justification are the satisfaction of a *proper function* condition and the absence of a *defeater*. The first requirement will be the focus of this chapter. The topic of defeaters will be taken up in Chapter 6. After explaining and giving arguments for this version of externalism, I will turn, in Chapters 7 and 8, to a consideration of two objections that have been proposed against all versions of externalism: that externalists are committed to permitting an objectionable kind of epistemic circularity and that the externalist response to skepticism is inadequate in a variety of ways. These two objections can be viewed as further motivations for internalism if one assumes that internalists can avoid them. I will argue that switching one's allegiance from externalism to a nonexternalist position doesn't enable one to avoid those two objections (unless one adopts a position that, for other reasons, is wholly implausible). This provides still more support for premise (V) from Chapter 1 because it shows that these proposed motivations for internalism are also inadequate replacements for the main motivation debunked in Chapter 1. Moreover, I will argue in Chapters 7 and 8 that although some of the charges associated with these two objections do apply to externalism, they don't provide us with a good reason to reject externalism.

The goal of *this* chapter is to defend a proper function account of justification. In an attempt to get my internalist opponents on the road to recognizing the plausibility of my conclusion, I will begin with the assumption, congenial to internalists, that the following evidentialist thesis is, initially at least, an attractive account of justification:

E_F: S's belief B is justified iff B is a *fitting* doxastic response to S's evidence.¹

¹ The more standard way of stating the evidentialist's thesis is to have the right-hand side of the 'iff' say something like 'B fits S's evidence' (see Feldman and Conee 1985). I'm assuming that doxastic response B is a fitting response to S's evidence only if (i) B fits S's evidence and (ii) B is formed in response to S's evidence. Clause (ii) is a way of adding a 'formed in the right way' requirement. Since justification, as I'm thinking of it in this book, is *doxastic* justification rather than merely *propositional* justification, the addition of such a requirement is quite unremarkable. (See Chapter 1, Section 1 for a discussion of the difference between doxastic and propositional justification.)

In Section 1 of this chapter, after saying a little about what the evidentialist position on justification is, I will argue that that position needs to be improved in three ways. Then, in Section 2, I will defend a proper function analysis of justification that captures the improvements recommended in Section 1. In the final section, I will respond to some objections to my proposed analysis of justification.

But before moving on, I'd like to highlight one additional benefit of this chapter: it shows that one of the standard objections to externalism is misguided. According to this objection, the unreliably formed beliefs of evil demon victims can be as justified as our own beliefs. This is supposed to create a problem for externalists because it is assumed that they impose a reliability condition on justification. But as I will make clear below, the proper function account of justification I give is an externalist account that doesn't impose a reliability condition on justification. In light of this we can see that this standard objection, while perhaps effective against reliabilists about justification, is not effective against externalism about justification.²

1 FROM EVIDENTIALISM TO PROPER FUNCTION

1.1 Evidentialism

Let's begin by trying to get a better understanding of evidentialism. What exactly counts as a subject's evidence? One proposal in the literature is that evidence is limited to things of which the subject is aware or potentially aware (i.e. to epistemically accessible things).³ Another proposal is that evidence is limited to the subject's mental states.⁴ Taking these as our guide, we may conclude that, in order for a belief to be a fitting doxastic response to a subject's evidence, it must be a fitting doxastic response to something of which that subject is (at least potentially) aware or to some of her mental states. I will assume that we have some grasp of what it is for a belief to *fit* one's evidence and that, with respect to certain clear cases, we have some fairly firm intuitions concerning whether the belief in question does or doesn't fit the subject's evidence. The very ease with which we make sense of the examples used by evidentialists to illustrate their position testifies to the plausibility of these last two assumptions. When, for example, Feldman and Conee (1985: 15) say that the belief that *there is something green before one* fits the evidence that a normal person has in ordinary circumstances when looking

² Nor is it effective against reliabilist accounts of warrant (that which makes the difference between knowledge and mere true belief) since demon beliefs are not typically supposed to be as *warranted* as our beliefs, even if they are typically supposed to be as *justified* as our beliefs. And given that reliabilists more often give analyses of warrant than of justification, this standard objection doesn't seem to apply to too many externalists.

³ See Feldman (1988a) for a discussion of this proposal.

⁴ Steup (2001a: 137). See also Conee and Feldman (2001: 231–9) where they suggest this reading.

at a plush green lawn in broad daylight or that the belief that *sugar is sour* does not fit our gustatory experience, we have no trouble understanding or accepting such claims.

A standard line taken by evidentialists involves endorsement of the following three claims:⁵

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

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

Necessity: the fittingness of doxastic response B to evidence E is an essential property of that response to that evidence.⁶

Nonreliability is extremely plausible. Just as it seems possible for a demon victim who is my mental twin to have justified beliefs despite the fact that, in the twin's case, those beliefs are unreliably formed, so also it seems possible for a demon victim's beliefs to fit her evidence. The demon victim's problem has to do with the fit between beliefs and evidence but with the connection between evidence and the world. So I join the evidentialist in accepting Nonreliability and in rejecting reliabilist accounts of justification. I also accept Objectivity (a position to be explained in greater detail below). Thus, I follow the evidentialist in thinking that there is an important sort of justification that is objective in the sense that innocently thinking all is well epistemically is insufficient for it. I made clear my agreement with evidentialists on Objectivity at the beginning of the book where I identified the sort of justification in which I was interested. Putting all this together, we can say that I join the evidentialist in denying that objective fittingness depends on reliability.

But then what *does* it depend on? The evidentialist's answer to this question is given by Necessity: the objective fittingness of doxastic response B to some evidence E is an *essential* property of that response to that evidence; it isn't a contingent property that depends on some other condition being satisfied. But notice that this answer isn't *forced* upon those who agree with evidentialists about Nonreliability and Objectivity. For the fact that the fittingness of doxastic response

⁵ Of these three claims—Nonreliability, Objectivity, and Necessity—the first is the one most obviously endorsed by evidentialists. Below I explain why I think Conee, Feldman, and other evidentialists believe in Objectivity and Necessity. Who counts as an evidentialist? Perhaps the clearest examples are Conee and Feldman (2004). But evidentialist sympathies can also be discerned in the writings of Chisholm (1977), Moser (1985, 1989), Fumerton (1995), Haack (1993, 1997), Russell (2001), and Steup (2001a).

⁶ The idea is that if B is a fitting response to E by itself, then, even if B could be an unfitting response to evidence that includes E *and more besides*, it couldn't be an unfitting response to E by itself.

B to evidence E isn't contingent upon E's being a reliable indicator of B's truth doesn't prove that it isn't contingent upon anything. So Nonreliability doesn't entail Necessity. I will be arguing that Necessity is false.⁷ In fact, the first of the three improvements to evidentialism that I'll be recommending below is that it drops its endorsement of Necessity.

In order to have a good understanding of my three recommended improvements to evidentialism, it will be helpful to recall the two questions employed in Chapter 3 for the purpose of classifying theories of justification:

- (A) What sorts of things can be appropriate inputs to our belief-forming processes?
- (B) Must the appropriateness of such an input be due to an essential property of that input?⁸

Evidentialists, by endorsing Necessity, give an affirmative answer to (B).⁹ My first recommended improvement to evidentialism is that it replace its affirmative answer to (B) with a negative answer to that question. The second recommended improvement arises out of my argument for the first improvement. The counterexamples I use to show that Necessity is false suggest that the fittingness mentioned in E_F should be understood in terms of proper function. Thus, my second recommended improvement is that we understand that fittingness in that way. The third improvement I will recommend makes it utterly clear, if it wasn't already, that if evidentialism is altered in accord with my recommendations, it will no longer be an *evidentialist* position. For my third recommendation is that we don't give an evidentialist answer to question (A). An evidentialist answer to (A) says that only evidential states can be appropriate inputs to our belief-forming processes. Given that only accessible states or only mental states can be evidence, the evidentialist answer to (A) says, in effect, that only accessible states or only mental states can be appropriate inputs to our belief-forming processes. I recommend that we reject those answers in place of a noninternalist, nonmentalist answer that fits with externalism. The remainder of Section 1 will be devoted to defending these three recommended improvements to the evidentialist position.

⁷ Notice that in rejecting Necessity I am not saying that the relation of supervenience holding between justification and its supervenience base is contingent. For even if it were true that belief B's *justification* supervened (of necessity) on the fittingness of doxastic response B to the subject's evidence (which happened to be E), that wouldn't entail that the *fittingness* of doxastic response B to evidence E was an essential property of that response.

⁸ For an input to be appropriate is for it to be capable of resulting in a justified belief.

⁹ Feldman is a prime example. See the discussion in last two paragraphs of Section 1.3.2 of Chapter 3 where I describe how he defends his version of mentalism by highlighting what he takes to be the virtues of his affirmative answer to question (B). And given Feldman's remarks about what mentalism and internalism are (I discuss these remarks in Section 1.2.2 of Chapter 3), it is clear that he thinks an affirmative answer to question (B) is widely endorsed, since he thinks it is a crucial component of being an internalist and a mentalist.

1.2 The First Improvement: Drop Necessity

My argument for dropping Necessity involves the use of a counterexample that is best understood in light of some of the implications of Objectivity. I'll begin, therefore, by considering what some of those implications are.

1.2.1 Objectivity

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

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

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

Upon learning this, Tim immediately forms the belief that

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

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

Now consider the view that the main evidence for our beliefs doesn't always constitute our total relevant evidence for them. For example, in the Jane case, there may be, in addition to the visual experience on which she bases her belief that the stick is bent, the strong *felt* inclination to take her visual sensations as indicative of the truth of the belief in question.¹⁰ Those of us who are completely familiar with the phenomenon of water distorting the appearance of a straight stick lack this felt inclination (at least it isn't as strong in us). As a result, our total evidence is somewhat different from Jane's. We can call Jane's strong

¹⁰ A felt inclination to do X is an experience of feeling inclined to do X. It isn't merely a disposition.

felt inclination a ‘connector’ between her main evidence and her belief based on that evidence. It’s a connector we lack. Another (related) difference is that we are inclined to think that the main evidence on which Jane relies is *not* indicative of the truth of the belief she bases on it. This provides us with a *disconnecter* between the main evidence and the belief in question—a disconnecter Jane lacks.¹¹ So if we think of these connectors and disconnecters as parts of one’s total evidence, then *our* total evidence when we view a stick looking bent in water is different from *Jane’s* when she considers such a sight (even if the main evidence—the way it visually appears—is the same).

In the Tim case there could be a connector too. Let us suppose it is the strong felt inclination to take the truth of (1) to entail the truth of (2). It is easy to imagine those less logically perceptive than Tim failing to see this sort of connection. We could describe this difference by saying that although these others have the same main evidence Tim has, they don’t have the same total relevant evidence since they lack a connector he possesses.¹²

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

Belief	Main Evidence	Connector
B1. The first person belief: ‘There is smallish hard round object in my hand’.	ME1. Tactile sensations of the type you experience when you grab a billiard ball.	C1. The strong felt inclination to take ME1 to be indicative of the truth of B1.
	ME2. Olfactory sensations of the type you experience when you smell a meadow full of flowers. ¹³	C2. The strong felt inclination to take ME2 to be indicative of the truth of B1.

¹¹ Thus, disconnecters will be a sort of undercutting defeater (see Chapter 6 for a discussion of undercutting defeaters). Notice that *lacking a connector* between a belief and one’s main evidence isn’t by itself sufficient for *having a disconnecter*.

¹² I am *not* suggesting that justification *requires* a connector between one’s beliefs and one’s main evidence. I’m just explaining what connectors are and noting that they may be present in some cases.

