

Undercutting Defeat

When it Happens and Some Implications for Epistemology

Matthew McGrath

It's commonplace to lose a justification for believing something. It can happen through simple forgetting. This paper focuses on another way: through defeat. When you lose a justification through defeat, the ground for your earlier justification remains but no longer justifies you in believing the target proposition, due to the operation of a “defeater” of that justification. Defeat strictly speaking encompasses not only loss of a justification but also reduction the degree of support provided by a justification.

John Pollock has distinguished between two types of defeaters: rebutting and undercutting.¹ Let's start with an intuitive gloss. Rebutters defeat by striking against the target proposition itself, undercutters by striking against the connection between the ground of a justification and the target proposition. For an example of a rebutter, suppose I obtain some justification from a *Buzzfeed* news report for believing that a certain public official suborned perjury, but later I lose my justification when I learn that *The New York Times* and *The Washington Post* independently reported later that this official did no such thing. In this case, the information I received the newspapers strikes against the claim that the official suborned perjury. It defeats the justification I obtained from *Buzzfeed*. For an example of undercutting defeat, we can vary the *Buzzfeed* example: suppose I obtain the same justification from the *Buzzfeed* report but subsequently learn that their reporters relied exclusively on the testimony of a single anonymous source known to have personal interests at stake. The new information doesn't strike against the claim that the official suborned perjury. It defeats, rather, by striking against the connection between my reason—that *Buzzfeed* reported that the official suborned

¹ See for instance, Pollock (1970, 1975, 1987); Pollock and Cruz (1999).

perjury—and the proposition I believe—that the official did this.² The category of undercutters is often thought to represent an important advance in epistemology, for it promises to help us explain the defeating power of considerations that don't bear on the truth of the target proposition, such as evidence that a testifier is unreliable or evidence that one is hallucinating.

Although there is disagreement about the details, the distinction between rebutting and undercutting defeat—defeat by striking against the target proposition vs. defeat by striking against the connection between a ground and the target proposition—is now part of the received wisdom in analytic epistemology. Not only that, but many epistemologists think Pollock's account of how these defeaters work is roughly on the right track. Recently, however, cracks have appeared in the consensus. While not questioning the existence of undercutting defeat, Scott Sturgeon argues that Pollock misconceives it: undercutting operates differently from rebutting in that it occurs only in conjunction with certain higher-order contributions, i.e., with beliefs about (or justifications to believe propositions about) the basis on which one does or would believe.

I will argue, contrary to Sturgeon, that in the case of inferential justification, undercutting defeat takes basically the shape Pollock suggests it does, not needing contributions from higher-order beliefs or justifications. However, I agree with Sturgeon that for noninferential justification, the Pollockian account is in trouble. I try to explain why there should be this difference. My interest, however, is not merely in determining what makes for a good account of defeat. This difference in types of defeat has important implications for other parts of epistemology. In a final section, I use the defeat-related difference between inferential and noninferential justification to argue that there is less immediate perceptual or testimonial justification than is commonly thought.

9.1. Pollock on Defeaters

Let's start by reviewing the basic elements of Pollock's mature account of defeat as developed in his 1999 book with Joseph Cruz. First, there is the notion of a *reason* to believe something:

² In principle, the same defeater could be at once a rebutter and an undercutter, if it strikes against both the target proposition and the connection between the reason and that proposition.

A state M of a person S is a **reason** for S to believe Q if and only if it is logically possible for S to become justified in believing Q by believing it on the basis of being in the state M .

As I understand the terminology here, ‘become justified in believing Q ’ ascribes doxastic justification, i.e., coming to have a justified belief in Q . In terms of this notion of a reason, Pollock and Cruz define defeaters in general:

If M is a reason for S to believe Q , a state M^* is a **defeater** for this reason if and only if the combined state consisting of being in both the state M and the state M^* at the same time is not a reason for S to believe Q .

Next, *rebutters*:

(Rebutters): if M is a defeasible reason for S to believe Q , M^* is a **rebutting** defeater for this reason if and only if M^* is a defeater (for M as a reason for S to believe Q) and M^* is a reason for S to believe $\sim Q$.

For *undercutters*, they give two definitions, one for doxastic reasons and another for nondoxastic reasons:

(The doxastic case): if believing P is a defeasible reason for S to believe Q , M^* is an **undercutting** defeater for this reason if and only if M^* is a defeater (for believing P as a reason for S to believe Q) and M^* is a reason for S to doubt or deny that P would not be true unless Q were true.

(The nondoxastic case): if M is a nondoxastic state that is a defeasible reason for S to believe Q , M^* is an **undercutting** defeater for this reason if and only if M^* is a defeater (for M as a reason for S to believe Q) and M^* is a reason for S to doubt or deny that he or she would not be in state M unless Q were true.

We can simplify these definitions, even by Pollock’s and Cruz’s own lights. They go on to tell us that undercutting defeaters are reasons to believe the appropriate subjunctive conditional is false (Pollock and Cruz 1999: 196). Their suggestion, I take it, is that it’s because undercutters are reasons to doubt or disbelieve such conditionals that they manage to be defeaters. This in turn suggests that the clause that “ M^* is a defeater for . . . as a reason for S to believe Q ” is redundant in the previous definitions. The same goes for rebutters. Thus, I attribute to them the following simplified definitions of undercutting defeat:

Rebutters:

If M is a defeasible reason for S to believe Q, M* is a **rebutter** for this reason if and only if M* is a reason for S to believe \sim Q.

Undercutters:

(The doxastic case): if believing P is a defeasible reason for S to believe Q, M* is an **undercutter** for this reason if and only if M* is a reason for S to doubt or deny that P would not be true unless Q were true.

