THIRD-FACTOR EXPLANATIONS IN EPISTEMOLOGICAL EXPLANATIONISM

Christopher NOONAN

ABSTRACT: According to explanationism about epistemic defeat, our attitude towards the explanation of our belief in P can sometimes defeat our justification for holding that belief. In this paper I argue for the superiority of a particular version of explanationism which is considered and rejected by Korman and Locke (2023). According to this position our belief in P is defeated if we are not entitled to believe it is either (i) explained by P (i.e e-connected), or (ii) explained by some third-factor that also explains P (i.e tf-connected). Korman and Locke call this position "EThirdFactor". I first argue that our belief in P can be justified if we are entitled to believe it is tf-connected but not e-connected, and argue against a recent proposal from Bogardus and Perrin that threatens this view (2019). I then outline some further rationale behind EThirdfactor, and my reasons for preferring it over alternative explanationist accounts that have been suggested in the literature. Finally, I defend this position against recent objections from Korman and Locke (2023) and others. I conclude that these objections do not threaten my claim that EThirdFactor is the most plausible version of explanationism on the market.

KEYWORDS: explanationism, epistemic, defeat, explanatory, constraint, Third-Factor

1. Introduction

Sometimes, we lose justification for holding a belief[1]. When this happens, epistemologists say that our belief has been 'defeated', and there are different views about when this process occurs. According to 'explanationism' about epistemic defeat, our beliefs can get defeated when we receive information about the explanation for why we hold that belief[2]. Explanationism is an increasingly popular position and there have been a number of recent attempts to defend different versions of it (Locke 2014; Lutz 2018; Korman 2019; Korman and Locke 2020, 2023; Barker 2023; Noonan 2023a). My aim in this paper is to defend one such version.

[1] Thanks to Fabienne Peter and Kirk Surgener for helping me to develop some of the ideas presented in this paper. Thanks also to Matthew Heeney and Jim Hutchinson.

[2] More precisely, it is our own attitude (or that attitude we would be entitled to have) towards the explanation of our belief that matters according to most forms of explanationism. More on this below.

Let us say that when our belief in P is explained by P our belief is 'e-connected'[3]. By contrast, if there is a third-factor that explains both our belief in P and P, then our belief is 'tf-connected'. My aim in this paper is to defend the following explanatory constraint on belief:

> EThirdFactor: If S is not entitled to believe that either (i) her belief that P is e-connected or (ii) her belief that P is tf-connected, then S's belief that P is defeated.[4]

In section 2 I clarify my understanding of e-connections and tf-connections. I then argue that our belief can remain justified even if we are not entitled to think it is e-connected, as long as we are entitled to think that it is tf-connected, which gives us reason to accept EThirdFactor. In section 3 I consider a recent proposal from Bogardus and Perrin that threatens this view (2019), and argue that their proposal fails. This discussion implies that we should reject Bogardus and Perrin's proposed analysis of knowledge (2019, *forthcoming*), as well as an explanatory constraint recently defended by Jonathan Barker (2023). In section 4 I outline some further rationale behind EThirdFactor and my reasons for preferring it over alternative constraints that have been suggested in the literature. Finally, in section 5, I respond to some objections to this constraint that have been raised by Korman and Locke (2023) and others. I argue that these objections will apply, in some form, to *any* adequate version of explanationism. As such, they fail to demonstrate any particular problem with EThirdFactor, and I therefore conclude that EThirdFactor remains the most plausible version of explanationism on the market.

But first, a caveat: in this paper I will be largely unconcerned with defending explanationism over alternative, non-explanationist views of defeat. This has been done at length elsewhere, and I advise the reader who is sceptical of the explanationist approach to take a look at this literature[5]. Here I am instead concerned with the following question: assuming explanationism is correct, which version of it should we accept? I think we should accept EThirdFactor. As will become clear, there is plenty to of work to be done when defending even this conditional position. In order to do this adequately I have therefore had to forgo a defence of explanationism more generally.

---

[3] The term e-connected is borrowed from Korman and Locke (2020), as is the term "explanatory constraint", used below.

[4] This constraint is taken from Korman and Locke (2023), who reject it for reasons to be discussed.

[5] See Lutz (2018), Korman (2019), Korman and Locke (2020), Bhogal (2023) and Noonan (2023b) for defences of explanationism over competing accounts of defeat.

2. E-connections, Indirect E-connections and TF-connections

Nothing in this paper depends on any particular view of the nature of explanation. I will be working with a minimal and intuitive understanding of the explanatory relation according to which, if X explains Y, then Y holds *because of* X. I take this to be consistent with any adequate account of explanation.

How does explanation relate to epistemic defeat? In some cases it seems that if we are given reason to think that our belief in P is not explained by P, then our belief in P becomes unjustified. For example, suppose I have a visual experience of a fire in my fireplace and therefore form the belief that my fireplace is lit. Later I find out that I was actually hallucinating when I 'saw' the fire. In this case I can no longer continue to justifiably hold my belief in the fire. Presumably, this is because I can no longer take my belief to be explained by there actually being a fire in front of me; I believe that my fireplace is lit because of my hallucination, and not because my fireplace is really lit.[6]

Cases such as these might motivate the following explanatory constraint on belief:

> (EC): If S is not entitled to believe that her belief that P is explained by the fact that P, then S's belief in P is unjustified[7].

In order to be plausible, such a constraint needs to be consistent with having justification for beliefs formed via inference. Say we see smoke emerging from our chimney and infer that our fireplace is lit. In order for (EC) to be consistent with this belief being justified, we must be entitled to believe that our belief in the fireplace being lit is explained by the fireplace being lit.

And, under usual circumstances, we will be. We will be entitled to believe that the smoke emerging from the chimney is explained by our fireplace being lit; that's exactly why we infer a belief in the fire on the basis of our belief in the smoke. We are also, under usual circumstances, entitled to think our belief in the fire is explained by the smoke; we only believe in the fire because we saw the smoke. And, in at least many cases, explanation is transitive such that if A explains B and B explains C, then A explains C[8]. So if we are justified in believing that the fire explains

---

[6] An alternative analysis of this situation appeals to modal, rather than explanatory, relations (Clarke-Doane 2020, 97-119). See Noonan (2023b) for some reasons for accepting explanationism over modalism.

