ORIGINAL ARTICLE

Internalism, phenomenal conservatism, and defeat

Christoph Kelp

COGITO Epistemology Research Centre, University of Glasgow

Correspondence
Christoph Kelp, COGITO Epistemology Research Centre, University of Glasgow.
Email: christoph.kelp@glasgow.ac.uk

1 | INTRODUCTION

This paper is about internalism, phenomenal conservatism, and defeat. It has three aims. The first is to develop an argument against internalism to the effect that the correct epistemology of defeat must be externalist (Section 1). The second is to show that cases involving defeat also cause trouble for phenomenal conservatism (Sections 2 and 3). The third is to cast doubt on the idea that phenomenal conservatism might work as a specific thesis about the status of seemings of particular kinds as justifiers. In particular, I will ask whether *perceptual* seemings might still be justifiers and provide some reason for pessimism (Section 4).

2 | INTERNALISM

Internalism in epistemology is the view that epistemic justification strongly supervenes on the internal in that any two possible agents who are exactly alike internally are exactly alike justificationally.[1] Externalism is the view that internalism is false. In this section, I will argue that the phenomenon of defeat means trouble for internalism. In other words, the correct epistemology of defeat is an externalist one.

Very roughly, a defeater for p undermines at least some reason one may have for believing that p.[2] There are at least two different kinds of defeater: rebutting and undercutting (there may be others as well). A rebutting defeater for p does its work by providing one with a reason for not-p. Suppose you believe that your colleague is currently on holiday in Italy because they told you last week that they would leave yesterday and would be away for two weeks. When today you see that they are at work, you have a rebutting defeater for believing that they are currently on holiday in

Philosophical Issues. 2020;30:192–204.

Italy. An undercutting defeater works by depriving one's reason for believing p of its supporting powers. Suppose you see that your colleague's suit is blue. When you are told that the coffee you just drank contains a drug that messes with your colour vision, you have an undercutting defeater for believing that your colleague's suit is blue.

Defeaters can themselves be defeated. A defeater defeater for p undermines a defeater one has for p. In the above colour case, finding out that the person who told you that your coffee contained a drug lied provides you with a defeater for your defeater for believing that your colleague's suit is blue.

One may have several reasons for believing p. While having a defeater for p means that some of one's reasons for believing p are undermined, it does not entail that one does not have any reason for believing that p. In fact, it may still be that one has plenty of reason for believing that p, enough, in fact, to know that p. In the above colour vision case, you may still know that your colleague's suit is blue because they told you so. And since one knows that p only if one has justification for believing that p, having a defeater for p is compatible with having justification for believing that p. In the colour vision case, you not only know that your colleague's suit is blue, you also have justification for believing that it is.

While it is worth noting that having defeaters for p is compatible with knowing that p and having justification for believing that p, there are plenty of cases in which defeaters undermine not only some reason one has for believing that p but also one's knowledge and justification for believing that p. For instance, in the holiday case, the defeater for believing that your friend is currently on holiday in Italy not only undermines some reason you have for believing that they are but also your knowledge and your justification for this belief. Since the cases I will be interested in here all involve defeat of the knowledge and justification undermining variety, I will set cases involving defeat with knowledge/justification aside here. As a result, for the purposes of this paper, it will be safe to assume that the relevant instances of the following principle are all true:

Justification-Defeat. If one has justification that p, then one does not have an undefeated defeater for p.

With these points in play, consider the following case:

> Case A. X has recently ingested a new recreational drug that X was promised to have an incredible high. The drug causes X to have a range of false beliefs, including that they are a doctor at the local hospital and that they have just received notice to perform an emergency surgery in OR1. Whilst preparing for surgery, X is approached by a person whom, thanks to the influence of the drug, X falsely believes to be a notorious drug abuser and disruptive element in the hospital. (In fact, the person is a doctor at the hospital.) The person tells X: "You are not a doctor. Please step away from the patient."

The testimony that X is not a doctor is a defeater for X's belief that they are a doctor. Here is the key observation about Case A: X's drug-induced false beliefs about being a doctor and the testifier's being a notorious drug abuser and disruptive element in the hospital do not defeat this defeater. By Justification-Defeat, X's belief that they are a doctor is not justified.

Consider next:

> Case B. Y is a doctor at the local hospital who knows that they are. Y has just received notice to perform an emergency surgery in OR1. Whilst preparing for surgery, Y is approached by a person whom Y recognises as a notorious drug abuser and disruptive

element at the hospital. The person tells Y: "You are not a doctor. Please step away from the patient."