(The nondoxastic case): if M is a nondoxastic state that is a defeasible reason for S to believe Q, M* is an **undercutter** for this reason if and only if M* is a reason for S to doubt or deny that he or she would not be in state M unless Q were true.

So revised, the definitions function as intended only if something's meeting the conditions for being a "rebutter" or an "undercutter" guarantees its meeting the conditions for being a "defeater." This will be important in the following.

Pollock and Cruz (1999: 197) suggest a simpler replacement for the complex proposition that it's not the case that P would not be true unless Q were true, namely *that P does not guarantee Q (in the circumstances)*. I follow them. Thus, undercutters in the doxastic case are reasons to think that P does not guarantee Q, and in the nondoxastic case they are reasons to think that one's being in state M does not guarantee Q. This is their attempt to capture the key intuitive idea of undercutting defeat as defeat that *strikes against the connection* between a reason and the target proposition. This striking-against-the-connection is explained in terms of having reasons to doubt or deny a certain sort of proposition about that connection.

One might raise a number of questions about Pollock's account of defeat, or the parts of it I have described.³ One could ask about the ontology of reasons. Why should we take reasons to be mental states (e.g., beliefs, experiences) rather than facts or propositions? One could ask about the normative conditions a defeater must meet to defeat: must doxastic defeaters be justified? If I form a silly unjustified belief that I am hallucinating, will that defeat my perceptual justification? One could also ask about the sort of conditional in terms of which Pollock defines undercutters. Is "P doesn't guarantee Q" the right one? Is that really equivalent to the *negation of the unless-statement* "it's not the case that P wouldn't be true unless Q"? Or should we build the account around the

³ I omit exposition of Pollock's discussion of defeated defeaters, as well as his full account of justification in terms of status-assignments assigning statuses of "defeated" and "undefeated" to elements of arguments (see Pollock and Cruz 1999: 197–200).

indicative conditional, or some other conditional? These worries concern, relatively speaking, matters of detail that likely won't sink the project of defending a Pollockian account of defeat. One might also worry that Pollock's account is incomplete. Are there other kinds of defeaters? And of course one can ask for clarifications about key notions such as the notion of its being possible to become justified in believing something based on a combination of states.⁴ My concern in this paper is whether the Pollockian picture is *approximately right as far as it goes*. We might need to change certain details. We might need to supplement it in a number of ways. My particular question, more precisely, is whether there is a serious error in the picture when it comes to undercutting defeat.

9.2. Sturgeon against Pollock

Sturgeon argues against Pollock's account of undercutting defeat, claiming it is both too strong and too weak. Here I focus on his argument that it is too weak.⁵ The argument in question is based on a case, the *Milk Taster*. I have revised the example to avoid a certain objection.⁶

The Milk Taster:

The milk taster thinks her conclusions about whether a batch of milk is spoiled are based solely on taste and not at all on smell. However, this is not true. It turns out that her conclusions are based on smell only and not taste. Relying on her olfactory experience, she concludes that a particular batch of milk is ok. She is then given misleading information to the effect that she is suffering a random olfactory hallucination. She comes to think that her olfactory experience does not guarantee that the milk is OK.

⁴ As Sturgeon (2014: 108–14) patiently shows.

⁵ Sturgeon's argument (using the Presupposer case) that it is too strong is convincingly criticized by Casullo (2018: 2900–1).

⁶ In the original, Sturgeon (2014, 114–15) uses the more realistic example of relying on a complex gustatory-cum-olfactory experience to believe the milk is OK, while getting evidence of olfactory hallucination. The subject in his case thinks she is relying on taste alone. Sturgeon says that after a bit of thinking the subject will deny the relevant *unless-statement*. She will deny that she wouldn't have the complex experience unless the milk was OK. But this seems wrong. The subject has no reason to think that the olfactory hallucination affects either the milk or whether the taste experience would covary with the state of the milk. So, she is in a position to reason like this: *I am relying on taste experience T; although I'm having an olfactory hallucination in having olfactory experience O, this is no reason to doubt that I wouldn't have T unless the milk was OK; given that O has no bearing on either T or the state of the milk, I wouldn't have T+O unless the milk was OK*. Of course, strengthening of the antecedent is not valid for subjunctive conditionals, but in cases where the proposition added to the antecedent is irrelevant, it may be fine to reason from the original to the strengthened conditional. My revised case avoids these issues and gives Sturgeon the result he wants.

Sturgeon thinks that the milk taster remains justified in her conclusion about the milk, in virtue of the same non-doxastic reason that justified her previously. No defeat occurs. Nevertheless, her belief that she is undergoing olfactory hallucination meets the (simplified) Pollockian definition of an undercutter. Putting things in terms Pollock's framework: although the olfactory experience o is a reason to believe the milk is OK, and the belief that one is suffering an olfactory hallucination (H) is a reason to believe that having o doesn't guarantee the milk is OK, the combination $\langle o, \text{belief in } H \rangle$ is still a reason to believe the milk is OK—i.e., one *can* become justified in believing that the milk is OK by believing the milk is OK on the basis of being in this combination of states, and the Milk Taster case is a case in point.⁷ Sturgeon concludes that Pollock's account of undercutters is too weak—it implies that there is an undercutter for a reason in this case when there is no defeat at all. However, Sturgeon thinks that there is defeat if we modify the case so that the milk taster has a correct belief about the source of her belief about the milk. This suggests to him that we need to revise Pollock's account of undercutting defeaters by giving an essential role to beliefs about one's mental states, in particular to beliefs about the sources of one's belief.