¹³ Since a meadow full of flowers usually has a number of different kinds of flower (as well as a variety of nonflowering plants) each of which has a distinctive odor, the type of olfactory

The subjectivist opponent of Objectivity might acknowledge that B1 fits ME1 and that it doesn't fit ME2. But she will add that B1 *does* fit the combined evidence of ME2 together with C2 because by adding C2 to ME2 (and adding no disconnectors¹⁴) we get subjective fittingness. The proponent of Objectivity rejects this suggestion. She says that since B1 doesn't fit ME2, it won't help merely to add C2 even if we stipulate that no disconnectors are present. The objective failure of B1 to fit the subject's evidence (ME2) isn't altered by adding C2 to the subject's evidence base; if B1 fails to fit ME2, it also fails to fit ME2 + C2 (the combination of sensation ME2 and connector C2).

Paradigm evidentialist Richard Feldman appears to be endorsing Objectivity so construed when he points out that one can't get justification for one's beliefs merely by thinking one has good reasons for them.¹⁵ Other evidentialists such as Bruce Russell, Paul Moser, and Richard Fumerton also seem to agree with Objectivity. For reliability isn't the only condition that they think is insufficient for E's being good evidence for B; they pay the same compliment to *thinking that E is good evidence for B*.¹⁶ In each case, the idea seems to be that if a belief doesn't fit one's main evidence, then merely adding a connector joining that belief with that main evidence won't help.

1.2.2 Learned and Unlearned Doxastic Responses

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

experience one has upon smelling a meadow full of flowers has many components. In this way it resembles the tactile experience one has upon grabbing a billiard ball, an experience that also has many components.

¹⁴ How could a person's total relevant evidence for B1 consist of only ME2 and C2 without any disconnectors? Wouldn't she glance at her hand or attempt to use the hand in which she thinks there is a small hard ball? Wouldn't others inform her that it is crazy for her to think, on the basis of ME2, that she is holding a small hard ball? These are certainly possibilities, likely ones even. But all we need is a possible example where no such disconnectors are present.

¹⁵ See Feldman (1988b: 411). See also Feldman and Conee (1985: sects. II–III) where they say that if B fails to fit one's evidence, it won't help to add subjective factors such as *trying one's best to hold only beliefs that fit one's evidence or being blameless in holding B*. They insist that even with those factors present, B still fails to fit the subject's evidence.

¹⁶ See Russell (2001: 37–8), especially his comments on the believers raised in the benighted religious community whom he thinks of as subjectively justified but not objectively justified. See also Moser (1989: 38–42, 47–52, 202–3) and Fumerton (1995: 8–20, 113–16, 183–224) where they emphasize that merely believing that one's evidence E is good evidence for B isn't sufficient for B's justification since E might not *in fact* make B objectively probable.

That concern about Objectivity strikes me as a sensible one. To handle it, we need to distinguish learned from unlearned doxastic responses. The distinction isn't easy to draw but it is something like this. *Learned* doxastic responses, such as an experienced birdwatcher's immediate bird identifications after a quick look (or listen), are ones a person comes to have only after first finding out independently (i.e. without relying in any essential way on other instances of that same type of doxastic response) that there is a correlation between the truth of such beliefs and the experiences to which they eventually become immediate responses. By contrast, an *unlearned* doxastic response to experience is a hardwired or automatic response that occurs (perhaps only after a certain level of cognitive development) without the subject first independently finding out that there is a correlation between the truth of the belief in question and the experience to which it is a response.¹⁷

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

But there's a related concern that remains—one prompted by the thought that someone might have an unlearned connector which she wrongly takes to be a learned connector. Suppose C is an unlearned connector that connects S's main evidence ME with S's belief B. But suppose also that it (mistakenly) seems to S that she learned to connect ME with B by first learning independently of a correlation between the truth of B and the occurrence of ME. (In actual fact, she didn't learn this. She just came to have the connector in question without any learning going on.) Then although C is an unlearned connector, it seems to S to be a learned connector. We can call these sorts of unlearned connectors 'seemingly-learned unlearned connectors'. It is arguable that seemingly-learned unlearned connectors between ME and B are as helpful (or nearly as helpful) as learned connectors between ME and B: in each case, adding the connector in question to ME seems to provide S with something to which B is a fitting response.¹⁸

¹⁷ This distinction is similar to and inspired by Reid's distinction between original and acquired perception. See Reid ([1785] 2002: 235–9, 1997: 171–2). It is basically the same as Goldman's (1986: 93–5) distinction between beliefs produced by native processes and beliefs produced by acquired methods.

¹⁸ Thanks to Michael Rea for drawing this objection to my attention.

To handle this last concern, we can explain Objectivity even more carefully as follows: some unlearned doxastic responses to one's evidence are unfitting and merely adding an *unlearned* connector to one's original evidence—one that (a) connects that doxastic response to that original evidence and (b) isn't seemingly-learned—won't change that. Recall the original statement of Objectivity:

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

Objectivity says that doxastic response B to E can seem appropriate to the subject—it can feel right—even when it is objectively unfitting. And the explanation at the beginning of this paragraph identifies the sort of 'feeling right' that is insufficient for objective fittingness. Sometimes a doxastic response to evidence feels right because it is a seemingly-learned response. Other times it feels right even though it isn't a seemingly-learned response. When an unfitting unlearned doxastic response comes (in an unlearned way) to feel right in a way that isn't seemingly-learned, that is not enough to make it a fitting doxastic response. That is the thought behind Objectivity. It opposes the view that a doxastic response to evidence is fitting so long as it feels right to the subject.

1.2.3 *A Counterexample to Necessity*

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

To do this, I will need to employ some assumptions implicit in Objectivity. In particular, I will be assuming that, just as there can be unlearned doxastic responses that are objectively fitting, so too there can be ones that are objectively *unfitting* (in the ways described in the previous two subsections). What follows is a description of a possible case in which the billiard ball belief B1 is a fitting unlearned doxastic response to ME2 and an unfitting unlearned doxastic response to ME1.¹⁹ Such a case is a counterexample to Necessity because in the

¹⁹ I call B1 a 'billiard ball belief' for convenience even though the concept of a billiard ball is not included in its content.

actual world, B1 is a fitting unlearned doxastic response to ME1 and an unfitting unlearned doxastic response to ME2.²⁰

Thomas Reid emphasized that there doesn't seem to be any logical connection between our sense experiences and the content of the beliefs based on them. For example, the tactile sensations we experience when touching a hard surface seem to have no logical relation to (nor do they resemble) the content of the hardness beliefs they prompt.²¹ In light of this he said that 'no man can give a reason why the sensations of smell, or taste, or sound, might not have indicated hardness' (Reid [1764] 1997: 57). Considering the matter in the abstract, tactile sensations do not seem to be any more suited than olfactory sensations to being indicators of hardness. Thus, it seems there could have been cognizers like us in outward appearance who experience, upon grabbing a billiard ball, a sensation that is qualitatively of the same type as one of our actual world sensations of smell.²² And it seems possible that the natural *unlearned* doxastic response of such a cognizer to that 'olfactory'²³ sensation is the first-person belief 'There is a smallish hard round object in my hand' (i.e. B1). There is nothing about the process of *grabbing a billiard ball then experiencing ME2 then holding B1* that makes it intrinsically less suitable (as a natural unlearned process for a cognizer to undergo) than the process of *grabbing a billiard ball then experiencing ME1 then holding B1*. In each case, there is a causal chain from an external stimulus to an experience to a belief. And in each case, the experience has the same functional role of connecting that stimulus with B1. The only difference is that in the one case, the experience playing this functional role is ME1 whereas in the other it is ME2. (I should add that the cognizers in this Reidian example are such that *all* of the experiences produced in them by the sorts of activities that produce tactile experiences in us are what we would call 'olfactory experiences'. So it isn't just

²⁰ Objection: Is B1 really an *unlearned* doxastic response to ME1 for normal humans? Perhaps psychological research will show (or has shown) that it is *learned*.

Reply: Maybe so. But suppose psychological research had shown that B1 is (as Reid believed) an *unlearned* doxastic response to ME1. Would that have given us a reason to take such beliefs to be unjustified? No. So it seems possible for there to be creatures for whom B1 is justified in virtue of being believed on the basis of ME1 even though B1 is (for them) an *unlearned* doxastic response to ME1. The possibility of such creatures is all I need to make my case.

²¹ Reid defends this view concerning hardness—along with a similar view concerning the relation of tactile (and proprioceptive) sensations to beliefs in extension, motion, and shape—in Reid ([1764] 1997: Chapter 5, Sections 2–6). For a superb discussion of Reid's views on perception see Wolterstorff (2001), especially Chapters 5 and 6.

²² Perhaps sensations supervene on brain states. Then we have to imagine, first, that sensations of the same qualitative type as our olfactory sensations could supervene on physical states (call them X-states) other than those on which ours supervene (cf. the Martians in Lewis 1980). Second, we imagine a creature who is like us in outward appearance and who, upon grabbing a billiard ball in its hand, causes itself to go into an X-state.

²³ I use quotation marks to indicate that although the sensation in question is qualitatively of the same type as *our* olfactory sensations, it may not, in the circumstances in question, be properly thought of as an olfactory sensation in the ordinary sense (since that sense may have functional role implications).

B1's ground that is different in this way. Likewise, none of the things that produce in us the experiences we call 'olfactory experiences' produce those sorts of experiences in them.)

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

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

What should we say of a species of cognizers for whom the natural unlearned response to grabbing a billiard ball is to experience *ME2* and then form *B1*? It seems we should say that for such cognizers, *B1* is a *fitting* unlearned response to *ME2* and an *unfitting* unlearned response to *ME1*. This shows that the fittingness of an unlearned doxastic response is a contingent feature of it, a feature that depends in some cases on the species of the cognizer who has the response.

²⁴ Cf. Greco (1999: 277; 2000: 173–4), Markie (2004: 530–3), and Plantinga (1993b: 54–63).

²⁵ See n. 20.

²⁶ See n. 14.

Necessity is, therefore, false. Hence my recommendation that the evidentialist drop it.²⁷

1.2.4 Defending the Counterexample Against Intentionalism

In my objection to Necessity in the previous subsection, I began by noting that for normal humans, the belief B1 is a fitting unlearned doxastic response to ME1 and an unfitting unlearned doxastic response to ME2. Then I gave a Reidian example in which B1 is a fitting unlearned doxastic response to ME2 and an unfitting unlearned doxastic response to ME1. And I claimed that that example is a possible one. In short, I claimed that it is possible that B1 is a fitting unlearned doxastic response to ME1 and also that it is possible that B1 is a fitting unlearned doxastic response to ME2. Philosophers who think sensory experiences have propositional content would take this to suggest the following proposal having to do with the *contents of experience*:

CE: It's possible (and actual) for normal humans to have experience ME1 with the propositional content of B1; and it's possible for there to be cognizers like those in the Reidian counterexample who have experience ME2 with the propositional content of B1.

But do sensory experiences like ME1 and ME2 *have* propositional content? If 'experience ME1 has the same propositional content as B1' is just shorthand for something like 'ME1 is an experience to which its subject is naturally inclined to form B1 as an unlearned automatic doxastic response' then I'd agree with CE. But I'm not convinced that sensory experiences have propositional content.²⁸ I agree, of course, that they can spontaneously elicit beliefs (which *do* have propositional content) as doxastic responses. But I'm not convinced that that is the same thing as the sensory experiences themselves having propositional content. It seems to me that the view that an experience E has propositional content *p* is

²⁷ The arguments in this section establish the following point on which I relied in Chapter 3: (iv) *The appropriateness of an input is due to a contingent feature, not an essential feature, of that input.* Cf. n. 32 in Chapter 3.

