

# Empirical Justification and Defeasibility

Juan Comesaña

## 1. Introduction

Two truisms about empirical justification are that experience plays a crucial role in it (hence the name) and that it is defeasible. There are, of course, different ways of developing these truisms into philosophical theories. I favor one particular view about the role of experience in empirical justification which may be thought to lead to problems in accommodating its defeasibility. My aim in this paper is to argue that the problems are illusory, based on an entrenched misconception how defeaters work.

## 2. Three Views About the Role of Experience in Empirical Justification

What is the role of experience in determining the rationality of our beliefs? One aspect of this question on which I will not spend much time is the issue of exactly which beliefs are the ones whose rationality is determined by our experience. I will not spend much time on that issue not because it is not intrinsically interesting, but rather because it is not the focus of disagreement between my view and its main competitors. I will in what follows simply assume that there are some propositions which can be the contents of experiences, and that the rationality of believing *those* propositions is somehow determined by an experience with that same content.<sup>1</sup> The question is, how does an experience with the content that *p* determine the rationality of a belief that *p*? From time to time, I will replace the generic *p* with specific propositions, but no deep theoretical commitment should be read in those examples. In this section I present three views about that question, and briefly indicate why I prefer one of them.

The first view is at least inspired by Williamson (2000), and I call it “Factualism.” According to Factualism, an experience that *p* makes a belief that *p* rational for a subject if and only if that experience provides the subject with knowledge that *p*. Thus, according to Factualism only some experiences with the content that *p* rationalize a belief that *p*. In particular, an experience with the content that *p* does not rationalize a belief that *p* when *p* is false—for, when *p* is false, an experience with the content that *p* cannot provide

---

<sup>1</sup>I'm talking about the representational content of experiences. I am neutral about the relationship between representational and phenomenal content.

knowledge that  $p$ . Thus, for the Factualist, we can divide the experiences with the content that  $p$  into the good ones and the bad ones, according to whether they provide the subject whose experience it is with knowledge that  $p$  or not. All experiences that  $p$  had when  $p$  is false will thus be bad experiences in this sense (although they may be perfectly good experiences in other senses of the term).

The second view is more traditional. Williamson calls it “the Phenomenal Conception of Evidence” (PCE).<sup>2</sup> According to PCE, an experience that  $p$  is itself evidence that  $p$  is true, and thus always provides some reason to believe that  $p$ . If that reason is not defeated (more on defeaters later on, of course), then the experience rationalizes a belief that  $p$ .

The metaphysics of PCE is not clear in the literature. There are several options about what basic perceptual reasons are, according to PCE. Suppose that I have an experience  $e$ . What evidence do I thereby have? Here are three salient options:  $e$  itself; the fact that I have  $e$ ; my having  $e$ . One is an object (or, at any rate, a state), another a fact, and the third an event.

There are a number of commonalities and differences between Factualism and PCE. Factualism and PCE agree that, when experience provides us with reasons, those reasons are true. For the Factualist, the reason provided by experience is the content of the experience, and it is provided only when that content is known (and, therefore, true). For PCE, the reason is the experience itself, and it is therefore always so provided when the experience exists. Thus, proponents of PCE and Factualists disagree about the metaphysics of basic empirical reasons—Factualists think they are known propositions, whereas proponents of PCE think that they are experiences—but they agree in rejecting the idea that experience can provide us false reasons. Another important difference between PCE and Factualism is that although both positions treat experience as providing non-inferential justification for beliefs, in that they do not conceive of experience doxastically, PCE treats experience as providing *evidential* justification. The experience itself is evidence for a belief in its content, according to PCE, whereas according to the Factualist it is the content itself which is the evidence. To put it in rough and ready terms, for the Factualist but not for the proponent of PCE it is possible for the proposition which is the content of the experience to be the first item of evidence that the subject ever receives.

Another crucial difference between PCE and Factualism is in their treatment of cases where a subject has an experience that  $p$  while the proposition that  $p$  is false. In that case, according to the Factualist no reason is provided for the subject to believe that  $p$ , so that even if he has no defeaters for the belief that  $p$  he will not be rational in believing that  $p$ . According to the proponent of PCE, on the other hand, the experience itself is still a reason for the subject to believe that  $p$  even when its content is false, and so if the reason thus provided is not defeated then the subject is rational in believing that  $p$  even when  $p$  is false. PCE is thus friendly to the view that there can be rationality false beliefs, whereas Factualism is incompatible with it.

---

<sup>2</sup>It is not easy to classify authors as clearly defending this position. Perhaps Conee and Feldman (1985) and Pollock (1986) come closest. Sometimes Pryor (2000) is also associated with the view, but it seems to me that Pryor’s views are neutral between the Phenomenal Conception of Evidence and my own Experientialism, to be introduced momentarily.

