

*Cognitive Penetrability
and Perceptual Justification**

SUSANNA SIEGEL

It is sometimes said that in depression, everything looks grey.¹ If this is true, then mood can influence the character of perceptual experience: depending only on whether a viewer is depressed or not, how a scene looks to that viewer can differ even if all other conditions stay the same. This would be an example of cognitive penetration of visual experience by another mental state. Here the influential cognitive state is a mood. Other putative examples of cognitive penetrability involve beliefs: to the reader of Russian, the sheet of Cyrillic script looks different than it looked to her before she could read it. When you know that bananas are yellow, this knowledge affects what color you see bananas to be (an achromatic banana will appear yellowish).² Or suppose that to a vain performer, the faces in the audience ranged in their expression from neutral to pleased, but remarkably no one ever looked disapproving, while to an underconfident performer, the faces in the audience ranged in their expression from neutral to displeased, but remarkably no one ever looked approving. Potential cognitive penetrators thus include moods, beliefs, hypotheses, knowledge, desires, and traits.

In some cases, cognitive penetration can be epistemically beneficial. If an x-ray looks different to a radiologist from the way it looks to someone lacking radiological expertise, then the radiologist gets more information about the world from her experience (such as whether there's a tumor) than the non-expert does from looking at the same x-ray. If Iris Murdoch and John McDowell are correct in thinking that having the right sort of character lets you see more moral facts than someone lacking that character sees when faced with the same situation, then there too, your perceptual experience becomes epistemically better, thanks to its being penetrated by your character.³

In other cases, however, cognitive penetration seems to make experience epistemically worse. The challenge to perceptual justification posed by cognitive penetrability seems related to a circular structure of belief-formation that

it introduces. In the simplest case, your experience is cognitively penetrated if it presents the world as being a certain way, only because that's the way the penetrating belief presents the world as being. For instance, suppose Jill believes that Jack is angry at her, and this makes her experience his face as expressing anger. Now suppose she takes her cognitively penetrated experience at face value, as additional support for her belief that Jack is angry at him (just look at his face!). She seems to have moved in a circle, starting out with the penetrating belief, and ending up with the same belief, via having an experience. From Jill's point of view, she seems to be gaining additional evidence from this experience for her belief that Jack is angry at her, elevating the epistemic status of that belief.

This situation seems epistemically pernicious. In general, visual experience purports to tell you what the world is like, allowing you to check your beliefs against reality. But if behind the scenes, the penetrating states are stacking the tribunal of experience in their own favor, then while experience will seem to let you check your beliefs against the world – to you, this will be just what's happening – really you'll just be checking your beliefs against your beliefs. The tribunal will be corrupted. On the face of it, epistemic elevation in such a circumstance seems illicit.

We can compare this situation to a gossip circle. In a gossip circle, Jill tells Jack that *p*, Jack believes her but quickly forgets that she's the source of his belief, then shortly afterward Jack tells Jill that *p*. It seems silly for Jill to take Jack's report that *p* as providing much if any additional support for *p*, beyond whatever evidence she already had. On the face of it, this looks like a feedback loop in which no new justification is introduced. Similarly, when beliefs are formed on the basis of cognitively penetrated experience, it is as if your belief that *p* told you to have an experience that *p*, and then your experience that *p* told you to believe that *p*.

If epistemic elevation is illicit in these cases, then a theory of perceptual justification shouldn't predict that such elevation occurs. More generally, we can ask: What epistemic roles can be played by cognitively penetrated perceptual experiences? And which theories of perceptual justification best explain the epistemic roles of such experiences?

This paper addresses these questions by concentrating on a simple and popular theory of perceptual justification known as dogmatism. I will argue that there are cases in which dogmatism predicts that a cognitively penetrated visual experience can elevate the subject from an epistemically bad situation to an epistemically better one, yet in which it is implausible to suppose that such epistemic elevation takes place.

Although the discussion will concentrate mainly on dogmatism, the purpose of this point of focus is to bring the contours of the issue into clearer view. The challenge to dogmatism posed by cognitive penetration applies to other theories of perceptual justification as well. My goal isn't to put one or another theory of perceptual justification to rest, but rather to reveal

the contours of a problem so that we might better see the contours of a solution, whatever the correct theory of perceptual justification turns out to be. The constraints that I'll argue are imposed by cognitive penetration can be met in a variety of different ways by different theories, though the form of dogmatism I'll focus on seems ill-equipped to respond to it.

The discussion will proceed as follows. Section 1 characterizes the phenomenon of cognitive penetrability in more detail, and section 2 introduces dogmatist theories of perceptual justification. Section 3 introduces two cases of cognitively penetrated experiences which *prima facie* do not justify corresponding beliefs. Section 4 argues that in cases like these, the elevation prediction is implausible. Section 5 explains how the challenge applies to theories of perceptual justification other than dogmatism, and discusses ways that the challenge generalizes beyond the scenario involving elevation to an improved epistemic situation.

1. What is Cognitive Penetrability?

We've been talking so far about the cognitive penetrability of visual experiences. But what are visual experiences? Visual experiences are conscious states typically had while seeing. Because they are conscious states, they have phenomenal features: there is something it is like to have a visual experience. When two visual experiences differ in their phenomenal features – as do, for instance, the visual experience you probably have while reading this paper, and the visual experience you'd have if you were looking at the horizon of the ocean with the moon shining on the water – there is a difference between what it is like to have each experience. Which phenomenal features a visual experience has depends not only on which scene (if any) the subject is looking at, but on where they're standing, their visual acuity, and what they're attending to.

For the purposes of this discussion, I'll be assuming that visual experiences have contents that can be true or false, and that the truth or falsity of the contents co-varies with the truth (veridicality) or falsity (falsidicality) of the experience. For a content to count as the content of a visual experience, the content must characterize how things look to the subject of the experience. Nothing will be lost in our discussion if we interdefine the phenomenal features and contents, as if the phenomenal features of experience just were the entertaining of contents in an experiential mode (rather than some other mode, such as a conative mode).

