

If Dogmatists Have a Problem with Cognitive Penetration, You Do Too

Chris Tucker

College of William and Mary

cstucker@wm.edu

(forthcoming in *dialectica*; finalist for *dialectica*'s 2012 essay prize)

Official version available at:

<http://onlinelibrary.wiley.com/doi/10.1111/1746-8361.12050/pdf>

Perceptual dogmatism holds that if it perceptually seems to S that P, then S thereby has prima facie perceptual justification for P.¹ But suppose Wishful Willy's desire for gold cognitively penetrates his perceptual experience and makes it seem to him that the yellow object is a gold nugget. Intuitively, his desire-manipulated seeming can't provide him with prima facie justification for thinking that the object is gold. If this intuitive response is correct, dogmatists have a problem.

Yet before you mock the speck in the dogmatist's eye, be sure that you don't have a plank in your own. Reliabilists, such as Alvin Goldman and Jack Lyons, have been quick to pounce on dogmatism for its implications regarding cognitive penetration, and Lyons has gone so far as to declare that reliabilism has special virtues on this topic. The truth, however, is that reliabilists have cognitive penetration problems that are even worse than those (allegedly) faced by the dogmatist. Many other views in the literature suffer from cognitive penetration problems that are at least as bad as those faced by the dogmatist.

I write this paper to defend dogmatism. Now, little of what I say shows that the cognitive penetration objection to dogmatism fails. But misery loves company. If I'm going down, I'm going to take as many down with me as I can. If these (alleged) problems are as widespread as I suggest, then very few epistemologists will be able to use cognitive penetration as a reason to prefer their view over dogmatism.²

In section 1, I clarify the problem that allegedly afflicts the dogmatist. In section 2, I explain why entitlement conservatism and the most popular versions of coherentism are committed to essentially the same problems with cognitive penetration. In section 3, I explain why the reliabilist faces cognitive penetration problems that are worse than those faced by the dogmatist. In section 4, I show that the reliabilist's problems generalize to virtue theories and proper functionalism. In section 5, I consider whether there are any views that avoid dogmatism's cognitive penetration problems.

1. Cognitive Penetration by Desire

¹ Prima facie justification for P can be had whether or not one has defeaters; (ultima facie) justification is prima facie justification in the absence of defeaters. Dogmatism is a theory of propositional justification, which one can have for a *proposition* P whether or not one believes P. The theories considered in sections 3 and 4 are generally put forward as theories of doxastic justification, which is justification one can have for a token *belief* that P. Whenever I talk about the justification of propositions, I'm talking about propositional justification. Whenever I talk about the justification of beliefs, I'm talking about doxastic justification. As a general rule, then, whenever I use the term 'justification' or its cognates in section 3 and 4, I am referring to doxastic justification, and whenever I use that term elsewhere in the paper I'm referring to propositional justification. Nothing interesting in this paper hinges on the distinction between propositional and doxastic justification (cf. Siegel 2013: 704-6).

² Elsewhere I argue more directly that the objection fails, but I suppress the reference for the sake of blind review.

For our purposes, we can say that a cognitive state C1 **cognitively penetrates** another C2 when C1 (partially) causes C2 in an internal way. The causal chain from C1 to C2 is **internal** just in case it is contained *entirely* within the subject.³ Suppose I desire for the light to be on and so turn it on. The light's being on then makes it seem to me that it is on. In this case, the causal chain from my desire to my seeming isn't internal because it involves links outside of me, such as my interaction with the light switch and the light's actually turning on.

Given this notion of cognitive penetration, both beliefs and experiences can be cognitively penetrated. Suppose I believe there is a goat on the basis of seeming to see one. Then my belief that there is a goat is internally caused, and so penetrated, by my perceptual seeming. The Wishful Willy case with which we began is a case in which a perceptual seeming *it's seeming that the object is a gold nugget* is internally caused, and so penetrated, by a desire for gold.

A variety of cases have been given to show that cognitive penetration poses a problem for dogmatism, but they generally appeal to penetration by desires or unjustified beliefs.⁴ In this paper, I focus solely on cognitive penetration by desire because these are the cases that most trouble me as a dogmatist. Despite my commitment to dogmatism, I do sometimes get the intuition that Willy's wishfully-produced seeming is not capable of producing prima facie justification. Lyons (2011: 299-300) agrees that such cases are the most embarrassing for the dogmatist, and there is a long list of people who use such cases to object to dogmatism.⁵

Before I turn to the main argument of the paper, three additional points are worth mentioning. First, not every instance of cognitive penetration by desire is thought to be problematic. Suppose the child really wants to find the Easter eggs. Her desire may cause her to attend to the features of her perceptual imagery that indicate Easter eggs, features that she otherwise would have missed and which make it seem to her that an Easter egg is nearby. On this description of the case, the causal chain is internal because it only involves the child's directing her attention to certain features of her imagery and not moving her eyes or head. This sort of cognitive penetration, however, isn't obviously problematic and may enhance perceptual recognition.

Let us say that a seeming is **desire-manipulated** relative to P just in case it is cognitively penetrated by a desire and this penetration intuitively prevents the seeming from providing prima facie justification for P. Notice that a seeming might be desire-manipulated relative to one proposition but not another. Even if Wishful Willy's seeming can't justify its content *that the object is gold*, it might nonetheless justify the proposition *that it seems to Willy that the object is gold*. Willy's desire-penetrated seeming is arguably desire-manipulated relative to the former but not the latter proposition.⁶ In the rest of the paper, I will talk as though seemings can be desire-

³ This definition of cognitive penetration is essentially the same as that of Stokes (Stokes forthcoming, sec 1; 2012: 479 and 482, nt 6), except that Stokes only defines the relation for when an experience is penetrated by a desire. This definition needs refinement in all sorts of ways, but it will suffice for our purposes (cf. Lyons 2011: 290). When MacPherson (2012, sec 2) and Siegel (2012, sec 1) clarify the relevant issue, they focus on defining cognitive penetrability rather than cognitive penetration, but the two issues are obviously related.

⁴ See my 2013a, sec 3 for an overview of the cognitive penetration objection to dogmatism.

⁵ See, for example, Brogaard (2013, sec 2); Goldman (2008: 72-3; 2009: 330); Lyons (2011); Markie (2005: 356-7; 2013); McGrath (2013); Siegel (2012); and Steup (2004: 415-6; 2013, sec 6).

⁶ Desire-manipulated seemings may be what Siegel calls "checkered experiences" (2013). If there is a difference, it is that, roughly, desire-manipulated seemings are those that intuitively cannot justify a given proposition whereas checkered experiences are those that intuitively *and* genuinely cannot justify a given proposition (Siegel 2013: 716). Just as I am allowing that a seeming could be desire-manipulated

manipulated simpliciter and will ignore the need to relativize the desire-manipulation to certain propositions. With this simplification, desire-manipulated seemings are those desire-penetrated seemings that allegedly pose a problem for dogmatism, and the Wishful Willy case is a paradigmatic example.

Second, the objection to dogmatism is *not* (merely) that dogmatism allows wishfully-produced seemings to provide *ultima facie* justification. In ordinary cases, a subject will have a defeater which makes it irrational for him to trust the desire-manipulated seeming. For example, on any natural filling-out of the Wishful Willy case, he will be aware that he lacks the relevant training to be able to distinguish gold from fool's gold. This defeater will prevent him from being *ultima facie* justified in believing that the object is gold. The objection is that the dogmatist implausibly allows Willy to have *prima facie* justification.

Although I do sometimes get the intuition that Willy doesn't have even *prima facie* justification, many others do not. On their view, the intuitive negative reaction to Willy is just that he has a defeater available to him. When you spell the case out so that he clearly doesn't have a defeater, there's nothing counterintuitive about saying he's justified, much less *prima facie* justified.⁷ I have some sympathy with this approach, but this paper is written for those who do get the intuition that Willy lacks even *prima facie* justification. In this context, we should assume that appealing to defeaters can't help the dogmatist or anyone else avoid cognitive penetration worries.

Third, I often find it convenient to say that some theory X has cognitive penetration problems that are *as bad as* or *worse than* the dogmatist's. Such expressions shouldn't be interpreted as saying that either view has a genuine problem. The points, respectively, are that X is *no better off* or is *worse off* than dogmatism. X might be no better off simply because it has implications similar to those of dogmatism but neither it nor dogmatism has a genuine problem. X might be worse off than dogmatism by having implications that are more counterintuitive than dogmatism's implications, and yet neither view may be genuinely problematic because we should reject all those intuitions.

With the above clarificatory remarks in place, I turn toward showing that, if cognitive penetration is a genuine problem for dogmatism, it is a far more ubiquitous problem than is currently recognized.

