The Structure of Defeat:

Pollock’s Evidentialism, Lackey’s Framework, and Prospects for Reliabilism

Peter J. Graham

University of California, Riverside

Jack C. Lyons

University of Glasgow

ABSTRACT: Epistemic defeat is standardly understood in either evidentialist or responsibilist terms. The seminal treatment of defeat is an evidentialist one, due to John Pollock, who famously distinguishes between undercutting and rebutting defeaters. More recently, an orthogonal distinction due to Jennifer Lackey has become widely endorsed, between so-called doxastic (or psychological) and normative defeaters. We think that neither doxastic nor normative defeaters, as Lackey understands them, exist. Both of Lackey’s categories of defeat derive from implausible assumptions about epistemic responsibility. Although Pollock’s evidentialist view is superior, the evidentialism per se can be purged from it, leaving a general structure of defeat that can be incorporated in a reliabilist theory that is neither evidentialist nor responsibilist in any way.

Epistemic justification is very often *defeasible*: the addition of new information can weaken or destroy what was a perfectly adequate justification. Normally, if an object looks red to you, then you’re justified in believing it’s red. Often such justification is good enough to yield knowledge. But if it looks red and you find out it’s illuminated with red light, then—absent further information—you’re no longer justified in believing it’s red. If you persist in believing so, you don’t know it. You might have *prima facie* justification, but you lack *ultima facie* justification: your initial justification is *defeated*.

Most discussions of defeat occur within either an evidentialist or a responsibilist framework. This is problematic. First, it is far from obvious what the resulting theories of defeat have to do with each other and whether they can even be rendered compatible. Second, both approaches seem to fit poorly with a reliabilist epistemology—a fact that is of much concern to us, since we endorse broadly reliabilist views.

In this chapter we first discuss evidentialist and responsibilist accounts of defeat before offering a reliabilist view. In §I, we discuss Pollock’s seminal evidentialist views on reasons and justification, defeat and defeaters. We then discuss Jennifer Lackey’s responsibilist inspired account in §§II-IV focusing on her influential claim that all defeaters are either “doxastic” or “normative.” [[1]](#footnote-1) Despite its influence, we argue her distinction is entirely unfounded. “Doxastic” and “normative” defeaters in Lackey’s sense simply do not exist. We then turn to the reliabilist treatment of defeat in §V. We think defeat fits naturally into a reliabilist epistemology, in a way that is neither responsibilist nor evidentialist.[[2]](#footnote-2)

I

Pollock provided the seminal treatment of defeat and defeaters.[[3]](#footnote-3) Pollock first defines (epistemic) reasons (to believe) in terms of justified belief.

A state M of a person S is a reason for S to believe P if and only if it is logically possible for S to become justified in believing P by believing it on the basis of being in state M.[[4]](#footnote-4)

Justified belief is then Pollock’s explanatory primitive. Other evidentialists take the notion of evidence or reasons as basic. Evidence is their explanatory primitive.

What do reasons have to do with defeat? The technical terms “epistemic defeat” and “defeater” then get their sense from the defeasibility of reasons and justification. Suppose Carol perceptually experiences an object as red and as having a particular shape. This is a reason to believe that is an apple.[[5]](#footnote-5) But then she picks it up and feels its texture and weight. It feels more like a polished rock than an apple. Perhaps it is an *objet d’art* made of stone, cleverly made to look just like an apple. Carol’s new experience is a reason to believe both that is a stone and that is not an apple. Her new experience defeats her old experience as a reason to believe that is an apple. Her old experience combined with her new experience is not a reason to believe that is an apple, but instead a reason to believe it is a stone, cleverly shaped to look like an apple. Given her new experience, the rational thing for Carol is to change her mind, if she hasn’t already.

Given one set of relevant reasons, you can have a good reason to believe p. Given another, more inclusive set, you might no longer have a good reason to believe p. Some reasons to believe p are defeated by other reasons. Those other reasons are *defeaters*. Those defeaters, in turn, might be defeated by further reasons. Those are *defeater-defeaters*. A reason that can be defeated is a *defeasible* reason. What you are ultimately justified in believing thus depends on your *total* evidence, not just part of it.[[6]](#footnote-6)

Pollock and Cruz (1999) formulate the notion of a defeater as follows:

If M is a reason for S to believe p, a state M\* is a **defeater** for this reason iff it is logically possible for S to be in the combined state consisting of being in both state M and state M\* at the same time, and this combined state is not a reason for S to believe p.[[7]](#footnote-7)

M and M\* range over the kinds of mental states that can serve as reasons for beliefs. Defeaters are then just reasons functioning in a destructive rather than constructive mode.[[8]](#footnote-8)

For Pollock, all defeaters are either rebutters and undercutters:

If M is a defeasible reason for S to believe p, M\* is a **rebutting** defeater for this reason iff M\* is a defeater (for M as a reason for S to believe p) and M\* is a reason for S to believe ~p.[[9]](#footnote-9)

If M is a defeasible reason for S to believe p, M\* is an **undercutting** defeater for this reason iff M\* is a defeater (for M as a reason for S to believe p) and M\* is a reason for S to doubt or deny that he or she would not be in state M unless p were true.[[10]](#footnote-10)

For Pollock, beliefs —doxastic states—are not the only possible defeaters, for he denies “doxasticism” in epistemology—the view that only beliefs can serve as reasons. M and M\* range over a variety of mental state types, including perceptual experiences, sensations, feelings, moods, and memories, as well as doxastic states. Pollock’s view is an evidentialist one: the justificational status of a belief is determined by the evidence--the reasons--for and against believing. Pollock’s evidentialism entails that you can’t have defeat without reasons/evidence; the only things that can produce defeat are things that can produce justification.

Reasons aren’t restricted to beliefs for Pollock, but they are restricted to things that you might base a belief on. Prima facie *doxastic* justification is limited to things that you *have* based your belief on. Ultima facie justification--even ultima facie *doxastic* justification--is importantly different, for it also depends on other reasons you possess but which haven’t positively or negatively causally influenced that belief. Pollock is surprisingly inexplicit about ultima facie justification, but it is fairly clear that he thinks that a belief is ultima facie justified just in case it is prima facie justified, and the agent doesn’t possess any ultimately undefeated defeaters for it. To *possess* a (for example, rebutting) defeater for the belief that p is to have a reason to believe ~p, i.e., to be prima facie propositionally justified in believing ~p. This makes ultima facie (doxastic) justification depend in part on propositional justification in a way that prima facie (doxastic) justification doesn’t. Thus, while having merely propositional justification can’t increase your doxastic justification, it can diminish it.

Though rightly influential, there are a few idiosyncratic features of his formulations worth pointing out.