[7] This is considered by Korman and Locke (2023). Korman and Locke consider it, as I do, on the road to presenting something more plausible.

[8] Some deny that explanation is always transitive (Schaffer 2012; Paul and Hall 2013). But transitivity is only supposed to break down in cases that exhibit specific structural features. This case (just like all the other cases discussed in this paper) exhibit no such features. As such, there is

the smoke, and that the smoke explains our belief in the fire, we'll be justified in thinking the fire ultimately explains our belief in the fire (Goldman 1967, 360-2, 1984, 103; Jenkins 2006, 159; McCain 2014, 164-5). Thus, (EC) is consistent with this kind of inferential justification.

We have stipulated that when our belief in P is explained by P, our belief is 'e-connected'. As we have just seen, (EC) can accommodate cases of inferential justification where we form a belief in P on the basis of our belief in some distinct fact Q that we are entitled to think is explained by P. Let us say that, in cases like these, we are entitled to think our belief is 'indirectly e-connected'. Indirectly e-connected beliefs are a subset of e-connected beliefs; they are cases where our belief is explained by P (and so the belief is e-connected), but P explains our belief via some other fact, Q, from which we infer belief in P.

Though (EC) can accommodate this kind of inferential belief, it falters when it comes to a different kind of inference. These are cases where we infer our belief in P from our belief in some distinct fact Q, and we are entitled to believe that Q explains P. Unlike cases where we take our belief to be indirectly e-connected, which involve inferring P from our belief in some fact Q which we take to be *explained by* P, this other kind of inference involves believing P on the basis of our belief in some distinct fact Q which we take to *explain* P. In cases like these we can be justified in believing P without being entitled to think that P explains our belief. These cases therefore constitute counterexamples to (EC).

Here is an example of such a case, adapted from Goldman (1967, 366)[9]. Brigid sees that her fireplace is lit and forms the belief that there is smoke emerging from her chimney. She forms this belief on the basis of her belief that her fireplace is lit in addition to her belief that, if her fireplace is lit, then smoke will be emerging from her chimney.

Brigid is entitled to believe that there is some third-factor (the fire) that explains both her belief that there is smoke emerging from her chimney and the fact that there is smoke emerging from the chimney. In the parlance I introduced above, Brigid is entitled to believe that her belief is tf-connected. But it seems she has no reason to think that the smoke emerging from the chimney explains her belief in the smoke (and so no reason to think that this belief is e-connected). Brigid does not see the smoke or have any other kind of 'direct' access to it. The smoke therefore does not explain her belief by, say, her seeing it and forming a belief in the smoke. Furthermore, the smoke does not explain the fire; according to any acceptable understanding of explanation, this gets things backwards. So it is *not* as if the smoke

no reason to think that transitivity fails for the cases discussed in this paper.

[9] I also use this example to illustrate tf-connections in my (2023a).

explains the fire and the fire explains her belief in the smoke, which would imply that her belief was indirectly e-connected. Thus, it is not in virtue of its explanatory relationship with the fire that the smoke explains Brigid's belief.

Might the smoke explain Brigid's belief in some other way? Maybe the smell of the smoke emerging from the chimney is what drew Brigid into the room and explained her looking at the fire in the first place. This would mean that, in addition to the fire, the smoke also explains her belief in the smoke. But it seems equally possible that no such additional connection holds. As such, it seems perfectly possible to describe the case such that Brigid is entitled to believe that her belief is tf-connected but not e-connected. This implies that we can be justified in believing P if we are entitled to think our belief is tf-connected, even when are not entitled to think our belief is e-connected. These cases therefore constitute counterexamples to (EC).

Most commentators agree that tf-connected beliefs are distinct from e-connected beliefs, and that they represent a class of belief that cannot be accommodated by more basic versions of explanationism like (EC) (Goldman 1987, 103; McCain 2015, 335-9; Korman 2019, 356; Korman and Locke 2020, 324). These commentators think that we should reject (EC) in favour of something that is not subject to such counterexamples, and this is exactly the move that I will eventually make myself. But Bogardus and Perrin have recently presented an argument which implies that (EC) *can* accommodate justified tf-connected beliefs. I consider this argument in the following section.

3. Bogardus and Perrin on Seeing and Deduction

Bogardus and Perrin have argued that some tf-connected beliefs are also e-connected. If their proposal is successful, and it can be applied to *all* tf-connected beliefs for which we agree that justification is possible, it would threaten my claim that tf-connected beliefs can be justified even if we are not entitled to think they are e-connected[10]. It would therefore threaten my view that (EC) is inadequate and remove the motivation for rejecting (EC) in favour of the alternative constraint that I defend below. In order to justify this move, then, I must first consider Bogardus and Perrin's reasons for thinking that some tf-connected beliefs are also e-connected.

Bogardus and Perrin discuss the following case, where it would be natural to posit a third-factor explanation between belief and truth. We know that if Neil

[10] I assume here that if justified tf-connected beliefs are also e-connected then we can sometimes be entitled to believe that these tf-connected beliefs are e-connected, in which case (EC) could accommodate justification for these beliefs.

comes to the party he will get drunk. We also know that Neil will come to the party. We infer that Neil will get drunk.

If we fill in the details in a normal way, what explains our belief that Neil will get drunk is the fact that Neil will come to the party, which also explains the fact that Neil will get drunk. Hence, our belief is tf-connected, and we are entitled to think so. But Bogardus and Perrin defend an analysis of knowledge according to which our belief in P counts as knowledge iff it is explained by P. They are therefore obliged to show how this kind of tf-connected belief is also e-connected in order to avoid the implausible result that such a belief cannot qualify as knowledge.