More generally, let U be the claim that source S is untrustworthy about whether P and $BOS-P$ be the claim that your belief in P is (or would be or likely would be) based on source S . Suppose you believe U and believe P . Sturgeon proposes that in such a situation:

Your belief in U undercuts your belief that P iff you believe $BOS-P$.⁸

⁷ As noted earlier, more information is needed about what it is to become justified in believing something on the basis of being in a certain combination of states. It would be too much to demand that the combination be used as one's grounds for the belief. When defeat occurs, it's not because it would be wrong to use the combination of states as one's ground in any intuitive sense. I wouldn't think of using the combination of *John told me that p and John was insincere in his remark to me on whether p* as my ground to believe p ; rather, I might use *John told me that p* as my ground to believe p despite being also in the state of believing that *John was insincere in his remark to me on whether p* . I will read 'on the basis of being in state M ' in such a way as to allow for such possibilities. To believe p on the basis of being in state M , then, as I will understand it within Pollock's framework, is to be in state M and to believe p using some component state of M (proper or improper) as one's ground.

⁸ As Casullo (2018) points out, it's implausible that the mere having of higher-order beliefs does much epistemological work. What counts is justification to have such beliefs. We could reformulate Sturgeon's proposal so that it concerns justification to believe rather than belief. But as I mentioned, my main concerns don't hinge on such tweaks, so I will not bother with reformulation. Mellis (2014: 438) usefully refines Sturgeon's proposal, bringing in the concept of a justificatory process in addition to the concept of a source of justification.

If we accept Sturgeon's understanding of undercutting, we must revise Pollock's general definition of a defeater to make it disjunctive. The needed revision would presumably be something like this:

If M is a reason for S to believe Q, a state M* is a **defeater** for this reason if and only if either (i) the combined state of being in M and being in M* is not a reason for S to believe Q, or (ii) the combined state of being in M, being in M* and *believing that one's basis for believing Q is or would be M* is not a reason for S to believe Q.

Finally, to handle undercutting defeat of reasons for beliefs one already has, we would need to adjust the definition of 'reason' slightly as well:

A state M of a person S is a **reason** for S to believe Q if and only if it is logically possible for S to become *or remain* justified in believing Q by believing it on the basis of being in the state M.

Presumably the issue for the milk taster isn't becoming justified but remaining justified, and Sturgeon's thought is that if you have a belief about the identity of the source of your belief that p, together with a doubt or disbelief that having that source guarantees P, you can't remain justified in believing P.

Undercutting for Sturgeon is thus a matter of believing both:

- My belief that P is (or will be) based on source S.
- Source S is (or has a good chance of being) unreliable in the circumstances on whether P.

Whether one's belief actually is or will be based on source S is neither here nor there, as Sturgeon himself notes (2014: 117). All that matters is what the believer thinks about her beliefs' sources. The resulting view implies that if you think you are using taste and you think that taste is unreliable in the circumstances, then even if you aren't using taste (and are using a source you believe to be reliable), still your belief is undercut.⁹

⁹ Thus, if Sturgeon is right about undercutting, Pollock's account is not only too weak; it is also too strong. An alternative view is that undercutting happens iff Pollock's conditions are satisfied *and* one believes one source is M. On this view, Pollock's account would not be too strong. I don't have the space to examine this view in any detail. But note that even if this view is correct, undercutting doesn't happen *solely in virtue* of having reason to believe something that strikes against the connection between one's reason and the target proposition—higher-order contributions are also part of the explanatory story.

There is something odd about Sturgeon's account of undercutting. The whole idea of such defeat was that it was to be defeat by virtue of striking against the connection between an *actual ground* of justification and the target proposition. What Sturgeon gives us under the banner of "undercutting defeat" has nothing to do with the actual ground, as Juan Comesaña (Forthcoming) notes. Sturgeon-undercutters strike at best against the connection between one's *perceived* basis or ground and the target proposition, whether or not this perceived basis is one's real basis.

What Sturgeon is pointing to, I think, is actually a third kind of defeat, what I'll call "higher-order" defeat. M^* is a higher-order defeater of M as a reason to believe P iff M^* is a reason to think that one's belief that P is (or would be) unreliably based. Something can be a higher-order defeater of a reason M without being a rebutter, i.e., without being a reason to believe not- P , and without being an "undercutter" in the intuitive sense of something that defeats by striking against the connection between one's actual reason and P .^{10,11}

Casullo (2018: 2904) comes to Pollock's defense. He questions whether Sturgeon has put his finger on an asymmetry between undercutters and rebutters:

Consider a cognizer A . Suppose that source S_1 justifies A 's belief that P to degree d_1 , that source S_2 justifies A 's belief that not- P to degree d_2 , and that $d_1 = d_2$. (S_1 and S_2 need not be different sources). Is this sufficient for A 's belief that not- P to defeat A 's justification for the belief that P ? There are three possible responses:

1. Weak Internalist: No, A must believe that $d_1 = d_2$.
2. Strong Internalist: No, A must believe that $d_1 = d_2$ and that belief must be *justified*.
3. Externalist: Yes.

¹⁰ We saw an example of this previously. You falsely think you are relying on taste, which you think (and even know) to be unreliable, but in fact you're relying on olfaction: you rely on an olfactory reason O . Then your belief that you are using the unreliable source of taste is a higher-order defeater of O as a reason to believe P , but it is not a reason to believe not- P (and so not a rebutter) and it does not strike against the connection between O and P (and so is not an undercutter).

¹¹ Note that higher-order defeat does not boil down to the same thing as "defeat involving defeaters with higher-order content." Perhaps some defeaters with higher-order content, such as *I have strong evidence against p* are rebutters, and perhaps others can be undercutters. "Higher-order defeat," as I am stipulatively using this terminology, refers to defeat by virtue of being a reason to believe that one's belief is unreliably based.

Since Sturgeon sees no role for higher-order beliefs or justifications in the case of rebutting defeat, Casullo thinks Sturgeon must give the externalist answer. Casullo's claim is then that there is no reason to think that if the externalist position is correct here it wouldn't also be correct for cases of undercutting defeat. He concludes that the two sorts of defeat work in the same way—they both work with—or without—the help of higher-order beliefs/justifications.