First, in Pollock’s classic formulations (though this is rectified in his 2001) the definition of defeat doesn’t admit of degrees and thus doesn’t allow for partial defeat: cases where a reason reduces justification without eliminating it. Second, this problem is exacerbated by the oddly conjunctive framing of the definitions for rebutting and undercutting defeat: a rebutting defeater is something that’s both a defeater and a reason to believe the negation of the target proposition. If we weaken the definition of defeat to allow for partial defeat, then the first conjunct will be entailed by the second and is thus unnecessary. If we don’t allow for partial defeat, then the conjunctive definitions of undercutting and rebutting defeat are too narrow.

Third, only reasons/evidence can be defeaters. This is very much a deliberate choice on Pollock’s part. Later, we will see it come under attack from two quite different quarters, responsibilist and reliabilist. It is worth highlighting for now.

Fourth, somewhat surprisingly, only reasons can be defeat*ed*. Of course, this has to be true in the case of undercutting defeat, but it’s not obvious why rebutting defeat attacks reasons for the belief that p rather than the belief’s prima facie justification. It seems more natural to think of rebutting defeaters as working foremostly on prima facie justification and only derivatively on reasons.

For these reasons, it’s worthwhile to simplify and reformulate the basic Pollockian outlook. Here’s a theory that’s squarely Pollockian in spirit:

d is a **rebutting** defeater for S’s belief that p iff d is a reason for S to believe not-p.

d is an **undercutting** defeater for S’s belief that p iff d is a reason for S to believe that her reasons for believing p are inadequate.

Since we haven’t defined undercutting and rebutting defeat in terms of the more general notion, we can do it the other way around:

d is a **defeater** for S’s belief that p iff d is a rebutting or an undercutting defeater for S’s belief that p.

Finally:

S’s prima facie justification for believing p is **defeated** iff S possesses a defeater for that justification, and that defeater is not ultimately defeated.

A prima facie justified belief is ultima facie justified iff it’s not defeated.

These definitions make clear the evidentialist presumption that defeat is simply a matter of having reasons/evidence, and that ultima facie justification is thereby a matter of what the balance of the evidence--what the whole of the evidence--supports. They allow defeat to come in degrees, like reasons more generally. On the evidentialist view, the basis for defeat is the same as the basis for justification: the subject’s evidence.

II

Jennifer Lackey has argued that Pollock’s distinction between rebutting and undercutting defeaters requires supplementation with another distinction between “doxastic” and “normative” defeaters.

According to Lackey, a doxastic defeater is “…a doubt or belief that is had by S that indicates that S’s belief that p is either false or unreliably formed or sustained.”[[11]](#footnote-11) By ‘indicates’ she means ‘obviously entails.’ In particular:

* A doxastic rebutting defeater for a subject’s belief that p is a further belief of the subject that q, where that belief, given its content and the contents of other beliefs held by the subject, very obviously entails ~p. A doxastic rebutting defeater creates an incoherent set of first-order beliefs. For example: S believes {p, q, if q then ~p}.
* A doxastic undercutting defeater for a subject’s belief that p is the subject’s further belief that my grounds do not make it likely that p (or something functionally equivalent, e.g. the source of my belief that p is not reliable). A doxastic undercutting defeater also creates an incoherent set of beliefs: S believes {p, my grounds do not make it likely that p}. The incoherence here does not lie solely in the contents, but in the subject’s attitudes towards those contents, like Moore-paradoxical thoughts generally.

Must doxastic defeaters themselves be justified to do their defeating work? Not at all. Lackey repeatedly emphasizes that doxastic defeaters need not be true, justified, or rationally held to do their defeating work.[[12]](#footnote-12) Doxastic defeaters are justified *or* unjustified beliefs that indicate that a subject’s belief that p is either false or unreliably formed or sustained.

That’s her account of doxastic defeaters. There are also instances of defeat where the agent doesn’t have a doxastic defeater, hence the need for another category of defeaters: normative defeaters. There are two very different uses of the phrase ‘normative defeater’ in the literature, Lackey’s use and another we will explain shortly. We begin with Lackey’s use.

According to Lackey, a normative defeater involves three parts. A proposition q is a normative defeater for a subject’s belief that p in Lackey’s sense iff:

1. S has evidence—experiences, beliefs, etc.—that justifies believing q for S;
2. q would be a doxastic defeater for S’s belief p if S were to believe q;
3. S ought to believe q. [[13]](#footnote-13)

Condition (1) explains why evidence grounds normative defeaters. Condition (2) explains why they defeat the subject’s belief that p; it explains why the evidence is counterevidence for it explains why normative defeaters are “counterparts” of doxastic defeaters.[[14]](#footnote-14) Condition (3) explains why Lackey calls them “normative” defeaters, for epistemic *duty* requires the subject to believe the normative defeater. (3) can look redundant given (1), but it’s actually much stronger. To say that you’re justified in believing p is to say that you’re epistemically *permitted* to believe p. To say that you *ought* to do something, however, is to say that you’re doing something wrong if you don’t do it; you have a duty or obligation to do it. Clearly one could have epistemic permission to believe p without having an epistemic obligation to believe it. We will explain why Lackey includes (3) below, in section IV.

What is the other sense of the phrase ‘normative defeater’? On the other use, a normative defeater is:

1. a worldly piece of evidence (like a letter hidden under the doormat);
2. that the subject is not aware of and so does not “possess” (the subject has no idea about the letter);
3. that if the subject were aware of, then the subject would possess a defeater for one of her beliefs (if she read the letter, she’d now have a defeater for a belief);
4. that the subject *should* be aware of, that the subject *should* possess (for some reason she should check under her doormat, and read the letter).

A normative defeater for a subject’s belief *in this sense* is a *potential* defeater that the subject *does not actually possess* but *should*.[[15]](#footnote-15) This is *not* Lackey’s sense. Lackey’s idea of normative defeat arises from beliefs the subject should have but does not, given evidence the subject *already possesses*.[[16]](#footnote-16) This other use of ‘normative defeaters’ is not what she has in mind.[[17]](#footnote-17)

An example from Lackey of undercutting normative defeater should help make her use of the phrase ‘normative defeater’ clear:

Alice is incorrectly told by an otherwise reliable optometrist that her vision is nearly completely unreliable, yet she refuses to accept his diagnosis, without having any rational basis for doing so. So, one might say, even though the optometrist's report is false, Alice should accept his diagnosis, given all of the evidence that she has available to her, and thus she has a normative defeater for her visual beliefs.[[18]](#footnote-18) Nevertheless, as she is walking out of the doctor’s office, [Alice] sees a car accident on Michigan Avenue, forms the corresponding true belief that there was such an accident. She rejects the report [that she has unreliable vision] for no rational reason.[[19]](#footnote-19)

Alice has good evidence from the doctor’s diagnosis that her vision is unreliable: she knows he told her that her vision is unreliable, and she believes he is a reliable informant on matters optometrical. Given this evidence she possesses, believing the proposition my vision is unreliable is justified for her. Given Lackey’s definition of doxastic defeaters, this proposition would be a doxastic defeater if she were to believe it. Furthermore, Lackey insists, she *ought* to believe it. That’s why it is a *normative* defeater. The proposition my vision is unreliable is then Alice’s normative defeater for her belief there was a car accident.[[20]](#footnote-20)

Lackey’s example involves a belief as the evidential basis for the normative defeater. This belief must be justified, given her definition of normative defeaters, for unjustified beliefs are not a part of one’s evidence that justifies further beliefs. Non-doxastic experiences can serve as the evidential bases for normative defeaters as well, for experiences too are a part of one’s total evidence.