Again, the testimony that Y is not a doctor is a defeater for Y's belief that they are a doctor. Here is the key observation about Case B: Y's knowledge about being a doctor and the testifier's being a notorious drug abuser and disruptive element in the hospital does defeat this defeater. Compatibly with Justification-Defeat, Y's belief that they are a doctor continues to be justified.

What the two key observations show is that there is an important asymmetry between Cases A and B. In Case B, Y does have justification that they are a doctor despite the testimony, whereas, in Case A, X does not. Unfortunately for internalists, they cannot accommodate this asymmetry. This is because we may assume that X and Y are internal duplicates. They are internally exactly alike (thanks to the effects the drug had on X in Case A). Accordingly, internalism entails that they must be justificationally alike also. Cases of defeat mean trouble for internalism.

One might wonder whether internalists couldn't respond as follows. Since X and Y are internal duplicates, they have the same evidence. In particular, in Case A, X will have ostensible memories that they are a doctor and that the person approaching them is a notorious drug abuser and disruptive element. But if so, X can dismiss the defeater after all. As a result, they continue to justifiably believe that they are a doctor after all.[3]

I do not mean to deny that X can go ahead and dismiss the defeater in the way indicated. But if X's ostensible memories provide X with a genuine defeater defeater, they must *justify* X in dismissing the defeater. The question remains, however, whether this is the case. My intuition is it is not. The fact that a knowledgeable doctor tells X that they are not a doctor provides a defeater that X's drug-induced ostensible memories do not enable X to justifiably dismiss.

If this case doesn't elicit this intuition in you, consider:

Case C. As a result of ingesting a new recreational drug with an incredible high, Z has ostensible memories and a corresponding belief that his Greek lover (who in fact is Z's Czech flatmate) wants to sleep with him. When Z attempts to initiate sexual intercourse with her, he interprets the Czech 'ne' which means 'no' as the phonetically very similar (and to Z indistinguishable) Greek 'nai' which means 'yes'.

There can be no question that, in this case, when Z's flatmate says no, this provides a defeater for Z's belief that she wants to sleep with him. More importantly yet, Z's drug-induced ostensible memories do not enable Z to justifiably dismiss this defeater. They do not constitute a genuine defeater defeater. As a result, Z does not justifiably believe that his flatmate wants to sleep with him.

Just as the Z's drug-induced ostensible memories don't constitute a genuine defeater defeater because they don't allow him to justifiably dismiss the defeater provided by his flatmate's word in Case C, so X's drug-induced ostensible memories don't either because they don't allow X to justifiably dismiss the defeater provided by the doctor's testimony in Case A. Just as Z's belief in Case C is not justified, X's belief that they are a doctor in Case A is not justified either.[4]

But doesn't that serve to point the way forward for the internalist? Since it turns out that X's belief is not justified, should internalists just hold that Y's belief isn't either?

No. If assertions by people whom we know to be deranged can defeat our best justified beliefs, defeat is much too easy to come by, and scepticism looms across a range of important subject matters, including climate change, the reality of systemic racism, and the seriousness of the Covid-19 pandemic, to name but a few. By the same token, holding that Y's belief that he is a doctor isn't

justified in Case B because the word of person whom Y knows to be a notorious drug abuser defeats their belief isn't an option either.[5]

In sum, what these considerations suggest is that, in Case A, X cannot justifiably dismiss the defeater for their belief that they are a doctor. They do not have a genuine defeater defeater and that, as a result, their belief is not justified. At the same time, in Case B, Y can justifiably dismiss the defeater for their belief that they are a doctor. They do have a genuine defeater defeater. As a result, their belief is justified. Since X and Y are internal duplicates, cases of defeat do mean trouble for internalism after all. By the same token, there is reason to believe that the correct epistemology of defeat will have to be an externalist one.

3 | CLASSIC PHENOMENAL CONSERVATISM

According to phenomenal conservatism it is a bedrock epistemic principle that seemings are justifiers, one that features in epistemic explanations of the most fundamental kinds. It is easy to see that phenomenal conservatism is compatible with internalism, as seemings are internalistically respectable in the relevant sense. Phenomenal conservatism offers an attractive solution to the regress problem in epistemology according to which seemings are the regress stoppers and offers the perhaps most popular general theory of basic justification among contemporary internalists.