2. The Same Cognitive Penetration Problems

Recall that dogmatism is the claim that, if it perceptually seems to S that P, then S *thereby* has *prima facie* perceptual justification for P. The "thereby" indicates that the (*prima facie*) perceptual justification *consists solely* in its perceptually seeming a certain way;⁸ for dogmatism, a perceptual seeming that P can be the *sole constituent* of one's perceptual justification. A *constituent* of justification is just a special sort of condition for justification. To be a constituent of X, X must obtain at least partly in virtue of that condition's being satisfied. Since $2+2=4$ is a necessary truth, that $2+2=4$ is a (trivial) necessary condition of my having justification for the claim that I'm

with respect to one proposition but not another, Siegel (2013: 703) allows that a seeming could be checkered with respect to one proposition but not another.

⁷ See, for example, Huemer (2013a: 343; 2013b, sec 1), Lycan (2013, sec 7), and Gage (2012: 373). Chudnoff (2011: 316) notes sympathy with this approach but then develops an alternative response to the cognitive penetration worries.

⁸ It's worth mentioning that my use of 'dogmatism' is different than that of Pryor, who is largely responsible for the popularity of the term. On his usage, dogmatism about perceptual justification is that some perceptual justification is immediate but underminable.

wearing a black shirt. Nonetheless, my justification for this claim does not consist, even partially, in the fact that $2+2=4$.

The dogmatist, then, is claiming that a person can have prima facie justification for a proposition that obtains solely in virtue of its seeming that the proposition is true. At the other extreme, a simple process reliabilist will deny that perceptual seemings ever count as constituents of perceptual justification. Yes, perceptual seemings can be a part of the process that provides perceptual justification, but the reliability of the process, not the perceptual seeming, is the (sole) constituent of one's perceptual justification. For the simple process reliabilist, seemings don't do any of the real justificatory work; they are just along for the ride.

My diagnosis of dogmatism's cognitive penetration problems invokes:

Filter Failure 1: A theory of justification is false if it allows a desire-manipulated seeming to count as a constituent of justification.

Since I endorse dogmatism, I think Filter Failure 1 is false. My diagnosis, though, is that dogmatism's penetration problems are genuine if and only if Filter Failure 1 is true. If this diagnosis is correct, any other view which satisfies (the sufficient condition identified by) Filter Failure 1 suffers from those same problems. I won't be arguing for this diagnosis directly; it is simple and intuitive enough to take as a starting point. But I will challenge an alternative diagnosis which, if correct, would incriminate dogmatism without incriminating two of its closest rivals, entitlement conservatism and experiential coherentism.⁹ In 2.1, I introduce these rivals and the alternative diagnosis. In 2.2 and 2.3, I explain why, at least for the time being, the alternative diagnosis should be rejected.

2.1. An Alternative Diagnosis of Dogmatism's Cognitive Penetration Problems

In the last decade or so, entitlement conservatism has emerged as a major player in epistemology, and its proponents talk as though it accommodates the promise of dogmatism without inheriting any of its problems.¹⁰ It allows a perceptual seeming to be a constituent, just not the sole constituent of perceptual justification. More specifically, **entitlement conservatism** holds that perceptual justification requires two constituents, the perceptual seeming that P and entitlement to discount skeptical hypotheses, such as that one is a brain-in-a-vat. This view is a kind of *conservatism*, because it holds that perceptual justification (partially) consists in having justification to discount skeptical hypotheses. It is distinguished from other conservatisms in two ways. First, it allows the seemings themselves rather than, say, justified beliefs about seemings to count as constituents of one's perceptual justification. Second, it allows the required justification to discount skeptical hypotheses to come in the form of entitlement, where an entitlement to P is a status that comes by default, without need of evidence or any achievement on the part of the subject. Crispin Wright (2004), Roger White (2006), Stewart Cohen (1999: 76–7), Allan Hazlett (2006), and Martin Smith (2013) have endorsed positions in the neighborhood of entitlement conservatism.

The most popular type of coherentism holds that perceptual justification requires two constituents, perceptual experience and coherence. Call this view **experiential**

⁹ Siegel (2012: 220) and McGrath (2013, sec 1; conversation) agree that both entitlement conservatism and experiential coherentism share dogmatism's cognitive penetration problems; however, neither argues for this conclusion, perhaps because they simply assume that something like Filter Failure 1 is the correct diagnosis of dogmatism's penetration problems.

¹⁰ White (2006) and Wright (2007) are perhaps the clearest examples of proponents who talk this way.

coherentism.¹¹ It has been endorsed by Elgin (2005: 165-6); Haack (1993); and Steup (2004, 2013), among others.¹² Experience is integrated into coherentism for at least two reasons. First, certain cases suggest that “fit” with experience is a necessary condition of justification.¹³ Second, some hold that coherence can only amplify some initial credibility (which may fall well short of justification); if the beliefs have *nothing* going for them, then coherence can’t help. Experience is thought to be a plausible source of this initial credibility.

Steup (2013: 144-8) argues that dogmatism’s penetration problems are genuine but that these problems do not afflict his own experiential coherentism. I’ve heard, in conversation, entitlement conservatives make the parallel claim about their own view. Since entitlement conservatism and experiential coherentism don’t impose any restrictions on the causal history of seemings, they allow some desire-manipulated seeming to be a constituent of perceptual justification and so satisfy the condition identified by Filter Failure 1. They must, therefore, reject Filter Failure 1 and find an alternative diagnosis of dogmatism’s penetration problems.

Like dogmatism, these conservatives and coherentists allow seemings to count as constituents of justification. Unlike dogmatism, they deny that seemings can count as *sole* constituents of justification. Perhaps, then, we should diagnose dogmatism’s difficulties by rejecting Filter Failure 1 (because it’s too strong) and then endorsing:

Filter Failure 1*: A theory of justification is false if it allows a desire-manipulated seeming to count as the *sole* constituent of justification.

If this diagnosis is correct, we’d have a good explanation of why dogmatism has problems with cognitive penetration when the relevant versions of conservatism and coherentism do not.

2.2. *The Failure of the Alternative Diagnosis*

To assess this alternative diagnosis, we need to clarify what entitlement conservatives and experiential coherentists are committed to, if they are to claim that cognitive penetration by desire poses a problem for the dogmatist but not for them.

Since Wishful Willy has the required seeming and background entitlement, the entitlement conservative must say that Willy’s desire-manipulated seeming provides *prima facie* justification that the object is gold. And, as Steup (2013: 142-3) appreciates, some descriptions of the Wishful Willy case allow that Willy has the required sort of coherence, so experiential coherentism is likewise committed to allowing Wishful Willy (on some ways of filling out that example) to have *prima facie* perceptual justification for the relevant claim. Consequently, while entitlement conservatives and experiential coherentists deny that desire-manipulated seemings can count as the *sole* constituent of justification, they are committed to:

Compatibility: a seeming’s being desire-manipulated is compatible with its being a constituent of justification.

¹¹ Pure coherentism, which isn’t very popular, holds that a new belief is justified if and only if it is coherent with one’s existing beliefs. Since this kind of coherentism prevents seemings from being constituents of justification, it doesn’t satisfy Filter Failure 1 and doesn’t have the same problems that dogmatism allegedly has. Its problems are even worse: its failure to impose any causal constraint on how beliefs are formed leaves it vulnerable to the problems that afflict reliabilism.

¹² Haack refers to this sort of position as “foundherentism” and Steup (2004: 419-20) refers to it as a “holistic foundationalism”. Obviously, whether people use the label “experiential coherentism” for the view I have in mind is not important. What’s important is that this view is a competitor to dogmatism which suffers from essentially the same cognitive penetration problems as dogmatism.

¹³ See, for example, Plantinga’s (1993: 82) case of the epistemically inflexible rock climber.

They must hold, in other words, that a desire-manipulated seeming can provide justification, even if they deny that a desire-manipulated seeming can provide justification by itself.

Since they are committed to Compatibility, they must reject Filter Failure 1 as the correct diagnosis of dogmatism's penetration problems. We are supposing that they endorse Filter Failure 1* and claim that only dogmatism has a problem because only dogmatism allows desire-manipulated seemings to count as the *sole* constituent of justification.

For this strategy to be successful, these conservatives and coherentists must think that, by imposing additional requirements on perceptual justification, they must have a plausible *explanation* of why desire-manipulated seemings can provide justification, but the dogmatist does not. If the dogmatist has a plausible explanation of why desire-manipulated seemings can provide justification, then their penetration problems aren't genuine after all. If these conservatives and coherentists don't have a plausible explanation either, then they too suffer from dogmatism's penetration problems. If the additional requirements they impose don't contribute to their allegedly superior explanation, what good are these additional requirements?

The problem with the strategy emerges when we realize what this allegedly superior explanation *can't* be: it can't be merely that seemings are, by themselves, insufficient for justification. If these conservatives and coherentists concede that desire-manipulation poses no special obstacle to perceptual justification—if they could have made the same point by appealing to a seeming caused by veridical perception—then their complaint against dogmatism isn't an objection *from cognitive penetration*. To be an objection from cognitive penetration, they must suppose that desire-manipulation (whether or not one is aware of it) is a special obstacle to perceptual justification. As we will see in the next sub-section, Steup misses this point.