Bogardus and Perrin claim that tf-connected beliefs like these are formed via deduction. We believe that Neil will go to the party, and also that if Neil will go to the party then he will get drunk. We deduce that Neil will get drunk. Then Bogardus and Perrin argue:

> Deductive arguments are a means by which we come to metaphorically "see" the truth of some proposition […] so long as your visual system is functioning well, it puts you in a position to see objects before you, and if one such is a chair, you believe there's a chair there because there's a chair there. The fact that there's a chair there is a crucial part of the explanation of why you believe there is. Similarly, when you see the premises are true and the inference valid, deduction helps you see the conclusion. (13)

So, according to Bogardus and Perrin, when we successfully deduce a conclusion from premises that deductively entail it, the premises allow us to 'see' the truth of the conclusion, just as we might see that there is a chair in front of us. When we believe in the chair because we visually see the chair, the chair explains our belief in it. Similarly, Bogardus and Perrin argue, when we believe in P because we deductively 'see' P, P explains our belief. Hence, Neil's tf-connected belief is also e-connected.[11]

If *all* justified tf-connected beliefs can be given a similar treatment, then they threaten my claim that we can have justification for believing P if we are entitled to think that P is tf-connected even though we are not entitled to think they are e-connected[12]. When applying their argument to all cases of tf-connected beliefs, Bogardus and Perrin would have to view Brigid's belief as the result of the following deductive argument:

[11] Jenkins gives an alternative account of how Neil's belief could be e-connected (2006, 160-161). But, as Bogardus and Perrin point out, there are serious issues with Jenkins' proposal (2020, 13). I therefore focus on Bogardus and Perrin's more recent argument.

[12] Again, assuming we are justified in believing in the additional e-connection.

P1) The fireplace is lit

P2) If the fireplace is lit, then smoke is emerging from the chimney

C) Smoke is emerging from the chimney

Furthermore, they could argue, deductive arguments like these allow us to 'see' the truth of the conclusion, and so Brigid's belief in the smoke is in fact explained by her 'seeing' the smoke emerging from the chimney. Her belief is, therefore, e-connected.

In my view, there are serious problems with Bogardus and Perrin's proposal. Firstly, the analogy between visual perception and deduction is obscure. In cases of visual perception there is a ready account of *how* the facts explain our belief. There is a physical causal chain that links the chair to our visual impression of the chair, on the basis of which we form our belief. This chain involves photons bouncing of the chair and entering our retina, and our brain processing the information. It is because of some such causal chain that we are inclined to say that we believe in the chair *because of* the chair, and that the chair explains our belief in it.

But it is very hard to see how there could be an analogous chain linking the smoke emerging from the chimney with Brigid's belief in the smoke[13]. Of course, it is obvious how the *fire* explains Brigid's belief the smoke; the fire causes Brigid's visual impression of the fire, from which she infers her belief in the smoke. But why should this imply that the smoke emerging from the chimney also explains Brigid's belief? What is the explanatory chain linking the smoke emerging from the chimney with Brigid's belief in the smoke?

Remember that Bogardus and Perrin cannot appeal to the explanatory chain between the fire and the smoke because this chain runs in the wrong direction. The smoke is explained by, and does not explain, the fire. Hence, it is *not* the case that the smoke explains the fire which explains Brigid's belief in the smoke which would, by the transitivity of explanation, imply that the smoke explains Brigid's belief in the smoke. This is the difference between tf-connected beliefs like Brigid's and the indirectly e-connected beliefs that were described above. Hence, we can grant that Brigid's belief in the smoke is explained by the fire without this giving us any reason to think her belief is also explained by the smoke. So why think that the smoke, as well as the fire, explains Brigid's belief?

[13] To be sure, I am not claiming that one fact only explains another if there is a physical causal chain linking the two. But it is in virtue of such a chain that we are, in the case of perceptual vision, inclined to say that our beliefs are explained by the relevant facts. So if we concede there is no such physical chain for beliefs formed *via* deduction, the analogy with perception gives us no reason to think that our belief is explained by the facts in cases of deduction.

Note that we cannot answer this question by simply insisting that Brigid 'sees' the smoke emerging from to chimney *in the sense that she realises it must be true, given her premises.* This is just another way of saying that Brigid deduces the conclusion from the premises, which I am not denying. The question is why we should think that this kind of deduction involves P explaining our belief. Why think that, when Brigid realises that the smoke must be emerging from the chimney, the smoke emerging from the chimney explains this realisation? Why not think it is entirely explained by the fire in the fireplace, as well as her background belief that if the fireplace is lit, then smoke will emerge from the chimney?

If I read them correctly, Bogardus and Perrin have an answer to this question. They write:

> in a sound argument, the conclusion is already 'there,' so to speak, in the premises. When one appreciates a sound argument, one is not merely seeing the truth of the premises and the validity of the inference; one also sees the truth of the conclusion thereby. In that case, you believe the conclusion because it's true. (13)

Bogardus and Perrin claim that the conclusion is already 'there' in the premises, and so when we see the premises we thereby see the conclusion. This is how the truth of the conclusion explains our belief in the conclusion.

What could it mean to say the 'conclusion' of an argument is already there in the premises? Well, according to Bogardus and Perrin's account of knowledge, our belief in P can only qualify as knowledge if *the fact that P* explains our belief[14]. If the conclusion of a deductive argument is the proposition that there is smoke emerging from the chimney, knowledge of this proposition is only possible if our belief is explained by *the fact* that smoke is emerging from the chimney. So when Bogardus and Perrin say the conclusion is already 'there' in the premises, they must mean *the fact that makes the conclusion true* is already 'there' in the premises. If they meant merely that the *proposition* that P is already there in the premises of the argument, then we could grant this point without it having any bearing on how *the fact* that P is supposed to explain our belief. So if Bogardus and Perrin hope to successfully demonstrate that *the fact that P* can explain our belief (and hence that

[14] Bogardus and Perrin sometimes write that our belief in P counts as knowledge iff 'the truth' of P explains our belief, but their defence of this position involves showing that 'the fact that P' explains our belief in P in various test cases (2019, 7-17). This implies that Bogardus and Perrin use 'the truth of P' interchangeably with 'the fact that P'. This holds specifically in their discussion of deduction where, in defending the analogy with perception, they write: "The fact that there's a chair there is a crucial part of the explanation of why you believe there is" (2019, 13). Clearly, then, Bogardus and Perrin think they must show that the fact that P explains our belief in P in cases of deduction in order to show that deductive knowledge is consistent with their account.

their analysis of knowledge is consistent with this kind of deductive knowledge), then they must mean that *the fact that P* is already there in the premises that deductively entail it.