While phenomenal conservatism is compatible with internalism and indeed typically endorsed by internalists, it is worth noting that one might be a phenomenal conservative without being an internalist. In particular, one might take the above argument to succeed in showing that the correct epistemology of defeat must be externalist and hold on to the key thesis of phenomenal conservatism that seemings are justifiers. In this way, phenomenal conservatism might allow its champions to hold on at least to a limited form of internalism. The aim of this section is to show that cases of defeat also serve to cause trouble for phenomenal conservatism.

To begin with, consider the following dummy version of the view:

Dummy PC. If it seems to one that p, then one has justification for believing that p.

Dummy PC is false. To see why, consider:

Case 1. A is an agent to whom it seems that every woman is less intelligent than any man (= P1), that Theresa May is a Chiquita banana (= P2) and that their only dining table is both round and square (= P3).

These seemings do not give A justification for believing the corresponding propositions. In fact, the case of A is as compelling a counterexample to Dummy PC as they get in philosophy.

It will come as no surprise, then, that phenomenal conservatives do not espouse Dummy PC. Instead they go for the following weaker principle:

PC. If it seems to one that p, then one has *prima facie* justification for believing that p.[6]

According to PC, seemings continue to be justifiers. Crucially, however, they are defeasible justifiers. If one has a seeming that p, one has justification for believing that p unless one's justification for believing that p is defeated. Let's grant phenomenal conservatives that, thanks to the additional clause, PC can get Case 1 right.[7]

Consider, next, the following case:

> Case 2. B is an agent who is exactly like A with one exception: it also seems to B that all defeaters for believing what seems to them to be the case are misleading.[8]

By PC, B has prima facie justification for believing that all defeaters for believing what seems to them to be the case are misleading. But given that this is so, it cannot be that it seems to B that p and B's prima facie justification for p is defeated. This is because the defeating power of any candidate defeater is undercut by the justificatory power of the seeming that all defeaters for believing what seems to B to be the case are misleading. In other words, the fact that it seems to B that all defeaters for believing what seems to them to be the case are misleading means that B has a defeater defeater for any defeater for believing what seems to B to be the case. This gives us the following:

> **PC's Key Commitment**. For B, prima facie justification will always convert to ultima facie justification: if it seems to B that p, then B has justification for believing p.

With PC's Key Commitment in play, let's return to the counterexamples to Dummy PC. Recall that A's seemings that P1, P2 and P3 do not give A justification for believing the corresponding propositions. Recall also that B is just like A except that it also seems to B that any defeater for believing what seems to B to be the case must be misleading. Crucially, there is excellent reason to believe that this additional seeming does not improve B's epistemic state compared to A's; if anything, it makes it worse. If this isn't immediately obvious, take a moment to reflect on whose epistemic state you'd rather be in. I, for one, take it to be a complete no-brainer that the choice is *not* going to be B; if anything, it's going to be A rather than B. But, of course, what one has justification for believing supervenes on one's epistemic state. Given (i) that B is just like A except for one additional seeming B has; (ii) that this additional seeming does not improve B's epistemic state compared to A's; and (iii) that what one has justification for believing supervenes on one's epistemic state, it follows that whatever B has justification for believing, A does, too. Since A doesn't have justification for believing P1, P2 and P3, neither does B. And since that's incompatible with PC's Key Commitment, PC is in trouble.[9]

Before moving on, I'd like to consider a few responses on behalf of champions of PC. Here is the first. If B's seeming that all defeaters for believing what seems to B to be the case are misleading is a defeater defeater, then it is a defeater itself. If so, isn't this defeater defeated by the rebutter *this is a defeater and it's not misleading*?[10]

No. This is because, while B's seeming is indeed a defeater, its content doesn't hold that all defeaters are misleading but only that all defeaters for believing *what seems to B to be the case* are. What's more, since the seeming is not a defeater for believing anything that seems to B to be the case, it will not be rebutted in the way envisaged.

Second, might champions of PC not resist the existence of 'supercharged' defeaters such as the one that B has in Case 2. Suppose you acquire evidence that you are in a state of global hallucination but then everything remains coherent and as expected for a while. In that case, can't that be evidence that you are not in a state of global hallucination after all?[11]

I agree that this might be the case. After all, you might have reasonable background knowledge about how long hallucinations typically last and how they unfold that might constitute evidence that you are not hallucinating after everything remains coherent and as expected for a while. At the same time, there is excellent reason to think that there are 'supercharged' defeaters. Consider

a case in which you are abducted by evil scientists who tell you that they will take your brain out of your skull, hook it up to a supercomputer and stimulate it such that you have deceptive experiences as of being in Cockaigne. When you wake up you do indeed have experiences as of being in Cockaige, with experiences as of roasted pigs walking around with knives in their backs to facilitate carving, grilled geese flying directly into your mouth, and so on. Suppose you remain in this predicament for the rest of your life, being painfully aware of this fact. You clearly do not arrive at a point at which you have evidence that you are not being deceived, simply because everything remains coherent and as expected for a while. On the contrary, it is clear that in this case you have a 'supercharged' defeater that doesn't go away. Calling the possibility of such defeaters into doubt is thus not something that champions of PC can reasonably hope to do in order to rescue their view.