How does Lackey understand the basis of defeat, for both doxastic and normative defeaters? “The underlying thought” behind both kinds of defeaters, she says,

...is that certain kinds of counterbeliefs [doxastic defeaters] and counterevidence [the grounds for normative defeaters] contribute epistemically unacceptable irrationality to doxastic systems and, accordingly, that justification and knowledge can be defeated or undermined by their presence.[[21]](#footnote-21)

On Lackey’s view, irrationality—the failure to conform to the requirements of rationality—is then the source of defeat. We will elaborate on this idea as we critically engage Lackey’s framework. Both categories, we shall argue, are bogus. There is no such thing as a doxastic or a normative defeater in Lackey’s sense.[[22]](#footnote-22)

III

We begin with doxastic defeaters.

Although Lackey’s examples of doxastic defeaters throughout her corpus involve justified doxastic defeaters, central to Lackey’s concept of doxastic defeaters is the insistence that they can be unjustified, or even irrational.[[23]](#footnote-23) We will argue that that is a serious problem.

To test the intuitive plausibility of this view, we should imagine cases involving *unjustified* doxastic defeaters. Here’s a recipe:

1. S believes that p on very strong total evidence.
2. S comes to believe q on absolutely no evidence, in a completely unreliable way. If any belief is unjustified, this one is. Let the belief occur through wishful thinking, emotional attachment, a brain tumor, mental illness, radiation, etc.
3. S’s background beliefs include if q then ~p.
4. The subject now has an incoherent set of beliefs: S believes {p, q, if q then ~p}.
5. By Lackey’s definition, S’s belief that q is then a doxastic defeater for S’s belief that p. Absent a defeater-defeater for q, then on Lackey’s view S’s belief that p is no longer justified.[[24]](#footnote-24) For Lackey, S should stop believing p.
6. But that is counter-intuitive, given that S has very strong evidence for p, and no evidence against it. (The only consideration that speaks against believing p is q, but she has, by stipulation, no evidence for q and shouldn’t believe it.)

We think that in such cases the subject should not believe q but should continue believing p. p is based on strong evidence, formed in a reliable way. q is neither. How could a belief that is well-justified on strong evidence have the force of its evidence defeated by an unjustified belief, based on no evidence at all?

A number of other authors have endorsed--sometimes offhandedly, sometimes deliberately--the idea of unjustified defeaters, i.e., the idea that an unjustified belief can nevertheless defeat the justification of a prima facie justified belief.[[25]](#footnote-25) But Lackey’s discussion of defeat disguises this very controversial claim as an innocuous terminological point. It’s nothing of the sort. Notice first that this view of doxastic defeaters is incompatible with any roughly evidentialist view. On an evidentialist view, nothing can be relevant to justification but evidence, and unjustified beliefs aren’t evidence.[[26]](#footnote-26) On the evidentialist view, defeat occurs when one’s *partial* evidence supports believing p but one’s *total* evidence (or even just a more inclusive body of evidence) does not. A defeater is then just another piece of evidence. On the Pollockian theory of defeat, nothing can provide defeat that can’t also provide justification, but unjustified beliefs can’t provide justification. Hence, an unjustified belief cannot be a defeater.

Lackey needs an argument that there might be unjustified doxastic defeaters. Without such an argument, her entire category of doxastic defeaters is unmotivated, for the evidentialist already recognizes that *justified* beliefs can play the role of defeaters, either rebutting or undercutting. Without an argument, there’s no need for Lackey’s *broader* category of *doxastic defeaters*.

Does she have an argument? She does not articulate one. But here is a suggestive passage:

[It] is important to notice that *what makes an undefeated doxastic defeater epistemically problematic* is that it is held *in conjunction with another belief*. The defeater itself need not be true, justifiedly believed, or rationally believed to have the power to defeat other beliefs precisely because it need not be true, justifiedly believed, or rationally believed *to render it* ***irrational*** *to hold certain other beliefs.*[[27]](#footnote-27)

Doxastic defeaters thereby contribute “epistemically unacceptable irrationality to doxastic systems.”[[28]](#footnote-28) That’s the idea we flagged at the end of the last section. She thus seems to be thinking something like the following:

1. Rationality requires coherence amongst one’s propositional attitudes. (Assumption)
2. When one holds a particular belief that p despite also holding a doxastic defeater for that belief—either a rebutter or undercutter, whether justified or unjustified—then one holds an incoherent set of propositional attitudes. (Definition of doxastic defeater.)
3. Hence, if one holds a belief that p despite a doxastic defeater (justified or unjustified), then one is being epistemically irrational. (From 1, 2)
4. If one is being irrational partly in virtue of holding a belief that p, then *that* belief that p is not justified. (Assumption)
5. Hence, if one holds a belief that p despite a doxastic defeater, whether justified or unjustified, then that belief that p is not justified. (From 3, 4)
6. If a belief that p was justified, but is no longer justified, due to the possession of a defeater, then its justification is defeated. (Assumption)
7. Hence, if one holds a belief that p despite a doxastic defeater, whether justified or unjustified, then its justification is defeated. (From 5, 6)

This argument looks valid. And it makes plain a connection between doxastic defeaters, irrationality, and defeat, while allowing that doxastic defeaters may be unjustified.

Justified or unjustified, doxastic defeaters induce incoherence; incoherence induces irrationality; irrationality induce defeat. That’s why unjustified doxastic defeaters defeat.

But is the argument sound? No. (4) is false. Notice that (3) says that if a subject holds a belief that p, then acquires a belief that q, where q is a doxastic defeater (possibly unjustified), but the subject continues to believe that p, then the subject has entered an irrational state, the state of holding an incoherent set of propositional attitudes. But (4) then condemns *one* of the beliefs in the set *in particular* as unjustified—the belief that p that the subject started with. And if rationality requires that we not believe propositions that are not justified for us, then rationality requires that the subject give up his belief that p. That’s the route the subject is required to take to exit his current state of irrationality. So (4) says if the subject starts with a *justified* belief that p, then acquires an *unjustified* belief that q, where that belief that q combined with the belief that p then puts the subject into an irrational overall state, then the subject is required to give up on the *justified* belief that p to exit the irrational state.