Plantinga's argument (1993b: 54–63) that the *warrant*-conferring power of a belief ground is a contingent feature of it is similar to my objection to Necessity. The main difference (aside from the fact that he focuses on warrant and I focus on justification) is that his examples are less convincing than the Reidian one I employed in this subsection. One problem with Plantinga's examples is that many of them involve *learned* doxastic responses. This leads to the concern discussed in Section 1.2.2—a concern that the Reidian example, which focuses on *unlearned* doxastic responses, sidesteps. Another problem is that one of Plantinga's examples (1993b: 62–3) is met with understandable resistance by Bruce Russell (2001: 45–6) who seems to think that Plantinga's example involves treating a doxastic response that must be *learned* as if it can be an appropriate *unlearned* doxastic response. I think Plantinga's example can survive Russell's criticisms. But the Reidian example is better because it involves a belief that clearly *can* be an appropriate unlearned doxastic response.

²⁸ See Alston (2005b) for some objections to the view that they do.

grounded in nothing more than the observation that a person having E is naturally inclined to form a belief that p as an automatic response to it. However, for the sake of convenience, I will proceed as if in giving my Reidian counterexample to Necessity, I am proposing CE, since that will make it easier to consider an important objection from intentionalism.²⁹

Intentionalism is the view that the phenomenal character of experience is entirely determined by its propositional content. One consequence of intentionalism is:

CI: No two possible experiences that differ in phenomenal character have the same propositional content.

Given the obvious fact that ME1 and ME2 are experiences with different phenomenal character, CI entails that CE is false. Hence, intentionalism entails that my counterexample to Necessity is metaphysically impossible.

There is a significant literature debating the truth of intentionalism.³⁰ Obviously I can't do justice to it all here. What I will do instead is consider an impressive recent argument for CI and respond to it. This will go some way toward defending CE, on which my objection to Necessity depends. The argument for CI that I have in mind is proposed by Alex Byrne (2001). The argument speaks of an idealized person, which is a human-like person with powers of introspection, memory, and perception. An idealized person is one of whom the following three things are true: she is *competent* at accessing memories (i.e. when she tries to access memories she succeeds), she *tries* to detect whether there has been a change in the phenomenal character of her experience, and her memory is *perfectly accurate* (Byrne 2001: 208–9). Here then is the argument.³¹

Byrne's Argument for CI

1. Necessarily, if the phenomenal character of an idealized person's experience changes, that person notices a change in the phenomenal character of her experience.³²

²⁹ In conceding this for the sake of convenience, I'll insist that experiences, like sentences, have their propositional content contingently (though only in the latter case is this contingent upon human convention). However, I think of beliefs as having their propositional content essentially. Beliefs are individuated by their propositional content whereas experiences are individuated by their phenomenal character.

³⁰ Those in support of intentionalism (also called 'representationalism') include Byrne (2001), Dretske (1995, 1996), Harman (1990, 1996), Lycan (1996), and Tye (1995). Those with anti-intentionalist sympathies include Alston (2005b), Block (1990, 1994, 1996, 1998, 1999, 2003), Burge (1997), Chalmers (1996), Levine (1997, 2001), Peacocke (1983, 1984), and Robinson (1998).

³¹ The formulation is my own, based on Byrne (2001: 206–17).

³² Byrne notes that one might object that, at best, this is true only for non-negligible changes in phenomenal character (i.e. only for changes in phenomenal character the subject can reliably discriminate). To that Byrne (2001: n. 19) replies that non-negligible changes in phenomenal character can be viewed (on a perfectly reasonable precisified sense of 'phenomenal character') as the

2. Necessarily, if a person notices a change in the phenomenal character of her experience, then the way things seem to her changes—i.e. the propositional content of her experience changes.
3. Therefore, necessarily, if the phenomenal character of an idealized person's experience changes, the propositional content of her experience changes. (From 1 and 2)
4. Necessarily, if a change in the phenomenal character of an idealized person's experience entails a change in the content of her experience, then a change in the phenomenal character of a nonidealized person's experience entails a change in the content of her experience.³³
5. Therefore, necessarily, if the phenomenal character of a person's experience changes (whether the person is idealized or not), the propositional content of her experience changes—in other words, same experiential content entails same phenomenal character. (From 3 and 4)
6. If it's possible for two people to have experiences with different phenomenal character but the same propositional content, then it's also possible for a single person to have changes in the phenomenal character of her experience without any change in the propositional content of her experience.³⁴
7. Therefore, it's not possible for two people to have experiences with different phenomenal character but the same propositional content. (From 5 and 6)
8. If it's not possible for two people to have experiences with different phenomenal character but the same propositional content, then no two possible experiences that differ in phenomenal character have the same propositional content.³⁵

only changes there are in phenomenal character. Whether this is a satisfactory response is something I won't take up here.

³³ Here is Byrne's defense of 4 (2001: 214–16). Suppose a nonidealized person has experiences e and e^* which differ in phenomenal character. It seems possible for an idealized person also to have experiences e and e^* with exactly the same propositional content and phenomenal character as, respectively, e and e^* had by the nonidealized person. Now suppose that 3 is true (i.e. suppose that if two experiences had by an idealized person differ in phenomenal character, then they differ in content). Then it follows that e and e^* had by the idealized person differ in propositional content. And since the nonidealized person has the same experiences with same phenomenal character and propositional content, her experiences e and e^* differ in propositional content too.

³⁴ Here is Byrne's defense of 6 (2001: 216–17). Suppose one person has an experience with phenomenal character P and propositional content p and another person has an experience with a different phenomenal character Q and propositional content p . Given some sort of recombination principle for determining what is possible, it seems to follow that it's also possible for a third person to have both experiences consecutively: first the one with character P and content p and then the other with character Q and content p .

³⁵ This premise is not made explicit or defended by Byrne. But he clearly relies on it. For (Byrne 2001: 216–17) argues for the conclusion that 'if two subjects enjoy e and e^* , respectively, and if the contents of e and e^* are the same, then so are their phenomenal characters'. This is equivalent to 7

9. No two possible experiences that differ in phenomenal character have the same propositional content. (From 7 and 8)

The conclusion of this argument would, by my lights, be very surprising if true. What it says is impossible seems perfectly possible to me—namely, that two possible experiences (perhaps in different possible worlds) that differ in phenomenal character can each naturally incline their subjects to form, as automatic responses, beliefs with the same content. Since the reasoning is valid, this makes me suspect that at least one of the premises is false. In my response to the argument, I'll be focusing my attention on premise 2 (which is not to concede that the other premises are true).

The first thing to notice is that premise 2 can be read in two ways:

- 2*. Necessarily, if a person notices a change in the phenomenal character of her experience, then there is a change in the way things other than the phenomenal character of her experience seem to her.
- 2**. Necessarily, if a person notices a change in the phenomenal character of her experience, then *either* there is a change in the way the phenomenal character of her experience seems to her *or* there is a change in the way things other than the phenomenal character of her experience seem to her.

2** is tautological (and, therefore, highly plausible) but not of any value for Byrne's argument. 2* is the version Byrne needs for his argument but it isn't nearly as plausible, though I'm sure Byrne thinks it is plausible enough. The reason I say 2** isn't of any value for Byrne's argument is that the contents of experience he has in mind are not propositions about the phenomenal character of the subject's experience but instead propositions about something else such as the subject's environment (Byrne 2001: 201–3). A *perceptual* experience (which will be our focus) represents the subject's *environment* as being a certain way; *that* is its content. It doesn't represent *itself* as having a certain phenomenal character. Henceforth, 2 will be understood as 2*.

My objection to 2* is that it is plausible *only if* it is restricted to a certain kind of perceptual experience. It isn't true in the fully general way it is stated.

above. After defending this conclusion, he then goes on to 'restate' it on (2001: 217) as Conclusion C++: 'For any two possible experiences e and e^* , if they differ in phenomenal character, then they differ in content'. This is equivalent to 9 above. It appears, therefore, that he must be relying on 8. But consider the contrapositive of 8 which can be stated as 8*: *If (a) it's possible for a subject to enjoy e and it's also possible for a subject to enjoy e^* (where e and e^* differ in phenomenal character and yet have the same propositional content), then (b) it's possible for there to be two subjects, one of whom enjoys e and the other of whom enjoys e^* (where e and e^* differ in phenomenal character and yet have the same propositional content)*. 8* is of the form: If (a) A is possible and B is possible then (b) A&B is possible. And of course not all claims of that form are true (e.g. if A is p and B is $\sim p$). So it seems we would need some special reason to accept 8* (which is equivalent to 8). But Byrne doesn't give us any such reason. Indeed, he never even acknowledges that he is relying on 8. Whether this will ultimately cause a problem for Byrne's argument is something I won't consider here.

This is important because it shows that Byrne's argument is unsuccessful if its conclusion, CI, is not likewise restricted to the same sort of perceptual experience. Moreover, my Reidian counterexample to Necessity depends on the falsity of the *unrestricted* version of CI, not on the falsity of the restricted version. Consequently, my objection to Necessity is not harmed by Byrne's argument for CI and intentionalism.

I need to defend my claim that 2* is plausible only when restricted to perceptual experience of a certain kind. The kind of experience I have in mind is sensory experience which prompts belief in *secondary* (rather than primary) qualities of objects in our environment. Reid ([1785] 2002: 200–5) describes the difference between primary qualities (such as extension, divisibility, figure, motion, hardness, softness, solidity, and fluidity) and secondary qualities (such as sound, color, taste, smell, heat, and cold) as follows. Our sensory experience gives us a direct and distinct notion of primary qualities:

Every man capable of reflection may easily satisfy himself, that he has a perfectly clear and distinct notion of extension, divisibility, figure, and motion. The solidity of a body means no more, but that it excludes other bodies from occupying the same place at the same time. Hardness, softness, and fluidity, are different degrees of cohesion in the parts of a body. It is fluid, when it has no sensible cohesion; and hard when it is strong. Of the cause of this cohesion we are ignorant, but the thing itself we understand perfectly, being immediately informed of it by the sense of touch. It is evident, therefore, that of the primary qualities we have a clear and distinct notion; we know what they are, though we may be ignorant of their causes. (Reid [1785] 2002: 201)

By contrast, our sensory experience gives us only a relative and obscure notion of secondary qualities:

It is otherwise with secondary qualities. If you ask me, what is that quality or modification in a rose which I call its smell, I am at a loss to answer directly. Upon reflection I find, that I have a distinct notion of the sensation which it produces in my mind. But there can be nothing like to this sensation in the rose, because it is insentient. The quality in the rose is something which occasions the sensation in me; but what that something is, I know not. My senses give me no information upon this point. The only notion therefore my senses give is this, that smell in the rose is an unknown quality or modification, which is the cause or occasion of a sensation which I know well. The relation which this unknown quality bears to the sensation with which nature has connected it, is all I learn from the sense of smelling: but this is evidently a relative notion. The same reasoning will apply to every secondary quality. (Reid [1785] 2002: 202)³⁶

³⁶ Although Reid's account of the difference between primary and secondary qualities is similar to Locke's in ways, the two accounts are importantly different. One important difference is that Locke believed that our ideas of primary qualities *resemble* the qualities which cause them whereas our ideas of secondary qualities do not resemble the qualities which cause them. Reid, by contrast, was impressed by Berkeley's point that no idea or mental state can resemble a nonmental cause. From this Reid concluded that neither sensations caused by primary qualities of objects nor sensations caused by secondary qualities of objects resemble the qualities that cause them. See Reid ([1785] 2002: 207–10).