The view that I favor, which I call “Experientialism,”<sup>3</sup> agrees with Factualism in claiming that experience provides us with reasons but is not identical to the reasons provided, but agrees with PCE in claiming that even bad experiences can rationalize belief. One way of thinking of Experientialism is as a generalization of Factualism. Factualism holds that experiences provide their content as a reason for belief only when the belief counts as knowledge; Experientialism holds that experiences *always* provide their content as a reason for belief. Thus, Experientialism holds not only that experience can rationalize false beliefs (in this, Experientialism sides with PCE and against Factualism), but also that experience can provide us with *false evidence*. For Experientialism as well as for Factualism, the proposition which is the content of the experience can be the first item of evidence ever received by a subject, but, as opposed to Factualism, that item can be false. Thus, Experientialism agrees with Factualism about the metaphysics of reasons, but disagrees with it about the epistemology.

The views presented so far can be summarized in the following table:

	What are basic reasons in the good case about?	Are basic reasons the same in the good and the bad case?	Can basic reasons be false?
Factualism	Content	No	No
PCE	Experience	Yes	No
Experientialism	Content	Yes	Yes

Let me illustrate the table with a pair of examples:

Good tomato: I open the refrigerator and have an experience with the content that there is a tomato in it. The experience is in fact veridical, and everything else about the case is normal and I acquire knowledge that there is a tomato in the refrigerator.

Bad tomato: I open the refrigerator and have an experience with the content that there is a tomato in it. This time, however, my experience is not veridical, for what I am looking at is a paper mache replica of a tomato, which found its way into my refrigerator by a series of unfortunate (and very unlikely) events.

According to both Factualism and Experientialism, that there is a tomato in the refrigerator is one of my basic reasons in Good tomato. According to PCE, by contrast, my basic reason is that I have an experience with the content that there is a tomato in the refrigerator. Notice that, in attributing this view to PCE, I have forced an interpretation of it according to which basic reasons are propositions—propositions about experience—rather than experiences. Thus, in this respect Experientialism agrees with Factualism and disagrees with PCE. On the other hand, both PCE and Experientialism have it that my basic reason is the same in Good tomato and Bad tomato, whereas Factualism disagree. Thus, to put it in a quick and dirty way, Experientialism sides with Factualism over PCE regarding the good case but sides with PCE over Factualism regarding the bad case. However, Experientialism is on its own when it comes to whether basic reasons

<sup>3</sup>In Comesaña and McGrath (2016) we called it “Propositionalism”. I don’t much like either name.

themselves can be false. Both PCE and Factualism have it that they cannot, whereas Experientialism has it that they can.

So far I have presented three views about basic perceptual beliefs. A fourth view is also suggested by Williamson. Considering a case of deception, he says:

In unfavorable circumstances, one fails to gain perceptual knowledge, perhaps because things are not the way they appear to be. One does not know that things are that way, and  $E = K$  excludes the proposition that they are as evidence. Nevertheless, one still has perceptual evidence, even if the propositions it supports are false. True propositions can make a false proposition probable, as when someone is skillfully framed for a crime of which she is innocent. If perceptual evidence in the case of illusions consists of true propositions, what are they? The obvious answer is: the proposition that things appear to be that way. Of course, unless one has reason to suspect that circumstances are unfavorable, one may not consider the cautious proposition that things appear to be that way; one may consider only the unqualified proposition that they really are that way. But it does not follow that one does not know that things appear to be that way, for one knows many propositions without considering them. When one is walking, one normally knows that one is walking, without considering the proposition. Knowing is a state, not an activity. In that sense, one can know without considering that things appear to be some way<sup>4</sup>

The view suggested in that passage is not PCE: the idea is not that the experience itself is evidence for its content, but rather that knowledge (albeit unconsidered knowledge) of the experience serves as an inferential base for belief in its content. The view is a version of Classical Foundationalism, according to which beliefs about one's experience are what justify beliefs about the external world. Williamson's complete view is therefore disjunctivist: Factualism about the good cases and Classical Foundationalism about the bad cases. There are different varieties of Classical Foundationalism: the Cartesian variety holds that beliefs about our own experiences are infallible, and that only deduction is allowed as a provider of inferential knowledge. A more modest variety lifts the infallibility claim and allows induction as well as deduction to aid in the acquisition of inferential justification. Williamson's variety seems to be in between these two: infallibilism about the basic beliefs (because they consist of known propositions), but with ampliative inferences allowed.

Classical Foundationalism has not been getting a lot of attention lately. One fundamental objection to it is that it seems to lead to a pretty severe form of skepticism. Suppose we grant that we do in every case have rational beliefs, and even knowledge, about our own experiences—something which, Williamson's appeal to implicit knowledge notwithstanding, I am not so sure we should grant. Still, the vast majority of us (and all of us the vast majority of the time) do not form our beliefs about external objects on the basis of those beliefs about our experience of them. And if it is only beliefs about our experiences of external objects which justify beliefs about them, then most of our beliefs about

---

<sup>4</sup>Williamson (2000), p. 198.

them will not be *ex-post* justified (they will not be based on the evidence which justifies them).<sup>5</sup> Given that knowledge requires *ex-post* justified beliefs, we know almost nothing about external objects. Williamson's combination of Factualism about the good case and Classical Foundationalism about the bad case is particularly puzzling. For this combination means that in order to be rational a subject would have to guess whether he is in the good case or the bad case, and then form his beliefs accordingly. For these reasons, I will not be further discussing Classical Foundationalism, but rather concentrate on PCE, Factualism, and Experientialism.