This, we think, is obviously false. That is not what rationality—even *coherence* rationality—requires of the subject. What coherence rationality requires is that the subject give up *at least one* of the beliefs in the incoherent set, but not any one belief *in particular*. The subject could exit irrationality by giving up q, or some other auxiliary belief. Rationality requires *some* accommodation. It does not require giving up the original belief *in particular*. That’s just the lesson of the Quine-Duhem thesis. It seems perverse to require that the subject abandon her *justified* belief that p when she acquires an *unjustified* belief that q, as if two wrongs made a right. One might rather think—and this is just the point we’ve made already—that if anything, rationality would have required the subject *not* to form the *unjustified* belief that q in the first place. Now that she has, the best option is to retract it.

Not only is this the intuitively obvious response, and the response that evidentialism requires, it’s the response that a coherence/responsibilist theory would endorse anyway. The coherence theory never said to get rid of every belief that’s implicated in a local incoherence; it says to get rid of (or modify) the less-well-connected beliefs to maximize coherence among the better-connected beliefs. We want a large, comprehensive, and coherent belief set, and the way to get that isn’t to just toss out everything that’s involved in any kind of conflict, but to selectively pin the blame for the conflict on whichever beliefs are less coherent with the rest of the set, independent of the conflict in question. The proponent of unjustified defeaters would have us do the opposite of what the coherence theory recommends: they would have us abandon the well-connected beliefs to accommodate the poorly-connected ones.

There’s another kind of argument one sometimes sees in defense of unjustified defeaters. The proponent will say something along the lines of “*given that* she believes her senses are unreliable, S should stop believing that p.”[[29]](#footnote-29) This would support a modus ponens argument for the desired conclusion (that S should stop believing p) if ‘given that’ functioned like ‘if’: i.e., if the starting point were “*if* S believes her senses are unreliable, S should stop believing that p.” But that’s far from obvious; in fact, it’s the very controversy we’re trying to resolve. More likely, the plausibility of the ‘given that’ claim derives from its introduction of a scope ambiguity between something like ‘if Bel(p) then Ought(~Bel(q))’ and ‘Ought(if Bel(p) then ~Bel(q))’.[[30]](#footnote-30) The argument requires the former disambiguation, while it’s only the latter that’s either intuitively obvious or supported by coherence considerations.

This kind of ‘given that’ talk is quite tempting. It’s natural to say that *given that* you believe that this tank of piranhas is a hologram, you ought to be willing to put your hand in it. It doesn’t follow that it’s rational for you to put your hand in the tank—just because you do (irrationally) believe it’s a hologram. It’s only conditionally rational, rational relative to a belief that turns out to be itself irrational. Relative to that crazy belief, that action wouldn’t be crazy. But since the belief in fact *is* crazy, so the action is too. The action is rational relative to the *rationality* of the belief, not relative to its mere existence. You don’t get to run a modus ponens-type inference in these cases.

Finally, Lackey frequently makes reference to *undefeated* doxastic defeaters: “ Millicent acquires an undefeated defeater for her visual powers via another person’s testimony, and *hence* fails to have the knowledge in question…. Millicent’s hearers [are justified] but Millicent is not, *because* she has the undefeated defeater”.[[31]](#footnote-31) Such repeated qualification suggests that even Lackey herself thinks that a *defeated* doxastic defeater doesn’t kill the *prima facie* justification of the belief it’s a defeater for. But why should that be? If defeaters don’t need to be justified, they shouldn’t need to be undefeated to do their defeating work. She thinks they can be irrational and still defeat, after all. Could it really matter whether a defeater is unjustified because defeated, or unjustified because the agent pulled it out of a hat? This suggests a tacit recognition on Lackey’s part that unjustified beliefs can’t serve as defeaters after all.

Lackey’s right that incoherence is a problem, and also right that so-called doxastic defeaters, as she has defined them, introduce incoherence. And we agree that a subject is required to make some adjustments. But we just don’t think incoherence *as such* is the same as epistemic defeat. Defeaters defeat propositional justification by introducing *counterevidence* into the subject’s *total* evidence, not by *merely* introducing incoherence at the level of the subject’s propositional attitudes.

This means that Lackey is not only wrong about unjustified doxastic defeaters, but also about why *justified* doxastic defeaters defeat. They don’t defeat because they introduce incoherence; they defeat because they negatively change the subject’s total evidence.

Lackey’s talk about doxastic defeaters is not a harmless and neutral bit of taxonomy. The idea of unjustified defeaters is flatly incompatible with evidentialism, which holds that only reasons or evidence can provide or play the role of defeaters; unjustified beliefs cannot play that role. The idea that “doxastic” defeaters (whether justified or unjustified) defeat in virtue of incoherence simply misunderstands the nature of defeat. There is no rationale to introduce Lackey’s category of *doxastic defeaters*.

IV

We now turn to Lackey’s category of normative defeaters.

Recall that for Lackey a normative defeater for a subject’s belief that p is a proposition that q where (1) S has evidence that justifies believing q for S; (2) q would be the content of a doxastic defeater for S’s belief that p if S were to believe q; and (3) S ought to believe q. Just as we see no rationale for Lackey’s category of doxastic defeaters, we also see no rationale for her category of normative defeaters, and not only because normative defeaters are understood in terms of doxastic defeaters.

Before turning to criticism, there’s an important difference between her two categories. Lackey’s concept of doxastic defeaters is supposed to describe a certain phenomenon--defeat by belief, whether justified or not--that simply doesn’t exist. The phenomenon that normative defeaters are supposed to describe, on the other hand--defeat that results from ignored evidence--is certainly a real thing. It’s Lackey’s explication of this phenomenon, in terms of duties to adopt doxastic defeaters, that is flawed, as we’ll argue.

There are three problems with Lackey’s theory of normative defeat.

First, it involves unnecessary complications to treat the cases of defeat it is designed to cover, while simultaneously relying on dubious category of doxastic defeaters. Normative defeaters, according to Lackey, are *grounded* in counterevidence possessed by the subject. But they do not *consist* in that counterevidence. Instead, Lackey makes that counterevidence relevant only through the circuitous route of doxastic defeaters that one should have, beliefs that you would have had, had you not been flouting your epistemic duties.

Why is this complication unnecessary? Look at things from the evidentialist’s perspective: the counterevidence *alone* already defeats the subject’s propositional justification. If the subject maintains the original belief in the face of such counterevidence, then the subject’s belief is not doxastically justified, for it is no longer propositionally justified. That’s all there is to it. Univocal evidence supports believing p; equivocal evidence (evidence that contains sufficient counterevidence) does not. There’s no need to postulate the obligation to believe the proposition that the counterevidence supports.

Recall Alice. Alice sees a car crash. Ordinarily that is sufficient evidence to justifiably believe there was a car crash. But Alice was also told by a reliable optometrist that her vision has been significantly impaired. The evidence from the testimony of her senses and the evidence from the testimony of her optometrist combined fail to justify believing there was a car crash. The combined evidence is equivocal and thus does not provide sufficient overall reason. Thus, when she does so believe, her belief is not doxastically justified. The testimony of her senses is defeated by the testimony from her optometrist. That’s it. There is no need to introduce an obligation to believe the proposition the evidence supports to explain why the evidence defeats.