In general, there is nothing fishy about saying that X is a special obstacle to Y, but also allowing that X and Y are compatible. A broken ankle poses a special obstacle to walking, an obstacle to walking that only some people must overcome. Nonetheless, broken ankles are compatible with walking—provided that one uses something or someone as a crutch. Those who wish to endorse Filter Failure 1* while rejecting Filter Failure 1 must view desire-manipulation along the lines of a broken ankle: it poses an obstacle to perceptual justification that only some seemings must overcome. They must then view entitlement or coherence or whatever as the crutch that allows desire-manipulated seemings to overcome their desire-manipulation and thereby provide perceptual justification.

The alternative diagnosis of dogmatism's penetration problems is conceptually coherent, as illustrated by the analogy with broken ankles; however, it is inferior to the diagnosis offered by Filter Failure 1. There is an available, plausible explanation of Filter Failure 1's truth, namely that certain causal histories are incompatible with perceptual justification (cf. Siegel 2012: 220). On the other hand, there is no available explanation of Compatibility that allows desire-manipulation to be a special obstacle to perceptual justification, so there is no explanation of how Filter Failure 1* could be true without Filter Failure 1 also being true.

Consider existing explanations of Compatibility. Huemer and I argue that the causal history of a seeming—whether defective or not—is relevant only to whether the seeming can provide warrant (the property that makes true belief knowledge), not to whether it can provide justification (Huemer 2013a: 344-5, 2013b: 747-8; Tucker 2010: 539-40, 2011: 70-72). Our explanation would vindicate Compatibility, but then Filter Failure 1 and 1* would both be false and desire-manipulation wouldn't be a problem at

all insofar as perceptual justification is concerned. Similar points apply to the alternative explanations of Compatibility provided by Chudnoff (2011) and Skene (2013). It's hard to see how, even in principle, we could explain Compatibility while allowing desire-manipulation to pose a special obstacle to perceptual justification.

Until we can find such an explanation, Filter Failure 1 will be a better diagnosis of dogmatism's penetration problems than Filter Failure 1*. For there is a plausible explanation of Filter Failure 1, namely that certain causal histories are incompatible with perceptual justification. Perhaps entitlement conservatism and experiential coherentism are superior to dogmatism for some reason or another. Yet, for the time being, conservatives and coherentists can't use cognitive penetration as a reason to prefer their views over dogmatism.

2.3. Cognitive Penetration Objections vs Subject's Perspective Objections

Since Steup has explicitly argued that cognitive penetration provides a reason to endorse his coherentism over dogmatism, we should examine where his argument goes wrong. Steup's main concern, especially in his 2013, is to address the subject's perspective objection, i.e. to explain how the subject's perspective is involved in justification. Bonjour's (1985) clairvoyance examples are often thought to show that the subject's perspective must be involved *somehow*, though there is considerable disagreement over how it must be involved.¹⁴ Steup (2013: 144-8) holds that for S to have justification for P, P must be likely from S's perspective; and for P to be likely from S's perspective, S must have (i) a seeming that P is true and (ii) "memory data" that affirm the reliability of the kind of seeming involved in (i). Steup complains that dogmatism fails to require (ii), and so it succumbs to the so-called subject's perspective objection.

Whether Steup's objection to dogmatism succeeds is, for the purposes of this paper, neither here nor there.¹⁵ Steup talks as though cases involving a seeming with a "bad origin like wishful thinking, drug-induced hallucinations, brain-lesions, and the like" play a crucial role in his objection (144-5). But the appeal to bad origins is just a red herring. To provide perceptual justification, on Steup's view, all perceptual seemings have to overcome the same obstacle—being insufficient for justification on their own—regardless of how good or bad their causal history is. And, regardless of their causal history, having the relevant sort of memory data is all it takes to overcome this obstacle (Steup 2013: 147).

If you want to show, as Steup does, that the subject's perspective must meet certain conditions *even in cases of veridical perception*, you must do what Bonjour did when he originally raised the objection: assume that belief in the relevant proposition has a good causal history (Bonjour assumed that the subject had reliable clairvoyant powers) and argue that the good causal history isn't enough for perceptual justification. In effect, Steup conflates the subject perspective and cognitive penetration objections to dogmatism. *Subject's perspective objections* contend that, even if everything is dandy with the causal history of a belief, the subject must have a perspective on the belief which satisfies certain conditions. *Cognitive penetration objections* contend that, even if everything is dandy regarding the subject's perspective, the casual history of the

¹⁴ In my (2013a, sec 5.3), I briefly survey a few proposals concerning how the subject's perspective must be involved in justification.

¹⁵ For the record, I think Steup's positive view is plausible, but I also think that dogmatism can adequately handle the subject's perspective objection. To show the latter, I need only generalize the position in my 2012.

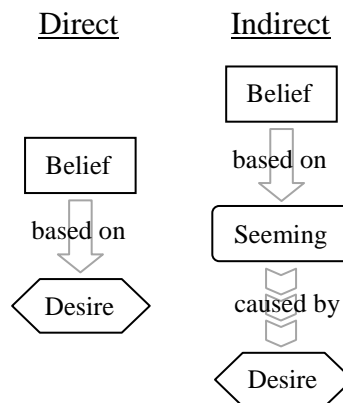
belief must satisfy certain conditions (e.g., not be based on a desire-manipulated seeming).

Filter Failure 1 provides a simple and intuitive diagnosis of dogmatism’s penetration problems: assuming dogmatism’s problems are genuine, a theory has those problems whenever it allows a seeming to be a constituent of perceptual justification, whether it is the sole constituent or not. We considered whether entitlement conservatives and experiential coherentists could make use of an alternative diagnosis that would incriminate dogmatism but vindicate their views. This alternative isn’t promising, because it is hard to see how desire-manipulated seemings could produce justification if desire-manipulation is a genuine obstacle to perceptual justification. Steup failed to see this point because he conflated the subject’s perceptive and cognitive penetration objections to dogmatism. At least tentatively, then, we should conclude that entitlement conservatives and experiential coherentists (including Steup) have the same cognitive penetration problems as dogmatism.

3. Cognitive Penetration Problems for Reliabilism

Reliabilists, such as Goldman (2008: 72-3; 2009: 330) and Lyons (2011: 299-300), have been quick to pounce on dogmatism for its approval of Willy’s desire-manipulated seeming. Lyons (2011: 300ff) also argues that reliabilism has special advantages in dealing with seemings that are penetrated by desires of the subject. The truth of the matter, however, is quite inimical to reliabilism: reliabilists suffer from cognitive penetration problems that are *worse* than those faced by the dogmatist. These problems arise because the reliabilist doesn’t put enough constraints on which mental states can be acceptable bases for a belief.

Contrast direct and indirect wishful thinking. The former occurs when one bases a belief that P directly on one’s desire (wish, lust, etc.) that P. The latter occurs in any other way. In this paper, we have been concerned so far with a very specific form of *indirect* wishful thinking, namely the sort of wishful thinking when our desires influence the way things seem. The difference between the two kinds is illustrated by the following diagram:



Both types of wishful thinking involve cognitive penetration by desire, because both involve a desire causally influencing another cognitive state in an internal way. I admit it’s bad, or at least looks bad, to allow typical cases of *indirect* wishful thinking to provide prima facie justification.¹⁶ It’s worse, in my mind, to allow possible cases of *direct* wishful thinking to provide prima facie justification. It’s one thing to allow a

¹⁶ As explained in note 1, I shift to talking about doxastic justification in secs 3 and 4.

belief to be justified on the basis of a seeming when your desires have an imperceptible effect on the way things seem. It's quite another to allow a belief to be justified when it is based solely on a desire for that belief's truth, when the subject treats the desire as a reason for the belief.

We can generalize the point as follows:

Filter Failure 2: A theory of justification is false if it allows a possible non-introspective belief that P to be (prima facie) justified solely on the basis of a desire.

Filter Failure 2 appreciates the general rule that it's bad to base a belief solely on a desire, but the 'non-introspective' qualifier also recognizes that there may be a certain kind of exception to this general rule. For many epistemologists (not me) hold that the belief *that I desire P* could be justified solely on the basis of my desire for P.¹⁷ The cases I focus on below are clearly not exceptions to the general rule, so even if Filter Failure 2 needs to accommodate some other exception that I'm neglecting, that wouldn't affect the main points in this paper.

Any theory that satisfies (the sufficient condition identified by) Filter Failure 2 has cognitive penetration problems that are *worse* than the problems faced by the dogmatist. Such problems are not only more counterintuitive but also more intractable. If approving of direct wishful thinking isn't a problem, it's hard to see why approving of the indirect kind would be. On the other hand, if approving of the indirect kind isn't a problem, approving of the direct kind still might be.