Cognitive penetrability is a kind of causal influence on visual experience. Not every kind of influence by a cognitive state on visual experiences is a case of cognitive penetrability. You can choose to move your head to see what's behind you, or to focus your attention in order to see something in more detail. Disturbed by the dead squirrel in the road, you may look elsewhere while bicycling by. Intent on spying on a man in the airport, you may pay

no attention to the billboards. Here intentions, desires, and aversions play a selective role in which visual experiences you will have, by selecting the location from which stimuli will give rise to visual experience in the first place. These are cases of relatively global selection from among possible stimuli. In the spy example, non-experiential mental states – the decision to come to the airport, the desire to keep track of the man, etc – help determine that you’re perceiving the airport rather than your house, and specifically the doorway from Customs rather than the paintings on the wall.

Although we could consider global selection to be a kind of cognitive penetration in a broad sense, it will bring the epistemic problems into sharper focus if we define cognitive penetrability more narrowly. By themselves, global selection effects do not obviously lead to any illicit feedback loops, as these effects simply determine where information will come from. Feedback loops get going when we introduce a kind of insensitivity to stimuli, so that the visual experience you end up with is unduly influenced by the penetrating states. The insensitivity could either take the form of relative indifference to the stimuli, or it could take the form of a selection bias. Both can be illustrated by extreme cases. In an extreme case of indifference to stimuli, no matter what you look at, you end up having a visual experience of an angry face. In an extreme selection bias, you’re not able to attend to anything other than angry faces, and nothing else registers with you.

We can distinguish between three aspects of cognitive penetrability: the penetrated aspects of visual experience; the potential penetrators; and the type of influence they have. I will concentrate on the sensitivity of the content of visual experience to doxastic states (including both beliefs and hypotheses), desire, mood, and emotion. Here is a first pass at a definition of cognitive penetrability:

Cognitive Penetrability (first pass):

If visual experience is cognitively penetrable, then it is nomologically possible for two subjects (or for one subject in different counterfactual circumstances, or at different times) to have visual experiences with different contents while seeing the same distal stimuli under the same external conditions, as a result of differences in other cognitive (including affective) states.

For all the first pass says, when the penetrating states influence the content of visual experience, they do so by affecting what parts or aspects of the distal stimuli the subjects fixate on or covertly attend to. For instance, the first pass would count the following as cases of cognitive penetrability:

Expertise-influenced fixation:

- Before and after X learns what pine trees look like, pine trees look different to her, and the visual experiences she has under the same external conditions differ in their content. But this is because gaining pine-tree-expertise makes her fixate on the shapes of the leaves on the trees. If a novice

fixated the way the expert did, then her experience would have the same contents. The expertise influences experience content, by influencing fixation points.

- You and moth expert X take a walk in the forest, looking for moths on tree bark. You look at the same piece of bark. X sees moths where you see none. That's because her familiarity with the exact shapes of moths lets her more easily fixate on moth-shaped pieces of bark. Sometimes, she sees through its camouflage.

Covert attention:

- X, a subject in a psychology experiment, fixates on the cross in the middle of the screen. When primed with hypothesis H1, X finds herself attending to the left side of the display, where she sees three green bars. Experiments with other subjects suggest that if X were primed with hypothesis H2, X would attend to the right side of the display, where she would see four red circles.

In the cases in which background expertise influences fixation points, the distal stimuli (pine trees, tree bark) is held constant in the sense that under the same conditions, expert and novice view the same trees and the same tree bark, and these things don't change. In the case of covert attention, the distal stimulus likewise stays the same: the dots and the bars are each there to be seen, no matter which hypothesis X is primed with. These are cases in which background state has a selective effect. It selects which part of the distal stimulus comes to be represented in subject's visual experience.

If one interpreted "distal stimulus" in a more fine-grained way, so that distal stimulus could be determined by fixation point and covert attention, then the first pass wouldn't after all count these as illustrations of cognitive penetrability, since the distal stimuli in that more fine-grained sense would differ in the relevant cases. The distal stimuli would differ for expert and novice in the expertise cases, and would differ depending on priming in the covert attention case.

These cases are illustrations of cognitive penetrability, considered broadly. But it will be simpler to avoid the complexities introduced by focal and non-focal attention, and define cognitive penetrability more narrowly, so that fixation points and non-focal attention are part of what is held constant, rather than part of what can vary with background state. Although some powerful potential examples of cognitive penetration involve influences on where attention is directed,⁴ the discussion will be more tractable if we set aside the complications introduced by counting attention as an effect of cognitive penetration. This suggests a second pass:

Cognitive Penetrability (second pass):

If visual experience is cognitively penetrable, then it is nomologically possible for two subjects (or for one subject in different counterfactual circumstances, or

at different times) to have visual experiences with different contents while seeing *and attending to* the same distal stimuli under the same external conditions, as a result of differences in other cognitive (including affective) states.

In most cases of cognitive penetration, the following counterfactual will hold:

If the subject were not in background state B but was seeing and attending to the same distal stimuli, she would not have an experience with content p.

If there is any cognitive penetration in the actual world, this counterfactual will hold much of the time. But it does not provide a definition of cognitive penetration, for the usual sorts of reasons. In some situations, a subject has an experience that p because of her background state B, but were she not in B, she would be in state B*, which would also lead her to have an experience that p. In other situations, a subject has an experience that p because of her background state B, but were she not in B, a higher power would cause her to have an experience that p.

With the second pass on the table, we can see how an opponent of cognitive penetrability might try to re-describe the putative cases of cognitive penetrability. They might appeal to any of these four alternatives:

1-Introspective error. When you're depressed, things don't really look grey. But you believe that they look grey.

Here the background state is influencing your beliefs about your experience, but not your experience contents themselves.

2-Influence limited to first-order beliefs downstream of experience.

When you learn what pine trees look like, your experience doesn't represent the trees as being pine trees – it just represents color, shape, illumination and motion properties. But you form the belief that they are pine trees.

Here the background state is influencing your first-order beliefs, without influencing the contents of experience itself.

3-Selection effect. Flowers really do smell nice, but you only notice this when you're in a good mood. Likewise, the decrepit house is sinister-looking, but you only notice this when primed with the hypothesis that the villain lived there.

Here the background state has a selection effect.

I'm going to assume that there are some genuine cases of cognitive penetration of visual experience – cases that cannot accurately be re-described in any of these ways. This assumption would be a substantial empirical

claim. Rather than defend the assumption in this paper, the main point is to explore the epistemic challenge that arises once the assumption is made.