Third, couldn't phenomenal conservatives simply embrace PC's Key Commitment and the consequence that B does have justification for the relevant propositions. Might they even add that, from where they stand, this consequence is not all that problematic?

Of course, they can. Any champion of any view whatsoever can simply accept whatever unpalatable consequences their view entails and claim that they are not all that problematic, at least from where they stand. This does not mean that their view isn't refuted by the unpalatable consequences, nor that we don't have good reason to take it to be so refuted. If this isn't immediately obvious, consider a champion of the view that there are true contradictions and that explosion holds (TC+E). Suppose that, when presented with the objection that TC+E has many unpalatable consequences, including that any proposition whatsoever is true, the champion of TC+E responds by telling us that they simply embrace these consequences. Suppose they were to add that, from where they stand, this consequence is not all that problematic. After all, it's just what one you'd expect if their view were true.

I expect sympathisers of PC to object that the case of PC is very different from TC+E in a way that renders PC much less problematic. Unfortunately, it's not even clear that this claim stands up to closer scrutiny. To see why not, consider:

> Case 3. C is an agent to whom it seems that (i) all defeaters for believing what seems to them to be the case are misleading, (ii) that there are true contradictions, (iii) that explosion holds, and (iv) that justification is closed under entailment.

While TC+E entails that any proposition whatsoever is true, PC entails that C is in a position to acquire justification for believing any proposition whatsoever. I take it to be clear that both consequences are equally unpalatable. Now, champions of PC and TC+E alike may tell us that they embrace the relevant consequence of their view and continue to insist that it is not really that problematic from where they stand. All that means, however, is that we won't be able to argue them out of their views. It does not mean that their view isn't in serious trouble. On the contrary, it clearly is.

4 | RESTRICTED PHENOMENAL CONSERVATISM

Even if PC can't be right, couldn't phenomenal conservatives hold out hope that some restricted version of PC might still work? More specifically, phenomenal conservatives might try to identify a further property, call it J, that seemings must have in order to confer prima facie justification and that the problematic kind of seeming B has in Case 2 simply doesn't have. Now, if it turned out

that there is such a property J, champions of PC would have reason to replace PC by the following general principle:

PC+J. If it seems to one that p and one's seeming that p has property J, then one has *prima facie* justification for believing that p.[12]

Recall that I considered phenomenal conservatism as a way for internalists to hold on at least to a limited form of internalism. It is versions of phenomenal conservatism that internalists might retreat to in the light of the arguments from Section 1 that I am interested in here. What this means for present purposes is that I will focus on internalist versions of PC+J. What we are looking for, then, is an internalistically respectable property, J, such that (i) seemings that have J confer prima facie justification and (ii) the problematic kind of seeming B has in Case 2 doesn't have J.

Unfortunately, there is excellent reason for thinking that there is no J that fits the bill. To see this, I will start with a slight detour. Its relevance will become clear in due course.

Consider the following case:

Case 4. As a side-effect of live-saving medication, D has suffered from total amnesia. When D wakes up, they see the sun shining through their window, the bed they are in, etc.

While D has lost their memories, upon seeing the sun shining through the window, D knows that the sun is shining. And upon seeing themselves in a bed, D knows that they are in a bed. If so, they must also have justification for believing what they know. All of this is, of course, entirely as phenomenal conservatism would have it. After all, A does have the relevant perceptual seemings.

Consider next:

Case 5. In their sleep, E is abducted by an evil scientist who takes E's brain out of their skull, places it in a vat with nutrients and hooks it up to a supercomputer. The evil scientist erases E's memory of the past and now stimulates their brain to make it seem to E that they wake up in the morning, that the sun is shining through their window, that they are in a bed, etc.

Suppose, as we may, that D and E are internal duplicates. Just like D, E has justification for believing that the sun is shining, that they are in a bed, etc. In fact, one of the central motivations for internalist views is that they can accommodate this data point.[13] And the reason why PC can do so is, of course, that D and E have exactly the same seemings. Now, I take it that a condition of adequacy on any internalistically respectable J that PC+J accommodate these intuitions also. If so, D and E's seemings must both have property J.