I've encountered some resistance to the claim that the reliabilist's problems are worse than those faced by dogmatism. One consideration motivating this resistance is that reliabilism plausibly gets the right result in ordinary construals of the Wishful Willy case and dogmatism doesn't. If, after reading this section, you don't agree that the reliabilist's problems are worse, I think you'll agree to this much: it's at least *unclear* whether reliabilism is overall any better off with respect to cognitive penetration by desire. And that much agreement is all I need to prevent the reliabilist from gaining any meaningful advantage from cognitive penetration by desire. For then the appeal to Wishful Willy would be no less lame than Mr Sleazy's appeal to Mr Davis' affair: "If you care about family values, you have a reason to vote for me over Mr Davis: Mr Davis' shameful affair with Ms Sinclair has been well-documented. Of course, as those meddlesome reporters won't let you forget, it's unclear that I'm overall more faithful to my wife than Mr Davis is to his. My point is just that, insofar as indiscretions with Ms Sinclair are concerned, you have a reason to vote for me." Mr Sleazy's point is probably correct, but only in a lame, meaningless sort of way.

In 3.1, I argue that indicator reliabilism satisfies Filter Failure 2. In 3.2, I argue that process reliabilism also satisfies Filter Failure 2. Reliabilists have tried to address these problems and, in 3.3, I criticize three of their attempts to do so.

3.1. Simple Indicator Reliabilism

To see the basic point, consider:

Simple Indicator Reliabilism (SIR): S's belief that P is (prima facie) justified iff S bases the belief on a mental state M which reliably indicates P.

For our purposes, we can say that a mental state M *reliably indicates* P just in case, in most of the closest possible worlds in which a subject has M, P is true. Alston (1989) and Comesaña (2010) endorse views like SIR.

¹⁷ I'm understanding "introspective" broadly enough so that the belief *that there exists at least one mental state* would turn out to be introspective if it were based solely on one's desire for P.

Since SIR doesn't put any other constraint on which mental states can provide prima facie justification, desires can provide prima facie justification provided that they reliably indicate a proposition. There are at least two ways in which desires can reliably indicate a proposition. The first is for the modal profile of a proposition to make *any* mental state, including a desire, a (trivially) reliable indicator of that proposition. Suppose, for the sake of an example, that Fermat's Conjecture is necessarily true, that I strongly desire it to be true (maybe I bet on its truth), and I believe it solely on the basis of the desire. Since the conjecture is necessarily true, it is true in every possible world. Trivially, then, my desire reliably indicates my belief: in every close possible world in which I desire the conjecture to be true, the conjecture is true. Consequently, SIR entails that my belief is at least prima facie justified solely on the basis of the desire.

The problem for SIR is not just that it approves of direct wishful thinking, as in the above case. Since the modal profile of necessary truths guarantees that every mental state reliably indicates them, given SIR, my belief in the conjecture would be justified if it were based on *any* desire, not merely a desire for the conjecture. My belief in the conjecture would be justified even if I based it on a desire for moldy cheese.

The modal profile of certain contingent truths also poses a problem for SIR. Take **similarity-constituting truths** for instance. Some contingent truth T is *constitutive of similar worlds* just in case T is true in all the most similar worlds. For example:

The Nomic Proposition: the proposition identifying all and only genuine laws of nature.

Given some respectable accounts of the laws of nature, the Nomic Proposition is contingently true. Nonetheless, it is widely assumed that there are no similar worlds in which the laws of nature are different. In all the closest possible worlds—not just the closest possible worlds in which I desire the Nomic Proposition to be true—the Nomic Proposition is true. SIR entails that my belief in the Nomic Proposition can be justified solely on the basis of the desire.

Or consider:

No Grammy: I do not win a Grammy.

There are no sufficiently similar worlds in which No Grammy is false. (Background info: I once won a bottle of coconut rum for being the *worst* karaoke singer in a group of very bad karaoke singers. No, I wasn't trying to sing badly; singing badly just comes naturally.) Maybe I desire not to win a Grammy—maybe I think it glorifies trivial and worthless pursuits. Since it is true in all the closest possible worlds—not just the closest possible worlds in which I desire it to be true—SIR entails that my belief in No Grammy can be justified solely on the basis of on my desire for its truth.

Again, the problem for SIR is not just that it approves of direct wishful thinking. Since the modal profile of similarity-constituting truths guarantees that every mental state (that I have in at least one close possible world) reliably indicates them, given SIR, my belief in No Grammy would be justified even if it were based solely on my love of brownies or my preference for footnotes over endnotes.

The second way to ensure that desires reliably indicate certain propositions is the manipulation of epistemic angels and demons (cf. Fumerton 2006: 80). Epistemic angels can bring about the required reliability by organizing the world to ensure that desires (within a certain domain) regularly come to pass. Epistemic angels might be a sufficiently stable feature of the environment that they exist and do their beneficent work in all the closest possible worlds. In such a case, SIR would hold that a belief that P could be justifiably based solely on a desire for P.

Epistemic demons can bring about the required reliability by organizing the world to ensure desires (within a certain domain) never come to pass. And they might be stable features of the environment and do their dirty work in all the closest possible worlds. In such a case, a belief that $\sim P$ could be justifiably based solely on a desire for P . Or suppose that every now and then I become terrified of snakes. The demon might ensure that whenever this terror strikes, I usually do encounter a snake. In such a case, given SIR, I could justifiably base my belief that there is a snake nearby on my becoming terrified of snakes (even if I had no reason to believe that there was a correlation between the terror and there actually being a snake).

Thanks to the modal profile of certain propositions and the manipulation of angels and demons, SIR allows certain non-introspective beliefs to be justified solely on the basis of a desire. So SIR satisfies Filter Failure 2 and thereby suffers cognitive penetration problems that are more counterintuitive and more intractable than those (allegedly) faced by the dogmatist.

My basic worry about reliabilism (both indicator and process) isn't new: I am not the first to point out that reliabilism has counterintuitive implications when you consider the modal profile of certain propositions or manipulated environments.¹⁸ Solutions to these problems have been proposed (e.g., Comesaña 2010, sec 8.2; Goldman 1979: 43-4). The main goal of this section does not require that I prove that these problems are insoluble: my main goal is to reveal how these somewhat familiar problems bear on the present discussion and to address a few of the existing "solutions". These problems bear on the present discussion insofar as they lead to counterintuitive implications concerning cognitive penetration by desire. To whatever extent it is unclear or doubtful that the reliabilist has a solution to these problems that doesn't also get the dogmatist off the hook, it is, to that extent, awkward for the reliabilist to use cognitive penetration as a reason to praise reliabilism and profane dogmatism. Imagine how awkward it would be for Mr Sleazy to lambaste Mr Davis' infidelity when Mr Sleazy's excuses for his infidelity are not clearly better than those of Mr Davis. Reliabilists have missed the way these problems undercut their ability to use cognitive penetration by desire as an objection to dogmatism. These problems may be old news, but their significance in this context is news to reliabilists.¹⁹

3.2. Process Reliabilism

Since indicator reliabilism suffered from extensive cognitive penetration problems, reliabilists might appeal to:

Simple Process Reliabilism: A belief is (prima facie) justified iff it is produced by a reliable process.

For now, we will assume that a process is *reliable* iff, in most of the closest possible worlds (to that of the subject), its outputs are mostly true. We will consider whether an alternative account of reliability will save the process reliabilist in 3.3. Does Simple Process Reliabilism, like Simple Indicator Reliabilism, allow a possible belief to be justifiably based solely on a desire? I think so, and some reliabilists think so too (e.g.,

¹⁸ We should not, however, conflate these problems with the problems raised by Bonjour's (1985) clairvoyance examples. Bonjour's point is not that clairvoyance leads to reliability in some sort of trivial way. Nor is it about clairvoyance as such. The problem is that the subjects in question fail to have the right sort of perspective on their clairvoyant beliefs (recall the relevant points in sec 2.3 above).

¹⁹ I defend the modal profile problems at much greater length in my ms. Among other things, I provide further examples of similarity-constituting truths, consider alternative accounts of reliability, and criticize Comesaña's (2010, sec 8.2) three ways of remedying reliabilism's problems with necessary truths.

Goldman 1979: 43-44²⁰ and Comesaña 2002: 262-3)—but the generality problem prevents there from being a decisive argument to this effect.

As the generality problem shows, the process reliabilist needs an account to tell us what counts as the process whose reliability determines the justification of the target belief. Let me be clear that my objection to process reliabilism is *not* the generality problem, and if anything, the generality problem makes it *harder* for me to press my objection. Since process reliabilists typically say very little about process individuation,²¹ it is difficult, if not impossible, to show decisively that process reliabilism satisfies Filter Failure 2. For they can always hold out hope that there is some account of process individuation that will allow them to avoid the cognitive penetration troubles. Despite this limitation, I can make it very plausible that they suffer from these problems.