This, in turn, implies that when they say that the conclusion of a deductive argument is already there in the 'premises', Bogardus and Perrin must mean the conclusion is already there in *the facts that make these premises true*. This is because the fact that makes the conclusion true cannot be "already there" in the *propositions* that make up the premises to the argument. So Bogardus and Perrin must be claiming that fact that makes the conclusion true is already there in the facts that make the premises true.

But now we must ask a further question. What does it mean to say the fact that makes the conclusion true is already 'there' in the facts that make the premises true? And how could this imply that Brigid's belief in the smoke is explained by the smoke? Bogardus and Perrin use the following analogy to clarify their position:

> It's like putting the pieces of a puzzle together, and then seeing the picture. In this analogy, the premises and inference are the pieces, and the conclusion is the picture. Now, when there's a completed puzzle of a pangolin in front of you, and you believe there's a picture of a pangolin there, a crucial part of the explanation of why you believe there's a picture of a pangolin there is because there's a picture of a pangolin there. In a similar way, when [one] knows those premises together and grasps the entailment, he's in position to see the truth of the conclusion. The conclusion is now "part of the picture," as it were, formed by the true premises and valid inference, and is thereby directly appreciable. And so it's correct that: [one] believes the conclusion because it's true. (13)

Bogardus and Perrin claim that when you put the premises of a deductive argument together you get the conclusion, just like when you put puzzle pieces together you a get the picture. When we are looking at a completed pangolin puzzle, our belief in the picture is explained by *both* the puzzle pieces being arranged in the correct way *and* there actually being a picture of a pangolin in front of us, precisely because the picture is just made up of the puzzle pieces. So if our belief is explained by the truth of the premises that deductively imply a conclusion, and the truth of the conclusion is just made up of the truth of the premises that deductively imply it, this could imply that our belief is thereby explained by the truth of the conclusion as well[15]. Applied to Brigid's belief in the smoke, the idea is that her belief can be explained by the smoke (and thereby e-connected) just if it is explained by the facts

[15] Here I am, like Bogardus and Perrin, using 'the truth of the conclusion' interchangeably with 'the fact that makes the conclusion true'.

that make the premises of her deductive argument true, because the smoke is just made up of these facts.

The reason that this argument seems plausible with the puzzle example is because the relation between the puzzle pieces being arranged in a particular way and the completed picture is one of *constitution.* It is this particularly intimate relationship that inclines us to say that if our belief is explained by the puzzle pieces, then it is also explained by the picture[16].

But is it plausible to say that the truth of Brigid's belief in the smoke is *constituted* by the truth of the premises that deductively entail it? The first of these premises is that there is a fire in the fireplace. But, crucially, the fact that there is smoke emerging from the chimney is not even partially constituted by the fact that there is a fire in our fireplace. The two are causally, but not constitutively, related to each other. The fire in the fireplace occupies an entirely different space as the smoke emerging from the chimney. It would be radical to suggest that the fire can even partially constitute the smoke emerging from the chimney at such a distance. Thus, the smoke emerging from the chimney is in no way "already there" in the lit fireplace.

This means that, if the fact that smoke is emerging from the chimney is constituted by the truth of the premises that deductively entail it, then it must be entirely constituted by the truth of the second premise alone (because it is not even partially constituted by the first). But the problem with this suggestion is that, if this is the case, then it follows that Brigid could legitimately reason as follows:

P2) If the fireplace is lit, then smoke is emerging from the chimney

C) Smoke is emerging from the chimney.

That is, Brigid could have knowledge the conclusion of this argument (that there is currently smoke emerging from the chimney) *entirely* on the basis of this second premise, without having any belief in the first premise at all. This is because Bogardus and Perrin imply that we can see the truth of the conclusion by seeing the truth of the premises which constitute the truth of this conclusion, and we are now supposing that the truth of this conclusion is constituted *entirely* by the truth of the second premise alone. Hence, we can see the truth of the conclusion by seeing only the truth of this second premise. As seeing the truth of the conclusion in this way is sufficient for knowledge (according to Bogardus and Perrin) it follows that Brigid

[16] One might object to the underlying assumption that, if X constitutes Y, then if X explains Z then so does Y. I grant this point for the sake of argument, in order to focus on a separate problem with Bogardus and Perrin's proposal that I go on to discuss.

can know that smoke is emerging from the chimney just if she has knowledge of this second premise alone.

But, obviously, Brigid *cannot* have knowledge of the conclusion just on the basis of her knowledge of the second premise alone. In order to avoid this result, Bogardus and Perrin must deny that the truth of Brigid's belief is entirely constituted by the truth of the second premise, which would allow them to deny that Brigid can see the truth of the conclusion just by seeing the truth of this second premise. But above I argued that the smoke is not even partially constituted by the truth of the first premise. It follows that the truth of the conclusion is not already 'there' in the conjunction of both these premises. The first premise does not even partially constitute the truth of this conclusion, and the second premise does not wholly constitute the truth of this conclusion. So when we put these premises together we do not get the truth of the conclusion. The truth of the conclusion is not already 'there' in the conjunction of the premises that entail it.

Hence, there is a crucial disanalogy between the example that Bogardus and Perrin use to illustrate their position and at least some tf-connected beliefs. It is true to say that the picture of a pangolin is already 'there' in the puzzle pieces being arranged in a particular way, and so it becomes plausible to say that if our belief is explained by the puzzle pieces it is also explained by the picture. But it is *not* true to say that the truth of Brigid's belief in the smoke is already 'there' in the truth of the premises that deductively imply it, and hence not true to say that her belief in the smoke is explained by the smoke if it is explained by the truth of these premises.

But the suggestion that the conclusion is already 'there' in the premises was precisely how Bogardus and Perrin hoped to show that some tf-connected beliefs were also e-connected. So once we realise that the truth of Brigid's belief is not constituted by the truth of the premises that entail it, we lose any reason for thinking that the truth of the conclusion explains Brigid's belief. We can grant that her belief in the conclusion is explained by the truth of the premises on which she bases her belief in the conclusion, but this does not give us reason to think that her belief is also explained by the truth of the conclusion itself. Brigid's belief in the smoke is explained by the fire, as well as her background belief that fires cause smoke, but this in no way implies that her belief is also explained by the smoke. In other words, Bogardus and Perrin give us no reason believe that Brigid's tf-connected belief is also e-connected.