Second, it’s not very plausible that defeat will typically involve *obligations* to believe anything at all, let alone to believe doxastic defeaters. Epistemic reasons typically produce permission, not obligation. And yet, any time an agent has permission to believe ~p, she thereby has a defeater for believing p. Lackey’s reliance on propositions you *ought* to believe will leave us with far too few defeaters.[[32]](#footnote-32)

Third, there’s no chance that standard cases of defeat will even involve permissions to believe doxastic defeaters, let alone obligations. In other words, Lackey’s view makes counterintuitive predictions about standard, straightforward cases of defeat. Suppose David is indoors and Edwina calls from outside and says it is raining. David believes her and is justified in this belief. But then Fiona calls from outside and tells David that it is not raining. David has no reason to trust Edwina over Fiona or vice versa, but he ignores Fiona’s testimony and continues to believe it is raining. The evidentialist would say that David’s belief that it’s raining is no longer propositionally justified, now that his evidence is equivocal and thus isn’t doxastically justified. Edwina’s testimony gives him a reason to believe it’s raining, but Fiona’s testimony gives David a rebutting defeater.

What would Lackey say about this case? David has a rebutting defeater, but he doesn’t *believe* a rebutting defeater, so on Lackey’s view he must have a normative rebutting defeater. But this means that he ought to believe a rebutting doxastic defeater for the belief that it’s raining. That is, he ought to believe that it’s not raining. But clearly this is wrong. In the face of such conflicting testimony, *David shouldn’t believe anything about the weather*. Not only is he not obligated to believe Fiona’s testimony, he’s not even permitted to believe it in the circumstances. This kind of case is easily and naturally handled by a Pollockian theory of defeat, while Lackey’s theory fails pretty decisively.

Why, then, did Lackey ever think there had to be normative defeaters? Her idea, as we’ve already noted for doxastic defeaters, seems to be that they follow from the requirements of rationality themselves. If a subject holds a belief despite a normative defeater, the subject is being irrational and so their belief is not justified. Here’s a relevant passage:

…a subject who holds a belief in the face of a normative defeater is being *epistemically irrational and/or irresponsible*. For instance, in discussing a classic case of normative defeat [the case of Samantha, who believes the President is in NYC but has massive counterevidence that she ignores that he is in Washington DC], Laurence BonJour says that such a subject is “…being thoroughly irrational and irresponsible” in **disregarding the evidence** in question.[[33]](#footnote-33)

BonJour goes on to say that the subject’s “irrationality and irresponsibility prevent[s] [Samantha’s] belief from being epistemically justified.” Along with BonJour, Lackey also approvingly cites similar passages in Michael Williams (e.g 1999:22-23). Here is a relevant passage from Williams:

[We] can ask whether, in forming a certain belief, I have **negligently ignored important counter-evidence**….beliefs reached [this way] would be irresponsibly held and would not count [as justified for] epistemic responsibility is the primary sense of justification…an essential component of rationality.”

So-called normative defeat must then occur when a subject *ignores or disregards* relevant counter-evidence they possess, for the subject is being irrational and/or irresponsible in so doing.

If the subject is being irrational, then the subject must be violating a requirement of rationality. That’s the “underlying thought,” after all, at least according to Lackey, when a subject’s belief is defeated. What then is the requirement of rationality that Lackey must have in mind? Clearly, it must be the requirement not to ignore one’s counterevidence.

With this requirement in mind, here is an argument for normative defeaters, in Lackey’s sense:

1. A subject’s belief that p is doxastically justified only if it is not irrationally held. Rationality is necessary for justification. (Assumption)
2. A subject’s belief that p is irrationally held if the subject ignores her counterevidence; rationality requires acknowledging counterevidence. (Requirement of rationality)
3. A subject ignores her counterevidence unless she believes the propositions that the relevant body of counterevidence supports. (An account of what it is to acknowledge counterevidence.)
4. Hence, if a subject who believes p does not believe the propositions that her body of counterevidence relevant to p supports, then her belief that p is irrationally held. (From 2, 3)
5. Hence, if a subject who believes p but does not believe the propositions that her relevant body of counterevidence supports, then her belief is not doxastically justified. (From 1, 4)
6. Hence, if a subject who believes p in the face of normative defeaters (as defined by Lackey), then her belief is irrationality held (from 4) and so not justified (from 5).

The first three premises seem to validly entail the conclusions, and so it looks like we have an argument that makes explicit connections between rationality, justification, defeat, and so-called normative defeaters. We’ve thereby made sense of Lackey’s view that when a subject “ignores” or “disregards” counterevidence, that entails “epistemically unacceptable irrationality” that defeats the subject’s justification for her belief. So if we are looking for an argument for why there must be normative defeaters, we have found one.

The argument, however, does not work. Without taking a stand on the first or second premise, we think it should be clear to everyone that the third premise is false. *One* way to regard or acknowledge counterevidence is to believe what it is evidence for. That is certainly true. But it is not the *only* way to regard or acknowledge counterevidence. When subjects like Alice, Carol and David have counterevidence, that evidence defeats their *prima facie* justification for their belief that p, so that their *total* evidence no longer propositionally justifies believing that p. How might they acknowledge or show due regard to their counterevidence? They might do it—and this is what the evidentialist recommends—by giving up their belief that p. And that is because their belief that p is no longer propositionally justified on their total evidence. Alice could and should acknowledge the force of her counterevidence *by not believing* there was a car crash. Carol could and should give due regard to her counterevidence *by giving up* her belief that is an apple. David could and should respect his counterevidence from Fred’s testimony *by suspending judgment* as to the outside weather.

Insofar as our characters violate a requirement of rationality if they continue to believe despite their counterevidence, it is *not* that they are *failing to believe* the propositions that their counterevidence supports. It is rather that they are believing a proposition that their *total* evidence *does not support*. The requirement of rationality at work in all of these cases—a requirement that supports the requirement to acknowledge one’s counterevidence—it is the all-too-familiar requirement *not* to believe what one’s *total evidence* does *not* support. Lackey’s view of normative defeat misunderstands the phenomenon of defeat and defeaters for it misunderstands not only how counterevidence defeats, but also what it takes to acknowledge one’s counterevidence. Defeaters qua counterevidence do not defeat by introducing obligations to believe propositions supported by the defeating evidence. Defeaters qua counterevidence defeat by changing what propositional attitudes a subject’s total evidence supports taking. The evidentialist has a handy treatment of all of these cases without introducing a requirement to believe--or even a permission to believe--what the counterevidence supports. Just as there are no doxastic defeaters, there are no normative defeaters.[[34]](#footnote-34)

Internalist epistemologists tend to be motivated by one or both of two distinct principles. The first is that your beliefs should always fit your evidence. The second is that when you believe something, that imposes on you an obligation to adjust your other beliefs accordingly if doing so is necessary to maintain coherence. We have seen the first--evidentialist--principle used to articulate a relatively promising theory of epistemic defeat. We have seen an attempt to use the second—responsibilist—principle to understand defeat, but that effort seems to us much less promising.