The narrower and more specific the process, the easier it is for the modal profile of a proposition or epistemic angels and demons to ensure that the process is reliable. Consider, for example, Comesaña's (2006) solution to the generality problem. His idea is that the relevant process is always in this form: *a belief that P's being based on X*, where X picks out the basis of the belief. This is a very narrow process, and it essentially collapses the distinction between indicator and process reliabilism. Consequently, all the counterexamples to Simple Indicator Reliabilism will work against Simple Process Reliabilism. For example, the modal profile of Fermat's Conjecture makes *believing Fermat's conjecture on the basis of my desire for moldy cheese* a reliable process for exactly the same reasons it makes my desire for moldy cheese a reliable indicator of the conjecture.

On the other hand, the broader and more general the process, the easier it is for a desire to be an input to a process that is overall reliable. To illustrate the idea, suppose a process has been used 100 times and only takes desires as inputs two times. Even if the process gets the wrong answer on both of those occasions, the process might be overall 98% reliable. Since justification is determined by the reliability of the process, those two false beliefs would be justified even though they are based solely on a desire.

At this point, the reliabilist might wish to modify her view and say that justification shouldn't track the reliability of the process (full stop), but the reliability of the process given a certain type of input. Call such a view *input-restricted reliabilism*.²² Input-restricted reliabilism will get the correct result in the case from the previous paragraph. Although the process is overall 98% reliable, it is very unreliable when it takes desires as inputs. Nonetheless, input-restricted reliabilism falls prey to the same sort of counterexamples that afflicted indicator reliabilism.

Suppose Fearful Freddy has some process that is overall reliable and that can take affective states as inputs. Freddy has a deeply ingrained intellectual character which virtually never allows his affective states to inappropriately influence his beliefs, but he has one flaw. When he becomes (irrationally) terrified of snakes, this terror tends to be an input to the relevant process and produces the belief that he will soon encounter a

²⁰ All page references to this paper are to the reprinted version.

²¹ The most detailed account of process individuation I know of is that of Beebe (2004), but I don't see how his account will help the reliabilist in this context.

²² A reliabilist might go even further and suggest that input-types affect process individuation, such that a difference in the type of input guarantees a difference in the process. On such a view, even if a single system takes both experiences and desires as inputs, the process the system employs when it takes desires as inputs is necessarily distinct from the process employed when it takes experiences as inputs. Such a view would be subject to the same sort of problems that afflict input-restricted reliabilism.

snake. Thanks to the work of the demon, his irrational fears come true and so the overall reliable process is still reliable when it takes affective states as inputs.

Or suppose Hateful Harry has some process that is overall reliable and that can take affective states as inputs. It might be that Harry's deeply ingrained intellectual character virtually never allows his desires to inappropriately influence his beliefs, but he has one weak spot. He hates me and he knows it's my lifelong dream to win a Grammy. He has never heard me sing and has no other reason to think I'm a terrible singer. His hate for me is an input to the relevant process which produces the belief that I won't win a Grammy. This process also produces general beliefs to the effect that I'm miserable, unhappy, and will never amount to anything. Harry has no reason to think these other beliefs are true, but, as it turns out, I'm deeply depressed (it's really upsetting to know that I'll never win a Grammy). These beliefs are, therefore, not only true, but also true in the closest possible worlds. Consequently, the relevant process is not only overall reliable, but also reliable when it takes affective states as inputs. The Freddy and Harry examples show that input-restricted reliabilism doesn't prevent a non-introspective belief from being justified solely on the basis of desires and the like.²³

Regardless of whether the process reliabilist individuates processes broadly or narrowly or something in between, it's doubtful that there is a principled account of process individuation that makes it *impossible* for *basing a non-introspective belief solely on a desire* to be (part of) a reliable process. And if it's not impossible, then process reliabilism satisfies Filter Failure 2 and thereby suffers from cognitive penetration problems that are worse than those (allegedly) faced by dogmatism.

Note that if the reliabilist is to show that she doesn't suffer from *any* cognitive penetration problems, she needs to show two things. She must show, of course, that her theory doesn't allow a non-introspective belief to be justified solely on the basis of a desire; however, she must also show that her theory doesn't allow a desire-manipulated seeming to be an input to a process that is overall reliable. In other words, she must also show that her theory doesn't satisfy (the sufficient condition identified by):

Filter Failure 1.5: A theory is false if it allows a belief to be justifiably based solely on a desire-manipulated seeming.

Filter Failure 1.5 is different than Filter Failure 1 because it doesn't include the condition that the desire-manipulated seeming be a *constituent* of justification. Since at least process reliabilism denies that seemings are constituents of justification, reliabilism isn't committed to satisfying Filter Failure 1. But, barring some surprising account of process individuation, it is committed to satisfying Filter Failure 1.5.²⁴ Just

²³ The Freddy and Harry counterexamples work, not by restricting the possible inputs/outputs of the relevant processes, but by restricting which possible inputs/outputs of the process make a difference to its reliability in the world under evaluation. Input-restricted reliabilism claims that a belief is justified iff the process that produced a belief is reliable *given a certain type of input*. We can assume that the relevant type of input is affective states. The deeply ingrained intellectual characters of Freddy (Harry) guarantee that, in the closest possible worlds, only one affective state is ever taken as an input to that process, the terror of snakes (hate of me). Hence, the reliability of the process *given that it takes some affective state as an input* is equivalent to the reliability of the process *given that it takes the terror of snakes (hate of me) as an input*. Since the manipulation of demons (modal profile of No Grammy and propositions affirming my misery and lack of prospects) guarantees the reliability of the process given that it takes terror of snakes (hate of me) as an input, it guarantees the reliability of the process given that it takes an affective state as an input. (See sec 5 of my manuscript for a more detailed use of a similar strategy in response to the Comesaña (2010)/Sosa (1991) suggestion that processes should be relativized to fields of propositions.)

²⁴ Why didn't I just rely on Filter Failure 1.5 in section 2? Mainly because the particular arguments I gave in that section are more plausible when we are talking about constituents of justification than when

as the modal profile of a proposition and the manipulation of angels/demons can guarantee reliability when desires are at issue, they can guarantee reliability when desire-manipulated seemings are at issue. If the reliabilist shows only that her theory doesn't satisfy Filter Failure 2, then her view would avoid *worse* cognitive penetration problems. Yet, if her view nonetheless satisfies Filter Failure 1.5, she would still suffer from problems that are *as bad* as the ones that afflict the dogmatist.

3.3. Reliabilist Replies

Suppose that Goldman, Comesaña, and I are correct that reliabilism is forced to approve of cases in which a subject bases her belief solely on some desire. How might a reliabilist try to save her view? In this section, I consider three strategies suggested by Goldman (1979). The first is to bite the bullet and to hold that when, e.g., direct wishful thinking is reliable, it does provide justification (cf. Goldman 1979: 43). I have no problem with this suggestion *provided that* two things are acknowledged. First, the reliabilist has a bigger bullet to bite than does the dogmatist. Second, and more importantly, the reliabilist who bites the bullet cannot use dogmatism's (alleged) cognitive penetration problems as a reason to prefer reliabilism. If reliabilists give themselves a free pass concerning cognitive penetration by desire, then, on pain of hypocrisy, they must extend that free pass to dogmatists.²⁵

The second strategy is to add a condition on reliabilism that demands a "non-manipulated environment" (Goldman 1979: 43). This fix is, at best, a partial solution. It avoids the counterexamples caused by demon manipulation (which is what Goldman and Comesaña were focused on), but not the ones caused by the modal profile of the believed truths. Furthermore, as Comesaña (2002: 262-3) seems to appreciate, a non-manipulation condition is *not* plausible as a condition on justification, even it is plausible as a condition on warrant (the property that makes true belief knowledge). Demon victims are in manipulated environments, but most assume that their perceptual beliefs can be justified. Furthermore, standard fake barn Gettier cases typically assume that you are in a manipulated environment, which robs your belief of knowledge but not justification.

The third strategy is to endorse Actual World Reliabilism (Goldman 1979: 43-4). Simple Process Reliabilism holds that it is reliability *in the subject's world* that matters: S's belief in possible world W is (prima facie) justified if and only if it results from a cognitive process that is reliable *in W*. Goldman's suggestion is that what matters is the reliability of the process in *the actual world*, and so we end up with something like this: S's belief in possible world W is justified if and only if it results from a cognitive process that is reliable *in the actual world*.²⁶ On such a view, the fact that wishful thinking is reliable in some remote world does not show that beliefs produced by wishful thinking in that world are justified. For it is assumed that wishful thinking isn't reliable in the actual world.

The first problem with this strategy is its assumption that there is no reliable process in the actual world that involves a non-introspective belief's being based solely on a desire. Just as those pesky problems with process individuation block me from decisively showing that process reliabilism satisfies Filter Failure 2, those pesky problems block the reliabilist from justifiably making this assumption. For example,

we are talking about whether a theory allows a belief to be justified solely on the basis of a desire-manipulated seeming.