With this point in play, consider:

Case 6. F is an agent who is exactly like E with one exception: The evil scientist also makes it seem to F that *not* all defeaters for believing what seems to them to be the case are misleading.

Now, F is just like D and E, except that F has one additional seeming. Note that this seeming is generated in the same way as the other seemings F hosts, i.e. by means of an intervention of the evil scientist. And there is nothing in F's perspective that could provide them with reason to think that

the seeming shouldn't be taken at face value. After all, F's other seemings don't speak to this one and, thanks to F's memories having been erased, their background beliefs are too impoverished to have any bearing on the issue. By the same token, there is reason to think that F's seemings have J also.

Finally, consider:

> Case 7. G is an agent who is exactly like E with one exception: The evil scientist also makes it seem to G that all defeaters for believing what seems to them to be the case are misleading.

The trouble for champions of PC+J is that if F's seemings have property J, then so will G's. After all, just as in the case of F, there is nothing in G's perspective that could provide them with reason to think that the seeming that all defeaters for believing what seems to them to be the case are misleading shouldn't be taken at face value. Again, G's other seemings don't speak to this one and, thanks to G's memories having been erased, their background beliefs are too impoverished to have any bearing on the issue. If so, there is reason to think that G's seemings have property J, just like F's.

But, of course, if G's seemings have property J, then G's seeming that all defeaters for believing what seems to them to be the case are misleading has property J. This means that any attempt to identify an internalistically respectable property, J, such that (i) seemings that have J confer prima facie justification and (ii) the problematic kind of seeming B has in Case 2 doesn't have J is bound to remain unsuccessful. In fact, we can now see that PC+J incurs the following key commitment:

> **PC+J's Key Commitment**. For G, prima facie justification will always convert to ultima facie justification: if it seems to G that p, then G has justification for believing p.

It goes without saying that this commitment just as problematic as PC's Key Commitment in Section 2. PC+J leaves phenomenal conservatives no better off than PC.

Again, before moving on, I will consider a couple of responses on behalf of champions of PC+J. First, couldn't champions of PC+J insist that seemings only give prima facie justification relative to a sufficiently rich stock of background beliefs? Since in the above cases, the agent's background beliefs are impoverished by memory loss, their seemings don't give them even prima facie justification.

Two points by way of response. First, this kind of move threatens to lead to the wrong verdict in ordinary cases of knowledge such as Case 4. Recall that, despite memory loss, D comes to know for instance that the sun is shining upon seeing it. If memory loss prevents seemings from conferring prima facie justification, then D's seemings don't give them justification that the sun is shining. If so, D doesn't know that it is shining either. Second, we can easily amend Case 7 such that we have our evil scientist implant a whole slew of background beliefs into G that are all highly coherent with each other and with G's seeming. Perhaps our evil scientist has G believe a rich story centred on the idea that they are an oracle, gifted by God with a unique capacity to host infallible seemings.[14] In that case, G's seeming that all defeaters for believing what seems to them to be the case are misleading will confer justification, even if we agree that seemings only do so given sufficiently rich stock of background beliefs. At the same time, none of this serves to reduce the extent to which PC+J's Key Commitment is problematic.

But couldn't internalist champions of PC+J insist that the above cases are at least highly abnormal and so don't carry much evidential weight?[15]

To answer this question, it may be worth reflecting on exactly how the kind of response we are considering might enable champions of PC+J to rescue their view. To this end, let's begin by asking how abnormality may undermine evidential weight in cases of radical deception. There are three ways in which this might happen: abnormality may get in the way of correctly gauging (i) whether the cases are really possible; (ii) how the target view applies to the cases; or (iii) whether the target view's prediction about the cases are really mistaken.

Now, radical deception cases are widely agreed to be possible. Moreover, as I already mentioned, radical deception cases are among the central motivations for internalist accounts of justification. As a result, we may expect internalist champions of PC+J, whom I am engaging here, to grant the possibility of these cases. Once we allow that cases involving evil scientists radically deceiving agents with the help of supercomputers are possible, there is every reason to think that all of the above cases involving radical deception are possible also. After all, the only differences between them concern the details of how the evil scientists deceives their victims. This means that (i) is not the problem.

Moreover, there is not really any issue with how PC+J applies in these cases. It is clear that, in Case 6, PC+J predicts that G has justification for believing the propositions that seem to G to be the case. As a result, (ii) isn't the issue either.