Casullo, as I read him, doesn't strictly take a stand on whether rebutting defeat needs higher-order help. But it is implausible to think that it does. If I have evidence for P, and I then gain equally good evidence for not-P, I don't need to have higher-order beliefs or justifications about these matters in order for the different justificatory forces to weigh up in one direction or another. This is intuitive as it stands. But we can add as well that if higher-order contributions were necessary, worries would arise about how creatures that don't grasp these higher-order propositions could have rebutting defeaters (e.g., young children or higher animals).¹²

There is also reason to worry that if rebutting defeat requires higher-order contribution, the same would hold for a single source's justificatory power. We could ask: in order for a source to justify a subject's belief to a certain degree, must the person believe that the belief is justified (by that source) to that degree? It is difficult to see why the facts about how justificatory forces weigh up should require beliefs about those forces while the existence of a single justificatory force doesn't. But it is implausible to think that in order for a source to provide justification for a belief the person would have to believe (or be justified in believing) some proposition about its doing so.

I conclude that rebutting defeat does not require higher-order contribution. This doesn't of course show Sturgeon is correct to distinguish the way rebutting and undercutting defeat do their work. For, as we saw with help from Comesaña, what Sturgeon describes is not really undercutting defeat at all—at least not in the intuitive striking-against-the-connection sense; it is what I have called “higher-order defeat,” which naturally requires higher-order contribution.

¹² Using a distinction from Audi (1993), we can distinguish two ways that a justificatory fact—such as my having justification from a piece of evidence or my justification being defeated—could depend on higher-order beliefs (or justifications). One way is negative dependence: the justifications obtain only if you do not have certain *negative* higher-order beliefs (justifications), such as a belief that your basis is a poor one. But this is not the kind of dependence at issue. At issue is positive dependence: in order for justificatory facts to obtain one must have a higher-order beliefs (justifications) about those facts.

9.3. In Defense of Pollockian Undercutting Defeat (in the Inferential Case)

We have seen that what Sturgeon says about the Milk Taster case is plausible: Pollock's conditions for undercutting are met but there is no defeat. I agree with Sturgeon that the failure is not a minor one: there are no defeaters of nondoxastic reasons that do their work by attacking the connection between one's actual reason and the proposition believed. The closest things are defeaters that work by giving one a reason to think one's belief is unreliably based. But Sturgeon draws a much broader conclusion: there are no defeaters—even for doxastic reasons—that work in this attacking-the-connection way. I now argue that this broader conclusion is false: we find just such defeat in the case of inferential justification, i.e., of defeat of doxastic reasons.

Consider an example of undercutting defeat described by Sturgeon (similar to one discussed by Pollock and Cruz):

The Polling Case

A pollster surveys 1000 voters in Texas at random, asking whether they will vote Republican or Democrat in the next election. The pollster believes G(oin)g = Roughly 87 percent of Texans are going to vote Republican.

On the basis of believing

T(ell) = 87 percent of respondents tell the pollster they will vote Republican.

Suppose the pollster then comes to believe

U = Respondents decided their answer by coin flip.

Belief in U fits the (simplified) Pollockian definition of an undercutter. Is there defeat? On Pollock's own definition of defeat, it appears the answer is yes: it seems impossible to become (or remain) justified by believing in G based on the combination of <belief in T, belief in U>. Moreover, this fits with the intuitive thought that there is defeat in this case. Notice that higher-order beliefs don't come into it. The pollster might even lack such a belief.¹³

¹³ What if higher-order beliefs are present? Suppose, despite believing respondents decided their answers by coin flip (U), and despite believing that roughly 87 percent of Texans are going to vote Republican (G) on the basis of the belief this is what 87 percent of the Texan respondents told you (T), you *think* you believe G on the basis of some other belief, H, which you think makes G likely true. In such a case could you become or remain justified in believing G based on <belief in T, belief in U>? I think not. These false beliefs about your basis even if justified do not seem to *restore your actual basis as a reason* to believe G by somehow defeating your defeater (U) for that reason. Intuitively, you're on the hook, epistemically, for the basis you have; and you can't get off the hook by having beliefs (even justified beliefs) about having some other basis.

Why is it not possible to become justified in believing G based on this combination? Having a plausible explanation will bolster the previously mentioned intuitive judgments. I want to explore two accounts.

The first appeals to the well-known *taking condition* on inference. This is the condition that one must *take* the premise(s) to support the conclusion in order to infer the conclusion from the premise(s). Once the pollster comes to think that the respondents were determining their answers by coinflip (U), it seems the pollster is no longer entitled to a number of crucial claims about the connection between T and G, such as:

- If T, then G
 - If 87 percent of Texas respondents told me they will vote Republican, then roughly that percentage of Texans are going to do so.
- It wouldn't be that T unless G
 - It wouldn't be that 87 percent of Texas respondents told me they will vote Republican unless roughly that percentage of Texans are going to.
- T indicates in the circumstances that G
 - 87 percent of Texas respondents telling me they will vote Republican indicates in the circumstances that roughly that percentage of Texas are going to do so.