²⁵ Recall that many think dogmatism's implications concerning cognitive penetration are intuitively correct and so deny that dogmatism has a bullet to bite at all (see nt 7 above).

²⁶ Comesaña's (2002) indexical reliabilism would be a variation on this theme.

without some account of process individuation at our disposal, how can we tell that there is no actual, broad reliable process that might take a desire as an input? And, remember, the reliabilist also needs to show that they avoid satisfying Filter Failure 1.5. So he needs to show that he doesn't allow any actual desire-manipulated seeming to be an input to a reliable process. But without an account of process individuation, he is in no position to make assumptions about whether there is an actual reliable process that takes desire-manipulated seeming as an input.

Let me stress again that my objection is *not* the generality problem. I'm not saying that tricky issues regarding process individuation show that reliabilism has some deep problem. Here I am merely pointing out that those tricky issues prevent the reliabilist from *justifiably claiming a certain advantage*. Except perhaps for the issues regarding angelic and demonic manipulation, the challenges I raised in 3.1 and 3.2 apply to both the actual and merely possible worlds. The reliabilist needs an account of process individuation, or an argument of some kind, to vindicate her assumption that there is no actual reliable process that involves a belief's being based solely on a desire (or solely on a desire-manipulated seeming).

Set that problem aside. If Actual World Reliabilism avoids approving of any actual cases of wishful thinking (direct or indirect), then I also think it avoids approving of merely possible cases of wishful thinking. If we set the first problem aside, then Actual World Reliabilism probably doesn't have any counterintuitive implications concerning cognitive penetration by desire. Yet the Actual World Reliabilist can't claim victory, because she secures this result by being absurdly *chauvinistic* with respect to the actual world. Consider some possible species of creature that is considerably smarter than we are, such that many of their belief-forming processes are not instantiated in the actual world. What makes them so much smarter than us might be the very fact that they rely on different belief-forming processes. Actual World Reliabilism is incompatible with those creatures having justification for their beliefs. Since they rely on cognitive processes that are not instantiated in the actual world, their cognitive processes are *ipso facto* not reliable in the actual world. Consequently, Actual World Reliabilism absurdly entails that their beliefs aren't justified.²⁷

4. Who Else Has a Plank in Their Eye?

In section 3, we saw that reliabilists point out the speck in the dogmatist's eye while having a plank in their own. The reliabilist's cognitive penetration problems are worse because they are more counterintuitive and reliabilism may have a problem with cognitive penetration even if dogmatism doesn't. Yet reliabilists aren't the only ones who have these worse cognitive penetration problems: virtue theorists and proper functionalists do too. In 4.1, we will see that their best defense against these problems appears to be a teleological requirement that demands a process aim at truth or reliability. In 4.2, we will find that the best defense is no defense at all.

4.1. Virtue Theory and Proper Functionalism

Although some virtue theories count as reliabilisms (including the two we'll consider below), I treat virtue theory separately because it isn't committed to holding that reliability is necessary or sufficient for justification.²⁸ Since virtue is generally cashed

²⁷ It's worth mentioning that actual reliabilists reject actual world chauvinism and hold that merely possible processes should yield justified beliefs. See, for example, Lyons (2013, sec 3) story about Grand Interworld Station.

²⁸ Virtue theory without reliabilism has some benefits, e.g., it arguably avoids the new evil demon problem for reliabilism.

out in terms of stable dispositions, a virtue account of justification typically understands justification in terms of stable dispositions of the agent. For example, Greco (1999) argues that we need to move from the claim that justification supervenes on the reliability of *processes* to the claim that justification supervenes on the reliability of *stable dispositions of the agent that are typically manifested when the agent is thinking conscientiously*.²⁹ By ‘thinking conscientiously,’ Greco has in mind “the usual state most people are in as a kind of default mode—the state of trying to form one’s beliefs accurately” (1999: 289). In section 3.2, we showed that processes could be reliable even when they involve a belief’s being based solely on a desire. If so, presumably a disposition can be reliable even when it is a disposition for an agent to base a belief solely on a desire, and even if it is a stable disposition that is typically manifested when the agent is trying to form beliefs accurately. Greco’s virtue theory, then, apparently satisfies Filter Failure 2 and suffers from worse problems than dogmatism.

Sosa’s virtue theory, in contrast to Greco’s, holds that a belief is justified only if the process that produces the belief is aimed at truth.³⁰ The appeal to teleology here has considerable intuitive appeal. If a process takes desires as inputs in such a way that a non-introspective belief is based solely on a desire, it can seem obvious that the process is *not* aimed at truth. Perhaps Greco’s virtue theory is so susceptible to reliabilism’s problems precisely because it doesn’t appeal to teleology.

Chapter 2 of Sosa’s *Knowing Full Well* is, in large part, an attempt to explain why his virtue theory doesn’t allow a certain sort of wishful thinking to yield knowledge, namely “the sort of wishful thinking that aims at the intellectual comfort of the believer” (2011: 15). In particular he considers cases in which wishful thinking aims *solely* at comfort (2011: 15-6) and wishful thinking that aims at truth but also aims at comfort, where the two aims compete with one another (2011: 29). But he never considers whether it’s possible for a non-introspective belief to be based solely on a desire and for this to be constitutive of the agent’s pursuit of truth. To avoid Filter Failure 2, Sosa, and perhaps virtue theorists more generally, needs an account of *aiming at truth* which makes it impossible for a process to aim at truth when it allows a non-introspective belief to be based solely on a desire.

Bergmann (2006, ch 5) and Graham (2012) defend proper functionalist accounts of justification. They contend, roughly, that a belief is *prima facie* justified just in case it is formed by a process that is (i) functioning properly, (ii) aimed at reliably forming true beliefs (alternatively: aimed at truth), and (iii) reliable in the environment for which it is designed (alternatively: reliable in normal conditions). On this view, although processes need to *aim* at reliability, they don’t need to achieve that aim, they don’t need to be reliable, in either the subject’s or the actual environment; they only need to be reliable in the environments for which they were designed. This feature of the view arguably allows it to escape reliabilism’s new evil demon problems (Bergmann 2006: 137-45).

²⁹ See Greco’s definition of ‘agent reliabilism’ on 287-88 and the further restrictions imposed by what he calls “VJ” (289). In some of his later work, Greco (e.g., Greco and Breyer 2008: 175) imposes an additional requirement on subjective justification, namely that these cognitive dispositions be *cognitively integrated* with the subject’s other cognitive dispositions. There’s no reason why dispositions involving basing a non-introspective belief solely on a desire couldn’t be integrated in the relevant sense with one’s other dispositions.

³⁰ Strictly speaking, Sosa is talking about knowledge, not justification, and he requires the belief, not the process, to be aimed at truth. Nothing significant hangs on these differences.

Suppose that processes can aim at truth or reliability even if they involve a non-introspective belief's being based solely on a desire. My arguments in section 3 could then be applied to show that those processes can be reliable in the environments for which they designed, and it's hard to see why such processes couldn't function properly. Hence, if proper functionalism avoids Filter Failure 2, it is because it requires that processes aim at truth. Like virtue theory, proper functionalism's hope for escaping reliabilism's problems depends on finding some explanation for why a process can't aim at truth when it (occasionally) involves basing a non-introspective belief on a desire. This explanation is harder to give than you might think.

4.2. Aiming to Avoid Reliabilism's Problems

Our brief consideration of virtue theory and proper functionalism suggests that their best hope for avoiding reliabilism's problems is by appealing to a teleological requirement which demands that processes aim at truth or reliability. We will consider three attempts to make good on this suggestion and find them all unsatisfactory.

Why could a subject be pursuing truth by basing a non-introspective belief on an experience but not a desire? A natural suggestion is that the experience, but not the desire, can be a reliable indicator of the perceptual belief's truth; however, we dispelled that suggestion in section 3 when we found various ways in which non-introspective beliefs could be reliably formed on the basis of desires. Indeed, the fact that desires can be reliable indicators is at least some reason to think that it is possible for a subject to pursue truth by basing a non-introspective belief solely on a desire.