There are two main points of controversy surrounding the assumption of cognitive penetrability. The first point concerns its extent. It is an open question to what extent visual experiences are influenced by other aspects of cognition. There are many suggestive experimental results,⁵ the brain area V1 is connected via thousands of neurons to the amygdala, so brain architecture does not rule out emotional influences on visual experiences. But facts about brain architecture alone will not settle the question, and it remains unclear whether cognitive penetration of visual experience is the exception or the norm.

Second, it is a potential consequence of cognitive influence on visual experience that its contents are 'rich', so that it can represent such properties as being sinister, or being a pine tree needle. Some philosophers think there are limits on how 'rich' the contents of experience can be. For instance, some would deny that being sinister is really way that a house (or anything else) can look, on the grounds that visual experience can represent only a quite limited class of properties, not going far beyond color, shape, illumination, and motion.

These two points of controversy are independent. Cognitive penetrability is a thesis about the etiology of experience contents, whereas theses affirming or denying richness concern what contents experiences can have, rather than their etiology. One might deny that experiences are cognitively penetrable, but hold that some experiential contents are 'rich'. For instance, presumably causation falls on the 'rich' side of the rich/thin divide, but it is coherent to suppose that experiences represent causation due to hard-wiring, not as a result of cognitive penetrability. Conversely, one might allow that experiences can be cognitively penetrated, while denying that this ever results in 'rich' contents, because there are limits to what contents experiences can come to have as the result of cognitive penetration.

It is not necessary to settle the controversy over the extent of cognitive penetration in order to get the epistemic challenge off the ground. Theories of perceptual justification make predictions about the circumstances under which visual experiences can justify beliefs. These predictions concern hypothetical cases as well as actual cases. To generate an objection to dogmatism, for example, it is enough if there is a hypothetical case in which it makes a false prediction about whether an experience that *p* in the hypothetical circumstances would immediately justify the subject in believing *p*.⁶ So for the purpose of understanding the epistemological issue, we can set aside the empirical question about the extent to which visual experiences are cognitively penetrated. As it happens, although the examples of cognitive penetration that I will discuss are probably not actual, they are also not far-fetched.

2. Dogmatism

Dogmatism is called ‘dogmatism’ because of the response to skepticism that it recommends. It consists of two main claims. First, absent defeaters, having a perceptual experience with content *p* suffices to give you justification for believing *p*. Second, when a subject *S*’s experience justifies believing *p*, the justification is immediate: there need be no further propositions that *S* must be justified in believing, in order for experience to justify her in believing that *p* – or if there are, being justified in believing these propositions need not play a role in *S*’s getting justification to believe *p* from her experience.⁷ The main target of the epistemic challenge from cognitive penetration is the first claim, and so the target is somewhat broader than dogmatism.

Dogmatism is motivated by the maximally simple structure it accords to perceptual justification. Such simplicity is appealing to the extent that perceptual justification seems offhand to be a straightforward affair. It is also motivated by the respect it pays to the presentational aspect of visual experience. It is part of the distinctive phenomenology of seeing that we are in contact with our immediate surroundings. Experience seems to tell us how things are in our environment. Arguably this is part of what makes it seem fit to be a tribunal that allows us to test beliefs against reality, so that if we want to know how long the stem of a rose is, or whether it has any thorns, or whether any mustard is left in the bottle, we can look and find out.

As stated, dogmatism seems to require that there are at least some contents of experience that can also be believed. Some thinkers deny this. According to them, experiences have truth-assessable contents, but these contents differ so fundamentally in their structure and nature that they cannot be believed. Sometimes such contents are said to be ‘non-conceptual’, where it is assumed that belief contents are ‘conceptual’. Many of these views, however, can nonetheless accept that experiences provide immediate justification for beliefs. They just have to provide an account of how ‘non-conceptual’ contents may be systematically related to belief contents. Providing such an account would need to be done anyway, regardless of views about the structure perceptual justification, in order to describe the differences between beliefs that are closer to the deliverances and those that are farther removed from it. For simplicity, I’ll be talking as if the same contents can be experienced and believed.

There are many possible versions of dogmatism. Dogmatism can be either pure or limited with respect to contents. Pure dogmatism places no limits on the values for *p* with respect to which experiences that *p* can immediately justify beliefs that *p*, absent defeaters, whereas content-limited dogmatism does invoke such limits. Dogmatism can also be limited or unlimited with respect to sensory modality. Modality-unlimited dogmatism places no limits on which sensory modalities host experiences that can provide immediate justification, absent defeaters, whereas according to modality-limited dogmatism,

only some kinds of experiences (e.g. visual experiences, or bodily experiences such as heaving a headache) provide such justification. Versions of dogmatism could also be expanded beyond perceptual modalities, to include other potential sources of justification, such as memory or testimony.

Finally, dogmatism is often discussed as a view about justification, where the notion of justification is supposed to be a generic one, tied to what is epistemically appropriate to believe. Views in the vicinity of dogmatism could also be defined for other epistemic notions, including knowledge (If you have a visual experience that p and no defeaters, then you know p), or for deontic notions (If you have a visual experience that p and no defeaters, then you're absolved of epistemic irresponsibility if you believe p). These views are variations on only the first of the two main elements of dogmatism: that absent defeaters, an experience that p by itself suffices to provide justification for p .

For our purposes, the relevant version of dogmatism is pure dogmatism about visual experience, using the generic notion of justification tied to what is epistemically appropriate to believe, where this is a binary rather than a degreed notion. The reason for considering pure dogmatism is that limits on the contents to which dogmatism applies are at odds with dogmatism's phenomenological motivation, and are irrelevant because the epistemic challenge can get going even with 'low-level' contents, such as color and shape. Later on (in sections 4 and 5), we will see how the challenge from cognitive penetrability applies to dogmatism when it uses a degreed notion of justification. This version of dogmatism involves a notion of epistemic improvement, rather than justification simpliciter. According to it, if you have a visual experience that p and there are no defeaters, then experience by itself suffices to give proposition p a little boost of justification,⁸ but that evidential boost may fall short of making belief in p epistemically appropriate.

3. Some cases of cognitively penetrated experiences

For the purpose of assessing dogmatism's predictions about the epistemic status of cognitively penetrated experiences, let us consider two putative cases of cognitively penetrated experience in a bit more detail. By stipulation, these are genuine cases of cognitive penetration, and so cannot be re-described in any of the ways reviewed earlier (introspective error, influence limited to states downstream of experience, or a selection effect).