Thus, my sense of touch gives me a clear and direct (nonrelative) notion of the hardness of a billiard ball in my hand. (Of course, it doesn't provide a scientifically informed notion of what hardness is. But it does provide a grasp of hardness—i.e. the strong cohesion of the object's parts—that makes no reference to my sensory experience.) However, my sight gives me only an obscure and relative notion of a thing's color, one that *does* make reference to my sensory experience: a color is that (pre-scientifically unknown) quality in a thing which causes a certain color sensation in me. Notice that this distinction is an *epistemological* distinction concerning our untutored pre-scientific grasp of these properties, not a *metaphysical* distinction concerning a difference in their intrinsic nature.³⁷

In elaborating upon this distinction, Reid points out that the sensations prompting our beliefs in secondary qualities of things are naturally and easily made objects of our attention while those prompting our beliefs in primary qualities of things are not:

The first [secondary qualities] are not only signs of the object perceived, but they bear a capital part in the notion we form of it. We conceive it [the secondary quality] only as that which occasions such a sensation, and therefore cannot reflect upon it without thinking of the sensation which it occasions. We have no other mark whereby to distinguish it. . . . But having a clear and distinct conception of primary qualities, we have no need when we think of them to recall their sensations. When a primary quality is perceived, the sensation immediately leads our thought to the quality signified by it, and is itself forgotten. We have no occasion afterward to reflect upon it; and so we come to be as little acquainted with it, as if we had never felt it. . . . They [the tactile sensations producing beliefs in the primary qualities of objects] carry the thought to the external object, and immediately disappear and are forgotten. Nature intended them only as signs; and when they have served that purpose, they vanish. (Reid [1785] 2002: 204–5)³⁸

Reid's point here is that we have a difficult time attending to the tactile sensations that prompt our beliefs in primary qualities of objects and that we rarely make those sensations the object of our attention. This supports his conclusion that our notion of primary qualities, unlike our notion of secondary qualities, is not a relative notion concerning the sensation that quality produces in us.

³⁷ This should go some way toward tempering objections (to drawing a primary/secondary quality distinction) from those who think there is no *metaphysical* difference between primary and secondary qualities. Jennifer McKittrick argues that Reid's view of the metaphysics of secondary qualities of objects is that they are the 'causal basis of a disposition to produce a sensation' in us (McKittrick 2002: 488). And she notes in the same passage that Reid's view of the metaphysics of secondary qualities is similar to Frank Jackson's even though Jackson (2000) calls his view 'the primary quality view of color'. This helps us to see that Jackson's insistence that color is *metaphysically* similar to primary qualities doesn't count against Reid's point that our pre-scientific grasp of color (and other secondary qualities) is different from our pre-scientific grasp of the qualities typically identified as primary qualities.

³⁸ See also Reid ([1764] 1997: 55–8) for a similar discussion.

Let's consider now the implications of this for 2*. In particular, let's consider how 2* applies to experiences which have as their content propositions attributing secondary qualities such as colors to things. Given the truth of what Reid has said about our pre-scientific grasp of secondary qualities, a proposition saying of a thing in my environment that it is red says of that thing that it has a property which causes a red sensation in me.³⁹ Now consider a case of color-spectrum inversion which goes undetected because both subjects use the same words to refer to the same objects. For example, pointing to the same tomato, they both say of it that it is red, even though one person bases her judgment on what we would call a 'red sensation' whereas the other bases her judgment on what we would call a 'green sensation'. Do the experiences and judgments of these two people have the same propositional content? It seems not. The content of the experience and judgment of one person says the tomato has a quality which causes *this* sort of sensation (pointing inwardly to what we'd call a 'red sensation') while the content of the experience and judgment of the other person says the tomato has a quality which causes *that* sort of sensation (pointing inwardly to what we'd call a 'green sensation'). If we didn't know their color spectra were inverted relative to each other, we wouldn't know that their judgments and experiences had different contents, since we couldn't know which experiences they were inwardly pointing to (and so couldn't see that they were different). Even so, the difference in propositional content remains. The difference remains because their color spectra are inverted and the phenomenal character of their perceptual experiences plays a role (in the way Reid says) in their attribution of secondary qualities to things. Hence, when we consider 2* restricted to experiences with contents attributing secondary qualities, it seems to give the right result: if we notice a change in phenomenal character, then the content of the experience changes too.

But this doesn't show that the unrestricted version of 2*, on which Byrne's argument depends, is true. The considerations in the previous paragraph are silent concerning perceptual experiences with contents attributing *primary* qualities to things in our environments. Moreover, Byrne's discussion in defense of premise 2* (what he calls 'Premise B') focuses solely on experiences with contents attributing *color* to things (Byrne 2001: 210–11). My suspicion is that the reason Byrne is taken in by premise 2* is that his focus is narrowed in this way to such secondary-quality experiences. In order to test the truth of 2* in its fully general

³⁹ But is what Reid says about secondary qualities—concerning our grasp of them rather than concerning their intrinsic nature—true? I must say that I find it exceedingly plausible given how strongly supported it is by reflection on my own grasp of the qualities in question. For some excellent discussions of Reid's account of primary and secondary qualities—discussions which include defenses of Reid's view against criticisms as well as comparisons highlighting its advantages over other accounts—see Nichols (2003) and McKittrick (2002). See also Smith (1990) for a good survey of views on primary and secondary qualities and a defense (in Section VI) of a position similar to Reid's.

form, we need to consider its plausibility when applied to perceptual experiences with contents attributing *primary* qualities to things.

In order to facilitate our discussion, let us consider three different kinds of creature (two of which aren't possible according to intentionalism). Each of the creatures I have in mind is a perceptual cognizer resembling us in important ways: each is similar to us in physical appearance, each makes perceptual judgments based on sensory experience, and each is able to reason and form beliefs and draw conclusions. Now suppose that unlike us, these three creatures have no vision, no sense of taste, and no sense of smell. (Let's suppose they differ from us in outward appearance only by having no eyes or noses.) The first kind of creature has perceptual abilities of hearing and touch that operate in the same way ours do. Hence, the sensory experiences one of them has when hearing and touching have the same phenomenal character as our own when we hear and touch.

The second kind of creature makes the same perceptual judgments as the first kind of creature, when in the same circumstances. In particular, when a creature of this second kind makes judgments based on *touch*, they're the same (when in the same circumstances) as those of the first creatures. The only difference is that the phenomenal character of the sensory experience had by this second type of creature when she physically touches things is not at all like our tactile experience. Instead, it is like our olfactory experiences. (In this way, these creatures are like the cognizers in the Reidian example given in the previous subsection.)⁴⁰

The third kind of creature is the one that is most interesting for our purposes here. A creature of this kind also makes the same judgments based on touch (when in the same circumstances) as those made by creatures of the first kind and the second kind. Sometimes when grabbing a billiard ball, she experiences ME1 just as the first kind of creature does. On such occasions, her natural, unlearned inclination is to form a belief like B1 (again, just as it is for the first kind of creature). We may conclude from this that her experience, ME1, has the same propositional content as B1. However, on other occasions, when grabbing a billiard ball, she experiences ME2 (what we would call an 'olfactory experience'), just as the second kind of creature does. On such occasions, the natural, unlearned inclination of this third kind of creature is the same as it is for creatures of the second kind: it's to form a belief like B1. We may conclude from this that her experience, ME2, has the same propositional content as B1. What determines

⁴⁰ Notice, by the way, that for this second kind of creature, a sensory experience with the same phenomenal character as our olfactory sensations prompts a belief in *primary* qualities of objects whereas that type of sensory experience in us prompts a belief in *secondary* qualities of objects. Above I quoted and discussed a passage from Reid ([1785] 2002: 204–5) in which he notes that we rarely attend to sensory experiences that prompt beliefs in primary qualities of objects whereas it's the norm for us to focus our attention on sensory experiences that prompt beliefs in secondary qualities of objects. In light of this, we can assume that, although we make our olfactory sensations the focus of our attention, this second kind of creature rarely attends to its sensory experiences that have the same phenomenal character as our olfactory sensations.

whether grabbing the billiard ball results in her having ME1 or ME2? Let's say it's due to some indeterministic physical process that occurs when she grabs a billiard ball, with the result that about half the time she experiences ME1 (when she grabs a billiard ball) and half the time she experiences ME2. Let's also add that all of her activities of touching things have similar indeterministic results: half the time they result in what we'd call 'tactile experiences' and the other half they result in experiences we'd call 'olfactory experiences' (though the same physical stimulation is always correlated with one or the other member of the same pair of sensory experiences).

Now, the key question is whether it's possible for both ME1 and ME2 to have the same content for a single creature. It seems to me that it is possible. It seems clearly possible for the first kind of creature to form B1 as a natural and unlearned response to ME1. It also seems possible for the second kind of creature to form B1 as a natural and unlearned response to ME2. This leads me to think it's also possible for the third kind of creature to form B1 as a natural and unlearned response to ME1 and also as a natural and unlearned response to ME2. We can add that the creature notices the difference in phenomenal character of ME1 and ME2. But though she notices this, her natural unlearned inclination to form judgments about her environment is the same in each case. Admittedly, this would seem odd to you and me (to have a change in phenomenal character of our experience without a change in the judgments we're naturally inclined to form about our environments). But it doesn't seem odd in the least to her since this is how it has always been for her.

Now suppose that while grabbing a billiard ball, her experience changes from ME1 at t_1 to ME2 at t_2 . And suppose we ask her if her environment seems different at the two times (we can stipulate that no other experiential changes have taken place in that interlude). Then she will say that her environment seems exactly the same to her at t_2 as it did at t_1 . She might add that the phenomenal character of her experience has changed; but even so, she'll insist that the world as represented to her by her experience didn't seem to change at all. Again, this will strike us as odd because this sort of thing doesn't happen to us.⁴¹ But it won't seem odd to her. Nor does it seem to be impossible.

Notice that because B1 attributes primary qualities to things in her environment, there isn't the problem noted earlier with secondary qualities since, in the case of primary quality attribution, there's no reference to the phenomenal character of the subject's experience. So we haven't got *that* reason to think it impossible for the content of ME1 and ME2 to be the same. Do we have any other reason? I don't see that we do. Byrne, in defending premise 2*, says 'the subject can only discover the phenomenal character of her experience by attending to the world . . . as her experience represents it' (2001: 211). The idea is that we couldn't

⁴¹ I assume that this oddness plays a role in Byrne's conclusion that it is impossible. See his discussion in Byrne (2001: 206–7).

find out that the phenomenal character of our experience had changed if what it represented didn't change. But this seems flatly false, especially in light of the example we're considering. If my natural and unlearned inclination in response to both ME1 and ME2 is to form B1, then it seems that both of those sensory experiences have the same content—it seems that both represent the world in the very same way. But this doesn't prevent me from noticing that ME1 and ME2 differ in their phenomenal character.

I see no reason to think it's impossible for there to be a creature of this third kind who operates in the way described above. And if it is possible, then premise 2* from Byrne's argument is false. I conclude, therefore, that Byrne's argument for CI fails and that it presents us with no good reason to think my Reidian counterexample to Necessity fails.

1.3 The Second Improvement: Add Proper Function

Having rejected Necessity, we must return to the question that prompted our discussion of it, namely: 'If objective fittingness doesn't depend on reliability, what does it depend on?' The counterexample to Necessity considered in Section 1.2.3 suggests that the fittingness of a doxastic response depends, in some cases at least, on the species of the cognizer who has it. What is it about the species of a cognizer that determines such fittingness in those cases? An answer that immediately suggests itself is that what makes a belief a fitting unlearned doxastic response to an experience has to do with the way the cognitive faculties of the cognizer in question are *supposed to function*. For clearly that is something that can vary from species to species. *Our* cognitive faculties are supposed to function so that when we experience tactile sensation ME1, our unlearned doxastic response is B1. Not so the cognizers described in the Section 1.2.3. *Their* faculties are, we may assume, supposed to function so that when those cognizers experience 'olfactory' sensation ME2, their unlearned doxastic response is B1. The sense in which our faculties and theirs are *supposed to function* in the ways just specified is the same as the sense in which our hearts are supposed to function so that they beat less than 200 times a minute when we are at rest. And the 'supposed to' of heart function is clearly connected with the notion of proper function or healthy function. This suggests that the fittingness of a doxastic response to evidence is contingent upon the *proper function* of the cognitive faculties of the person in question. And this, in turn, suggests that the evidentialist claim E_F could be illuminatingly improved if it were changed to say something like the following (where a *PF-induced* doxastic response is one produced by the proper functioning of the subject's cognitive faculties):

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

That we replace E_F with E_{PF} is my second recommendation. It amounts to the suggestion that epistemic fittingness should be understood at least partly in terms of proper or healthy cognitive functioning.