Going back to the three main views presented in this section, there is much to be said about each one of them. I myself favor Experientialism, and have argued for it elsewhere.<sup>6</sup> I will not rehearse those arguments here. Rather, I turn now to the objection to Experientialism that it cannot handle the defeasibility of empirical justification.

### 3. The Problem of Defeaters

The objection in question can be put simply, although unravelling it in all its detail will take some time. Suppose that I have an experience with the content that there is a red wall in front of me. If everything goes well, then everyone agrees that I can thereby come to be justified in believing, and even know, that there is a red wall in front of me. But suppose now that, in addition to having that experience, I am also justified in believing that there are red lights shining on the wall. In this case, the traditional thought is, I am no longer justified in believing that there is a red wall in front of me.

Now, some versions of PCE have been construed explicitly to take into account this kind of defeaters. It is therefore thought that PCE can account for the defeasibility of empirical justification in unproblematic ways. We'll see below that matters are not so straightforward. But whatever happens with PCE, the objection in question is that Experientialism cannot deal with that example. For remember that for Experientialism the basic reason I have is that there is a red wall in front of me, and the information that there are red lights shining on the wall is irrelevant to that proposition. Therefore, Experientialism has it that I continue to be justified in believing that there is a red wall in front of me even when I am justified in believing that the wall is bathed in red light.

Spoiler alert: my response to this objection will be that indeed Experientialism has the consequence that I am justified in believing that the wall in front of me is red even when I am justified in believing that the wall is bathed in red light, but I hold that that is the right consequence. This example is typically taken as paradigmatic of one kind of defeat, called by Pollock "undercutting defeat". I will argue that it is by no means obvious that there is such a thing as undercutting defeat, at least as Pollock conceived of it.

---

<sup>5</sup>The distinction between *ex-post* and *ex-ante* justification is also sometimes referred to as the distinction between propositional and doxastic justification. When talking about *ex-post* justification, the bearer of justification is the mental act of believing a proposition, whereas propositions themselves are the things which are *ex-ante* justified for a subject.

<sup>6</sup>See Comesaña and McGrath (2014), Comesaña and McGrath (2016), and Comesaña (2016).

### 3.1 Pollock on Reasons and Defeaters

Pollock thinks that all reasons are mental states—some are beliefs and some are experiences. The primitive in Pollock’s system is the notion of *possible ex-post justification*. More precisely, it is the notion that *it is logically possible for a subject  $S$  to be justified in believing a proposition  $P$  on the basis of mental state  $M$* . In terms of this primitive, Pollock defines the notion of a reason:<sup>7</sup>

**Reason:** 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 state  $M$ .

An understandable question at this point is: what does Pollock mean by saying that it is (logically) possible for  $S$  to be justified in believing  $Q$  on the basis of  $M$ ? Does he mean on the basis of  $M$  alone, or on the basis of  $M$  together perhaps with some other mental states? The answer *has* to be “on the basis of  $M$  alone”, because if we allow  $M$  to interact with other mental states, this overgenerates justification implausibly. For assume that  $M$  is a (perhaps conjunctive) justified belief. Is it logically possible for a subject  $S$  to be justified in believing an arbitrary  $Q$  on the basis of  $M$ , together with other mental states? Except perhaps for the cases where  $Q$  is incompatible with  $M$ , the answer will be “Yes”, for we can always add to  $M$  the justified belief *If  $M$ , then  $Q$* , and it is of course possible to be justified in believing  $Q$  on the basis of  $M$  plus this other mental state. So, Pollock must mean that for  $M$  to be a reason for  $S$  to believe  $Q$  it must be possible for  $S$  to be justified in believing  $Q$  on the basis of  $M$  alone. This atomism of Pollock’s system is not particularly plausible, but I won’t press this objection here.

There is another way in which Pollock’s definition of a reason badly overgenerates. Even if one thinks that it is beliefs themselves which justify, rather than their contents, it should only be *justified* beliefs which justify. This is the venerable principle of inferential justification.<sup>8</sup> But Pollock’s official definition of a reason allows unjustified beliefs to be reasons. For of course it is logically possible for an unjustified belief to be justified. Take, then, an unjustified belief, and plug it in for  $M$  in Pollock’s definition. For instance, to use an example from Pollock himself to which we will return, suppose that  $M$  is the belief that 87% of a random sample of voters from Indianapolis intend to vote Republican in the next election. On the basis of this belief, the subject concludes that 87% of the voters in Indianapolis will vote Republican in the next election. But suppose that the subject in question has the initial belief out of wishful thinking, and not on the basis of anything that remotely justifies it. It is still obviously logically possible for the subject to be justified in believing that 87% of the voters in Indianapolis will vote Republican in the next election, for it is logically possible for  $M$  to be justified for the subject (even if it actually isn’t).