As I mentioned above, Bogardus and Perrin defend the view that our belief in P counts as knowledge iff P explains our belief. But I have argued that there is no reason to think that Brigid's belief in the smoke emerging from the chimney is explained by the smoke emerging from the chimney. This implies that, on pain of

claiming that Brigid's belief (and tf-connected beliefs in general) cannot qualify as knowledge, we have to reject Bogardus and Perrin's account of knowledge. Similar comments apply to (EC). Unless we want to insist that Brigid's belief in the smoke is unjustified, we need to amend (EC) such that it can allow for justified tf-connected beliefs.

This discussion also has implications for the following explanatory constraint on belief, which has been recently proposed by Barker (2023):

> Truthmaker Explanationism (TE): If some facts, the xs, are p's truthmaker and S believes that her belief that p is not explanatorily connected to the xs, then S has a defeater for her belief that p. (20)

According to Barker, two facts are explanatorily related if one of them explains the other (2023, 4). Given the plausible assumption that the truthmaker for the proposition that there is smoke emerging from the chimney is the worldly fact that there is smoke emerging from the chimney, it follows from Barker's Truthmaker Explanationism that Brigid's belief in the smoke cannot be justified[17]. Brigid's belief is *not* explained by, and does not explain, the fact that there is smoke emerging from the chimney and so is not 'explanatorily connected' to the truthmaker of her belief given Barker's understanding explanatory connectedness. According to Truthmaker Explanationism, then, Brigid's belief cannot be justified once she recognises this situation. But this is the wrong result, and Truthmaker Explanatinism should therefore be rejected in its current form.

In the next section I will present a constraint that *is* consistent with Brigid's belief in the smoke being justified, and the rest of the paper will be a defence of this alternative explanatory constraint on belief.

4. EThirdFactor

Here is an explanatory constraint on belief, formulated by Korman and Locke (2023):

> EThirdFactor: If S is not entitled to believe that either (i) her belief that P is e-connected or (ii) her belief that P is tf-connected, then S's belief that P is

[17] As Barker observes, some people think there are cases where the truthmaker for a proposition P is *not* the worldly fact that [P], because there is no such worldy fact (2023, 13-16). Cases include negative propositions (i.e, there are no unicorns) and disjunctive propositions (i.e, either unicorns exist or Biden is president). But these examples give us no reason to doubt that the truthmaker for the proposition that there is smoke emerging from the chimney is the worldy fact that [there is smoke emerging from the chimney], this proposition being neither negative nor disjunctive. If the truthmaker for some proposition P is *ever* the worldly fact that [P] (something which Barker accepts) then there is no reason to doubt that this holds for the proposition that there is smoke emerging from the chimney.

defeated. (9)[18]

I think that this constraint is correct. My reasoning is as follows. As the Brigid-example illustrates, our beliefs can be justified if we are entitled to believe that they are tf-connected but we are not entitlted to believe that they are e-connected. This means that the correct explanatory constraint has to allow for beliefs to be justified even if we are not entitled to think they are e-connected.

At this point we have a choice. One option is to adopt a constraint that straightforwardly allows tf-connections into the realm of explanatory connections that can support justification. This is what EThirdFactor does. Alternatively, we can adopt an explanatory constraint on belief that does allow for the possibility of justified tf-connected beliefs, but attaches different strandards of justification for tf-connected beliefs as compared with e-connected beliefs. Here is an example of such a constraint, defended by Korman:

> (EC5) S is justified in believing p only if: either (i) S doesn't withhold belief that [S's belief that p is e-connected to p], or (ii) S's belief that p is based on beliefs Bp1 ... Bpn that prima facie justify S in believing p and S doesn't withhold belief that [S's beliefs that p1... pn are e-connected to p1... pn]. (Korman 2019, 357)

This constraint implies, roughly, that any belief that we deny to be e-connected (like those we take to be tf-connected) must be ultimately based on beliefs that we take to be e-connected. But it does *not* imply that beliefs we take to be e-connected must also be based on beliefs that we ultimately take to be e-connected. Beliefs that we take to be e-connected escape defeat by satisfying clause (i) of (EC5) and so do not have to satisfy clause (ii) in order to remain justified.

Hence, the standards of justification for beliefs we take to be e-connected belief are different from those we take to be tf-connected, according to (EC5). Tf-connected beliefs need to be based on beliefs that we do not withhold belief in being e-connected in order to be justified, but e-connected beliefs do not. Because indirectly e-connected beliefs are a subset of e-connected beliefs, this implies that beliefs we take to be indirectly e-connected have different standards of justification from those we take to be tf-connected.

But, in my view, there is no significant epistemological difference between indirectly e-connected beliefs and tf-connected beliefs. Both kinds of belief involve inferring from our belief in some fact P to a belief in a distinct fact Q which we take to either explain or be explained by P. Why think the validity of this inference is

[18] Korman and Locke formulate this constraint with slightly different terminology. But given my stipulated definitions of 'e-connected' and 'tf-connected', the two constraints are identical. As discussed in section 5, Korman and Locke reject EThirdFactor.

affected by the direction of the explanatory relationship that we take to hold between the two facts? An inference from Q to P seems just as good if it is based on a background belief that Q explains P, rather than a background belief that P explains Q (as long as these background beliefs are equally justified). The direction of the believed explanatory connection does not raise significant differences in how these two kinds of belief can be justified. It is not the case that, if we are inferring Q from P because we think Q explains P then our belief needs to be ultimately based on some belief we take to be e-connected, but if we are inferring P from Q because we think P explains Q then our belief is no longer subject to this requirement.

As a result, I think that the correct form of explanationism will treat indirectly e-connected beliefs and tf-connected beliefs equally; it will not attach different standards of justification to these two kinds of belief. And this is true of EThirdFactor; it does not attach any additional requirements that beliefs that we are entitled to view as e-connnected or tf-connected need to meet in order to be justified. It therefore does not posit higher standards of justification for beliefs we are entitled to view as tf-connected compared with those we are entitled to view as indirectly e-connected; it treats these two kinds of belief equally.