1.4 The Third Improvement: Drop Mentalist and Internalist Implications

E_{PF} says that a belief is justified only if it is a PF-induced response to S 's *evidence*. As I noted earlier, this could suggest that some form of mentalism or internalism is true. For if evidence is something to which I have special epistemic access, then the proponent of E_{PF} will answer the question:

- (A) What sorts of things can be appropriate inputs to our belief-forming processes?

by saying that only states that are accessible on reflection alone can be appropriate inputs. But to say this (while asserting E_{PF}) is to impose an awareness requirement on justification. And I've argued at length in Part I of this book against that position. If, on the other hand, evidence must be a mental state of the subject, then the proponent of E_{PF} will answer question (A) by saying that only mental states can be appropriate inputs. But there seems to be no good reason for that answer to question (A). Moreover, as I noted in Chapter 3, there are examples of possible scenarios in which beliefs produced without any mental states in their causal ancestry seem to be justified beliefs.⁴² These examples give us a reason to *reject* the mentalist answer to (A).

In order, therefore, not to give any hint of endorsing a form of internalism or mentalism, E_{PF} should be replaced with:

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

This is my third recommendation. The only restrictions I_{PF} places on what counts as an appropriate input to one's belief-forming processes are (a) that the input is something to which the belief is a (causal) response and (b) that the input is something to which the subject is *supposed to* have the doxastic response in question (where the 'supposed to' has to do with proper function).⁴³

We began with the admittedly attractive E_F understood in accord with Nonreliability, Objectivity, and Necessity. Call that package 'the evidentialist position'.

⁴² See Section 1.3.4 of Chapter 3.

⁴³ In Bergmann (2004b: 43) I said that what count as candidates for appropriate inputs are all things that Pollock and Cruz call 'internal states' (see Section 2.1.1 of Chapter 3 where I explain what Pollock and Cruz say about internal states). I am now less confident that I know enough about what Pollock and Cruz have in mind by 'internal states' so I've decided not to use their term to explain my position.

After my recommended improvements, we now have I_{PF} understood in accord with Nonreliability, Objectivity, and the denial of Necessity. Call that package ‘the proper function position’. Clearly the proper function position is significantly different from the evidentialist position. But despite this, it retains the attractiveness of the evidentialist position in addition to having the advantages of the three recommended improvements discussed above. The reason it retains the initial attractiveness of the evidentialist position is that things counting as evidence are perhaps the most prominent sort of input to *our* belief-forming systems. Consequently, there will be a wide range of *standard* cases in which the different implications of the evidentialist position and the proper function position are not noticeable. The result is that the *prima facie* plausibility of the evidentialist position, which arises from a consideration of these standard cases, carries over to the proper function position.

2 A PROPER FUNCTION ACCOUNT OF JUSTIFICATION

In Section 1 I argued that the initially appealing evidentialist position is made more plausible if it is improved to become the proper function position which denies Necessity and replaces E_F with I_{PF} . In this section I will lay out and defend an analysis of justification that can be viewed as a more careful statement of the idea behind I_{PF} .

2.1 The Proposed Analysis

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

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

- (i) the cognitive faculties producing B are *functioning properly*
- (ii) the cognitive environment in which B is produced is sufficiently similar to the one for which S’s faculties were ‘designed’⁴⁴
- (iii) the modules of the ‘design’ plan governing the production of B are directly aimed at the production of *true* beliefs
- (iv) there is a high objective probability that a belief formed in accord with those modules in that sort of cognitive environment is true.⁴⁵

⁴⁴ The quotation marks are to indicate that the design in question needn’t involve a literal designer. For discussion of how to make sense of design talk without postulating a literal designer, see Section 3.1 below.

⁴⁵ This is *not* a direct quotation; for his exact wording, see Plantinga (1993a: 19, 46–7). Plantinga insists that this is still only a first approximation. For further details, qualifications, and emendations see Plantinga (1993a, 1996, 1997, 2000: 156–61). In particular, see Plantinga (1996,

He adds condition (iv) because he recognizes that the first three conditions don't entail that the belief was reliably formed and he thinks warrant should entail reliability.

A similar line of reasoning leads me to think he should add a fifth condition requiring the absence of a defeater:

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

The next chapter will be devoted to discussing defeaters in general and this condition in particular. For now I'll say only that just as it seems that warrant should entail reliability and that proper function conditions don't require reliability, so also it seems that warrant should entail the absence of a defeater and that proper function conditions don't do that. There could be cognizers designed by a literal creator to form beliefs in a reliable way but also to take each of their reliably formed beliefs to be defeated. Furthermore, they could have been designed so that although they took their beliefs to be defeated, they ignored this fact and continued to hold them. Such beliefs would *not* be warranted even if they satisfied Plantinga's four conditions.⁴⁶

My proposed account of justification is just my modified version of Plantinga's account of warrant (i.e. his account together with the fifth condition I just mentioned) *without* Plantinga's condition (ii)—the environmental condition. Once that condition is dropped, his account ceases to be reliability-entailing. Thus, to a first approximation,⁴⁷ we may say that:

J_{PF}: S's belief B is justified iff (i) S does not take B to be defeated and (ii) the cognitive faculties producing B are (a) functioning properly, (b) truth-aimed and (c) reliable in the environments for which they were 'designed'.⁴⁸

To see how this analysis is a refinement of I_{PF}, it's important to notice that clause (ii) is to be understood so that it entails that S's doxastic response to the

2000) where he notes that, in addition to the conditions (i)–(iv) noted above, we need to add a fifth condition called the 'resolution condition'. The purpose of this resolution condition is to guarantee that warrant entails truth, thereby enabling it to handle Gettier problems by ruling out cases of warranted belief that is *accidentally* true. On the question of whether warrant entails truth, see Merricks (1995).

⁴⁶ Plantinga says that warrant depends on the proper functioning of our defeater systems (1993a: 41). So he is aware of the concern I'm raising. However, what he doesn't acknowledge is that just as his conditions (i)–(iii) don't entail the satisfaction of a reliability condition, so also his conditions (i)–(iv) don't *entail* the satisfaction of condition (v).

⁴⁷ A more careful account would also explain how justification comes in degrees. For an attempt to use a proper function analysis to explain how *warrant* comes in degrees, see Plantinga (1993a: 7–9). Something similar could be said, *mutatis mutandis*, using the J_{PF} account of *justification*.

⁴⁸ Clauses (ii)(b) and (ii)(c) of J_{PF} are shorthand for their correlates in Plantinga's account of warrant given above. Thus, to say that the faculties producing B are truth-aimed is to say that 'the modules of the design plan governing the production of B are directly aimed at the production of *true* beliefs'. And to say that the faculties producing B are reliable in the environment for which they were designed is to say that 'there is a high objective probability that a belief formed in accord with those modules (i.e. those modules of the design plan governing the production of B) in that sort

input to her belief-forming systems is due to the proper functioning of S's faculties. As Plantinga points out (1993a: 22–4), when we say cognitive faculties are functioning properly, the basic idea is that their functioning results in the cognitively healthy doxastic response to the circumstances in which they are operating (which will include, rather prominently, the input to the subject's belief-forming systems).

The motivation for clause (i) of J_{PF} was discussed briefly above and will be discussed at length in the next chapter. The motivation for clause (ii)(a) of J_{PF} —the proper function condition—was discussed in Section 1 where I argued against Necessity and in support of understanding fittingness in terms of proper function. But nothing said so far addresses the motivation for adding clauses (ii)(b) and (ii)(c). So why does J_{PF} require that the belief be produced by *truth-aimed* cognitive faculties that are *reliable* in the environments for which they were 'designed'? Why not require, in addition to the absence of any believed defeater, merely that the belief be produced by *properly functioning* cognitive faculties?

The answer here is very much like the explanation Plantinga (1993a: 11–20, 26–8) gives for including clauses like (ii)(b) and (ii)(c) in his account of *warrant*. The main difference is that when Plantinga adds such clauses to his account of warrant, they entail that the belief is produced by a cognitive faculty that produces mostly true beliefs whereas they do not have that consequence when added to my account of justification. The reason that the addition of such clauses doesn't make my account reliability-entailing is that I've dropped the 'right environment' clause included in Plantinga's account. J_{PF} , therefore, is not a *reliabilist* account of justification despite the fact that it includes a clause having to do with reliability. Because it isn't a reliabilist account, it is able to avoid the standard evil demon objection to reliabilist accounts of justification. According to that objection, the fact that victims of evil demons could have justified beliefs despite the fact that their beliefs are unreliably formed shows that reliabilist accounts of justification are mistaken to say that reliability is necessary for justification.⁴⁹

To return, then, to the question at hand, let me explain why I've added clauses (ii)(b) and (ii)(c) which have to do with the cognitive faculty in question being truth-aimed and reliable in the right environment. My inclusion of clause (ii)(b), requiring that the faculty be truth-aimed, is due to the following sort of counterexample to the suggestion that clauses (i) and (ii)(a) are sufficient for justification. Suppose that the cognitive faculties of some cognizer were created by a being

of cognitive environment is true'. See Plantinga (1993a: 22–4) for an account of how to think of design plans in terms of the circumstance–response pairs (or the circumstance–response–purpose triples) they dictate. And see Plantinga (1993a: 38–40) for a discussion of the difference between being directly aimed at truth and being indirectly aimed at truth.

⁴⁹ The remarks in this paragraph establish the following point on which the argument of Chapter 3 relied: (iii) *Unlike reliabilism, the proper function account of justification handles the intuition that a belief can be justified even though formed in an unreliable way.* Cf. n. 32 in Chapter 3.

interested in that cognizer's comfort. But suppose also that the particular faculty producing the belief we are considering is intended by its creator not to produce *true* beliefs but, rather, to produce beliefs that will minimize psychological trauma (even if that involves regularly producing false beliefs). Then it seems that beliefs being produced by such a cognitive faculty won't be *epistemically* fitting responses to the input to the subject's belief-forming system though they may be appropriate in some other sense. This would be a case of a belief that isn't justified (since it isn't an epistemically fitting doxastic response) even though it is produced by a properly functioning cognitive faculty.⁵⁰

Or suppose some evil demon intentionally designs creatures to form mostly false beliefs. For example, suppose that this malevolent creator decides to create Vic in such a way that his natural unlearned doxastic response to ME2 (the 'olfactory' experience) is to form the billiard ball belief B1. Vic's cognitive faculties are, therefore, designed to function in much the same way as are the faculties of the possible cognizers described in my Reidian counterexample to Necessity in Section 1.2.3. The difference is that the beliefs of those cognizers are mostly true if their cognitive faculties are functioning properly in the environment for which they were 'designed' whereas Vic's beliefs (including B1) are not. Now we noted earlier that the beliefs of those possible cognizers mentioned in 1.2.3—beliefs like B1—are (epistemically) fitting responses to their evidence. But consider Vic who has been designed (by the demon who is intent on having Vic form mostly false beliefs) to form the false billiard ball belief B1 in response to the 'olfactory' experience ME2. Is Vic's belief B1 an *epistemically* fitting response to ME2? It seems not. It's true that we can't blame Vic for holding the belief. And his epistemic blamelessness might give his belief a *subjective* sort of epistemic justification. But that won't be enough to make his belief *objectively* epistemically fitting. Nor will the fact that he was created by a designer who *intended* that he respond in this non-truth-conducive way to his experiences. To avoid counterexamples like those given in this paragraph and the preceding paragraph, it is important to add clause (ii)(b) to J_{PF} requiring that the belief be formed by a properly functioning faculty that is *aimed at truth*.