I foresee two ways for Pollock to solve this issue. One would be to hold that beliefs have their epistemic status essentially—i.e., that an unjustified belief is just a different state

---

<sup>7</sup>I take the formulations from Pollock and Cruz (1999). A previous formulation can be found in Pollock (1970).

<sup>8</sup>But cf. Schroeder (2007), who argues that all beliefs justify, but unjustified beliefs are guaranteed to have their justificatory power defeated.

from a justified one. This move seems obviously ad-hoc to me. Another solution is to give an explicitly recursive definition of a reason, where a mental state is a reason if and only if it is either an experience or it is justified by a reason (where justification is defined as having ultimately undefeated reasons). It is a delicate matter how exactly to formulate the recursive definition, but given that I think it is pretty clear that this last option is the one that better fits the spirit of Pollock's theory, I will proceed as if this is the view.

Pollock then defines the general notion of a *defeater*:

**Defeater:** 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 of being in both the state  $M$  and the state  $M^*$  at the same time is not a reason for  $S$  to believe  $Q$ .

Unpacking the definition of reason gives us the following:

**Defeater, unpacked:** If it is logically possible for  $S$  to become justified in believing  $Q$  by believing it on the basis of being in state  $M$ , then state  $M^*$  is a *defeater* for this reason if and only if it is not logically possible for  $S$  to become justified in believing  $Q$  by believing it on the basis of being in the combined state  $M$  and  $M^*$ .

The definition is somewhat awkwardly stated in a conditional form, but what Pollock is really doing here is defining the notion of  $M^*$ 's *being a defeater for  $M$  as a reason for  $S$  to believe  $Q$* . For that to happen,  $M$  has to be a reason for  $S$  to believe  $Q$ , and the combination of  $M$  with the defeater state not be a reason for  $S$  to believe  $Q$ . The fact that what Pollock is defining is this four-place relation will become important to our discussion.

Famously, Pollock distinguished two kinds of defeaters. One of them, he says appealing to the analogy between reasons and arguments we used before, "is a reason for denying the conclusion" (Pollock and Cruz (1999), p. 196):

**Rebutting defeater:** If  $M$  is a 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  $\neg Q$ .

Thus, *believing that Lucy is in her office and that most of the times when she is in her office Lucy turns on the lights ( $P$ )* is a reason for believing that *the lights in Lucy's office are on ( $C$ )*, but believing that *Joe was just outside Lucy's office and he reports seeing through the window that the lights were off ( $D$ )* is a rebutting defeater for our belief in  $P$  as a reason to believe  $C$ . For we cannot become justified in believing  $C$  by believing it on the basis of our beliefs that  $P$  and that  $D$ , and belief in  $D$  is by itself a reason to believe  $\neg C$ .

The second kind of defeater discussed by Pollock "attacks the connection between the premises and the conclusion rather than the conclusion itself" (Pollock and Cruz (1999), *ibid.*). Here we must pause to notice the awkwardness in Pollock's theory which results from treating all reasons as mental states. The example he gives of a rebutting defeater is that of a pollster attempting to predict what proportion of residents of Indianapolis will vote for the Republican candidate in the upcoming election:

She randomly selects a sample of voters and determines that 87% of those polled intend to vote Republican. *This* gives her a defeasible reason for thinking that approximately 87% of all Indianapolis voters will vote Republican. But then it is discovered that purely by chance, her randomly chosen sample turned out to consist exclusively of voters with incomes of over \$100,000 a year. *This* constitutes a defeater for the inductive reasoning... (Pollock and Cruz (1999), *ibid.*, my emphases).

Notice the unclarity about the referent of the two emphasized appearances of ‘this’. What is the initial reason, and what is the defeater? The official story must be that they are mental states of the pollster, but the natural way to read the sentences is as referring to the *propositions* that 87% of those polled will vote Republican and that all those polled have incomes of over \$100,000 a year. And it is not just the natural interpretation of those sentences. After the passage quoted Pollock gives the following definition:

**Undercutting defeater:** 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.

There is a lot to go through in that definition. First, notice that Pollock sticks to the official line that reasons are mental states when it comes to the initial reason and the defeater. The initial reason is *believing that  $P$* , and the defeater is some unspecified mental state  $M^*$ . Instantiating the definition to Pollock’s own case, the initial reason is the Pollster’s belief that 87% of those polled will vote Republican, and  $M^*$  is presumably the pollster’s belief that all of those polled have incomes of over \$100,000 a year. Notice also that the choice for the initial reason is a poor one. As I already noted, presumably an *unjustified* belief that 87% of those polled will vote Republican is not a reason to believe that 87% of the relevant population will vote Republican—but bracket that issue. Another remarkable feature of the definition is that undercutting defeaters are reasons to doubt or deny a counterfactual conditional. The obvious way of defining an undercutting defeater (setting aside for a moment the problem to be presented in the next paragraph) would have been the following:  $M^*$  is an undercutting defeater for  $M$  as a reason for  $S$  to believe  $Q$  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 think that  $M$  is not a reason for  $S$  to believe  $Q$ . Presumably, Pollock appealed to the counterfactual instead of to this explicitly epistemic definition because he thought the latter definition “too intellectualistic”, in that it would only apply to subjects who themselves have the notion of a reason. I come back to this important issue later in this section.