Case 1: Angry-looking Jack. Jill believes, without justification, that Jack is angry at her. The epistemically appropriate attitude for Jill to take toward the proposition that Jack is angry at her is suspension of belief. But her attitude is epistemically inappropriate. When she sees Jack, her belief makes him look angry to her. If she didn't believe this, her experience wouldn't represent him as angry.

What exactly are the contents of visual experiences that represent a person as angry? There are several sub-questions here. First, do the contents pertaining to anger attribute the property of anger to a person, or do they attribute the property of expressing anger to a face, or to the geometrical configuration of the face? So that we have a specific proposal to work with, I'm going to assume that they attribute anger to the person. In the end, it won't matter if the only contents in the vicinity concern the expression of the face.

Second, to specify the contents of the penetrated experience, we need a way of representing the angry person, and here there are a number of prima-facie plausible options, including mental analogs of the second-person pronoun, the third-person pronoun, or a demonstrative such as 'that person'; and further options still in the nature of these mental analogs themselves. Since none of these differences matter for our purposes, we can just use the variable X in characterizing the content to stand for any of these options.

With these assumptions in hand, we can label the content of the cognitively penetrated experience:

E1: X is angry.

By hypothesis, Jill believes E1 before she sees Jack. And by hypothesis, E1 is a content of Jill's experience, and wouldn't be a content of the visual experience she has upon seeing Jack, if she didn't believe E1. These two features of the case stem from the status of the experience as cognitively penetrated by the belief.

Let's suppose that the epistemically appropriate attitude for Jill to take to E1, prior to seeing Jack, is suspension of belief. Given all the relevant information about Jack's mental states prior to seeing him, an epistemically flawless person would not believe that she is angry, and Jill should not believe this. She should suspend belief until given more information. However, in the example, Jill's attitude toward E1 (before she sees Jack) is not epistemically appropriate.

Although we're supposing that the contents of the penetrating belief and the penetrated experience are the same (E1), the states of believing E1 and the visual experience with content E1 are quite different mental properties. The visual experience has a phenomenal character, and there are plenty of other contents to the experience as well, such as contents characterizing other ways that Jack's face looks. Let us call the belief with content E1 the anger-belief, and the experience with content E1 the anger-representing experience. (If the simplifying assumption that states as different as belief and experience can share contents is false, then the anger-representing experience won't have E1 among its contents, but will have some other contents such that in having those contents, the anger-representing experience represents that he's angry.)

If the anger-representing experience provides justification for the anger-belief, then that experience can elevate Jill's epistemic standing from one in

which she has an epistemically inappropriate belief, to one in which she has a justified (epistemically appropriate) belief. Prior to seeing Jack, the evidence available to Jill neither justified her in thinking that he was angry nor justified her in thinking that he wasn't angry. Once the anger-representing experience comes into the picture, according to dogmatism it becomes epistemically appropriate to hold the anger-belief, provided there are no defeaters.

The elevation prediction for the anger case is that an experience with content E1 provides justification for believing E1.

The second case has the same structure.

Case 2: Preformationism. Many of the first users of microscopes favored preformationism about mammalian reproduction. Some of them claimed to see embryos in sperm cells that they examined using a microscope.⁹

Prior to looking at sperm cells under the microscope, our (perhaps fictional) preformationist favors the hypothesis that there are embryos in healthy sperm cells. At this time no theory of mammalian reproduction is well-confirmed, and the epistemically appropriate attitude to take toward preformationism is suspension of belief. But our preformationist does not suspend belief. When he looks under the microscope, he has an experience with E2 as its content.

E2: There's an embryo in the sperm cell.

The elevation prediction in the preformationism case is that an experience with content E2 provides justification for believing E2. When combined with the assumption that the particular case of an embryo in the sperm supports the general thesis of preformationism (e.g. by abduction), this elevation prediction results in justification for believing preformationism.

In these cases, the contents of the cognitively penetrated experience are the same as the contents of the penetrating state. In many cases of cognitive penetration, however, the contents will diverge. The preformationist case itself would be like this, if it weren't oversimplified, as it is above. Prior to looking under the microscope, the preformationist isn't in a position to identify any particular sperm cell, and so isn't in a position to believe E2. In other cases, the penetrating state may have more general content than the content of experience. For instance, in depression, the penetrating state is a general mood, whereas the experience will concern specific items.

In his paper defending dogmatism, James Pryor briefly discusses cognitive penetration:

The claim 'observation is theory-laden' might mean that what theory you hold can causally affect what experiences you have. . . For instance, if you believe that the object you're looking at is a . . . carrot, you're likely to experience it as being more orange than you would if you lacked that belief. . . Does this. . . show that

your justification for believing that object is orange cannot be immediate? It does not. I'm concerned with which transitions from experience to belief would result in justified belief. The present claim concerns how one comes to have the experiences, in the first place. These are independent issues.¹⁰

Pryor says cognitive penetration itself doesn't impede immediate justification, because it need not introduce justificatory intermediaries. This seems correct. He also suggests that it doesn't impede justification at all, on the grounds that etiology and justification are independent issues. But the cases just described suggest that the etiology introduced by cognitive penetration does sometimes impede justification, not because it forces the structure of justification to be mediate rather than immediate, but because some kinds of etiology seem to place constraints on when experience can justify beliefs at all – a fortiori, on when experiences can immediately justify them.

4. The challenge for dogmatism

The dogmatist can get off the hook in the problematic cases in two ways. First, if there is a defeater, then no elevation prediction is made. Second, even if there is no defeater, perhaps the elevation prediction is more plausible than I've suggested. Let us consider each of these responses.

4.1 *Is the elevation prediction plausible?*

If cognitive penetration is on par with getting zapped by an outside force, that can make the elevation predictions seem okay. Compare a situation in which a random zap leaves you with a visual experience representing a red cube in front of you. According to the dogmatist, this experience could still be a source of justification for believing that there's a red cube in front of you. It is a case where an accidentally caused experience – a psychological mishap – nonetheless elevates you epistemically.

Perhaps cognitive penetration is just like being zapped into having an experience that *p*, except the zap comes from within one's own cognitive system. If zaps from without can allow epistemic elevation, one might think, then cognitively penetrated experiences can too. If the analogy between cognitive penetration and the zap case holds, then the whole process by which experiences are cognitively penetrated is not under one's rational control.