I think that the motivation behind attaching different standards of justification to e-connected and tf-connected beliefs might be as follows. If we entitled to view our belief in P as tf-connected, this might presuppose that we already have some beliefs about P. Specifically, perhaps we must already believe that there is an explanatory relationship between the fact on which we base our belief in P and P itself. We cannot base our belief in P on Q if we do not already believe in some relationship between P and Q. So if we base our belief in P on some distinct fact Q which we take to explain P, this rationally commits us to some prior belief in a relationship between P and Q[19].

This implies that a belief we take to be tf-connected can only be justified if we have the correct background beliefs. By contrast, one might suppose, when we have 'direct' access to P we need have no such background beliefs about P to be justified in believing in P. So the conditions under which beliefs we take to be tf-connected can be justified are different from those we take to be e-connected; the former cannot be justified without these supporting background beliefs whereas the latter can.

I will not comment on the success of these arguments, and will instead point out that they apply just as much to indirectly-e-connected beliefs as they do tf-connected beliefs. When we take our belief to be indirectly e-connected, we take

[19] I do not mean to endorse this line of thought. I just mean to express it so I can comment on its significance to the current discussion further down.

ourselves to be forming a belief in P on the basis of some distinct fact Q. Just as we did for tf-connected beliefs, we could therefore reason that those beliefs we take to be indirectly e-connected presuppose certain background beliefs about the relationship between the basis and the object of our belief. As such, beliefs we take to be indirectly e-connected can only be justified if we have the appropriate background beliefs, whereas having direct contact with P justifies our belief in P regardless of whatever background beliefs we have.

The considerations that appeared to support different standards of justification for e-connected beliefs over tf-connected beliefs therefore actually support, at most, different standards of justification for a particular kind of "direct" e-connected beliefs over both indirectly e-connected beliefs and tf-connected beliefs alike. Hence, they give us no reason to treat e-connected beliefs *in general* differently to tf-connected beliefs. I therefore think we should adopt an explanatory constraint that treats indirectly e-connected and tf-connected beliefs equally, in the sense that it does not imply different standards of justification for each kind of belief to be justified. This is why I think EThirdfactor is preferable to Korman's (EC5), and to any other constraint that implies lower standards of justification for indirectly e-connected beliefs as compared with tf-connected beliefs[20].

Let us say that if our belief in P is explained by P, and there is no intermediate fact Q from which we infer belief in P, then our belief is "directly e-connected". We might now wonder: *can* our indirectly e-connected and tf-connected beliefs only be justified if they are ultimately based on beliefs which are directly e-connected?

My answer to this question is an underwhelming, but appropriate, 'perhaps'. It could be that all justified beliefs must ultimately bottom out in justification that we get from our 'direct' experiences. But the truth of this position would give us no reason to think that any of our beliefs can be justified *without* our being entitled to think they are e-connected or tf-connected. It would therefore give us no reason to reject EThirdFactor. We would just have a particular account of when we are entitled to view our belief as being either indirectly e-connected or tf-connected. On the other hand, if we are instead inclined to *reject* this picture of justification, then EThirdFactor should be supplemented with an alternative account of when its requirements are satisfied.

As such, the line of argument that appears to support different standards of justification for directly e-connected beliefs compared with indirectly e-connected and tf-connected beliefs is irrelevant for the purposes of this paper; neither its success nor its failure would give us reason to reject EThirdFactor. But some

[20] Locke's "Cognition Defeat" is one such constraint (2014). See footnote 24 for a discussion.

commentators have rejected EThirdfactor for other reasons, and I turn to these objections in the final section.

5. Objections to EThirdFactor

Though they are the ones to articulate it, Korman and Locke reject EThirdFactor. They do so in part on the basis of the following example:

> Cloud Chamber
>
> Jack, a physics enthusiast, has a more or less automatic tendency to believe, upon finding out that a proton has gone through a cloud chamber, that a certain kind of streak has appeared in the chamber. Jack has just learned from a reliable source that a proton has just been fired through cloud chamber C. Without looking at chamber C, he spontaneously forms the belief that there is a streak in C. However, he also learns that this proton-to-streak disposition was implanted in him by a mad neuroscientist who knows nothing about physics but simply likes the sound of 'proton' and 'streak'. (2023, 9)

According to Korman and Locke, EThirdFactor does not imply that Jack's belief that there is a streak in C is unjustified, and so this gives us reason to reject it. They write:

> […] by Jack's lights, the fact that a proton was fired through C serves as a third factor, explaining both the fact that there is a streak in C and his belief that there is a streak in C. Relying on his antecedent proton beliefs, Jack can reason his way to that third-factor explanation, and EThirdFactor provides no account of why he is not entitled to do so. (9-10)

But I think that EThirdFactor *can* account for the irrationality of Jack's belief, and so this apparent reason for rejecting EThirdFactor fails. Given a normal set of background beliefs, Jack lacks entitlement to take his *other* belief, that protons cause streaks in cloud chambers, to be either e-connected or tf-connected. Presumably, he has no background beliefs that would entitle him to think that the mad neuroscientist's behaviour is either explained by, nor explains, the fact that protons cause streaks in cloud chambers[21]. As such, when he finds out he only believes that protons cause streaks in cloud chambers because of the mad neuroscientist, he loses justification for believing that this belief is either e-connected or tf-connected. This

[21] If Jack *is* somehow justified in holding any such belief, then EThirdFactor implies that Jack's beliefs remain justified. But in my view this is exactly the right result; Jack would be justified in treating the mad-neuroscientist's behaviour as evidence of the fact that protons cause streaks, because he would justifiably take this behaviour to either explain, or be explained by, the fact that protons cause streaks in cloud chambers.

means that Jack's belief that protons cause streaks is defeated once he learns about the genealogy of that belief, according to EThirdFactor.