The most remarkable feature of the definition of an undercutting defeater, however, is that it just does not fit the informal gloss of the notion given by Pollock himself before the definition. To fit the informal characterization of an undercutting defeater as “attacking the connection between premises and conclusion”,  $M^*$  would have to be a reason for thinking that the premises would not be true unless the conclusion were true. Both the premises and the conclusion, sticking to the official story, are beliefs of the pollster. Therefore, the last clause of the definition would have to be “ $M^*$  is a reason for  $S$  to doubt or deny that  $S$  would believe  $P$  unless  $S$  believed  $Q$ .” But, of course, that clause

wouldn't make sense. What  $S$  would or wouldn't believe under different conditions is irrelevant to whether he has a defeater. Hence the weirdly disjointed definition, which sticks to the official story for the original reason and the defeater, but defects to thinking of reasons as the propositions believed, rather than the believing of them, in the last clause.

Without batting an eye, immediately after giving that definition Pollock says: "When the reason is a nondoxastic state, we must define undercutting defeat slightly differently", and gives the following definition:

**Nondoxastic undercutting defeat:** 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.

*This* is the real reason Pollock takes reasons to be states rather than their contents, and the reason for all the awkwardness in his definitions I have been pointing to: to take care of the case of justification by experience and its defeat. This difference with the doxastic case deserves extended discussion rather than what amounts to barely an acknowledgement that there is a difference. I take these infelicities in Pollock's system to be an indirect argument in favor of my own Experientialism, which treats the doxastic and the experiential case alike.

But, as anticipated, some people think that Experientialism cannot account for undercutting defeaters.<sup>9</sup> I turn to that objection next.

### 3.2 Seeking a Formulation of the Defeaters Objection

Ultimately, as I will argue, there is no coherent formulation of the defeaters objection to Experientialism. But the idea is the following. According to Experientialism, when I have an experience with the content that there is a red wall in front of me, I thereby have as a reason the proposition that there is a red wall in front of me. Now suppose that I am justified in believing that there are red lights shining on the wall. Is that belief of mine an undercutting defeater? The objection is that it isn't, but it should be. It should be because, intuitively, I am no longer justified in believing that there is a red wall in front of me. It isn't because the fact that there are red lights shining on the wall is irrelevant to its being a red wall.

But recall that, according to Pollock's definition, defeat is a four-place relation:  $M^*$  is a defeater for  $M$  as a reason for  $S$  to believe that  $Q$ . Taking this into account, how exactly should the defeater objection be formulated? The claim is that my belief that there are red lights shining on the wall is not, according to Experientialism, an undercutting defeater. But there is no such thing as a free-floating undercutting defeater. When the objector claims that it is not an undercutting defeater we must ask: what are the  $M$  and the  $Q$

---

<sup>9</sup>I first heard the objection from Stew Cohen in conversation. Mark Schroeder presents a version of it in Schroeder (forthcoming).

such that my belief that there are red lights shining on the wall is not a defeater for  $M$  as a reason for me to believe  $Q$ ? The only plausible  $Q$  in the vicinity is that the wall is red. So the claim must be that there is some  $M$  such that my belief that there are red lights shining on the wall is not an undercutting defeater for  $M$  as a reason for me to believe that the wall is red, although it should be. But there is no such  $M$ . According to Experientialism, that the wall is red is a basically justified belief, where this means not only that it is not justified on the basis of any other *belief*, but also that it is not justified on the basis of any other *evidence* or *reason*. Reasons start with the proposition that the wall is red, according to Experientialism. Of course, this doesn't mean that those beliefs are not justified, just that they are not justified by evidence.

To see the problem from a different angle, notice that Pollock's version of PCE also posits basic reasons in the sense just characterized. That is to say, according to Pollock there are some reasons, namely experiences, such that they are not justified on the basis of anything else—not beliefs, certainly, but also not any other mental state. In this case, these reasons themselves are not justified, but not because they are unjustified but rather because they do not admit of justification. Perhaps what is driving Pollock and other proponents of PCE is the idea that if a state admits of justification it must be justified *by something*—by experiences if not by beliefs. Proponents of Experientialism such as myself can happily accept this, of course. Where proponents of PCE go wrong, from the perspective of Experientialism, is in insisting that this kind of justification is just like inferential justification, in that it is justification by evidence. But regardless of these differences between PCE and Experientialism, they agree that some reasons are themselves not justified by any other evidence. Take now those states which are basic reasons for PCE—experiences. Can they be defeated? Not according to the official definitions of defeaters. For the defeating relation is four-place, and there is no state to play the role of reason for the basic reasons. There are no reasons for the basic reasons—that is what their basicity consists in.