In fact, the pollster gains reason to believe these are false. Notice that each of these conditionals seems to be a plausible way of spelling out what *T guaranteeing G* amounts to, and similarly for *T supports G*.¹⁴ Assume, then, that the taking condition is a genuine condition on inference. Then there is an explanation available for why you cannot become justified in believing G based on <believing T, believing U>, as follows:

This combination provides at best *inferential* justification from T to believe G. So, in coming to be justified in believing G based on <believing T, believing U> you must be epistemically appropriate to infer G from T,

¹⁴ Isn't guaranteeing much stronger than mere support? As the terms are used in ordinary parlance, yes. I use 'support' here only because it is the standard term used for the taking condition on inference. However, its ordinary meaning is not sufficiently strong for the purposes of giving an account of inference. I might take the presence of some clouds in the sky to support rain, but I certainly won't infer it will rain based just on such slim support. By contrast, if I take the particular clouds to *establish, ensure, imply, indicate*—guarantee—rain then I will make the inference. If I don't so take them, I may well not make the inference. There are complexities here about whether 'guarantee' is too strong, but I don't think they matter to our discussion, since the undercutters at issue are reasons not merely to deny that there is a guarantee but also to deny that there is strong support. To keep things simple, then, I use 'support' and 'guarantee' interchangeably.

despite believing U; but for your inference to be epistemically appropriate, since inference in this case requires taking T to support G, your taking T to support G must be epistemically appropriate, too, which it is not, because U gives you reason to doubt that T guarantees or supports G.¹⁵

Believing U spoils the epistemic appropriateness of taking T to support G, which prevents one from becoming justified in believing G based on <believing T, believing U>.

This argument relies on the assumption that for an inference to be epistemically appropriate, the relevant *taking* must be epistemically appropriate. Suppose you have good reason not to take P to support Q, say because you have reason to think that P does not support Q. Then if inferring involves taking, it seems you should not make the inference.

But is the taking condition a plausible condition on inference? There are *prima facie* reasons to accept it. It helps to explain, as Hlöbil (2014) notes, why “P, so Q, but P does not support Q” seems incoherent. In affirming the second conjunct you are explicitly denying something which you are taking to be the case insofar as you are making the inference expressed in the first conjunct. The taking condition would also help to explain, as Boghossian argues, how inference differs from mere association: we can be responsible for the inference itself and not merely for the output belief—and we can be directly responsible, rather than merely responsible for it in virtue of being responsible for doing something or omitting doing something which could foreseeably have led to the inference, as is the case for associations. The taking condition is not perhaps the only way to explain these facts, but it is one way.

But the taking condition brings with it a host of problems as well. What is taking? Is it believing? If so, then presumably for inference to transmit justification from premises to the conclusion, the taking state must itself be justified, which raises at least the threat of a regress.¹⁶ If taking isn’t believing,

¹⁵ One could alternatively recast the previous explanation in terms of *remaining* justified in believing G on the basis of the combination <believing T, believing U>. In place of the language of “being appropriate to infer G from T” one could substitute “being appropriate to continue to base inferentially one’s belief that G on one’s belief that T.”

¹⁶ However, one might claim that both this linking belief and its justification are derivative from the disposition to make the inference and the transmissibility of justification available through the inference. For an example of someone who argues that the linking belief is derivative from the reasoning, consider Broome (Forthcoming): “my account of reasoning entails that, if you reason from premises to conclusion, you implicitly have the linking belief. You believe the conclusion as a result of a reasoning process starting from your believing the premises. This shows you implicitly believe that the conclusion follows from the premises. This implicit belief is constituted by your disposition to derive the conclusion from the premises by a process of reasoning.” If one accepts this view of linking beliefs, it’s natural to claim that their justification derives from the justificational properties of the available inference.

what is it? And why should taking P to support Q be epistemically inappropriate if you have good reason to think P doesn't support Q? The taking state presumably cannot be a seeming, for instance, since it is not in general epistemically inappropriate for it to seem to you that something is the case when you have good reason to think it isn't the case.

Given such worries about the taking condition, it might seem I'm hitching my wagon to a controversial position about inference. Fortunately, a similar, second explanation doesn't presuppose the taking condition. In essence, we can just omit the taking step in the previous explanation. Why is it not possible to become justified in believing G based on <belief in T, belief in U>? The answer is:

This combination provides at best *inferential* justification from T to believe G. So, in coming to be justified in believing G based on <believing T, believing U> you must be epistemically appropriate to infer G from T, despite believing U, which you are not, because U gives you reason to doubt that T guarantees G.

Continuing to assume that 'guaranteeing' and 'supporting' are interchangeable in this context, the crucial claim is that it cannot be epistemically appropriate to infer G from T because U gives you a reason to doubt that T supports G.

This explanation does not depend on the taking condition. The fact that inferring from the premise T to conclusion G would be epistemically inappropriate if one has reason to doubt that the premise wouldn't be true unless the conclusion was true—or if you prefer that T supports G—is plausible on its own without assuming the taking condition. In fact, its plausibility provides at least some reason to accept the taking condition, insofar as the taking condition would help explain why inference would be epistemically inappropriate in such situations.

A key background assumption in these two explanations is that *inference* from premise(s) to a conclusion is something that itself can come up for assessment as epistemically appropriate or not, *ex ante* and *ex post*. This seems plausible.¹⁷ You can be epistemically *inappropriate* to infer even when you know the premises, and epistemically appropriate to infer—to draw the inference—even if your beliefs in the premises are epistemically inappropriate or unjustified. Once you've inferred, if you weren't epistemically appropriate to infer, we can criticize your actual inference as epistemically inappropriate (*ex post*).

¹⁷ See also Balcerak Jackson and Balcerak Jackson (2013: 115). They focus on normative evaluation of inferences people draw, and so on *ex post* rather than *ex ante* epistemic appropriateness.

I conclude that Pollock-style undercutting defeat can occur inferential cases, with no need for higher-order contributions.

9.4. Undercutting Defeat for Noninferential Justification?

The conclusion of the previous section raises a question. Given that there is genuine undercutting defeat in the Polling case and in other cases of inferential justification, why wouldn't the same hold for cases of noninferential justification such as the Milk Taster case? Why, in particular, wouldn't the explanations discussed in the last section apply just as well to non-inferential cases? This section aims to answer these questions.