This leaves (iii), i.e. that abnormality gets in the way of correctly gauging whether PC+J's predictions about the cases are mistaken. The thought is that our judgement is less trustworthy in radical deception cases because they are so abnormal. That's how radical deception cases might not carry the evidential weight needed to cause serious trouble for the target view.

With this point in play, I'd like to focus on what is required to develop this point into a successful defence of PC+J. Champions of PC+J will have to hold that the problematic predictions of their view are actually true. After all, what PC+J has to accommodate are the epistemic facts on the ground. Merely establishing that a certain judgement isn't as trustworthy as we would like it to be won't rescue the view if the judgement is nonetheless true. Holding that the problematic predictions are true is what the result that certain judgements are less trustworthy prepares the ground for. In other words, what the kind of attack on the evidential weight of abnormal cases we are considering here is meant to do is to make 'biting the bullet' palatable.

The trouble is that there is excellent reason to think that biting the bullet is not an option of champions of PC+J. To see why not, consider:

> Case 8. H is an agent who is exactly like E except: The evil scientist also makes it seem to H that (i) all defeaters for believing what seems to them to be the case are misleading, (ii) that there are true contradictions, (iii) that explosion holds, and (iv) that justification is closed under entailment.

As we have already seen, PC+J entails that agents like H to whom it seems that (i) – (iv) hold are in a position to acquire justification for believing any proposition whatsoever.

To develop the point about evidential weight into a defence of their view, champions of PC+J will have to hold that this prediction is actually true, i.e. that H is indeed in a position to acquire justification for believing any proposition whatsoever. The point about evidential weight of abnormal cases is meant to make this commitment palatable, here by calling into doubt the trustworthiness of our judgement to the contrary.

Unfortunately, there is excellent reason for thinking that this strategy carries little promise here, for two reasons. First, our judgement that H isn't in a position to acquire justification for believing any proposition whatsoever is not grounded in reasoning about Case 8. Rather, it is a direct consequence of the widely recognised fact that there is no human agent such that they are in a position to acquire justification for believing any proposition whatsoever. But since our judgement doesn't require reasoning about the case, there is simply no reason to think that it may have been led astray by its abnormality. Second, even if we set this aside, the fact remains that this result is simply not palatable (for reasons set out in Section 2). Some bullets are just too big to bite. This is one of them.

5 | PHENOMENAL CONSERVATISM ABOUT PERCEPTUAL JUSTIFICATION

But couldn't phenomenal conservatives at least retreat to a version of the view according to which seemings of particular kinds are justifiers? While we may of course consider a wide variety of kinds of seemings, in view of the fact that this issue is about perceptual evidence, I will here focus on the question of whether phenomenal conservatives might not be able to retreat to the claim that perceptual seemings are justifiers. In other words, I want to ask whether, even if PC and PC+J are abandoned, the following principle might not be viable instead:

> **Perceptual PC**. If it *perceptually* seems to one that p, then one has *prima facie* justification for believing that p.

Now, it is hard to deny that the kind of seeming that causes the trouble for PC—the seeming that all defeaters for believing what seems to them to be the case are misleading—is not a perceptual seeming. In fact, it is hard to see how this seeming could be a perceptual seeming. As a result, the above problem simply won't arise if we retreat from PC to Perceptual PC.

I agree that this retreat does avoid the problem. However, it does raise an important challenge for champions of Perceptual PC. They will have to answer the question what distinguishes perceptual seemings, which do provide justification, from the kinds that don't. (To keep things simple, I will hereafter refer to kinds of seemings that don't provide justification as X-seemings.) Of course, the answer cannot appeal only to properties that perceptual seemings have in virtue of being seemings in general. After all, any such property would be shared by X-seemings and so would not serve to explain the difference we are now looking for. Moreover, the answer cannot appeal to a general property (such as J) that might be thought to further restrict the conditions under which seemings confer prima facie justification, for reasons set out in Section 3. Instead, phenomenal conservatives will have to appeal to properties that perceptual seemings have in virtue of being *perceptual* in particular.

While I don't have any conclusive reason for thinking that this cannot be done, here is some reason for pessimism. Even if we agree that perceptual seemings are justifiers and that X-seemings aren't, the question as to whether there are kinds of seemings other than perceptual ones that are justifiers should remain open. Suppose, then, that there are other kinds of seemings that are justifiers. For instance, let's suppose that memorial seemings are justifiers as well. In that case, it would have to be properties that memorial seemings have in virtue of being memorial in particular in virtue of which memorial seemings are justifiers. But now notice that it would be a metaphysical coincidence of fantastic proportions that perceptual and memorial seemings (as well as perhaps

other kinds) have the right kinds of properties to make them justifiers, while X-seemings don't. The reason I am pessimistic about the retreat to Perceptual PC is that it would commit phenomenal conservatives to just this metaphysical coincidence.