Consider, for instance, the following objection. Suppose that, as before, I have an experience as of a red wall in front of me. Suppose now that my usually reliable friend tells me that my coffee has been spiked with a drug whose effect is to make it seem to me as if I have experiences as of a red wall in front of me when I don't. Suppose also that, although my friend is right about the drug, in this particular case I do have a genuine experience as of a red wall in front of me. In that case, it may be natural to suppose that my experience has been defeated. But, of course, Pollock cannot accept this, for the same structural reasons that makes it hard to formulate the defeaters objection to Experientialism. That is to say, given that defeat is a four-place relation, and given in particular that only states as reasons for believing propositions can be defeated, my experience itself cannot be defeated, for its role in Pollock's system is not that of a proposition belief in which is justified by some other mental state. Of course, my experience as a reason for believing some proposition can be defeated, but that is not the intended target here. Just as the defeaters objection to Experientialism has it that what Experientialism takes to be basic reasons can be defeated, so too here the idea is that what PCE takes to be basic reasons can be defeated. In both cases, the fact that the defeating relation is a four-place relation makes even the formulation of the objection hard to pin down.

Perhaps the objection comes down to no more than this. Avoiding talk of reasons and defeaters, we can describe the case in question as follows: I have an experience as of a red wall in front of me, and I am also justified in believing that the wall is bathed in red light. Of that case as described, we should say that I am not justified in believing that there is a red wall in front of me. PCE has the resources to deliver this verdict—for it can say that, although my experience does provide a reason for believing that there is a red wall in front of me, my belief that wall is bathed in red lights is an undercutting defeater for my experience as a reason to believe that there is a red wall in front of me. Experientialism, by contrast, has the consequence that I am justified in believing that there is a red wall in front of me. Talk of undercutting defeaters and basic reasons simply obscure what the basic objection is: it is simply that the case is a counterexample to Experientialism.

In what follows I explore two different answers to the objection formulated in those terms. One is concessive: it grants that in the case as described I am not justified in believing that the wall is red, and also grants that Experientialism as so far formulated has the consequence that I am, but argues that there is a fix for Experientialism. The other answer is dismissive: it holds that Experientialism is fine as it is, for I am justified in believing that the wall is red in the case as described. My inclination is to go with the second answer, but I develop both.

### 3.3 Undercutting Defeaters for Experientialism

Experientialism and PCE are not too far apart from each other, as I have already had occasion to note. Both views have it that experience plays a crucial role in the justification of beliefs about external objects, of course, but they also agree that those beliefs can be rational even if false. They differ on how best to describe the role of experiences in the justification of those beliefs. PCE has it that experiences themselves are evidence for the beliefs, whereas Experientialism has it that experiences provide those beliefs as reasons, but they are not themselves reasons for the beliefs.

We can use this structural similarity to formulate a version of Experientialism which is able to deliver the verdict that I am not justified in believing that there is a red wall in front of me in the case in question. Just as PCE has it that experiences are defeasible reasons, the defender of Experientialism can hold that experiences provide their content as reasons defeasibly.

To formulate the revised version of Experientialism explicitly, let us first pause to notice the distinction between *prima facie* and *all things considered* justification at the heart of Pollock's version of PCE. An experience is a reason for a belief in its content, but the justification that such reason provides is *prima facie* in that it can be defeated. But the relation between *prima facie* reasons and defeaters is not simple. It is not the case, for instance, that if an experience is defeated then it doesn't justify a belief in its content. This is because the defeater for the experience can itself be defeated. And this defeater defeater can itself be defeated, and so on and so forth. So, it's possible for a defeated experience to provide all things considered justification for a belief in its content, provided that it is *ultimately* undefeated. It is not straightforward to define the notion of

*ultimately undefeated reason*. Pollock does it in terms of graph theory, where defeaters cut reasons relations between reasons and beliefs, and when a defeater is defeated the reason it cut is restored—and further defeaters flip the cuts and restorations effected by the defeater they defeat. An experience then provides all things considered justification for a belief in its content provided that, taking into account all of the mental states of the agent in question, the link between the experience and the belief is present.

Experiences, according to Pollock, provide *prima facie* justification by being reasons for belief in their contents. Experientialism can also hold that experiences provide *prima facie* justification for beliefs in their content, although it must hold that they do not do so in virtue of being reasons for such beliefs, but rather in virtue of providing those contents as reasons. The idea is that experiences *presumptively* provide their contents as basic reasons. That presumption can be defeated, either by rebutters or by undercutters. The notion of an undercutting defeater is the same as that used in Pollock's view: a reason to doubt that I would not have the experience unless its content were true. An experience provides its content as a reason *all things considered* just in case it is ultimately undefeated.

To exemplify with the case of the wall, my experience as of a red wall in front of me presumptively provides the proposition that there is a red wall in front of me as a basic reason. But my belief that the wall is bathed in red lights defeats this presumptive reason-giving power, and so I am no longer justified in believing that the wall in front of me is red. A further modification to Pollock's theory can be made to make it more Experientialism-friendly, where the defeater is not a belief but its content, and it is had as a defeater only if the subject is justified in believing that content.