Is there a taking condition on experience-to-belief transitions (of the sort that can result in non-inferential justified belief)? This might seem implausible on its face (how could a taking be required for noninferential transitions; wouldn't that make them inferential?¹⁸), but it's worth looking into the details. If a taking condition holds for these transitions, then either the taking state must be explanatory of the transition or it needn't be. Let's look at both possibilities in turn.

A taking condition for experience-to-belief transitions cannot be explanatory of these transitions. Consider Boghossian (2014) on inference. He thinks that it's *because* you take your premises to support your conclusion that you "draw" your conclusion. But this clearly depends on your already believing the premises. If you take P to support Q but don't believe P, you will not be in a position to infer Q. In the same way, your taking some fact about your experience to support Q, or even your taking your experience itself to support Q, can't by itself explain why you come to believe Q. Takings are relevantly like commitments to conditionals, and for a commitment to a conditional to explain belief in the consequent, one needs a commitment to the antecedent, so that the content of this state together and that of the taking state "lock together" to explain the transition to the belief in Q. Thus, one must be committed to that which one takes to support Q—the fact about the experience or the experience itself. Plausibly, a belief is needed for commitment to

¹⁸ Balcerak Jackson and Balcerak Jackson (2013) claim that where there is no gap between the contents of states involved in a transition, it is implausible to posit inference. One might think that in genuinely perceptual justification, the transition is from an experience with the content P to a belief that P.

the antecedent. I must believe this fact about the experience obtains or that the experience itself exists.¹⁹

What if a taking state is required for the relevant transitions but isn't explanatory of them? John Broome (Forthcoming) is sympathetic to the view that if one reasons from beliefs in P_1, \dots, P_n to a belief in Q then one has a *disposition* to reason from the former to the latter, and that this disposition is an implicit belief that Q follows from $P_1 \dots P_n$. This implicit "linking belief" doesn't help explain inference; rather, the nature of inference guarantees that one has such a belief. Could something like this hold for experience-to-belief transitions? I don't think so. The relevant disposition exercised in a noninferential experience-to-belief transition doesn't seem, by itself, to constitute such a linking belief.

Why not? It's instructive to consider why Broome himself denies that every instance of reasoning requires a linking belief. He gives the example of reasoning involving transitions between intentions, such as instrumental reasoning in which one starts with an intention for an end and a belief about the means to that end and concludes with an intention concerning the means. Unlike the case of inference involving only beliefs, in which the contents for the linking belief would not go significantly beyond the contents of the beliefs involved, adding only whatever content is needed for the conditional, the contents of the linking beliefs corresponding to the instrumental reasoning with intentions would contain new material going far beyond the contents of the intentions/beliefs involved and beyond the contents needed for a conditional; the linking belief would have to be *about* intentions—it would have to be a belief that if I had such and such intention, then The same goes for experience-to-belief transitions; the linking belief, here too, would need to be about experiences. But it isn't plausible that your having the disposition to make experience-to-belief transitions ensures that you have a linking belief with such a rich content. Not only that, but as Broome notes, it's implausible to think that young children have such linking beliefs, but not implausible to think that they engage in the relevant instrumental reasoning with intentions. The same goes for the experience-to-belief transitions.

So, the explanation based on the taking condition, even if successful in the case of inferential justification, doesn't carry over to explain how there can be

¹⁹ It's *seeming* that a fact obtains or that an experience exists is not enough. Seemings do not ensure the needed commitment of the subject. The point is familiar: if it seems to me that one rod being longer than another, I need not be committed to its being longer.

undercutting defeat in the case of noninferential justification from experience. But, perhaps the second explanation, which omits reference to a taking condition, does carry over.

For the second explanation to extend to the noninferential case, we would have to be able to see how having reason to believe that being in experiential state E doesn't guarantee Q could, without higher-order help, make it epistemically inappropriate for you to make a noninferential transition from being in E to believing Q. Of course, with the right higher-order help, we can explain the inappropriateness of forming a belief. Supposing you knew that if you believe Q, it would be on the basis of experience E, you could reason like so:²⁰ *there is a good chance that I could have E without Q being the case; now, if I believe Q it would be on the basis of E; and so, if I believe Q there is a good chance I would be basing my belief on a source that is unreliable in the circumstances.* And if you (justifiably) believed the conclusion of this reasoning, this would explain why it would be epistemically inappropriate for you to form that belief. All this is with the higher-order help. But without it, how can we devise any reasoning, available to the subject, that concludes with something belief in which, or justified belief in which, makes the formation of the belief epistemically inappropriate? The conclusion *there is a good chance that I could have experience E without Q being the case* doesn't seem by itself to make forming a belief in Q (based on E) epistemically inappropriate, for one might have no idea that one would be relying on E. The best we can get is this:

If I rely on E in believing Q, my belief that Q has a good chance of being formed in a way that is unreliable in the circumstances.

But this by itself doesn't make the transition from experience E to believing Q epistemically inappropriate. Whereas in the inferential case a reason to doubt that P wouldn't be true unless Q seems by itself to make the transition from the belief that P to the belief that Q inappropriate, a reason to doubt that having experience E supports Q at most makes it inappropriate to transition from the *belief* that you have E to the belief that Q, but doesn't make it inappropriate to transition from the experience E itself to the belief that Q. The Milk Taster case exemplifies this contrast well. If the milk taster believed she was basing her belief on smell, it would be inappropriate for her

²⁰ I assume here that a doubt could appear in thought as an affirmation of a statement about chance. One might question this. However, these issues could be bypassed by thinking of a case in which the subject believes (justifiably) that it's not true that she wouldn't have E unless Q. We could then just load this negation into the reasoning.

to keep the belief; but if not, and especially if she believes (justifiably) that she isn't basing her belief on smell, it doesn't seem inappropriate.