To see the importance of clause (ii)(c), consider a creature designed by one of Hume's infant deities. And suppose that, although this incompetent creator was trying to make a believer with reliable faculties, it instead created one whose faculties produce mostly false beliefs when placed in the environment in which it was intended by its creator to produce true beliefs. For example, suppose this infant deity intentionally created Ric in such a way that, like the cognizers described in Section 1.2.3, his natural unlearned doxastic response to ME2 is

⁵⁰ Overly optimistic beliefs about chances of recovery in response to evidence of one's poor health are often suggested as examples of beliefs that might be the result of proper function even though they aren't epistemically fitting doxastic responses to the available evidence. Such beliefs might aid recovery, perhaps by reducing stress. They would, therefore, serve a purpose even if they aren't aimed at truth.

to form B1. But, contrary to what this bumbling creator had hoped, this belief (like most of the other beliefs Ric forms) is false when produced in the intended environment by faculties functioning as they were designed to function. Is Ric's belief B1 an epistemically fitting response to ME2? It seems not. And the same would be true if the infant deity had designed a more human-like creature to form B1 in response to ME1 rather than ME2, intending (but utterly failing to bring it about) that such a belief is true when formed in that way in that creature's intended environment. Despite the failed design attempt, the belief in question may be, in some sense, the output of properly functioning cognitive faculties (since those faculties are, in some sense, operating in the way they were designed to operate).⁵¹ But because the plan, when implemented as intended, didn't work, the belief doesn't seem to be an epistemically fitting response to the subject's evidence.

What the considerations in the previous paragraph suggest is that beliefs produced by properly functioning truth-aimed cognitive faculties do not result in epistemically fitting doxastic responses to one's evidence *if* the faculties in question aren't *successfully* aimed at truth—i.e. if they aren't likely to produce true beliefs when operating in the environment for which they were 'designed'. Hence, clause (ii)(c) of J_{PF}.

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

⁵¹ See Plantinga (1993a: 26–8) for a related discussion of two senses of working in accord with a design plan.

hence, justified, despite the fact that the experiences in question are illusory ones produced by the demon. However, when the demon victim is designed and created by the demon specifically to hold unreliably formed beliefs, things are different. For it is no longer clear that her beliefs are *epistemically* fitting responses to her evidence, even if there is some sense in which they are doxastic responses she is *supposed to* have.⁵²

2.2 The Virtues of the Analysis

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

2.2.1 Handling of Cases

One virtue of J_{PF} is that it handles certain examples better than reliabilism or Necessity-endorsing internalism handles them. Some of these examples involve the reliable cognizers described in Section 1.2.3 whose natural unlearned response to ME2 is B1 and whose natural reaction to grabbing a billiard ball is to experience ME2. I'll call the complete description of the natural way of functioning for those alien cognizers 'design plan B'. And I'll call the complete description of our natural way of functioning 'the human design plan'.⁵³

Consider the following the six cases (in each case, the doxastic response is unlearned):

Case I: A human forms B1 in response to the tactile sensation ME1. B1 is a reliably formed belief produced by properly functioning faculties in an appropriate environment.

Case II: A human forms B1 in response to ME1. However, since this human is a demon victim whose perceptual beliefs are all artificially produced, ME1 is not produced by actual contact with a billiard ball (nothing like such contact occurs for this person). The result is that B1 is not reliably formed.

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

⁵² What about a case where an incompetent demon sets out to create a cognizer that holds unreliably formed beliefs but bumbles things so badly that the result is a duplicate of an ordinary human in a standard human environment? For a discussion of examples like this one, see section 3.4 at the end of this chapter.

⁵³ As Plantinga points out (1993a: 22–4), we can think of 'design' plans as sets of circumstance–response pairs (or circumstance–response–purpose triples).

Case IV: A nonhuman cognizer with ‘design’ plan B forms B1 in response to ‘olfactory’ sensation ME2. This belief is a reliably formed belief produced by that cognizer’s properly functioning faculties in an appropriate environment for the operation of those faculties.

Case V: A nonhuman cognizer with ‘design’ plan B forms B1 in response to ME2. This belief is a properly functioning unlearned doxastic response for such a creature to ME2. However, since this cognizer is a demon victim whose perceptual beliefs are all artificially produced, ME2 is not produced by actual contact with a billiard ball (nothing like such contact occurs for this cognizer). The result is that B1 is not reliably formed.

Case VI: Due to cognitive malfunction caused by a radiation overdose, a non-human cognizer with ‘design’ plan B comes to have the tendency to form B1 in response to ‘tactile’ sensation ME1. The same overdose also prevents her from ever noticing anything wrong with forming this belief in this way. Since in this environment (which is an appropriate one for this cognizer) ME1 isn’t a reliable indicator of the truth of B1, the result is that B1 is not reliably formed.

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

We can summarize all six cases as follows:⁵⁴

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

⁵⁴ The six cases can also be described so that the belief ground in each case includes an *unlearned* connector (see Sections 1.2.1 and 1.2.2 of this chapter for an account of what an unlearned connector is). Just replace all appearances of ‘ME1’ (in the above descriptions of the six cases as well as in the table below) with ‘ME1 + C1’. Likewise, replace all appearances of ‘ME2’ with ‘ME2 + C2’. In the demon cases, the unlearned connector in question is caused by the demon; in the malfunction cases, the unlearned connector is caused by the radiation overdose.

Reliabilist accounts get cases II and V wrong. They say the beliefs should be unjustified because they are unreliably formed. Internalist accounts endorsing Necessity get cases IV, V, and VI wrong. For the belief ground in cases IV and V is the same as the belief ground in case III and yet, contrary to what such internalist accounts entail, B1 is *not* unjustified in cases IV and V even though it *is* unjustified in case III. Likewise, despite the fact that the belief ground in case VI is the same as the belief ground in cases I and II, B1 is justified in cases I and II but not in case VI. By contrast, J_{PF} gets all six cases right.⁵⁵

I noted in Chapter 3 that Feldman and Conee defend their evidentialist position by describing six pairs of examples their position handles accurately.⁵⁶

Example Pair 1: Bob and Ray both believe that it will be warm today. Ray bases his belief on reading yesterday's newspaper in an air-conditioned hotel lobby while Bob bases his belief on his doing that same thing together with his walking outside and feeling the heat. Bob's belief is more justified than Ray's.

Example Pair 2: Consider these two stories about Ray getting further information. Both begin with Bob, who is a pillar of integrity, coming in from the heat and telling Ray that it felt hot outside. The first story says that Ray thinks it just as likely that Bob is deceiving him as that he is telling the truth. The second story says that Ray has excellent reasons to think Bob is a pillar of integrity. Bob's belief about it being warm is more justified in the second story than in the first.

Example Pair 3: A novice bird watcher and an expert are together looking at a woodpecker and have (we'll stipulate for simplicity) the same visual experience of it. Both believe it is a woodpecker. But the expert has fully reasonable beliefs about what woodpeckers look like and the novice doesn't (nor does he have any other evidence—such as the expert's testimony—for thinking this is a woodpecker). The expert's belief is justified and the novice's is not.

Example Pair 4: A logic TA and a beginning logic student both are looking at the same sentence—one they both know is true. They are also considering a second sentence and they both believe it too is true. The logic TA can easily tell that the second sentence is logically equivalent to the first (this is why he believes it too) but the beginning logic student cannot and has no other information pertaining to the truth or falsity of the second sentence (not even a reason to think the TA accepts the second sentence). The logic TA's belief that the second sentence is true is justified whereas the beginning logic student's belief in that same claim is not.

⁵⁵ Given how easy it is to multiply examples like cases IV, V, and VI, the remarks in this paragraph establish the following point on which the argument of Chapter 3 relied: (ii) *There are many examples which are better handled by the proper function account of justification than by their mentalist view.* Cf. n. 32 of Chapter 3.

⁵⁶ See Conee and Feldman (2001: 236–8). For the sake of brevity, the descriptions given here use my wording, not theirs.

Example Pair 5: Smith believes, on the basis of good evidence presented to him by his officemate Jones, that Jones owns a Ford and that someone in the office owns a Ford (the latter conjunct is deduced by Smith from the first conjunct). In fact, the evidence was deceptive: Jones doesn't own a Ford though, unbeknownst to Smith, someone else in the office does. Later Smith comes to find out that Jones was only pretending to own a Ford but he persists in believing (now without any good evidence) that Jones owns a Ford and that someone in the office owns a Ford (again, the latter conjunct is deduced from the first conjunct). Smith's belief before the discovery that Jones is pretending is justified and Smith's belief after the discovery is unjustified.

Example Pair 6: Consider these two stories about Hilary, an envatted human brain recently abducted from a fully embodied ordinary environment. In both stories, his life now seems normal and he is artificially given the false memory that he had oatmeal for breakfast yesterday. In the first story, the artificial memory is faint, lacking in detail, and it seems to him an unlikely breakfast for him to have had. In the second, the memory is vivid and it seems a typical breakfast for him to have had. The former belief is unjustified and the latter belief is justified.

The first five pairs of examples each include one member that is basically like case I above: it is a reliably formed belief produced in accord with proper function in response to the subject's main evidence and, at least in some cases, a connector.⁵⁷ The other member in each of these first five pairs of examples is either less justified or not justified at all. And in each case this is due either to the absence of some main evidence (example pair 1), the absence of a connector (example pairs 2–4), or the addition of a disconnector (example pair 5). There is no reason to think a proper function account of justification would have any difficulty handling any of these cases. In each case, the justificatory difference would be due, at least in part, to the fact that the doxastic response is no longer in accord with proper function when the main evidence or connector is dropped or a disconnector is added.⁵⁸ In the last of Conee and Feldman's six pairs of examples, the belief is not reliably formed because it is produced in a brain in a vat via artificially induced experiences. But it is a human brain that is still responding to experience as a properly functioning human brain does. The well-justified member of this pair is like case II above. The other member is not well-justified due to the fact that the main evidence is different and there is a disconnector present. Although a reliabilist account of justification might have difficulty with this last

⁵⁷ See n. 54 above where I explain how my six cases can be read to include connectors.

⁵⁸ In Conee and Feldman's first two examples, the problematic member of the pair is still justified, but it is less justified. This is because the more strongly justified belief in each pair has two sources of justification and, in the case of the less justified belief, only one of these sources is missing (due to the absence of the main evidence or a connector). JPF has no difficulty with appealing to the fact that the second source of justification is missing in order to explain the justification differences in these cases.

pair of examples (since it includes a case of an unreliably formed belief that seems justified), J_{PF} will have no trouble explaining the justificatory differences between the members of this pair.⁵⁹

J_{PF} differs from a reliabilist account of justification in at least two ways. First, unlike reliabilism, J_{PF} denies that reliability is *necessary* for justification. Because of that difference, J_{PF} handled cases II and V above better than reliabilism did. Second, unlike reliabilism, J_{PF} denies that reliability is *sufficient* for justification. It is for this reason that proponents of J_{PF} needn't be embarrassed by the examples of strange but reliable ways of forming beliefs that are so often employed against reliabilism. BonJour's examples of reliable clairvoyants which he uses against externalism won't work against J_{PF} unless it is stipulated that the cognizers in question form these clairvoyant beliefs when their faculties are functioning properly. But once that is stipulated, we have to admit that their clairvoyant beliefs are no more strange than our a priori or memory or perceptual beliefs. And then we lose the intuition that the beliefs in the example aren't justified. Similar remarks apply to Jonathan Vogel's (2000: 612 f.) example of a 'gas tank' belief. In this case, a person forms beliefs about how full the gas tank is by looking at the gauge. The gauge is a perfectly accurate indicator of how full the tank is but the person has no reason whatsoever to think so. It's difficult to take seriously the stipulation that the person has no reason at all to think the gauge is a reliable indicator of the amount of gas in the tank. But if we make an effort to take that stipulation seriously and to ponder the example as described, then we have a good reason to agree with Vogel that the beliefs so produced aren't justified. But this doesn't count against J_{PF} , even if it does count against reliabilism. For our assumption is that the person forming these beliefs is like us insofar as the proper functioning of her faculties does *not* dictate that she form beliefs about the fullness of the gas tank based (in an unlearned way) on what the gauge reads. Thus, J_{PF} handles these cases used to show reliability isn't sufficient for justification better than reliabilism handles them.