One way to supplement this idea is to assimilate justification to epistemic blamelessness. Suppose blamelessly formed beliefs are always justified beliefs. Then if you're blameless for having your cognitively penetrated experience in the first place, and blameless for forming a belief on its basis, then if the elevation prediction will be plausible – even if the penetrating state was not itself justified.

Leaving aside the controversial question whether justification should be assimilated to blamelessness,¹¹ it is doubtful that all cases of cognitively

penetrated experience are analogous to the zap case in the relevant way. Arguably some (putative, potential) cognitive penetrators are under your rational control. For instance, we hold people responsible for some personality traits, such as being over- or underconfident. If vanity leads a performer to experience the neutral expression on the face of any audience member as approving, then the relationship between his vanity and his experience is not much like a zap. We can see other un-zaplike instances of cognitive penetration by considering some variants of the preformationism case. Let us say that neutrality-factors are the factors that make suspension of belief the epistemically appropriate attitude to take toward a proposition *p*. Neutrality factors figure in these cases:

Case A. (Confusion) The preformationist is confused about the relevant neutrality-factors, wrongly taking them to support preformationism.

Case B. (Dogma) The preformationist is aware of the neutrality-factors but intent on holding preformationism anyway (e.g., out of faith or dogma).

Just as we hold people responsible for personality traits like over- or underconfidence, so too we often hold people responsible for being confused, for beliefs formed on the basis of confusion, and for known failure to adjust beliefs in accordance with evidence. In cases where confusion, dogma, vanity or underconfidence are penetrators, the zap comparison does not hold, and will not make the elevation prediction more plausible.

A different attempt to vindicate the elevation predictions come from access internalism, which limits the factors that determine how justified a belief is to factors that are accessible to the subject.¹² It is in the spirit of access internalism to endorse the following supervenience claim about which propositions experiences by themselves can justify:

Access Supervenience: the facts about which propositions a subject's experience by itself justifies supervene on factors that are accessible to the subject.

Accessible factors are typically taken to include on the subject's phenomenal states, such as her experiences. For instance, compare two situations involving Jill. In both, she has a visual experience when seeing Jack that represents him as being angry, but in case one this experience is penetrated by her antecedent and unjustified belief that he is angry, whereas in case two it isn't, and her experience puts in her contact with Jack's actual expression of anger. Let us suppose that all other accessible factors in both cases are the same. In particular, in the case where Jill's antecedent belief penetrates her experience, Jill has no access prior to her experience to the fact that she believes that Jack is angry at her, or to its influence on her experience.

If Access Supervenience is true, then Jill's cognitively penetrated experience can provide no less justification for believing that Jack is angry than her

perceptual contact with Jack's anger can provide. Assuming that perceptual contact provides independent justification for believing that Jack is angry, the cognitively penetrated experience must do so as well.

This pair of cases involving Jill is not exactly a poster child for Access Supervenience. The fact that this case involves a feedback loop between Jill's antecedent belief that Jack is angry, her experience, and her resulting confidence that Jack is angry at least puts some pressure on theories that assimilate the justificatory power of the cognitively penetrated experience to cognitively unpenetrated one. There is something ridiculous about a gossip circle, especially one where the perpetrator and recipient are one and the same.

In any case, since this supervenience claim is at least as controversial as access internalism, it is unlikely to provide the kind of support for the elevation prediction that would firmly convince anyone who started out agnostic about the issue.

Perhaps the elevation prediction becomes more plausible, the smaller the elevation is. We can develop this idea by shifting away from the binary notion of justification we have been considering so far, to a version of dogmatism where justification of a proposition p by an experience that p comes in degrees. Consider the view that absent defeaters, all experiences that p by themselves give you justification for believing p , but they only ever give you a little boost of justification, so that the evidential boost you get without relying on any other factors is only ever a small one. For instance, in a normal perceptual case, where intuitively you are justified to degree $N+$ in believing p , the boost you get just from experience by itself only takes you to degree N . In those cases, you end up justified to degree $N+$ thanks to other factors (such as background beliefs, or the status of your experience as the upshot of a suitable process), which combine with the experience to provide you with the difference between $N+$ and N . It is compatible with this view that some cases of cognitive penetration epistemically compromise experience, without compromising the evidential boost provided by experience by itself, since that evidential boost is always small to begin with.

The proposal to minimize the degree of justification provided by experience by itself can be seen a compromise between the position that the elevation prediction is implausible, and the position that it is okay.

If the evidential boost provided by experience alone is the same in all cases, then lessening the boost to accommodate epistemically bad cases of cognitive penetration will weaken the justification provided by experience in straightforward cases, such as when one learns that the mustard jar is in the fridge by seeing it. The main justificatory role of experience will not be played by experience alone, and will not be immediate.¹³ In contrast, if the degree of justification provided by experience alone is supposed to vary depending on background conditions including epistemically bad cognitive penetration, then this supports the point that some kinds of cognitive penetration

compromise the justification for believing *p* that an experience that *p* by itself can provide. Finally, the comparison with gossip circles and other feedback loops suggest that very little if any justification is provided by experience alone, in at least some epistemically bad cases of cognitive penetration.

4.2 Is there a defeater?

When we ask whether there is a defeater for the justification provided by the experiences in the anger case or the preformationism case, we are asking whether there is an undercutting defeater. In general, an undercutting defeater for putative source of justification for a proposition *p* is a factor that removes the putative source of justification for *p*. In contrast, a rebutting defeater for a proposition *p* is a factor supporting the negation of *p*. In our cases, whether there are rebutting defeaters for the propositions E1 or E2 (or for preformationism generally) is not relevant to whether experiences with contents E1 or E2 can provide justification for believing E1 or E2.¹⁴

The distinction between undercutting and rebutting defeaters is cross-cut by the distinction between propositional defeaters, which can be outside the 'ken' of the subject, and evidential defeaters, which are not. It should be granted that some cases of cognitively penetrated experiences will clearly involve evidential defeaters. For instance, there would be an undercutting evidential defeater for your experience that *p*, at least in some cases, if:

- (i) you believed that you wouldn't be having an experience that *p* if you hadn't antecedently believed/hoped/expected/desired that *p*, and
- (ii) you believed that prior to having the experience, *p* was not justified.