One might balk here, and indeed balk at Sturgeon's (and my) judgments about the Milk Taster case. Isn't there something epistemically inappropriate about basing a belief on a source while believing that there is a good chance that that source is unreliable, even if you don't have believe it's your source? But just *what* is inappropriate about doing this? It's true that an ideal cognitive agent would do no such thing, but that is because an ideal cognitive agent would know its sources (Mellis 2014). It doesn't follow that we lack justification from the fact that we are non-ideal in this way. It might be claimed that the real problem is that you're conducting yourself in such a way as to guarantee that your resulting belief fails to be knowledge *if* your belief about your source's unreliability were true and you still believed on the basis of that source. If the milk taster's belief that olfaction is unreliable in the circumstances were true, and if she still based her belief about the milk on olfaction, it's true that her belief wouldn't be knowledge. But it's hard to see what the bearing of these facts is on whether the milk taster is in fact justified in believing the milk is ok.

Thus, while there may be senses in which the milk taster's belief is epistemically non-ideal and senses in which the belief would be problematic if the belief about the unreliability of olfaction were true, I don't think these defects are ones a theory of defeat needs to capture, because they do not eliminate or reduce justification.

9.5. A Tool for Distinguishing Inferential from Non-inferential Justification

The conclusions from the previous two sections imply that there is no such thing as Pollock-style undercutting defeat of noninferential justification, only of inferential justification. The nearest thing to genuine Pollockian *attack-the-connection* defeat of a noninferential justification is what I've called higher-order defeat: defeat in virtue of having reason to think that one's belief is unreliably based, a kind of defeat that requires beliefs about one's bases or sources and so about oneself and one's mind. If these conclusions are correct, we might use them to help us determine whether a justification in a certain case is inferential or not. This section considers the implications for perceptual and testimonial justification.

First, perception. Suppose there can be no undercutters for noninferential justification, only higher-order defeaters. And suppose you have noninferential justification from experience to believe that a wall is red. Then, if you learn that the lighting is abnormal, and this defeats your justification, it must be because this belief, together with a belief about the basis for your belief about the color of the wall, give you a reason to think that your belief is (or would be) unreliably formed. That is, the defeat occurs only because you have reason to think something about your own mind, about your belief and its formation. Or suppose you have noninferential justification from experience to believe that the thing before you is an eggplant, and then you learn that there are look-alike fake eggplants about. This, too, would serve as a defeater for your justification only if, together with your belief about the source of your belief, it gives you reason to think your belief is or would be unreliably formed. This is surprising. It is surprising that defeat in such a case should require beliefs about oneself.

Where you have a certain perceptual justification for believing P, borrowing terminology from Pryor (2004), let's say that a *non-perceiving possibility* for you for P is a possibility in which you have this same justification but you are not be in a position to perceive that P. Some such possibilities are objective: they are conditions of the world around that preclude you from perceiving that P solely by affecting the stimuli available to you. Objective non-perceiving possibilities include the abnormal lighting or the abundance of look-alike fakes. It is implausible that evidence for objective non-perceiving possibilities should require evidence about oneself. This is not so implausible in the case of defeat through learning of subjective non-perceiving possibilities such as hallucination or some defect of one's eyes. (I don't claim these categories exhaust non-perceiving possibilities.) It wouldn't be surprising to find out that the way evidence of the obtaining of these possibilities defeat is by helping give you reason to believe your belief is or would be unreliably formed.

To further confirm these claims, consider what it would take to reply adequately to a persistent challenger, when the two of you are standing in front of something that looks like an eggplant:

CHALLENGER: "Why don't you conclude it's an eggplant?"

YOU: "Because there are look-alike non-eggplants all about."

CHALLENGER: "What does there being look-alike non-eggplants all about have to do with it?"

YOU: "Because if there are, then my experience would be unreliable, and since if I believed it would be on the basis of experience, my belief would be unreliable."

It might be surprising that you start talking about *yourself* in the last line, but it's these facts about oneself that articulate the source of your justification for not concluding it's an eggplant. Contrast this final reply with a more natural one:

YOU: "Because if there are look-alike fakes around, then, even though this looks like an eggplant, it could easily be one of the fakes."

Similarly, in reply to a challenge about why you don't conclude the wall is red in a case of strange lighting, you might say:

YOU: "Because if the lighting conditions are abnormal, then, even though the wall looks red, it could easily be some other color."

These more natural final replies don't appeal to one's own mental states. The proposition that the thing looks like an eggplant does not entail anything about my mind, or arguably about other people's minds. It is a fact about eggplants that can survive my demise and that of the rest of humanity.

Is there a way to block the "mentalization" of such defeaters for perceptual justification? The classical foundationalist takes one's justification in the relevant cases to be inferential but still takes the justification to come from equally mentalized premises, about which experiences one has. But there is another sort of inferential approach available, suggested by the more natural replies to the previous Challenger character. Suppose the structure of my perceptual justification for believing *this is an F* (e.g., this is an eggplant) is as follows: I'm inferentially justified in believing this is an F in virtue of being justified in believing it looks like an F.²¹ If such an account is correct, information about rampant look-alikes can defeat my inferential justification for thinking a thing is an F without the assistance of higher-order contributions. The information about look-alike fakes can defeat in the way that genuine undercutters do, by being a reason to doubt a connection holds between one consideration—this thing has a certain look—and another—

²¹ There would need to be a story of how the visual experience contributes to the justification of the belief about the thing's looking like an F. Are the latter beliefs noninferentially justified, or are they themselves inferentially justified? Eventually, we must get down to beliefs that are not inferentially justified. For them, undercutting defeat would be impossible. I am sympathetic to the view that the noninferentially justified beliefs are beliefs about things having certain looks, and I suspect that such beliefs aren't subject to undercutting defeat, only higher-order and rebutting defeat of an object justifies me in believing it looks like an F, and this, perhaps together with some background information or a default assumption, inferentially justifies me in believing that it is an eggplant.

this thing is an eggplant. Such a doubt makes it epistemically inappropriate to make the inference. A similar explanation would hold for how information about lighting conditions defeats. Information about lighting condition spoils the connection between a thing's having a certain look that distinctive of red and a thing's being red, making the inference epistemically inappropriate.