6 | CONCLUSION

This paper has argued that the correct epistemology of defeat must be externalist and so internalism must be false. What's more, cases of defeat also mean trouble for phenomenal conservatism, which is the perhaps most popular general view of basic justification amongst contemporary internalists and one that internalists might still have hoped to hold on to. Finally, there is even reason to be pessimistic about retreating further by restricting phenomenal conservatism to specific kinds of seeming, notably perceptual seemings. As a result, today's news for internalists and phenomenal conservatives is bad.[16]

ENDNOTES

[1] Classic defences of internalism include BonJour (1985), Chisholm (1977), and Conee and Feldman (2004). For more on the internalism/externalism divide, see Kornblith (2001).

[2] For more on defeat see e.g. Bergman (2006), Lackey (2008) and Pollock (1984).

[3] Thanks to Juan Comesaña and Matt McGrath for bringing this to my attention.

[4] Suppose that X disregards the defeater and operates on the patient. It is entirely appropriate to admonish X afterwards as follows: "What happened there? The doctor told you that you are not a doctor and to step away from the patient." Moreover, it makes sense for X to respond with an apology and an excuse, rather than stand their ground. "I'm sorry. I had taken this new drug and was under the impression that I was a doctor and that the doctor who told me this was a notorious drug abuser." is an appropriate response. "What are you talking about? I had excellent reason to think that I am a doctor and that the person who told me otherwise was a notorious drug abuser." isn't. This provides further evidence that X's belief that they are a doctor isn't justified.

[5] Suppose that Y takes the defeater at face value and doesn't operate. It is entirely appropriate to admonish Y afterwards as follows: "What happened there? You knew that you are a doctor and that the person who told you otherwise is a notorious drug abuser." Again, it makes sense for X to respond with an apology and an excuse, rather than stand their ground. "I'm sorry. I don't know what came over me." is an appropriate response. "What are you talking about? I had excellent reason to think that I am not a doctor." isn't. This provides further evidence that Y's belief that they are a doctor is justified.

[6] Champions of PC include Huemer (2001, 2007), McGrath (2013), Skene (2013) and Tucker (2010). For useful overviews see Huemer (2014) and Moretti (2015). For further discussion of PC see e.g. Tucker (2013). Note that PC also often features an additional *pro tanto* proviso, i.e. that seemings give *some* justification for believing. Since this is of little consequence for present purposes, I will set it aside here.

[7] This is not to say that there aren't any further problems. See e.g. Markie (2006), Siegel (2013) for the perhaps most important problem stemming from cases of cognitive penetration.

[8] Note that what seems to one to be the case and what evidence one has is largely a contingent matter. This is why we may, in addition, assume that B has no seemings inconsistent with and no other evidence bearing on the proposition that all defeaters to what seems to B to be the case are misleading. Note also that BonJour's (1980) famous clairvoyant case makes similar assumptions about its protagonist's evidence. But even if you think that there are important differences, it is easy enough to fill in the details of the case to make the above plausible. I will give one way of doing so in Section 3 below.

We may also assume that B's seeming is as high quality as seemings get (which is why the pro tanto proviso does not matter for present purposes). As a result, we may even assume that B justifiably believes that that all defeaters for believing what seems to them to be the case are misleading because they form this belief based on this seeming. Now, I am not convinced that these additional assumptions are needed to get the argument against PC off the ground. However, the fact that they can be made may help to forestall some looming worries.

⁹ Suppose someone tells you that all defeaters for believing what they tell you are misleading. Given that testimony is a source of justification in the sense that you have prima facie justification for what you are told, you now have a defeater defeater for every defeater for believing what this person tells you. If they go on to tell you that P1 – P3, you are justified in believing P1 – P3. However, that's implausible. By the same token, the above argument against PC proves too much. It generates a problem for everyone, not just for champions of PC.

In response, note that any view of testimony that does entail that you are justified in believing P1 – P3 does face a serious problem. Note also that not every view of testimony delivers this result. Strikingly, more robustly externalist views such as e.g. Goldberg (2010), Greco (2020) and knowledge first views such as e.g. Sutton (2007) don't. In this way, there is reason to believe that the argument doesn't generate a problem for everyone. By the same token, rather than exculpating PC, the argument inculpates views of testimony according to which you come out as justified in believing P1 – P3.