Cases with evidential defeaters like this one would pose no challenge to dogmatism as it is standardly formulated, because standardly dogmatism is formulated with evidential defeaters. Version of dogmatism that used propositional defeaters would result in a notion of justification that is broadly externalist, rather than broadly internalist.¹⁵

These sorts of evidential defeaters will not always be present in cases where the elevation prediction is implausible. They don't seem to apply to cases A and B above (involving confusion and dogma). Nor do they apply to a third variant of the preformationism case, where elevation likewise seems implausible:

Case C. (Hope) The wannabe preformationist hopes that preformationism is true. He is not under any illusion about the (lack of) evidence for it. He just hopes that it is true.

The defeater we have been considering (claims (i) and (ii) above) does not apply to these cases for two reasons. First, although you might be aware that your experience depends on your hope, confusion, or dogma, you need

not be. Second, some penetrating states, such as moods (anxiety, depression) and traits (under/overconfidence) cannot easily incorporate this style of defeater, because there may be no relevant proposition *p* to plug into the schema.

In contrast, the defeater we have been considering does seem to be present in cases where the subject uses cognitive penetration to manipulate his own perceptual evidence. Consider case D:

Case D. (Evidence-manipulation) Albert resents Bea's good fortune. He wishes she had some flaw. Through initially willful misinterpretation of her past behavior, he convinces himself that she is an angry person and expects that she will look angry when he sees her. By the time he has convinced himself of this, he half-regards it as an insight. Albert also believes that Bea will look different to him, post-insight. He thinks he'll be able to see in her face what he has 'learned' about her character by reflecting on her past behavior. This pleases him, because he thinks he has put himself in a situation that will improve his epistemic situation: the angry look on Bea's face that he expects to see will confirm his insight when he sees her.

Albert's psychological complexities involve manipulating his own evidential situation. But at some level, through the complexity, he may still meet conditions (i) and (ii), in which case his experience representing Bea as angry is arguably undercut as a source of justification for his belief.

Finally, consider a variant of the anger case involving amnesia:

Case E. (Amnesia) Before seeing Jack, Jill forms the belief that he's mad at her, but she's jumping to conclusions. Maybe she is confused about what the right conclusion is, or maybe she is pathologically intent on believing that Jack is angry. (So far, this is just like cases A and B). But by the next time she sees Jack next, she has forgotten that this is what she thinks. (Cf. the amnesiac preformationist).

In this case, conditions (i) and (ii) are clearly not met, so there is no defeater of that sort. (Though it should be noted that an access internalist might say that the elevation prediction becomes plausible with the onset of amnesia, on the grounds that amnesia removes a potential defeater – namely, the subject's access to the penetrator, and with it her access to its role in producing the experience. This suggestion was addressed at the end of section 4.1).

Other potential evidential defeaters in the problematic cases

According to a first proposal, the circular structure introduced by cognitive penetration undercuts experience that *p* as a source of justification for *p*, so long as you're aware of it or in a position to be aware of it. For instance,

in cases A - D above – i.e., all but the forgetting cases – Jill is in a position to notice that her experience conforms to what she already believes. Perhaps this should give her pause. Maybe the fact that her observation confirms her antecedent belief raises the bar for perceptual justification. We can formulate this proposal as the following principle:

Double-check-1: If you notice or are in a position to notice that you have an experience that *p* when you antecedently believe *p* or favor *p* as a hypothesis, then your experience that *p* by itself does not suffice to justify the belief.

We can compare Double-check-1 to the rationale for triple-blind studies, in which the interpreter of experimental data (e.g., a statistician) does not know which hypothesis the data were collected to test. A rationale for triple-blind studies is that knowledge of which hypothesis was being tested that knowledge might influence the analyst's interpretation of the data.

Double-check-1, however, is not a promising principle for the dogmatist to appeal to. Before the skeptic challenges you, you believe you have hands. But if Double-check-1 were true, then experience could not provide immediate justification for believing that you have hands. Since this scenario is meant to showcase the simplicity of perceptual justification as dogmatism construes it, Double-check-1 would significantly dilute the strength of dogmatism.

Even putting aside skeptical challenges, Double-check-1 makes perceptual justification less frequent than dogmatism advertises it as being. Before entering a classroom for the first time, you may expect that it will have chairs in it. But if Double-check-1 were true, then your experience upon seeing the chairs in the classroom could not provide immediate justification, or justification all by itself, for your belief that the classroom contains chairs. More generally, we nearly always have expectations about what we'll see, and if these expectations weakened the justificatory force of experiences, then experiences would only infrequently provide the kind of justification for belief that dogmatists say is characteristic of them.¹⁶

In response to these difficulties, one might revise the principle to make the need for double-checking less pervasive. Perhaps the need to double-check arises not merely when your experience manifestly conforms to antecedent expectations, but when your experience is manifestly influenced by those expectations. A principle that captures this idea is Double-check-2:

Double-check-2: If you notice, suspect, or are in a position to notice that: you have an experience that *p* when *and because* you antecedently believe *p* or favor *p* as a hypothesis, then your experience that *p* by itself does not suffice to justify the belief.

Double-check-2 entails that if you suspect cognitive penetration by a doxastic state, then your experience will be undercut as a source of immediate justification. As such, it would prevent experiences by themselves from justifying beliefs (mediately or immediately), even in cases where cognitive penetration seems epistemically good or neutral. For instance, suppose you remember how different the trees in the endangered forest looked before you learned that they are Eucalyptus trees. Intuitively, this should not prevent your experience from justifying the belief that the trees are Eucalyptus trees. Or suppose you are a reformed villain, and when you see a thief pick someone's pocket on the subway, you find that whereas before you would have admired the pickpocket's grace, now your most salient reaction is disapproval, and you attribute this shift to your reformation. On the assumption that your visual experience represents that this act of theft is wrong and does so as a result of cognitive penetration by freshly acquired virtue, mere awareness of this sort of transformation from villainy should not prevent your experience from justifying you in believing that the act is wrong.