### 3.4 There are no Undercutting Defeaters for Experience

Recall how we got here: the objection was that, if I believe that the wall is bathed in red light my experience as of a red wall in front of me does not justify me in believing that there is a red wall in front of me. In previous section I played nice, and explained how to fix Experientialism to have this consequence. But I don't really think that Experientialism needs fixing, because I think that it is perfectly possible for my experience as of a red wall in front of me to justify me in believing that there is a red wall in front of me even if I am also justified in believing that the wall is bathed in red light. And even in those case where I am indeed unjustified in believing that there is a red wall in front of me, it is not because of undercutting defeaters understood *à la* Pollock.

Notice the difference between rebutting and undercutting defeaters. Rebutting defeaters work in a simple way, which can be explained by appealing to mathematical or physical analogies. There are reasons I have to believe a proposition  $P$ , and there are reasons I have to disbelieve  $P$  itself. We can think of the force these reasons have as being given by a vector. Whether I am all things considered justified in believing  $P$  can then be thought of simply as a result of adding these vectors together: if the result is positive then I am justified in believing  $P$ , otherwise I am not. Or we can think of the reasons as weights on a scale, and I am justified in believing  $P$  if the scale tilts to the  $P$ -side, otherwise I am not.

Undercutting defeaters are different. They do not push or pull in the *P*-direction—rather they work off-stage to interfere with the reason-giving power of the reasons themselves.

It is essential to PCE that my experiences can be my evidence even if I do not believe that I have them. This is what distinguishes PCE from Classical Foundationalism. Suppose, then, that I do not believe that I have an experience as of a red wall in front of me. Indeed, I do not even know what experiences are supposed to be.<sup>10</sup> Indeed, suppose that I am not aware of which sense modality it is that gives me justification for believing that there is a red wall in front of me. In that case, the information that the wall is bathed in red light may well be interesting, but, as far as I am concerned, it is irrelevant to whether the wall itself is red. Of course, normally, if the question arises and you know that there is a red wall in front of you by looking, you will also know that you are seeing that there is a red wall in front of you—and the same goes for the other sensory modalities. In that case, having reason to think that you didn't actually see that there is a red wall in front of you will indeed defeat the rationality of your belief that there is, but that will be because you have an additional belief as to where your belief came from. This additional belief will be, from the point of view of your rationally believing that there is a red wall in front of you, superfluous.

Indeed, if the situation is not normal, then you might have no belief whatsoever on whether you see (or have an experience) that there is a red wall in front of you, you might suspend judgment on whether you see that there is a red wall in front of you, or you might even disbelieve that you see that there is a red wall in front of you, all of which are compatible with your still being rational in believing that there is a red wall in front of you in virtue of the fact that you have an experience that there is a red wall in front of you. If adopting any attitude whatsoever towards the proposition that you see that there is a red wall in front of you involves having the concept of *seeing* (and adopting any attitude whatsoever towards the proposition that you have an experience that there is a red wall in front of you involves having the concept of an experience), then of course there are subjects who are rational in believing that there is a red wall in front of them without having any attitude whatsoever towards those other propositions—namely, all those subjects, such as infants and animals, who lack the required concepts. Even if you do have the relevant concepts, you might be unaware of the origin of your belief, and thus suspend judgment, or even disbelieve, propositions as to where the belief comes from. These will be abnormal circumstances, of course, but they will not rob your belief of its rationality.

Consider the phenomenon of forgotten evidence, where I know a certain fact but I forget how I learned about it. Some people might take this to be a counterexample to evidentialism (because you are now rational in believing a proposition for which you now have no evidence), and some other people might insist that you have evidence for the proposition. But even for these latter philosophers, the evidence in question will not take the form of specific hypotheses about the source of the belief, but more generic hypotheses such as the claim that you learned it in school or you read it somewhere. Suppose, for instance, that you see that there is a beetle in my box, but later you forget whether you know that

---

<sup>10</sup>It is relevant here that there are philosophers who do not think that there are experiences in the sense we posit here.

there is a beetle in my box because you saw it or because you touched it. Provided that I agree that it is still rational for you to believe that there is a beetle in my box even when you forget how you know it, having reason to believe that (say) you didn't see it will not, in general, make it irrational for you to continue to believe that there is a beetle in my box. In this case, it will just make it rational for you to believe that you learned that there is a beetle in my box by touching it. It doesn't matter, of course, how fast you forget the origin of your experience—indeed, it doesn't really matter whether you ever knew it. The “forgetting” can be instantaneous and we would still credit you with rational belief—after all, having a good memory cannot be a condition on acquiring empirical knowledge.

In that case, you do have some quite precise beliefs as to the origins of your belief: you think that you either saw or touched the beetle. But it is also conceivable that you have absolutely no idea as to the origin of your belief, other than a very generic one that you somehow learned it. In that case there will be even less pressure to think that learning that you didn't see a beetle in the box makes it less rational for you to believe that there is a beetle in the box. Notice that, in both of these cases, the Experientialist will say that it is still your visual experience that there is a beetle in the box that gives you the proposition that there is a beetle in the box as a basic empirical reason, even if you do not have a particularly high credence in the proposition that you had such a visual experience.