There are various worries one might raise about such an account.²² What are these "objective looks"? Isn't the resulting view of perceptual justification too intellectualized? Relatedly, do we really go through these reasoning processes when we see an eggplant? I will not try to answer these objections here. I only note that this view has one point in favor of it against the standard view according to which one's perceptual justification for believing this is an F is noninferential. On that view, if the conclusions of this paper are right, there is no such thing as simple attack-the-connection defeat—Pollock-style undercutting defeat—of one's perceptual justification. Even objective-condition defeaters about lookalike fakes or lighting conditions only defeat because they join up with beliefs about one's beliefs' sources to generate a higher-order defeater. And that is implausible.

An easier case, though, is testimony. One familiar epistemological theory of testimony holds that being told that p provides noninferential justification for believing that p (Burge 1993; Graham 2006). If this view is correct, then information that one's testifier is insincere or incompetent cannot by itself defeat one's testimonial justification; rather one must build up a higher-order defeater such as "if I believe p, it would be on the basis of this testimony, which is unreliably in the circumstances, and therefore my belief would also be unreliable in the circumstances." But again, this seems wrongly to assimilate objective and subjective *non-learning* possibilities. Non-learning possibilities for me for P are possibilities in which I have the same P-related testimonial justification that I actually have but in which I'm not in a position to learn that P through this testimony. Non-learning possibilities such as insincerity or incompetence of a testifier are objective ones, whereas possibilities such as my now suffering a momentary bout of receptive aphasia are subjective ones. The subjective ones do seem to defeat only because they contribute to a higher-order defeater, but this doesn't seem true of the objective ones. Again, it doesn't seem that adequate replies to challenges invoking objective non-learning possibilities must ultimately appeal to considerations about *oneself and one's beliefs*. It seems adequate to point out, e.g., that if the testifier is being insincere then they could easily have told me that P despite not-P.

²² For some development of this sort of account, see McGrath (2017, 2018).

We can avoid these consequences if we take testimonial justification to be inferential. For then we can appeal to Pollockian undercutting. The information that the testifier is insincere gives you a reason to doubt that the testifier word guarantees the truth, and this is a reason not to infer its being so from the testifier's saying it.

9.6. Conclusion

Contrary to Sturgeon, I've argued that Pollock-style undercutting defeat—defeat via attacking-the-connection between a reason and the proposition believed—is possible in cases of inferential justification. But, I've agreed with him that there is no such thing in the case of noninferential justification, and that the nearest cousin to it is what I've called higher-order defeat, i.e., defeat by having reasons to believe the source of one's belief is unreliable. Finally, I've explored how we can use this difference between the ways inferential and noninferential justification can be defeated to make progress on questions on whether a certain sort of justification is inferential or not, my test cases being perceptual and testimonial justification.²³

References

- Audi, R. (1993). *The Structure of Justification*. Cambridge: Cambridge University Press.
- Balcerak Jackson, Magdalen and Brendan Balcerak Jackson. (2013). Reasoning as a source of justification. *Philosophical Studies*, 164: 113–26.
- Boghossian, Paul. (2014). What is inference? *Philosophical Studies*, 169: 1–18.
- Broome, John. (Forthcoming). Linking beliefs. In: Balcerak Jackson, Magdalen and Brendan Balcerak Jackson (eds.) *Reasoning*. Oxford: Oxford University Press.
- Burge, Tyler. (1993). Content preservation. *Philosophical Review*, 102(4): 457–88.
- Casullo, Albert. (2018). Pollock and Sturgeon on defeaters. *Synthese*, 195: 2897–906.
- Comesaña, Juan. (Forthcoming). Empirical justification and defeasibility. *Synthese*,

²³ I am grateful for comments from an audience at the University of Pavia, especially Tommaso Piazza, James Pryor, and Crispin Wright. For comments on written drafts, I thank Jessica Brown, Al Casullo, Juan Comesaña, and Jack Lyons.

- Graham, Peter J. (2006). Testimonial justification: Inferential or noninferential? *Philosophical Quarterly*, 56: 84–95.
- Hlöbil, Ulf. (2014). Against Boghossian, Wright and Broome on inference. *Philosophical Studies*, 167: 419–29.
- McGrath, Matthew. (2017). Knowing what things look like. *Philosophical Review*, 126(1): 1–41.
- McGrath, Matthew. (2018). Looks and perceptual justification. *Philosophy and Phenomenological Research*, 96(1): 110–33.
- Mellis, Giacomo. (2014). Understanding undermining defeat. *Philosophical Studies*, 170: 433–42.
- Pollock, John. (1970). The structure of epistemic justification. *American Philosophical Quarterly*, 4: 62–78.
- Pollock, John. (1974). *Knowledge and Justification*. Princeton: Princeton University Press.
- Pollock, John. (1987). Defeasible reasons. *Cognitive Science*, 11: 481–518.
- Pollock, John and Cruz, Joseph. (1999). *Contemporary Theories of Knowledge*. 2nd ed. Lanham, MD: Rowman and Littlefield.
- Pryor, James. (2004). What's wrong with Moore's argument. *Philosophical Perspectives*, 14: 249–378.
- Sturgeon, Scott. (2014). Pollock on defeasible reasons. *Philosophical Studies*, 169: 105–18.