Putting aside the cases of epistemically good cognitive penetration, Double-check-2 will not provide a defeater in all the cases where elevation is arguably implausible. Double-check-2 is similar to condition (i) in the first evidential defeater we considered, and the application of both defeaters is limited in similar ways. Neither will it apply to cases in which the subject isn't in a position to notice the dependence of their experience on a penetrating state, such as the amnesia case, or cases in which a personality trait of which the subject is unaware penetrates their experience. As we noted earlier, under-confident or overconfident people are often unaware of those traits, and so would remain unaware of the trait's influence on their experience. Finally, since the extent of cognitive penetration is a substantive and open empirical question which requires experimentation to settle, many subjects will not be in a position to notice it when it occurs.

In the cases of cognitively penetrated experiences in which epistemic elevation is implausible, it is hard to see what the undercutting defeater would be. To avoid being stuck with the implausible prediction of elevation, it looks as if dogmatism and other theories analyzing justification in terms of defeaters would need propositional defeaters.

Propositional defeaters might help dogmatism in these cases, as there is no requirement that the defeating factor is within the subject's ken. For instance, if we could pinpoint the epistemically bad kind of cognitive penetration, then a propositional defeater could be formulated to reflect this specific kind of etiology for the experience. But this result would depart from the general spirit of dogmatism, which accords experience itself with justificatory power largely independently of its etiology.

5. How the challenge generalizes

We've been discussing a challenge to a simple version dogmatism posed by cognitive penetration. But the challenge applies much more widely than this.

First, the challenge is not avoided by excising experiences from the epistemic picture altogether. Suppose the very idea that there are experiences is called into doubt, or suppose one rejects the idea that experiences have contents.¹⁷ The challenge could be reframed in terms of uptake from perceptual stimuli instead of visual experiences. When uptake is influenced by problematic cases of cognitive penetration, beliefs that there is an embryo in the sperm or that Jack is angry resulting from perceptual uptake intuitively are not justified.

Second, the challenge applies to coherentism as well as to dogmatism. Consider a set of beliefs (or beliefs plus experiences) that are on the threshold of cohering in a way that would make the belief that *q* justified. With the addition of an experience that *q*, the set will cross the threshold into coherence, and the belief that *q* will be justified. Now suppose the experience that *q* is cognitively penetrated in a way that intuitively should diminish its justificatory force. The coherentist view will then predict epistemic elevation when intuitively there should be none.

Third, I have described the epistemic challenge posed by cognitive penetration in terms of elevation across a threshold to epistemically appropriate belief. Elevation across this threshold is a dramatic way of presenting the challenge posed by cognitive penetrability, but we shouldn't think of the challenge as tied exclusively to this sort of scenario. In some cases cognitive penetration, the penetrating state is justified, but arguably, experience doesn't provide independent justification for the corresponding belief. Here is a potential example:

Angry note: Jack left Jill an angry note, causing her to believe with justification that he was angry at her. The belief penetrated her experience, so when she saw him, her experience represented him as angry. But her experience would represent him as being angry, whether his expression is angry or neutral.

In this case, it is already epistemically appropriate for the Jill to believe that Jack is angry, before she has her cognitively penetrated experience. Does Jill's experience provide additional justification for this belief? Dogmatism would seem to predict that it does, so long as there are no defeaters. If the elevation predictions in our earlier cases are wrong, it would seem that this prediction is wrong too. This case shows that the challenge posed by cognitive penetrability does not exclusively take the form of accusing theories of perceptual justification with falsely predicting epistemic elevation across the threshold to epistemically appropriate belief.

Another way to generate a troublesome case without elevation is to assume epistemic conservatism. According to conservatism, if you already believe *p*, then it is epistemically appropriate to keep believing *p*. When we apply epistemic conservatism to cases in which an experience that *p* is penetrated by a belief that *p*, it looks as if it can't be epistemically inappropriate for the subject to believe *p* by the time their penetrated experience comes along. But we can still ask whether experience provides independent justification for believing *p*. Our discussion suggests that there are cases in which it does not.

Fourth, the challenge from cognitive penetration applies to theories that either reject or modify the idea that experiences that *p* can provide immediate justification for *p*. For instance, Silins (2008) modifies the idea, by suggesting that an experience that *p* can immediately justify a belief that *p*, only if certain background conditions are met. Versions of this theory that do not exclude the epistemically bad kinds of cognitive penetration from these background conditions will be subject to same challenge. In contrast, Wright (2007) rejects the idea that experiences can provide immediate justification, in favor of the idea that for an experience that *p* to justify believing that *p*, ancillary background entitlements are needed.¹⁸ According to Wright, these include entitlements to believe general propositions, such as the proposition that there is an external world. In cases of cognitive penetration, the penetrating states play an etiological role both with respect to experience, and to the beliefs based on experience. If entitlement theory accorded penetrating states an epistemically mediating role, then presumably the elevation prediction could be avoided, because an unjustified (confused, etc) belief can't transmit justification. But if the ancillary entitlements don't include entitlements to believe the contents of the penetrating states, and if those entitlements plus experience are sufficient for justification, then the challenge will still arise.

Finally, the challenge from cognitive penetration applies to versions of dogmatism that use a degreed rather than a binary notion of justification. According to one such version, experience that *p* provides you with an evidential boost for *p*, and whether the boost takes you across a threshold to appropriate belief depends on the degree of justification for believing *p* that you had prior to having the experience. Our challenge will arise in cases where a cognitively penetrated experience takes you across such a threshold. Even if the very idea of such a threshold is rejected, significantly raising the degree of belief seems implausible in our cases of confusion, dogma, hope, under- and overconfidence, and perhaps amnesia. (A more complex version of dogmatism using a degreed notion of belief was discussed in section 4.1.).

For theories of perceptual justification to meet the challenge posed by cognitive penetration, what's needed is a way for such theories to incorporate an etiological constraint informed by a distinction between the epistemically bad kind of cognitive penetration, on the one hand, and the epistemically good or neutral kind, on the other. Once the distinction is in hand, there

seem to be multiple ways to incorporate the constraint. Perceptual justification may still sometimes be immediate, but it seems better either to embrace propositional defeaters, or else to move to a less rigid theory with the structure proposed by Silins (2008), so that experience immediately justifies only under certain conditions, including the condition that it is not cognitively penetrated in an epistemically bad way. The etiological constraint could also take the form of a supplement to a coherence relation, or a refinement of the exact sort of reliable process that is needed to result in justified beliefs.

Our discussion of cognitive penetrability has gestured at the idea that some kinds of cognitive penetration of experience compromise perceptual justification, while other kinds of cognitive penetration do not. To press the epistemic challenge further, we need to know what makes cognitive penetration epistemically bad when it is.