In particular, the doxastic/normative defeat framework should be avoided. Talk about doxastic defeaters gives a free pass to the implausible view that unjustified beliefs can defeat beliefs that have otherwise very strong prima facie justification. Normative defeat is closer to what the rest of us would recognize as defeat--defeat by counterevidence. But if we’re going to understand defeat in terms of reasons/evidence, it should be in terms of reasons permitting or failing to permit the belief whose justification is at issue (as Pollock does), not in terms of reasons producing duties to believe propositions.

V

So far we’ve been looking at attempts to understand defeat in evidentialist and responsibilist frameworks. But these are not the only frameworks for understanding defeat. We think that the work we’ve done points the way toward a distinctively reliabilist account.

This is important. Reliabilism is primarily a theory of prima facie justification and as such, it needs a theory of defeat. If the only way to understand defeat is in evidentialist or responsibilist terms, this seems to present a problem, with the reliabilist invoking reliability in her treatment of *prima facie* justification, but then pivoting to the quite alien concepts of evidence or responsibility in her treatment of defeat and thus ultima facie justification.

Although extant versions of reliabilism have perhaps failed so far to offer adequate theories of defeat, we think this is not a problem intrinsic to reliabilism. The various forms of reliabilism do have sufficient resources for illuminating the nature of epistemic defeat, without deviating from the basic spirit of reliabilism. We argued above that Pollock’s evidentialist framework is much more promising than Lackey’s responsibilist one. This is good news for reliabilism, for although neither framework sits comfortably with reliabilism without modification, Lackey’s framework is much more deeply inhospitable to reliabilism. For example, it’s hard to see a rationale in terms of *reliability* for demanding that an agent ignore her highly reliable visual processes, simply because she has the false and irrational belief that those processes are unreliable.[[35]](#footnote-35) In contrast to this deep and probably constitutional incompatibility between responsibilism and reliabilism, the conflict between evidentialism and reliabilism might be relatively superficial--at least in the present context. We can get a reliabilist-friendly account of defeat--admittedly a sketchy and preliminary one--by taking the Pollockian structure articulated above and purging it of any reliance on reasons or evidence. Let’s address the why of this before the how.

Although several authors have developed evidentialist versions of reliabilism,[[36]](#footnote-36) reliabilism is typically taken to be a *competitor* to evidentialism. One reason for this is that standard forms of reliabilism state sufficient conditions for justification (e.g., process reliability) that don’t entail that the agent has any reasons, or evidence. Additionally, some reliabilists think reasons by their very nature are too internalistic and will deny that *any* justification is a matter of reasons. More modestly, some beliefs, especially so-called “IIPM beliefs” (introspection, intuition, perception and/or memory beliefs) might be justified in a way that doesn’t have anything to do with being based on reasons. Recall that reasons and evidence, on the broadly Pollockian view, are things on which a belief is based; on the nonevidentialist view, however, we “just know” the deliverances of perception, memory, and so on, without basing them on anything. We won’t argue against evidentialism here, but we do want to show how a reliabilist theory of defeat can avoid evidentialism.

Suppose you have it on good authority--and justifiedly believe--that the previous tenants left no furniture in the apartment. Later, when you open the door, you see furniture. This gives you a defeater for your belief that there’s no furniture in the apartment. If we’re assuming a kind of reliabilism where perception gives justification without the involvement of reasons, we’ll need a way to describe the situation without invoking reasons. We’ll say instead that perception gives you a warrant for believing that there’s furniture in the room: it gives you *prima facie* propositional justification for that belief, whether you end up believing that or not. Warrants are like reasons, but without any assumption that there’s some other mental state, introspectible or otherwise, on which the belief is *based*. Having a warrant for believing p is thus compatible with having reasons for p, but if there is justification that doesn’t derive from reasons, it’s also compatible with lacking any reasons for p. Thus, the concept of having a warrant is a more general and inclusive concept than that of having a reason. Any time you have a reason you have a warrant, but you might have a warrant without having a reason

This allows us to articulate a general view about defeat that is very much in the Pollockian spirit. We can use Pollock’s view about the *structure* of defeat and simply drop the assumption that all warrants involve reasons. We simply take what Pollock says in terms of reasons and say it instead in terms of warrants:

* S has a rebutting defeater for her belief that p iff she has a warrant to believe not-p.
* S has an undercutting defeater for her belief that p iff she has a warrant to believe that her warrants for believing p are inadequate.
* S has a defeater for her belief that p iff she has a rebutting or an undercutting defeater for her belief that p.[[37]](#footnote-37)

Having a warrant to believe p, we take it, is nothing more than having prima facie propositional justification to believe p. Thus, all the reliabilist needs to do to have a theory of defeat is to develop a theory of propositional justification. This is non-trivial, since reliabilism is first of all a theory of doxastic justification, but it’s far from hopeless.[[38]](#footnote-38) Pollock himself offers what’s clearly a theory of propositional justification, where propositional justification is derived from doxastic justification: a reason, for Pollock, is something that could give you doxastic justification.

Representing a more pessimistic view, Bob Beddor argues in a recent paper that the standard reliabilist accounts of defeat fail, and suggests that reliabilists appeal to reasons to correct this deficiency.[[39]](#footnote-39) While we want to concede that his argument is effective against a standard extant reliabilist theory of defeat, we will insist that it poses no insuperable problem for reliabilism more generally. And it does nothing to show that defeat needs to be understood in terms of reasons.

A (perhaps *the*) standard reliabilist theory of defeat is Goldman’s Alternate Reliable Process theory (ARP), according to which S’s *prima facie* justified belief that p at t is defeated just in case S has another reliable process available to her at t, which if used in addition to or instead of the one actually used, would have resulted in S’s not believing that p.[[40]](#footnote-40) ARP formulates defeat in terms of counterfactual *non*-belief rather than counterfactual disbelief, largely because, as we saw above, the proper response to defeaters is often agnosticism, rather than doxastic reversal.

Against ARP, Beddor offers the following cases:

S sees a tree but has a reliable skeptic’s-advice-predicting mechanism, which S normally trusts when she uses it. The skeptic would have advised not believing there’s a tree, and in this case, had S used this mechanism, S wouldn’t have believed there was a tree. Nevertheless, S’s visual *prima facie* justification is, intuitively, undefeated by the availability of this mechanism.[[41]](#footnote-41)

This case aims to show that, contrary to ARP, the agent’s having an available reliable process that would have led to non-belief, is insufficient for defeat.