Since Graham (2012) provides the fullest account of the relevant teleology that I can find, it is worth considering whether his account can allow his (or any other) theory to avoid reliabilism's problems. On his view, the process needs to aim at reliability. A process aims at reliability just in case, roughly, the process is reliable and the process exists because it's reliable (2012, especially secs 2.1, 2.3). We have the perceptual processes that we do because they were reliable and this reliability led to the survival and reproduction of our ancestors. These processes, therefore, aim at reliability. Notice, though, that a creature might have a certain belief-forming system, not because it was an ideal system, but because, of the systems a group of creatures started with, it was the only one reliable enough to make it likely enough that the creature would survive and reproduce in that environment. It's surely possible that some reliable system was selected, even though every now and then it allows a non-inferential belief to be based solely on a desire. In short, there is nothing in Graham's account that makes it impossible for the relevant sort of process to be aimed at reliability.³¹

Finally, one might appeal to desire's direction of fit. Beliefs typically aim at truth and so have world-to-mind fit; desires aim at realization, at making the world the way it

³¹ An anonymous referee suggested that, intuitively, the proper functionalist has more plausible implications regarding cognitive penetration by desire than both reliabilism and dogmatism. What was intuitively problematic to the referee was allowing a desire or desire-manipulated seeming to justify a belief when the creature isn't designed to function that way. Since only proper functionalism avoids those problematic cases, it comes out on top. My own intuition is that any design plan that allows desires to be the sole basis of a non-introspective belief is a bad design plan, a plan that gives desires a role they epistemically ought not have. I suspect many will share my intuition. But if you share the intuitions of the referee, then you might conclude that cognitive penetration by desire ultimately provides a reason to endorse proper functionalism. I wouldn't be too disappointed. First, I would still have shown that many other popular views can't be defended by appealing to dogmatism's cognitive penetration problems, such as entitlement conservatism, coherentism, reliabilism, and virtue theory. Second, I will have shown that three epistemologists—Steup, Goldman, and Lyons—complain about dogmatism's problems with cognitive penetration by desire when they have similar problems of their own.

is desired, and so have mind-to-world fit (cf. Platts 1979: 257). One might try to avoid reliabilism's problems with cognitive penetration by arguing for:

Direction of Fit Principle: a process can't aim at truth if it takes an input which has a mind-to-world direction of fit.

Any process that involves a subject's basing a belief solely on a desire will be a process that takes an input which has a mind-to-world direction of fit. If the Direction of Fit Principle is true, the requirement that a process aim at truth will enable a theory to avoid both satisfying Filter Failure 2 and reliabilism's worse cognitive penetration problems.

There are two problems with appealing to the direction of fit principle. First, it is a little on the ad hoc side. Some inputs in perceptual processes, e.g., the excitation of rods and cones, have no direction of fit and perception is still aimed at truth. But if a process can be truth-aimed when it has an input with no direction of fit, why can't it be truth-aimed when it has an input with a mind-to-world direction of fit?

Second, the principle is subject to two kinds of counterexamples. The first involves introspection: a desire that P is typically an input to the process that produces the belief that I desire P; however, introspective processes typically aim at truth (at least, if the proponent of the teleological requirement wants introspective beliefs to be justified, they had better aim at truth). The second kind of counterexample involves cases in which one bases a belief on a desire but not *solely* on a desire. Suppose I base my belief that P on my desire that P and a justified belief to the effect that, when I desire P, P is usually true. The Direction of Fit Principle mistakenly entails that the associated process couldn't be aiming at truth. Perhaps such processes can't produce justified beliefs for some reason or another, but they clearly can be aimed at truth. Of course, we could try to find a more restricted direction of fit principle that avoids these counterexamples, but the danger here is that we would just push the problem back and would need some explanation of why these restrictions hold.

To avoid reliabilism's problems, virtue theory and proper functionalism must make it impossible for a non-introspective belief to be justifiably based solely on a desire. We saw that their best chance to do so is to demand that a process aim at truth or reliability. There is intuitive merit to the idea that a process can't aim at truth (reliability) if it involves basing a non-introspective belief on a desire, but the three most promising ways of cashing it out aren't very promising. Furthermore, the fact that desires can be reliable indicators is at least some reason to think that they can be involved in processes that aim at truth. Proper functionalism and virtue theory, like reliabilism, have a plank in their eye: they satisfy Filter Failure 2 and suffer from cognitive penetration problems that are worse than those of dogmatism.

5. Does Everyone Have a Problem?

Let's survey the damage. If dogmatism has a genuine problem with cognitive penetration by desire, the following views have cognitive penetration problems that are worse (and may have those problems even if dogmatism doesn't have genuine problems with cognitive penetration):

- Reliabilism, both indicator and process;
- Virtue Theories; and
- Proper Functionalism.³²

³² It's worth mentioning two other views will satisfy Filter Failure 2 and so suffer from reliabilism's worse problems. The first is sometimes known as 'doxastic conservatism', though this use of 'conservatism' is different from the use of the term elsewhere in the paper. The basic idea of the view is that, if S believes that P, then S thereby has prima facie justification for P. William Lycan (1988, ch 8),

The following views would have cognitive penetration problems that are at least as bad:

- Entitlement Conservatism
- Experiential Coherentism

Given how popular these views of justification are, there's a good chance that if you are reading this paper, your view falls in one of the above camps.

To be sure, I do think some views may avoid dogmatism's cognitive penetration problems. Consider, for example, **Inferential Conservatism**. On this view, perceptual justification is inferential justification. One must infer that P is true from justified belief, say, in the conjunction that *P perceptually seems true and seemings are reliable indicators of reality* (e.g., Fumerton 1995). For our purposes the key contrast with entitlement conservatism is that this view denies that seemings are ever, strictly speaking, constituents of perceptual justification; it is only *justified beliefs about seemings* that are constituents of one's justification. Even if wishfully-produced seemings are garbage, justified beliefs about such seemings needn't be. Suppose, for example, I justifiably believe the premises of the following argument, and I see that they support their conclusion:

1. It seems to me that the object is gold.
2. My seemings about whether an object is gold are highly reliable.
3. Therefore, the object is gold.

Even if it turns out that my gold-seeming is caused by wishful thinking, I have justification for the conclusion. In fact, I think it's a problem for any theory that says otherwise (cf. my 2010: 541-2). Consequently, those who endorse something like inferential conservatism, such as Fumerton (1995), successfully avoid cognitive penetration problems, at least insofar as perceptual justification is concerned. Of course, these views have problems of their own (e.g., they notoriously struggle with skepticism), so they don't automatically deserve our endorsement because they come out on top on this one issue. But let's give credit where credit is due.

Two other positions are worth mentioning as possible ways to avoid dogmatism's cognitive penetration problems. First, perhaps we can avoid dogmatism's cognitive penetration problems by holding that only *factive* mental states can be constituents of one's justification, as is suggested by Williamson (2000) and various versions of disjunctivism. On this view, it isn't an experience of P, but *perceiving P* that provides justification for P. The appeal to factivity will seem attractive in this context if it is assumed that desires can't partly constitute the relevant sort of factive mental state. I worry that this assumption can't be vindicated. If experiences can partly constitute justifying factive mental states, such as perceiving that P, then why can't desires? Perhaps my worry can be addressed, but I won't try to settle the matter here.

Second, some philosophers have tried to restrict dogmatism with the explicit aim of avoiding (unrestricted) dogmatism's cognitive penetration problems.³³ Someday, hopefully soon, I'll argue that these attempts still suffer from cognitive penetration problems. Nonetheless, I think there is a lot going for this general strategy and it's worth exploring further.

Since I acknowledge that some views may avoid dogmatism's (alleged) cognitive penetration problems, when I say that "If dogmatists have cognitive penetration problems, you do too", I'm not quite talking to everyone; but I'm sure talking to a lot of

Richard Swinburne (2001), and Gilbert Harman (1986), among others, have endorsed views in this neighborhood. The other view that falls prey to reliabilism's worse problems are versions of coherentism that put no causal constraints on which beliefs are justified (see nt 11 above).

³³ See, e.g., Brogaard 2013; Markie 2013; McGrath 2013, forthcoming; and Siegel 2013.

you. And I'm talking to some of you (cough, cough, reliabilists, Steup) who picked on us poor dogmatists when you had problems at least as bad, and that's just not cool.³⁴

Conclusion

Dogmatism allows wishfully-produced seemings to provide prima facie justification for their contents, but this seems counterintuitive even to a dogmatist like me. But before you mock the speck in the dogmatist's eye, be sure that you don't have a plank in your own. Reliabilists, such as Jack Lyons and Alvin Goldman, have been quick to denounce dogmatism because of its consequences for cognitive penetration by desire. We've seen, however, that these denouncements are nothing more than hypocrisy, for they suffer from cognitive penetration problems that are even worse: reliabilism allows a non-introspective belief to be justified solely on the basis of a desire. Proper functionalism and virtue theory join reliabilism in suffering from implications worse than those of dogmatism. Entitlement conservatism and many versions of coherentism suffer from implications that are as bad.

If cognitive penetration is a genuine problem for dogmatism, it is a much more ubiquitous problem than is currently recognized. If dogmatism has a problem, you do too (well, most of you anyway). None of this shows, of course, that the dogmatist *doesn't* have cognitive penetration problems after all. What it shows is that many epistemologists can't use dogmatism's cognitive penetration troubles as a reason to prefer their own views. Once those epistemologists recognize that they have cognitive penetration problems of their own, I hope they'll join the dogmatist's attempts to explain away the intuition. But maybe the intuition shouldn't be explained away. Maybe it should lead us to revise our theories. Only time will tell.³⁵

References

- Alston, William. 1989. "An Internalist Externalism." Reprinted in Alston, William *Epistemic Justification: Essays in the Theory of Knowledge*. Ithaca: Cornell University Press, pgs. 227-245.
- Beebe, James. 2004. "The Generality Problem, Statistical Relevance, and the Tri-Level Hypothesis." *Nous* 38: 177-195.
- Bergmann, Michael. 2006. *Justification without Awareness: A Defense of Epistemic Externalism*. New York: Oxford University Press.
- Bonjour, Laurence. 1985. *The Structure of Empirical Knowledge*. Cambridge: Harvard University Press.
- Brogaard, Berit. 2013. "Phenomenal Seemings and Sensible Dogmatism." In Tucker, Chris (ed.).