2.2.2 Justification and Truth

Another virtue of J_{PF} is that it provides an account of the justification–truth connection—one that doesn't lead to the usual troubles. Stewart Cohen (1984) has argued that although there seems to be a nontrivial connection between justification and truth, there seems to be no good account of that connection. He considers and, for good reasons, rejects two types of accounts: (i) reliabilist accounts, according to which there is an objective connection between justification and

⁵⁹ The remarks in this paragraph establish the following point on which the argument of Chapter 3 relied: (i) *The pairs of examples Conee and Feldman use in support of their mentalist view are handled easily by the proper function account of justification.* This, together with nn. 27, 49, and 55 in this chapter, completes the book-keeping that was promised in n. 32 of Chapter 3.

truth, and (ii) subjectivist accounts that say that what matters is that there is a justification–truth connection *from the subject's perspective*.⁶⁰

But there are several other accounts he doesn't consider—understandably since they hadn't been proposed at the time. One is J_{PF}. According to it, the connection between justification and truth comes through the notion of an appropriate environment: there is a high objective probability that a justified belief will be a true belief *if* the properly functioning faculties that produce it are operating in the environment *for which they were 'designed'*. This explains the connection between justification and truth while avoiding the evil demon counterexamples to reliabilism. It also explains our intuitions that the beliefs of the nonhuman alien cognizers with design plan B (discussed earlier in this chapter) are justified when formed in accord with proper function (and in a reliable way) in the environment for which they were 'designed', despite the fact that their way of forming beliefs is unreliable in our environment.

Another account is Sosa's virtue reliabilism, as it is developed in some of his earlier work. According to that early development of his position, 'justification is relative to environment' (Sosa 1991: 144). Thus, the demon victim's beliefs are justified relative to *our* environment because the virtuous dispositions producing her belief are reliable in our environment; but those same beliefs aren't justified relative to the victim's environment because the dispositions in question aren't reliable in the victim's environment. Presumably something similar can be said about the alien cognizers: their beliefs are justified relative to their environment but not relative to our own environment.⁶¹ But the advantage of J_{PF} over this version of Sosa's virtue reliabilism is that the proponent of J_{PF} can recognize a nonrelativized concept of justification whereas this position of Sosa's cannot. But it seems that we do have the notion of justification *simpliciter*. This is important because although it is natural to think that the beliefs of human victims of an evil demon aren't reliably formed in their environments, it isn't very natural at all to think of their beliefs as *unjustified relative to their environments*. Likewise, it isn't natural to think that the alien cognizers' beliefs are *unjustified relative to our environment*. Those relativized concepts of justification don't seem to be ones we have in our pretheoretical repertoire of concepts. What we think of the demon victims and the alien cognizers is just that their beliefs exemplify justification *simpliciter*.⁶² The problem is that Sosa needs, but doesn't seem to have, a principled way of selecting *the* environment in which reliability matters for nonrelativized justification. J_{PF} *has* a principled way of doing that: the environment that matters is the one for which the cognizer's faculties were 'designed'.

⁶⁰ Cohen (1984: sect II) objects to reliabilist accounts of justification on the grounds that they fail to attribute justification to demon victims. He also objects (Cohen 1984: sect III) to subjectivist accounts on the grounds that they implausibly require for justification what we rarely have, namely, further beliefs about the reliability of our beliefs.

⁶¹ This view is developed in Sosa (1991: 140–4, 284–90).

⁶² For a similar concern about views like Sosa's, see Goldman (1992: 161–3).

A third account not considered in Cohen's 1984 paper is Sosa's more recent view (BonJour and Sosa 2003: 156–61) according to which a belief is *adroit-justified* if the method by which it is formed is reliable in the actual world and it is *apt-justified* if that method is reliable in the subject's world. Thus, the demon victim's beliefs are adroit-justified but not apt-justified whereas the alien cognizer's beliefs are apt-justified but not adroit-justified. (Our own beliefs are justified in both senses.) What seems implausible about this position of Sosa's is that, according to it, there is no single sense of 'justification' in which our beliefs, the demon victims' beliefs, and the alien cognizers' beliefs are all justified. So, in order to explain our intuitions that all are justified, Sosa must attribute to us an unnoticed switch between using the term 'justified' to mean apt-justified and using it to mean adroit-justified. But it seems doubtful that this is what is going on when we make such justification attributions. According to J_{PF}, on the other hand, there *is* a single sense of 'justification' in which our beliefs, the demon victim's beliefs, and the alien cognizers' beliefs are justified.

A fourth account is proposed by Juan Comesaña (2002). Comesaña's account is interesting and complicated. I cannot do justice to it here. However, the upshot of it seems to be this: when we attribute justification to our beliefs, the demon's victim's beliefs, and the alien cognizers' beliefs, we are, each time we say 'their beliefs are justified,' making two claims. To use Sosa's terminology, we are saying that they are adroit-justified and that they are apt-justified. Thus, in our own case, we are completely right. But in the demon victim case we are partly mistaken since we assert both a truth (the demon victim's beliefs are adroit-justified) and a falsehood (they are apt-justified). Likewise, in the alien cognizer case we are partly mistaken since of that case too we assert both a truth (the alien's beliefs are apt-justified) and a falsehood (they are adroit-justified). These are the implications of Comesaña's position as I understand it. But J_{PF}, which accounts for the plain truth of each of these three univocal justification attributions (to us, to the demon victim, and to the alien cognizers) seems more plausible than Comesaña's position according to which we are at best partly right in such attributions due to the fact that with each attribution we're unwittingly making two assertions (one of them false) when we thought we were making only one.

3 OBJECTIONS

There are many objections that can be directed at the above proper function account of justification. Here I will address four that I take to be among the most important.⁶³

⁶³ For other ways to handle a variety of objections to talk about proper function and its role in analyzing epistemic concepts, see Plantinga (1993a, 1993c, 1995, 1996).

3.1 Proper Function and Naturalism

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

There is a large and growing literature on functions and goals.⁶⁴ Many naturalists are quite confident that the notion of a goal and of a function can be understood naturalistically and applied to living organisms (or their parts) even if these organisms are not the product of intelligent design. But when it comes to the idea of a function a thing is *supposed to* have or a goal to which it is *supposed to* contribute, some naturalists are more skeptical. Let’s examine some of the possible views one could hold concerning the normativity implied by such ‘supposed to’ talk.

First, consider the following three claims:

- the human heart is supposed to function so that it beats less than 200 times a minute when the person is at rest;
- the human heart is supposed to contribute to the goal of circulating blood and, ultimately, to the survival of the human in question;
- the type of environment in which a heart’s functioning that way is supposed to result in the survival of the human of which it is a part is an environment relevantly similar to the one in which we find ourselves.

These ‘supposed to’ claims seem very sensible and natural. But can they survive as more than just ordinary ways of speaking? Is it literally true that there is a function a human heart is *supposed to* have, that there is goal to which it is *supposed to* contribute and that there is such a thing as the type of environment in which its functioning in this way is *supposed to* contribute to that goal? And are such claims, taken literally and seriously, compatible with naturalism?

⁶⁴ See Nissen (1997), Boorse (unpublished), and Allen, Bekoff and Lauder (1998) for summaries of the relevant literature.

There are (at least) three positions one could take in response to these questions. Let's say that 'supposed to' talk is *reducible* to naturalistically acceptable talk just in case it is possible to give truth conditions for 'supposed to' claims (applied to parts of living organisms) using only naturalistically acceptable concepts. The *reductivist* position is that 'supposed to' talk is reducible to naturalistically acceptable talk. According to this view, the serious 'supposed to' claims mentioned in the previous paragraph are true and compatible with naturalism.⁶⁵ The other two positions—nonreductive realism and eliminativism—are nonreductivist views insofar as they agree that 'supposed to' talk is not reducible to naturalistically acceptable talk. According to *nonreductive realism*, serious 'supposed to' claims are true and incompatible with naturalism (so naturalism is false).⁶⁶ According to *eliminativism*, serious 'supposed to' claims are neither true nor compatible with naturalism.⁶⁷ The eliminativist with respect to the serious 'supposed to' claims says that, strictly speaking, there is no such thing as the *proper* function of organisms (or their parts) or the *right* environment for them; these normative notions need to be eliminated from our careful conversation which will instead employ notions that are more friendly to naturalism— notions such as *statistically normal* or *evolutionarily selected* functions or environments (or even functions *valued by us*).

Each of these three positions has its proponents; each has a legitimate claim to be taken seriously even by those who reject it. Which position is the correct one? Obviously, I can't resolve this issue here. But I will say this. In Sections 1.2.3 and 1.3 of this chapter I have defended the view that the notion of fittingness is conceptually tied to the idea of there being a *right* way for one's cognitive faculties to function—a way its faculties are *supposed to* function. If I'm right about *that*, then whatever position one takes with respect to the normativity of 'supposed to' claims about an organism's (or its parts') functions and goals—whether reductivism, nonreductive realism, or eliminativism—one ought to take that same position with respect to fittingness. But Objectivity is opposed to eliminativism with respect to fittingness. And Nonreliability is opposed to at least one strand of reductivism with respect to fittingness. So those who find Objectivity and Nonreliability plausible are forced to be nonreductive realists with respect to fittingness or to come up with some nonreliabilist way of being a reductivist about fittingness. Such people should, therefore, be either nonreductive realists or a certain kind of reductivist about the commonsense 'supposed to' claims concerning functions and goals; eliminativism is not an option for them.

⁶⁵ Karen Neander (1991a, 1991b) holds this view, as do Fred Dretske (1994, 2000) and Robert Koons (2000).

⁶⁶ This view is endorsed by Alvin Plantinga (1993a) and perhaps also by Bedau (1993).

⁶⁷ This view seems to be endorsed by John Searle (1992: 238) and by Christopher Boorse (unpublished).

In sum, my response to those naturalists who find the notion of proper function naturalistically unacceptable (i.e. my response to eliminativists about this sort of ‘supposed to’ talk) is twofold. First, my claim that the notion of epistemic fittingness is conceptually tied to the notion of proper function is compatible with eliminativism about the normativity of proper function. So at least that part of my argument avoids this naturalistic challenge. Second, upon accepting the conclusion that there is that conceptual tie between fittingness and proper function (and, therefore, between justification and proper function), one can take the plausibility of Objectivity and Nonreliability to provide us with a reason for rejecting eliminativism and certain versions of reductivism about proper function.

3.2 The Supervenience Thesis

The supervenience thesis (as it is sometimes called⁶⁸) says that normative properties supervene on nonnormative ones. Applied to epistemology, and, in particular, to justification, it says that justification supervenes on nonnormative properties—that there couldn’t be a difference in justification without a difference in nonnormative properties. The plausibility of the supervenience thesis might give rise to the following complaint about proper function accounts of epistemic justification: ‘The goal in understanding epistemic concepts such as justification is to give analyses of them in terms of nonnormative properties; but since *proper function* is a normative concept, analyses in terms of it are unsatisfying or uninteresting or in some other way unacceptable.’