A further question is whether there are any internalist views of testimonial justification that can avoid this problem. While this is a fascinating question, I will not try to settle it either way here. Fortunately, I don't have to. After all, for present purposes, all that needs to be shown is that the problem isn't a problem for everyone. The fact that there are externalist accounts of testimony that avoid the problem establishes as much.

¹⁰ Thanks to Matt McGrath for raising this issue.

¹¹ Again, thanks to Matt McGrath for this.

¹² McGrath (2013) and Siegel (2013) can be viewed as pursuing this kind of strategy.

¹³ This is the new evil demon problem. For the original statement as a problem for process reliabilism see Lehrer and Cohen (1983) and Cohen (1984). For an argument that the problem arises for externalist accounts of justification more generally see Wedgwood (2002).

¹⁴ It may be worth bearing in mind that we are still considering internalist versions of PC+J. This means that E's beliefs can only be constrained by properties that are internalistically respectable – the fact that these beliefs are all false or unreliably formed is irrelevant. Of course, champions of PC+J might insist that these beliefs must be justified. But, of course, we are free to describe the case such that the evil scientist furnishes E not only with the beliefs but also with the corresponding memorial seemings that justify them.

¹⁵ Thanks to Matt McGrath for pressing me on this.

¹⁶ Thanks to Juan Comesaña, Matt McGrath, Luca Moretti and Mona Simion for helpful comments on this paper.

REFERENCES

Bergman, M. (2006). *Justification without awareness*. Oxford: Oxford University Press.

BonJour, L. (1980). Externalist theories of empirical knowledge. *Midwest Studies in Philosophy*, *5*, 53–73.

BonJour, L. (1985). *The structure of empirical justification*. Cambridge/MA: Harvard University Press.

Chisholm, R. (1977). *Theory of knowledge* (2nd ed.). Englewood Cliffs: Prentice Hall.

Cohen, S. (1984). Justification and truth. *Philosophical Studies*, *46*, 279–296.

Conee, E., & Feldman, R. (2004). *Evidentialism: Essays in epistemology*. Oxford: Oxford University Press.

Goldberg. S. (2010). *Relying on others: An essay in epistemology*. Oxford: Oxford University Press.

Greco, J. (2020). *The Transmission of knowledge*. Cambridge: Cambridge University Press.

Huemer, M. (2001). *Skepticism and the veil of perception*. Lanham: Rowman and Littlefield.

Huemer, M. (2007). Compassionate phenomenal conservatism. *Philosophy and Phenomenological Research*, *74*, 30–55.

Huemer, M. (2014). *Phenomenal conservatism*. In J. Fieser & B. Dowden (Eds.) Internet Encyclopedia of Philosophy.

Kornblith, H. (Ed.). (2001). *Epistemology: Internalism and Externalism*. Malden/MA: Blackwell.

Lackey, J. (2008) *Learning from words. Testimony as a source of knowledge*. Oxford: Oxford University Press.

Lehrer, K., & Cohen, S. (1983). Justification, truth, and knowledge. *Synthese*, *55*, 191–207.

McGrath, M. (2013). Phenomenal conservatism and cognitive penetration: The 'bad basis' counterexamples. In C. Tucker, (Ed.) *Seemings and justification: New essays on dogmatism and phenomenal conservatism*. Oxford: Oxford University Press.

Markie, P. (2006). *Epistemically appropriate perceptual belief. Noûs*, *40*, 118–142.

Moretti, L. (2015). Phenomenal conservatism. *Analysis*, *75*, 296–309.

Pollock, J. (1986). *Contemporary theories of knowledge*. Lanham: Rowman and Littlefield.

Siegel, S. (2013). The epistemic impact of the etiology of experience. *Philosophical Studies*, *162*, 697–722.

Skene, M. (2013). Seemings and the possibility of epistemic justification. *Philosophical Studies*, *163*, 539–559.

Sutton, J. (2007). *Without justification*. Cambridge/MA: MIT Press.

Tucker, C. (2010). Why open-minded people should endorse dogmatism. *Philosophical Perspectives*, *24*, 529–545.

Tucker, C. (Ed.). (2013). *Seemings and justification: New essays on dogmatism and phenomenal conservatism*. New York: Oxford University Press.

Wedgwood, R. (2002). Internalism explained. *Philosophy and Phenomenological Research*, *65*, 349–369.