³⁴ Since evidentialism is so popular, one may wonder why I didn't address it in this paper. Insofar as evidentialism is merely the claim that justification supervenes on evidence, it doesn't take a stand on what counts as evidence. Some, but perhaps not all, accounts of evidence will lead to cognitive penetration problems. Versions of evidentialism that endorse dogmatism, entitlement conservatism, or coherentism have counterintuitive implications concerning cognitive penetration; however, an evidentialism that endorses inferential conservatism may avoid cognitive penetration problems, at least insofar as perceptual justification is concerned.

³⁵ Helpful comments on the paper were provided by Nathan Ballantyne, Parker Crutchfield, Logan Gage, Thomas Grundmann, Joachim Horvath, Chris Kelp, Peter Markie, Matthew McGrath, Jonathan McKeown-Green, Glen Pettigrove, anonymous referees, and the audiences at the University of Glasgow, University of Edinburgh, University of Birmingham, the 2012 meeting of the *Australasian Association of Philosophy*, and the 2012 meeting of the *German Society for Analytic Philosophy*. The generous support of the Royal Society of New Zealand and John Templeton Foundation made the publication of this article possible. The views defended in this paper are mine and do not necessarily reflect those of either institution.

- Chudnoff, Elijah. 2011. "The Nature of Intuitive Justification." *Philosophical Studies* 153: 313-333.
- Cohen, Stewart. 1999. "Contextualism, Skepticism, and the Structure of Reasons." *Philosophical Perspectives* 13: 57-89.
- Comesaña, Juan. 2010. "Reliabilist Evidentialism." *Nous* 44: 571-600.
- _____. 2006. "A Well-Founded Solution to the Generality Problem." *Philosophical Studies* 129: 27-47.
- _____. 2002. "The Diagonal and the Demon." *Philosophical Studies* 110: 249-66.
- Elgin, Catherine. 2005. "Non-Foundationalist Epistemology: Holism, Coherence, and Tenability." In Steup, Matthias and Ernest Sosa (eds.). *Contemporary Debates in Epistemology*. Malden: Blackwell Publishing.
- Fumerton, Richard. 2006. *Epistemology*. Malden: Blackwell Publishing.
- _____. 1995. *Metaepistemology and Skepticism*. Boston: Rowman and Littlefield Publishers, Inc.
- Gage, Paul Logan. 2012. "Review of *Evidence and Religious Belief*." *American Catholic Philosophical Quarterly* 86: 372-75
- Goldman, Alvin. 2009. "Internalism, Externalism, and the Architecture of Justification." *Journal of Philosophy* 106: 309-38.
- _____. 2008. "Immediate Justification and Process Reliabilism." In Smith, Quentin (ed.). *Epistemology: New Essays*. Oxford: Oxford University Press.
- _____. 1979. "What is Justified Belief?" in Pappas, George (ed.). *Justification and Knowledge*. Dordrecht: Reidel. Reprinted in Goldman, Alvin. 2012. *Reliabilism and Contemporary Epistemology*. Oxford: Oxford University Press.
- Graham, Peter J. 2012. "Epistemic Entitlement." *Nous* 46: 449-482.
- Greco, John. 1999. "Agent Reliabilism." *Philosophical Perspectives* 13: 273-296.
- Greco, John and Daniel Breyer. 2008. "Cognitive Integration and the Ownership of Belief: Response to Bernecker." *Philosophy and Phenomenological Research* 76: 173-84.
- Haack, Susan. 1993. *Evidence and Inquiry: Towards Reconstruction in Epistemology*. Malden: Blackwell Publishing.
- Harman, Gilbert. 1986. *Change in View*. Cambridge: MIT Press.
- Hazlett, Allan. 2006. "How to Defeat Belief in the External World." *Pacific Philosophical Quarterly* 87: 198-212
- Huemer, Michael. 2013a. "Phenomenal Conservatism Über Alles." In Tucker, Chris (ed.).
- _____. 2013b. "Epistemological Asymmetries between Belief and Experience." *Philosophical Studies* 162: 741-748.
- _____. 2001. *Skepticism and the Veil of Perception*. Lanham: Rowman & Littlefield Publishers, Inc.
- Lycan, William. 2013. "Phenomenal Conservatism and the Principle of Credulity." In Tucker, Chris (ed.).
- _____. 1988. *Judgment and Justification*. Cambridge: Cambridge University Press.
- Lyons, Jack. 2013. "Should Reliabilists Be Worried About Demon Worlds?" *Philosophy and Phenomenological Research* 86: 1-40.
- _____. 2011. "Circularity, Reliability and the Cognitive Penetrability of Perception." *Philosophical Issues* 21: 289-311.
- Macpherson, Fiona. 2012. "Cognitive Penetration of Colour Experience: Rethinking an Indirect Mechanism." *Philosophy and Phenomenological Research* 84: 24-62.
- Markie, Peter. 2013. "Searching for True Dogmatism." In Tucker, Chris (ed.).

- _____. 2005. "The Mystery of Perceptual Justification." *Philosophical Studies* 126: 347-73.
- McGrath, Matthew. 2013a. "Siegel and the Impact for Epistemological Internalism." *Philosophical Studies* 162: 723-32.
- _____. 2013b. "Phenomenal Conservatism and Cognitive Penetration: the 'Bad Basis' Counterexamples." In Tucker, Chris (ed.).
- Plantinga, Alvin. 1993. *Warrant: The Current Debate*. Oxford: Oxford University Press.
- Platts, Mark. 1979. *Ways of Meaning*. London: Routledge and Kegan Paul.
- Pryor, James. 2000. "The Skeptic and the Dogmatist." *Noûs* 34, 517-549.
- Siegel, Susanna. 2013. "The Epistemic Impact of the Etiology of Experience." *Philosophical Studies* 162: 697-722.
- _____. 2012. "Cognitive Penetrability and Perceptual Justification." *Nous* 46: 201-222.
- Skene, Matthew. 2013. "Seemings and the Possibility of Epistemic Justification." *Philosophical Studies* 163: 539-59.
- Smith, Martin. 2013. "Entitlement and Evidence." *Australasian Journal of Philosophy* 91: 735-53.
- Sosa, Ernest. 2011. *Knowing Full Well*. Princeton: Princeton University Press.
- _____. 2007. *A Virtue Epistemology*. New York: Oxford University Press.
- _____. 1991. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge: Cambridge University Press.
- Steup, Matthias. 2013. "Does Phenomenal Conservatism Solve Internalism's Dilemma?" In Tucker, Chris (ed.).
- _____. 2004. "Internalist Reliabilism." *Philosophical Issues* 14: 403-25
- Stokes, Dustin. Forthcoming. "Cognitive Penetrability of Perception." *Philosophy Compass*.
- _____. 2012. "Perceiving and Desiring: A New Look at the Cognitive Penetrability of Experience." *Philosophical Studies* 158: 477-92.
- Swinburne, Richard. 2001. *Epistemic Justification*. Oxford: Oxford University Press.
- Tucker, Chris. Manuscript. "On What Inferentially Justifies What: The Vices of Reliabilism and Proper Functionalism."
- _____. 2013a. "Seemings and Justification: An Introduction." In Tucker, Chris (ed.).
- _____. (ed.). 2013b. *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism*. Oxford University Press.
- _____. 2012. "Movin' On Up: Higher-Level Requirements and Inferential Justification." *Philosophical Studies* 157(4), 323-40.
- _____. 2011. "Phenomenal Conservatism and Evidentialism in Religious Epistemology." In VanArragon, Raymond and Kelly James Clark (eds.). *Evidence and Religious Belief*. Oxford University Press, pgs 52-73.
- _____. 2010. "Why Open-Minded People Should Endorse Dogmatism." *Philosophical Perspectives* 24: 529-45.
- White, Roger. 2006. "Problems for Dogmatism." *Philosophical Studies* 131: 525-557.
- Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
- Wright, Crispin. 2007. "The Perils of Dogmatism." In Susana Nuccetelli and Gary Seay (eds.). *Themes from G. E. Moore: New Essays in Epistemology and Ethics*. New York: Oxford University Press.
- _____. 2004. "Warrant for Nothing (and Foundations for Free)?" *Aristotelian Society Supplementary Volume* 78: 167-212.