Although this objection seems to be taken seriously by some people, I fail to see what the trouble is. The proponent of a proper function account of epistemic justification is like the Chisholmian who thinks epistemic concepts are reducible to ethical concepts.⁶⁹ Both think that epistemic evaluation is reducible (at least in part) to another sort of evaluation. But that in no way commits them to a rejection of the supervenience thesis. If epistemic properties are reducible to proper function (or ethical) properties and proper function (or ethical) properties are reducible to nonnormative properties, then the supervenience thesis holds. So I see no reason to think the supervenience thesis conflicts in any way with the analysis of justification given in this chapter. It’s true that it would be nice to conjoin my proper function analysis of justification with an analysis of proper function in terms of nonnormative properties and that I haven’t done that here.⁷⁰ But that doesn’t make it uninteresting to explain one normative notion in terms of another in the way Chisholm does—not unless the *only* sort of analysis of interest

⁶⁸ See Steup (1996: 30–6).

⁶⁹ See Chisholm (1986a, 1991). For objections to Chisholm’s position, see Firth (1978).

⁷⁰ But see the references in nn. 65 and 66 for attempts at and discussions of such analyses.

in epistemology is an analysis of epistemic properties in terms of nonnormative (rather than merely nonepistemic) ones.

3.3 Swampman

Against proper function analyses of epistemic properties, Ernest Sosa (1993) uses Donald Davidson's example of Swampman who comes into existence by accident as a result of a lightning strike. Swampman is supposed to be a molecule-for-molecule replica of Donald Davidson complete with Davidson's beliefs, memories, experiences, and dispositions. Given his origin, it seems that Swampman has no design plan, not even in the sense in which faculties 'designed' by evolution have a design plan. Thus, according to this objection, we can't say that his beliefs satisfy the proper function requirement. But since Davidson's beliefs are justified why not say the same of Swampman's?

Before responding to this objection, let's consider another example. Let's suppose that a lightning strike in a swamp results in a configuration of matter that in outward appearance and behavior almost exactly resembles a human woman. But Swampwoman differs from the Davidson-like Swampman not only because she resembles a woman rather than a man but also because, despite her outward appearance, she has a silicon-based brain and nervous system rather than a carbon-based one like our own. Now consider Swampwoman's unlearned doxastic response when grabbing a billiard ball in the dark. Upon grabbing the billiard ball she has tactile experience ME1 and forms the billiard ball belief B1 discussed earlier. Consider the following two ways to think of this case: (i) Swampwoman has our design plan and is believing in accord with proper function and (ii) Swampwoman has design plan B and is malfunctioning. The fact that forming B1 in response to ME1 seems to Swampwoman to be completely appropriate doesn't distinguish her from a person with design plan B who forms B1 in response to ME1 as a result of a malfunction that *also* causes this response to seem to that person to be entirely appropriate (despite the fact that for that person it is not a fitting one). She is just fortunate that, in addition to having this malfunctioning doxastic response to experience, she also has a malfunctioning experiential response to physical stimuli (having ME1 in response to grabbing a billiard ball is evidence of this latter sort of malfunction).

When we hear the Swampwoman case, we immediately impose our design plan on her and conclude that B1 is a fitting response to ME1. But suppose there were in fact silicon-based creatures, resembling humans in outward appearance and behavior, who had design plan B. What would they think of such an example? Perhaps they would conclude that Swampwoman, who they are told forms B1 in response to ME1, is malfunctioning. Why? Because they might impose *their* design plan on Swampwoman and according to that design plan, B1 does not fit ME1. In the same way, if *we* heard the Swampwoman story told

so that she formed B1 in response to *ME2* (rather than *ME1*), we would not be inclined to think that Swampwoman's belief fits her evidence.

So our judgments about the fittingness of Swampwoman's beliefs depend on which design plan we impose on her. What should we conclude about this case? One plausible-sounding response is to say that once we (a) see that our judgments about fittingness depend on which design plan we impose and (b) remember that Swampwoman doesn't seem to have a design plan, we should conclude that there is no doxastic response to *ME1* that is fitting for her—none that she is supposed to have. From this it follows that Swampwoman's belief B1 does not fit her evidence. And from that it follows that Swampwoman's belief B1 is not justified.⁷¹ The illusion that her beliefs are justified is, according to this response to the Swampwoman case, due to our imposing a design plan we have no good reason to impose given the description of the case.

Now return to the Swampman case. If we take the lesson learned from the above discussion of the Swampwoman case (i.e. that since there is no design plan, there is no sense in which the creature's belief fits the evidence and hence no sense in which her beliefs are justified) we should say that Swampman has no design plan either and that his beliefs aren't justified, despite the fact that he is a molecule-for-molecule duplicate of Davidson. If my position forced me to endorse this conclusion, this would, I think, strike many readers as a *reductio* of my position.

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

And if we find it plausible to say that Swampman's heart and lungs are healthy, it is reasonable also to conclude that his heart and lungs are functioning the way they are *supposed to* function. And if we can say that about his heart and lungs, we can say the same about his cognitive faculties. It's true that there is no literal designer and no evolutionary origin. But because of the physical similarities (down to the DNA level), we are inclined to think of Swampman as being human. And that leads us to think that Swampman is supposed to function in the way humans are supposed to function.⁷²

⁷¹ Given I_{PF} and that evidence is one sort of input to one's belief-forming systems. See the explanation, in Section 2.1 of this chapter, of how J_{PF} is a refinement of I_{PF} .

⁷² Cf. with Plantinga's response to the Swampman objection in Plantinga (1993c: 76–8).

What our discussion in this subsection shows is that our intuitions pull us in two different directions. We are, on the one hand, inclined to think that DNA matching indicates membership in the same natural kind. And we are inclined to think that what counts as proper function is essential to kind membership. Thus, we are inclined to think that what counts as proper function for Swampman is the same as what counts as proper function for us. That explains our intuition that Swampman's beliefs are as justified as Davidson's are.

But on the other hand, when we consider the Swampwoman case and imagine what silicon-based, human-resembling creatures with design plan B would think of her case, we realize that our reasons for assuming that Swampwoman has our design plan are no better than reasons they might have for assuming that Swampwoman has design plan B. And we are then led to conclude that the most plausible account of the Swampwoman example is that she has no design plan. This fits with our intuitions that creatures coming into existence by accident via lightning strikes have no design plans. And it has the consequence that Swampwoman is neither healthy nor unhealthy because there is nothing that counts as proper function for her. From that we may go on to conclude that none of her beliefs *fit* her evidence, in which case none of them are justified.

So which way should we go in interpreting the Swampman case? Should we say that none of his beliefs are justified or should we say his beliefs are as justified as Davidson's? I won't try to answer that question. However, I will point out that either way, we are preserving the connection between proper function and justification. For our intuitive judgments about whether Swampman has lots of justified beliefs or none waver together with our intuitive judgments about whether Swampman has a healthy heart and lungs or no healthy parts at all. If it makes sense to say that Swampman has a healthy heart and lungs, then it makes sense to say the concept of proper function applies and to say that his beliefs are fitting and justified. If it doesn't make sense to say that Swampman is a healthy creature with our design plan rather than an unhealthy creature with another design plan, then the concept of proper function doesn't apply and we shouldn't say his beliefs are either fitting or justified. Either way, the Swampman case fails to provide us with a counterexample to J_{PF}.

3.4 Skeptical Scenarios

There are a number of proposed counterexamples to proper function analyses of epistemic properties that employ strange skeptical scenarios. Some of them are dealt with elsewhere.⁷³ I'll take the time here to consider and respond to just two. One came up earlier in this chapter. The other has been proposed recently by Conee and Feldman.

⁷³ See responses by Plantinga to such objections applied to his proper function account of *warrant* in Plantinga (1993a).

Suppose an incompetent malevolent demon sets out to create a cognizer to hold unreliably formed beliefs but bungles things so badly that the result is a duplicate of an ordinary human in a standard human environment.⁷⁴ You might think that this creature's beliefs are justified even though there is an important sense in which, given the demon's wicked intentions, the creature's faculties aren't aimed at truth. But here, as in the Swampman case, we are pulled in two directions. On one way of looking at this case, it seems that the faculties aren't truth-aimed and that they are malfunctioning and only accidentally resulting in true beliefs. And that suggests that the beliefs aren't justified. But, if we stipulate that the creature turns out, by sheer accident, to be a molecule-for-molecule duplicate of a properly functioning human, then we are inclined to think that, perhaps due to DNA considerations, this creature is a human and has the same design plan humans have, no matter how it came into existence. We are inclined to think that it is *supposed to* function the way we are. In that case, insofar as its beliefs are formed after the manner of cognitively healthy human belief formation, we are inclined to think its beliefs are truth-aimed and justified.

Which is the right approach to the example? I won't try to answer that question. I'll simply note, as I did in my discussion of the Swampman case, that no matter which answer we give, we can see how our intuitions follow J_{PF} in linking justification with production by properly functioning truth-aimed faculties. For if it makes sense to say that the faculties are properly functioning truth-aimed faculties with the same design plan ours have (that the victim is forming the beliefs it is epistemically supposed to form), then it also makes sense to say that the beliefs of the lucky victim of the bungling evil demon are justified. And if it makes sense to say that the faculties are not properly functioning truth-aimed faculties, then it also makes sense to say that these beliefs aren't justified.

The other counterexample to my account of justification is proposed by Conee and Feldman:

If future cognitive scientists create human-like brains in vats for cognitive research, and design them to think as ordinary people do but to live in vat environments, then, as best as we can understand the situation, when these designed brains function properly their processes will not be reliable in the environments for which they were designed. This, we think, makes no difference to the justificatory status of their beliefs.⁷⁵

The idea is that this is supposed to count against J_{PF} because the beliefs of these brains can be justified even though they aren't produced by faculties that are reliable in the environments for which they were designed by the scientists.

The question we need to consider is how the brains in question are supposed to operate. More specifically, the question is how they are *epistemically* supposed

⁷⁴ This is the example that came up earlier in the chapter, in n. 52.

⁷⁵ This was part of their reply to my comments on their symposium paper 'Some Virtues of Evidentialism' presented at the Central Division Meeting of the American Philosophical Association in April 2005.

to operate. How something is supposed to operate is tied to its proper function; how it is *epistemically* supposed to operate is tied to the *truth-aimed* proper function of its *belief-forming faculties*.⁷⁶ So, how are these human-like brains in vats epistemically supposed to function? To consider that, we need to consider what the truth-aimed proper function of their belief-forming faculties is. And here, once again, we are pulled in two directions. On the one hand, if we think they are human-like enough to count as humans (perhaps because they have the same DNA as humans), then we think that since they are of the same natural kind as we are, the truth-aimed proper function of *their* belief-forming faculties is the same as the truth-aimed proper function of *our* belief-forming faculties. But if, on the other hand, we think that their proper function is determined by the cognitive scientists who created them, we must conclude that they do not have a *truth-aimed* proper function for their belief-forming faculties.

So which is it? Perhaps we can't say with certainty. But either way, J_{PF} seems to come out fine. If we think that what counts as their proper function is determined by the cognitive scientists who designed them, then it seems they do not have a *truth-aimed* proper function for their belief-forming faculties, in which case there isn't any way they are *epistemically* supposed to operate (even if there is some way they are supposed to operate). From this we may conclude that their beliefs don't admit of justification. But if we think that what counts as proper functioning and the right environment for *them* matches what counts as proper functioning and the right environment for *us*, then we can agree with Conee and Feldman that the beliefs of these brains can be justified despite the fact that they are unreliably formed and false. However, in that case, the beliefs will be formed by properly functioning truth-aimed faculties that are reliable when operating in their right environment. So there is no problem here for J_{PF}.

⁷⁶ See Section 2.1 of this chapter for a defense of the claim that epistemic fittingness is tied to *truth-aimed* proper function.