But Jack is only entitled to take his subsequent belief, that there is a streak in C, to be e-connected or tf-connected *if* he is entitled to hold this background belief that protons cause streaks. This was his only reason for thinking his belief was tf-connected, and he presumably lacks any additional reason for thinking his belief is also e-connected[22]. This means that if Jack lacks justification for believing that protons cause streaks in cloud chambers then he also lacks justification for thinking that his belief that there is a streak in C is e-connected or tf-connected, and his belief that there is a streak in C is thereby defeated according to EThirdFactor[23]. Hence, EThirdFactor successfully delivers the correct verdict that Jack's belief is defeated, and Korman and Locke's main reason for rejecting EThirdFactor does not hold up.[24]

Korman and Locke cite some other issues that an advocate of EThirdFactor will have to address. These issues stem from the supposition beliefs can escape defeat if we are entitled to view them as tf-connected. These problems therefore give us potential reason to reject EThirdFactor, which implies that beliefs we are entitled to view as tf-connected can remain justified. Partially on the basis of these problems, Korman and Locke therefore suggest we should reject EThirdFactor in favour of an alternative explanatory constraint on belief.

I will not attempt to solve these additional problems raised by Korman and Locke, though I am hopeful that they can be solved. Instead, I will point out that these problems do not arise specifically from the supposition that beliefs can escape defeat if we take them to be tf-connected, but in fact also arise from the supposition

---

[22] If Jack is in fact entitled to believe that his belief that there is a streak in C is e-connected, then in my view continues to have justification for his belief. See footnote 21.

[23] See my (2023a) for a similar argument.

[24] Locke also uses a version of this proton/streak example in an earlier paper, where he takes it to show that our default entitlement for some belief P is defeated if we withhold belief that our disposition to form beliefs about P are explained by P-facts (2014, 227-232). As Locke suggests, this position effectively entails different standards of justification for those beliefs we take to be tf-connected versus those beliefs we take to be e-connected, with the former (but not the latter) only being justified if we have reason for accepting the relevant explanatory connection that is independent of our default entitlement for P (2014, 232-4). My above discussion of this example shows why I think this line of argument fails. The proton/streak example does *not* imply that our default entitlement for P depends on our taking our disposition to form P-beliefs to be e-connected, and can be accommodated by a proponent of EThirdFactor who would instead think that our default entitlement for P is only defeated if we are not entitled to take our P-disposition to be *either* e-connected or tf-connected. This example therefore does not give us reason for adopting the subsequent position that there are different standards of justification for e-connected and tf-connected beliefs.

that beliefs can escape defeat if we take them to be indirectly e-connected. This means that the problems that Korman and Locke identify do not arise specifically for EThirdFactor in virtue of its allowing justification on the basis of tf-connections. Instead, they arise for *any* version of explanationism that allows even our indirectly e-connected beliefs to be justified. In other words, these problems will arise for any form of explanationism that allows for inferential justification. And because any version of explanationism will have to allow for inferential justification in order to be plausible, the problems that Korman and Locke identify will arise for, and will have to be addressed by, *any* plausible explanationist account of defeat. They therefore do not give us reason to reject EThirdFactor in favour of some alternative version of explanationism, *contra* Korman and Locke. Thus, they do not threaten my position that, to the extent that we are committed to some form of explanationism, EThirdFactor is its most plausible version.

Onto the supposed problems. When assessing EThirdFactor, Korman and Locke write:

> we must somehow disallow cheap conjunctive explanations. For if we are allowed to simply conjoin that which explains Y with that which explains Z to get a thirdfactor explanation of both Y and Z, then EThirdFactor is in trouble: there will be essentially no cases where it implies defeat. (2023, 9).

This concern is also raised (and responded to) by Faraci in relation to tf-connected beliefs (2019). Here's a concrete version of the problem. We believe that there is smoke emerging from the chimney because our friend has told us there is smoke emerging from the chimney. We also justifiably believe that if there is smoke emerging from the chimney, this must be explained by our fireplace being lit. We then find out that our friend has no idea whether there is smoke emerging from the chimney, and only told us so to mess with us. Utilising conjunctive explanations, we can now take our belief in the smoke to be explained by the conjunctive fact that [our friend lied to us and the fireplace is lit], which we can also take to explain the fact that there is smoke emerging from the chimney. Hence, we can take our belief to be tf-connected regardless of its explanation, and our belief trivially avoids defeat.

The problem arises because, unless we have a principled reason for ruling out conjunctive explanations, we can always combine the explanation for our belief with something that we justifiably take to explain the truth of our belief. This conjunctive fact therefore potentially explains both our belief and the fact that makes it true. Hence, a constraint that allows justification on the basis of tf-connections threatens to make epistemic defeat impossible.

The point I now want to make is that, until we rule out conjunctive explanations, it is equally possible to combine any fact that explains our belief in P

with some fact that we justifiably take to be *explained by* P. And, once we do this, we can trivially construct indirect e-connections between our belief in P and P, with the conjunctive fact being explained by P and also explaining our belief in P.

Say we believe that our fireplace is lit because our friend told us so. We also justifiably believe that, if our fireplace is lit, it will explain the smoke emerging from our chimney. We then find out that our friend only told us the fireplace is lit to mess with us, and has no actual idea whether or not it is lit. Allowing for conjunctive explanations, we could take the conjunctive fact that [our friend lied to us and there is smoke emerging from our chimney] to both explain our belief in the fire, and also be explained by the fact that there is fire. Hence, the fire explains the conjunctive fact, which explains our belief, and our belief is thereby indirectly e-connected.

The problem of conjunctive explanations will therefore have to be solved by any explanationst who, at a minimum, wants to allow justification for indirectly e-connected beliefs. As long as Korman and Locke want to allow for justified indirectly e-connected beliefs, they must have an account of how their own preferred constraint escapes this problem. And once a solution is on the table, it can be applied to the problem of conjunctive tf-connected beliefs as well. There is nothing specific about the structure of tf-connected beliefs that generates this issue, and so this is not a problem that arises specifically for an explanatory constraint like EThirdFactor in virtue of its allowing justification on the basis of tf-connections. As a result, this problem does not constitute reason to reject EThirdFactor in favour of some alternative explanatory constraint on belief.

A final problem with EThirdFactor, raised by Korman and Locke, involves "cheap cosmological explanations" (2023, 9). Let's say we believe that the big bang explains everything. In that case, we will trivially take there to be a third-factor explanation between our belief in P and P, because we will take them both to be explained by the big bang. Hence, all of our beliefs become trivially immune from defeat, given EThirdFactor. Jenkins also raises this issue in relation to tf-connected beliefs (2006, 140).