Notes

* Thanks to audiences at MIT, St Andrews, Rutgers, Dubrovnik, Stockholm, OSU, Toronto, Kentucky, SMU, NYU, and Oxford, and to Maja Spener, Farid Masrouf, Jesse Prinz, Doug Lavin, Alvin Goldman, Benj Hellie, Jessica Wilson, Ned Block, Al Galaburda, Tim Schroeder, Alex Byrne, Justin Broackes, Charles Siewert, Scott Sturgeon, Jason Stanley, Karen Jones, Jim Pryor, Geoge Pappas, Paul Pietroski, Selim Berker, Mohan Matthen, John Hawthorne, and Janet Browne for helpful criticisms and suggestions. For extended discussion, special thanks to Justin Fisher, Jonathan Vogel, Ned Hall, Robert Howell, Roger White, Nico Silins, Sean Kelly, Declan Smithies and David Chalmers.

¹ “Color sensitivity and mood disorders: biology or metaphor?”. Barrick et al., *Journal of Affective Disorders* 68 (2002) 67–71.

² “Memory Modulates Color Experience”. Gegenfurtner et al, *Nature Neuroscience*, vol 9, no. 11. 2006. See also R Goldstone, “Effects of Categorization on Color Perception.” *Psychological Science* 5, 298–304.

³ For instance, according to this view, a rash person will not perceive the danger in a situation where a courageous person would. I. Murdoch (1970) “The Idea of Perfection”, in *The Sovereignty of Good*, Schocken Books; and J. McDowell (1979) “Virtue and Reason”, *The Monist* lxii.

⁴ For instance, Eberhardt et al (2004) presents evidence using a dot-probe paradigm that when primed with crime-related words or objects, the attention of white subjects is more readily captured by black faces than by white faces (“Seeing Black: Race, Crime and Visual Processing”, in *Journal of Personality and Social Psychology*, Vol. 87, No. 6, 876–893.)

⁵ In addition to the studies cited in earlier notes, other suggestive studies include Levin, D. T. & Banaji, M. R. (2006), which includes evidence that categorization of a racially ambiguous face as black or white influences how light subjects see it to be (“Distortions in the perceived lightness of faces: The role of race categories.” *Journal of Experimental Psychology: General*, 135, 501–512.). In a study by Payne (2001), people exposed to black faces were more likely to misidentify a tool as a gun under time pressure (“Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, 81, 1–12.) Eberhardt et al (op cit) found that white subjects primed with images of black male faces more readily detect guns in fuzzy images in a perceptual threshold task, compared with subjects primed with faces of white male faces. J. Broackes (2010) discusses a case in which expectations about what color a thing should be influences the color experience of

color-blind subjects (“What Do the Color-Blind See?”, in J. Cohen & M. Matthen, eds. (2010), *Color Ontology and Color Science* MIT Press).

⁶ Huemer, a proponent of dogmatism, writes: “Phenomenal conservatism. . . says that when it seems as if *P* and there is no evidence to the contrary, it is *reasonable to believe P*. . . Phenomenal conservatism is a necessary truth, not a contingent one. There is no possible world in which phenomenal conservatism is false.” *Skepticism and the Veil of Perception*, Ch. 5.3.

⁷ The first claim does not entail the second, since there might be a proposition *q* (or a range of them) that we have justification in believing, no matter what, and upon which any experience that *p* must in turn rely to justify believing *p*. (Wright’s entitlements fall into this category). If so, then absent defeaters, having an experience that *p* by itself suffices for the subject to have justification for *p*, but the justification will not be immediate. Conversely, the second claim does not entail the first. A reliabilist could agree that when a experience that *p* justifies believing *p*, the justification is immediate, while denying that absent defeaters, an experience that *p* always provides justification for believing *p*. (Whether an experience provides justification for believing *p*, they might hold, depends on whether the experience is embedded in the right kind of reliable process).

⁸ Immediate justification, though for us this won’t matter.

⁹ For discussion of this amusing episode in the history of embryology see C. Pinto-Correia (1998), *The Ovary of Eve*. University of Chicago Press.

¹⁰ Pryor (2000), “The Skeptic and The Dogmatist”. *Nous* 34.

¹¹ For an argument that justification should not be assimilated to blamelessness, see Pryor (2001), “Highlights of Recent Epistemology”, *British Journal for the Philosophy of Science*, vol 52.

¹² E. Conee and R. Feldman (2001) “Internalism Defended”. Reprinted in E. Conee and R. Feldman, Eds. *Evidentialism*. Oxford University Press, 2004.

¹³ A minimal evidential boost in the form of *immediate* justification may offer a reply to the skeptical position that experiences are deprived of justificatory force altogether. But it does not illuminate the full epistemic role of perception, and does not by itself vindicate what we’re supposing is an ordinary classification of certain perceptual beliefs to degree *N+*.

¹⁴ This version of the distinction between undercutting and rebutting defeaters is crude but will do for our purposes. For recent discussion of the distinction see M. Bergmann, *Justification Without Awareness*, Oxford University Press, 2006, and for classic discussion see J. Pollock, *Contemporary Theories of Knowledge*. (Towota, NJ: Rowman And Littlefield Publishers). 1st edition, 1986.

¹⁵ Huemer (op cit, Chapter 5.4) makes explicit that his notion of justification is meant to be internalist. See also Pryor (2000), footnote 9.

¹⁶ We could also formulate a principle much like Double-check-1 according to which it is enough if you suspect that your experience conforms to your belief, even if it does not actually so conform. This proposal would replace non-factive suspecting with factive noticing. It would face the same difficulties.

¹⁷ For discussion of whether experiences have contents, see Travis (2004), “The Silence of the Senses” *Mind*; Siegel (2010), “Do Experiences Have Contents?” in B. Nanay, ed. *Perceiving the World*. Oxford University Press, and A. Byrne 2009, “Experience and Content”, *Philosophical Quarterly* 59.

¹⁸ “The Perils of Dogmatism”, in *G. E. Moore: New Essays in Epistemology and Ethics*, Eds. S. Nucatelli and G. Seay. Oxford University Press, 2007, and “Skepticism and Dreaming: Imploding the Demon,” *Mind* 100 (1991), 87–115.