Here’s Beddor’s other case:

S hears a defeater for her belief that p, from a reliable source, but one for whom S holds an inveterate and irrational distrust. S would never take this person’s testimony to heart, so she persists in believing p. Intuitively S’s belief is defeated, even though her distrust means that the process of taking the testimony on board is not available to her.[[42]](#footnote-42)

This case aims to show that, contrary to ARP, the agent’s having an available reliable process that would have led to non-belief, is unnecessary for defeat.

There are serious objections one could level against the kind of thing Beddor counts as a process and what he takes availability to be. We won’t pursue these objections, for we think that Beddor’s cases reveal a genuine flaw with ARP, whether or not they refute ARP as they’re currently stated.

Beddor thinks that the problem with ARP is that it omits consideration of reasons. We think the problem with ARP is that it is formulated in terms of what the agent *would* have believed, or not believed, had she used another process. Roughly, an adequate theory of defeat needs to be formulated in terms of what the agent *should* have believed or not believed had she used another process. More precisely, since we’ve seen how badly that fared for Lackey’s normative defeaters, what she would or would not have had (prima facie) *warrant* to believe, had she used another process. Counterfactual non-belief only entails defeat if all the relevant components of the counterfactual situation go normatively correctly, but this isn’t guaranteed merely by stipulating that the process is reliable. This is because the counterfactual situation will typically involve all kinds of factors beyond the immediate consequences of the use of the alternate process. That alternate reliable process might feed into an unreliable one, so that the agent forms a justified belief and then concludes something unjustified from it. Or the justified belief might be one to which the agent would respond in an irrational way that’s in no way sanctioned by reliabilism. Beddor’s cases trade on just these kinds of possibilities, on the gap between what the agent *would* have believed (/not believed) and what the agent *should* have believed (/not believed) or is licensed in believing (/not believing). ARP is flawed because it focuses on the way alternative available processes yield counterfactual non-belief, when it ought instead to be focused on the way alternative available processes yield *warrants*.

Does this appeal to warrants, and to what the agent does or doesn’t rightly believe in the counterfactual situation, amount to a significant concession to Beddor, or to the evidentialist, or to the opponent of reliabilism? Not at all! Reliabilism is a normative epistemological theory, after all. Its raison d’etre is to give an account of warrants, of what beliefs are prima facie normatively appropriate. If it can’t do that, then defeat is the least of reliabilism’s problems.

Here, then, is the *sort* of thing a reliabilist can and should say about possession of warrants and thus about propositional justification:

S has (*prima facie*) warrant for believing p at t iff a cognitive process that satisfies the general theoretical requirements for *prima facie* (doxastic) justification (a) is available to S, and (b) if used at t, taking as inputs only states that S is already in, does or would likely produce p as output.

Again, this is a sketch and is intended primarily for the purpose of illustrating that reliabilism has promising resources for accommodating defeat. Nevertheless, several comments about the proposal are in order.

Notice that the formulation is very general, leaving open what the requirements are for *prima facie* justification. This is for two reasons. First, it allows the present theory of warrant possession (and consequently of defeat) to be easily imported into reliabilist theories that differ in other respects. Second, it makes quite explicit the fact that the present theory of defeat is making use of *old* theoretical apparatus, not importing *new* machinery that might not fit the general reliabilist ethos. It will therefore inherit the virtues (and vices) of that apparatus. Plug in a naturalistic and reductive and nonevidentialist theory of prima facie doxastic justification, and you’ll get a naturalistic and reductive and nonevidentialist theory of propositional justification, and thus of defeat, and thus of ultima facie doxastic justification.

Clause (b) restricts the relevant outputs of the available processes to those that would result from the system’s taking as inputs only states S is already in. This is perhaps already entailed by a reasonable individuation of cognitive processes (see Lyons 2019), but it’s worth making it explicit. We don’t want the fact that S has an unopened encyclopedia nearby to give S warrant for everything written in the encyclopedia.

Next, the formulation mentions the process producing “p as output”, rather than producing the “belief that p”. We trust that it is intelligible to say that a belief-forming process might be producing an output, even when that output isn’t a belief. A visual process might output the proposition that the lines are of unequal length, even though the viewer is convinced that it’s an illusion and thus does not believe they’re of unequal length. More generally, we want to say that having a process available can give you warrant (/propositional justification) for some proposition, even though you don’t believe that proposition, and perhaps never would.[[43]](#footnote-43)

Finally, the formulation is vague in some crucial respects. In particular, it relies on the *availability* of a process that does or would *likely* produce p. Some might object that the vagueness connected to availability and likelihood here are unacceptable. On the contrary, these might just be unavoidable features of propositional justification, on any reasonable theory. If I justifiedly believe p and that p implies q, then I’m propositionally justified in believing q; and all this gives me a defeater for my belief that ~q. But not merely because my justified beliefs entail q—they entail every necessary truth, many of which I’m not propositionally justified in believing and which don’t provide defeaters. Why? Because these consequences aren’t obvious to me, where obviousness is cashed out by the reliabilist in terms of process availability. Since obviousness comes in degrees, availability will have to as well, if this move is to have any hope of working. For similar reasons, we can’t demand of an available process that it *would* lead to an output that p. For those logical consequences that are not screamingly obvious but yet obvious enough to count as propositionally justified, it’s plausible that the use of the relevant process wouldn’t guarantee the output of that consequence, but would only make it quite likely.

Suppose Chuck believes that he’s never owned a weapon, although if he thought about it for a while, he’d remember that his parents bought him some Chinese throwing stars for a birthday back in the 1980s. Is Chuck’s belief that he’s never owned a weapon defeated? It depends, we think, on how hard it would be for Chuck to recall this: how long and hard Chuck would have to probe memory, how robust or fragile the cueing conditions would need to be in order to retrieve the memory, how likely a memory search would be to trigger recall, etc. Given that defeat comes in degrees, we might even claim--with some plausibility--that Chuck’s *degree* of ultima facie justification is affected by these considerations: if three seconds of thinking about it, under pretty much any conditions, would fetch the relevant memory, then Chuck’s prima facie justification for believing he’s never owned a weapon is greatly reduced; if only a great deal of focused concentration, when primed with ninja movies, would result in retrieval of this memory, then Chuck’s initial justification is only slightly reduced. In this way, what Goldman (1986) calls *power* (a ratio of truth to queries, where reliability is a ratio of truths to falsehoods) might be directly relevant to propositional justification and defeat, even if it’s not directly relevant to positive, prima facie justification.[[44]](#footnote-44)

Some people don’t like the idea that possession of reasons, or warrants, or evidence, might come in degrees. In order to avoid this result, Feldman (1988) goes so far as to claim that the evidence a person possesses at a given time is restricted to those conscious mental states she’s in at that time. The worry for this proposal is that it would seriously diminish the body of evidence anyone possesses at a given time. How many items can be in consciousness at once? Serious estimates range from *one* (Dehaene 2014) to 7 +/-2 (Miller 1957), with most people thinking it’s about 4. Not enough, it seems. We saw above that Pollock largely elides this important issue; although he articulates a very detailed theory of ultimate defeat, none of it applies directly to questions of an agent’s actual epistemic status, unless we add the claim that an agent’s belief is (prima facie) defeated iff she is in possession of a reason that serves as a defeater for that belief.