Once again there is an analogous problem pertaining to indirectly e-connected beliefs. The problem is not that all our beliefs are trivially justified, but rather that we are trivially entitled to believe that our belief in the all-explaining fact is indirectly e-connected regardless of our grounds for that belief. Suppose again that we believe that the big bang explains everything. We believe this because a scientist friend of ours told us so, and they are usually reliable about this kind of thing. Then we come to learn that the scientist was just saying random statements in order to gauge our response. Obviously, our belief in the big bang gets defeated. But, given our background belief that the big bang explains everything, it seems we

are also entitled to believe that the big bang explains the scientist's saying random things to gauge our response, and that our belief in the big bang is indirectly e-connected.

In my view, this problem is analogous to the problem of cheap cosmological explanations that arises in relation to tf-connected beliefs. Both cases arise because we believe that there is some all-explaining fact. Once we accept the existence of such a fact, explanatory connections between our belief in P and the truth of P come trivially once we have formed belief in P. As such, even if our initial reasons for believing P are debunked, there remains a cheap explanatory connection between our belief in P and P furnished by this all-explaining fact. This renders our belief in P potentially immune to defeat, which is unacceptable.

To be sure, the problem as it applies to indirectly e-connected beliefs is far more local, making just one belief immune from defeat (our belief in the all-explaining fact). By contrast, once we allow for justified tf-connected beliefs, the problem threatens to make *all* our beliefs immune from defeat. But any explanatory constraint will be unacceptable unless it can solve the local problem and account for cases in which our beliefs in all-explaining facts get defeated. Once such a solution is on the table, I think that it will provide the resources for solving the analogous problem for tf-connected beliefs. If there is some reason for thinking that our belief in the all-explaining fact gets debunked despite the presence of this trivial explanatory connection between our belief and truth, I think this can be used to show how our other beliefs can get debunked despite the presence of a trivial third-factor explanation between belief and truth.

As any acceptable explanatory constraint will have to solve this problem as it relates to indirectly e-connected beliefs, I see no reason for thinking it constitutes reason for rejecting EThirdFactor in favour of some such alternative constraint. EThirdFactor incurs no additional costs in relation to these cosmological explanations; the cost is to show how the indirectly e-connected beliefs in all-explaining facts can get defeated, and for this price we will get an answer to the problem of cheap cosmological tf-connections as well.

I therefore conclude that the apparent reasons for rejecting EThirdFactor are either not real problems or will apply with equal force to any plausible explanatory constraint on belief. As such, if *any* explanatory constraint on belief is to be accepted, it should be EThirdFactor. We lack reason for thinking that a superior constraint that has been suggested thus far.

6. Conclusion

In order to defend EThirdFactor, I first had to defend the view that our beliefs can be justified when we take them to be tf-connected, even if we are not entitled to think they are e-connected. This involved considering and rejecting a recent argument from Bogardus and Perrin that threatened this view. This discussion implied we should reject Bogardus and Perrin's account of knowledge, as well as a recently proposed explanatory constraint from Barker. I then argued that the standard of justification for (at least some) e-connected beliefs is no different from the standard of justification of tf-connected beliefs. This suggests we should accept an explanatory constraint like EThirdFactor, which straightforwardly allows tf-connections among the realm of explanatory relationships that are relevant for justification, rather than something akin to Korman's (EC5). Finally, I have defended EThirdFactor against a number of objections raised by Korman and Locke and others, and argued that they do not imply we should eschew EThirdFactor in favour of some alternative explanationist view about defeat. If we do accept explanationism then I conclude that we should accept EThirdFactor as its most plausible version.

References

Barker, Jonathan. 2023. "Genealogical Defeat and Ontological Sparsity." *Midwest Studies in Philosophy* 47: 1-23.

Bhogal, Harjit. 2023. "Explanationism versus Modalism in Debunking (and Theory Choice)." *Mind* 132(528): 1005-27.

Bogardus, Tomas, and Will Perrin. 2022. "Knowledge Is Believing Something Because It's True." *Episteme* 19(2): 178-96.

Clarke-Doane, Justin. 2020. *Morality and Mathematics.* Oxford University Press.

Faraci, David. 2019. "Groundwork for an Explanationist Account of Epistemic Coincidence." *Philosophers' Imprint* 19.

Goldman, Alan H. 1984. "An Explanatory Analysis of Knowledge." *American Philosophical Quarterly* 21(1): 101-8.

———. 1967. "A Causal Theory of Knowing." *Journal of Philosophy* 64(12): 357-72.

Jenkins, C. S. 2006. "Knowledge and Explanation." *Canadian Journal of Philosophy* 36(2): 137-64.

Korman, Daniel Z. 2019. "Debunking Arguments in Metaethics and Metaphysics." In *Metaphysics and Cognitive Science*, edited by Alvin Goldman and Brian McLaughlin, 337-63. Oxford University Press.

Korman, Daniel Z., and Dustin Locke. 2020. "Against Minimalist Responses to Moral Debunking Arguments." *Oxford Studies in Metaethics* 15: 309-32.

———. 2023. "An Explanationist Account of Genealogical Defeat." *Philosophy and Phenomenological Research* 106(1): 176-95.

Locke, Dustin. 2014. "Darwinian Normative Skepticism". In *Challenges to Moral and Religious Belief: Disagreement and Evolution*, edited by Michael Bergmann and Patrick Kain, 220-236. Oxford University Press.

Lutz, Matt. 2018. "What Makes Evolution a Defeater?" *Erkenntnis* 83(6): 1105-26.

McCain, Kevin. 2014. *Evidentialism and Epistemic Justification*. Routledge.

Noonan, Christopher. 2023a. "An Explanationist Account of What Goes Wrong with Third-Factor Replies to Debunking Arguments." *Synthese* 202(3): 1-26.

———. 2023b. "Evolutionary Debunking Arguments, Explanationism and Counterexamples to Modal Security." *Erkenntnis*: 1-23.

Paul, L. A., and Ned Hall. 2013. *Causation: A User's Guide*. Oxford University Press UK.

Schaffer, Jonathan. 2012. "Grounding, Transitivity, and Contrastivity." In *Metaphysical Grounding: Understanding the Structure of Reality*, edited by Fabrice Correia and Benjamin Schnieder, 122-38. Cambridge University Press.