Sturgeon (2014) argues against Pollock's notion of an undercutting defeater very similarly. He gives the following example: Suppose that Susan one day tastes some milk and comes to the conclusion that it is ok. In fact, the basis for Susan's belief about the milk is a complex gustatory-olfactory experience. Susan, however, thinks that taste has nothing to do with smell. Now Susan is told that her nose is plugged (this is in fact false, but Susan is justified in believing that it is true). This fits the definition of Pollock's notion of an undercutting defeater: that Susan's nose is plugged is a reason to doubt or deny that she would have the complex gustatory-olfactory experience only if the milk were ok. But, Sturgeon suggests, given Susan's belief that taste has nothing to do with smell (a belief which may well be justified), her justification for believing that the milk is ok is unaffected by her belief that her nose is plugged.

I repeat that it might well be that, in the normal case, I do have beliefs about the origin of my beliefs. I do not want to commit here one way or the other on that point. If I do normally have beliefs about the origin of my beliefs, for instance if I believe that I believe that the wall is red because I am seeing it, then of course the information that the wall is bathed in red lights will make it that case that I am no longer justified in believing that there is a red wall in front of me. But the explanation for my lack of justification in this case will not be Pollock's. For the explanation in this case appeals to some version of the plausible principle that if I believe that  $P$  and I justifiedly believe that my belief that  $P$  is based on the operation of faculty  $F$ , then if I am also justified in believing that the situation is not propitious for the functioning of  $F$ , then my belief that  $P$  is unjustified. If I have beliefs about the origin of my beliefs, then I better be justified in also believing that that origin is reliable. Thus, when I believe that my belief that the wall is red is based on vision, my belief that the wall is bathed in red lights defeats my justification for believing that the wall is red only because telling colors by vision under unusual lightning conditions is unreliable. According to Pollock, on the other hand, my belief that the wall

is bathed in red lights can *by itself* defeat the justification I have for believing that the wall is red, regardless of whether I have beliefs about the origin of my first-order beliefs.<sup>11</sup>

To recap: I have argued that there are two different cases of basic empirical justification. In one kind of case, the subject has beliefs not just about the external world, but also about the provenance of these beliefs about the external world. Gaining information that the source of these beliefs is unreliable will then of course defeat the subject's justification for the first-order beliefs in an ordinary way. In the second kind of case, the subject has beliefs about the external world unaccompanied by beliefs about their provenance. Information relevant to the reliability of the source of these beliefs will in this kind of case be irrelevant to their justification.

## 4. Conclusion

My aim in this paper has been to clear one obstacle in the way of Experientialism. The obstacle is the claim that the view cannot adequately deal with the defeasibility of experiential justification. My argument has been that the objection rests on a particular understanding of the defeasibility of empirical justification, in particular it rests on Pollock's notion of an undercutter defeater. In reply I have first pointed out some infelicities of Pollock's system, and then argued that Pollock's notion of an undercutting defeater is itself subject to a serious objections. Once the role of higher-order beliefs in so-called undercutting defeat is clarified, it is clear that Experientialism can indeed account for the defeasibility of empirical justification.<sup>12</sup>

## References

- Comesaña, Juan. 2016. "Normative Requirements and Contrary-to-Duty Obligations." *Journal of Philosophy* 112 (11): 600–626.
- Comesaña, Juan, and Matthew McGrath. 2014. "Having False Reasons." In *Epistemic Norms*, edited by Clayton Littlejohn and John Turri. Oxford University Press.
- . 2016. "Perceptual Reasons." *Philosophical Studies* 173 (4): 991–1006.
- Conee, Earl, and Richard Feldman. 1985. "Evidentialism." *Philosophical Studies* 48.
- Pollock, John. 1970. "The Structure of Epistemic Justification." *American Philosophical Quarterly*, no. 4: 62–78.
- . 1986. *Contemporary Theories of Knowledge*. 1st. ed. Rowman; Littlefield.

---

<sup>11</sup>An anonymous reviewer for *Synthese* points out that this is not an ordinary case of rebutting defeater. This is true insofar as the proposition that the wall is bathed in red lights is not a reason to think that it is not red, but it is a reason to think that the source of my belief is unreliable. Everyone, reliabilist or not, thinks that justification for believing that the source of your belief is unreliable undermines justification for that belief.

<sup>12</sup>Many thanks to Stew Cohen, Tim Kearl, Sarah Raskoff, Carolina Sartorio, Eyal Tal, Jason Turner, Robert Van't Hoff and three anonymous referees for *Synthese* for very helpful comments on previous drafts of this paper.

- Pollock, John, and Joseph Cruz. 1999. *Contemporary Theories of Knowledge*. Second edition. Rowman; Littlefield.
- Pryor, James. 2000. "The Skeptic and the Dogmatist." *Noûs* 34: 517–49.
- Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford University Press.
- . forthcoming. *Reasons First*.
- Sturgeon, Scott. 2014. "Pollock on Defeasible Reasoning." *Philosophical Studies* 169 (1): 105–18.
- Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford University Press.