Propositional justification was going to come in degrees anyway, because reasons and warrants vary in their strength. But strength, availability, and likeliness of output are orthogonal dimensions of variation: a piece of evidence, for example, can be weakly supportive but fully obvious/available, strongly supportive but far from obvious/available, etc. The fact that the current proposed reliabilist theory of defeat reveals and highlights some of the vagaries surrounding propositional justification is no criticism of that proposal, not if these vagaries were really already a part of propositional justification.

This proposal is surely in need of further refinement, and different reliabilists may want to tweak it in different ways. A full treatment of this issue demands a paper of its own. The point we wanted to make with it here was this: a roughly Pollockian view about defeat can be smoothly accommodated in a reliabilist epistemology, without the reliabilist having to embrace any evidentialist or responsibilist elements that are in any way foreign to the naturalistic spirit of reliabilism. A belief is prima facie defeated, roughly, when the agent has available to her a “good” cognitive process (the kind that would yield prima facie doxastic justification) that would output the content that that belief is false or unwarranted. The proposal bears some obvious similarities to the well-known ARP (which we take to be a selling point), although it fixes some serious problems by trading warrants talk for talk of counterfactual non-belief. At the same time, this shift from reasons to warrants serves to highlight what was right about Pollock’s seminal work on defeat. Although he himself focused on reasons, rather than warrants more generally, there’s nothing stopping the nonevidentialist from taking on board the vast majority of the Pollockian view. Having defeaters is having prima facie propositional justification, and having prima facie propositional justification is having warrants. Having warrants is something that can be understood in whatever general epistemological framework--evidentialist, reliabilist, perhaps even responsibilist--you like.

The recent literature has been dominated by two influential discussions of defeat. One of these has been healthy and productive; the other has clouded some relevant issues and taken as uncontroversially true something that we think no one should even believe, but certainly no one should simply assume: namely, that unjustified beliefs can defeat justified beliefs. The good strand of the standard view of defeat is the evidentialist strand, but it’s not good *because* it’s evidentialist. It’s good because it sees defeat as proceeding (only) from warrants, which is true. Evidentialists make the further claim that only evidence provides warrants. Those of us who deny that can still find a great deal of common ground with the evidentialist theory of defeat.[[45]](#footnote-45)

References

Alston, William. (1988) “An Internalist Externalism” *Synthese* 74: 265-283.

Bergmann, Michael (2006) *Justification without Awareness* (Oxford: Oxford University Press).

Beddor, Bob (2015) “Process Reliabilism’s Troubles with Defeat” *The Philosophical Quarterly* 65: 145-159.

-----. (this volume) “Reasons for Reliabilism”.

BonJour, Laurence. (1985) *The Structure of Empirical Knowledge*. Cambridge: Harvard University Press.

Broome, John (2013) *Rationality Through Reasoning* (Wiley-Blackwell).

Carter, J. Adam 2015. ‘Group Knowledge and Epistemic Defeat’ *Ergo* 2: 711-735.

----- (2018) “Meta-Epistemic Defeat” *Synthese* 195: 2877-2896.

Carter, J. Adam and Jesus Navarro (2017) “The Defeasibility of Knowledge-How” *Philosophy and Phenomenological Research* 95: 662-285.

Chisholm, Roderick. (1977) *Theory of Knowledge*, 2nd Ed. (Englewood Cliffs, NJ: Prentice Hall).

----- (1989) *Theory of Knowledge*, 3rd Ed. (Englewood Cliffs, NJ: Prentice Hall).

Comesaña, Juan (2010) “Evidentialist Reliabilism” *Nous* 44: 571-600.

Dehaene, Stanislas (2014) *Consciousness and the Brain* (New York: Penguin).

Feldman, Richard (1988) “Having Evidence” in *Philosophical Analysis*, D. Austin, ed. (Dordrecht: Kluwer): 83-104.

Feldman, Richard and Earl Conee (1985) “Evidentialism” *Philosophical Studies* 48: 15-34.

Frances, Bryan (2014) *Disagreement* (Cambridge: Polity Press).

Goldberg, Sanford (2018) *To the Best of Our Knowledge* (Oxford: Oxford University Press).

Goldberg, Sanford and Jon Matheson (2018) “The Impossibility of Mere Animal Knowledge for Reflective Subjects” *Erkenntnis*.

Goldman, Alvin (1979) “What is Justified Belief?” In G. S. Pappas, Ed., *Justification and Knowledge* (Dordrecht: Springer): 1-23.

----- (1986) *Epistemology and Cognition* (Cambridge: Harvard University Press).

Green, Christopher (n.d.) “Epistemology of Testimony” Internet Encyclopedia of Philosophy. http://iep.utm.edu/ep-testi/.

Grundmann, Thomas (2011) “Defeasibility Theory” in *The Routledge Companion to Epistemology*, S. Bernecker & D. Pritchard, eds. (New York: Routledge): 156-166.

Kornblith, Hilary (2017) “Doxastic Justification is Fundamental” *Philosophical Topics* 45: 63-80.

Lackey, Jennifer (1999) “Testimonial Knowledge and Transmission” *The Philosophical Quarterly* 49: pp. 471-490

----- (2003) “A Minimal Expression of Non-Reductionism in the Epistemology of Testimony” *Nous* 37: 706-723

----- (2005a) “Testimony and the Infant/Child Objection” *Philosophical Studies* 126: 163-190.

----- (2005b) “Memory as a Generative Epistemic Source” *Philosophy and Phenomenological Research* 70: 636-658.

----- (2006a) “Learning from Words” *Philosophy and Phenomenological Research* 73: 77-101.

----- (2006b) “Knowing from Testimony” *Philosophy Compass* 1/5: 432-448.

----- (2006c) “Introduction” *The Epistemology of Testimony*, J. Lackey & E. Sosa, eds. (Oxford: Oxford University Press): 1-21.

----- (2008) *Learning from Words* (Oxford: Oxford University Press).

----- (2011a) “Testimonial Knowledge” in *The Routledge Companion to Epistemology*, S. Bernecker & D. Pritchard, eds. (New York: Routledge): 316-325.

----- (2011b) “Testimony: Acquiring Knowledge from Others” in *Social Epistemology: Essential Readings*, A. Goldman & D Whitcomb, eds. (Oxford: Oxford University Press): 71-91.

----- (2014a) “Socially Extended Knowledge” *Philosophical Issues* 24: 282-298.

----- (2014b) “A Deflationary Account of Group Testimony” In *Essays in Collective Epistemology*, J. Lackey, ed. (Oxford: Oxford University Press): 64-96.

----- (2018) “Credibility and the Distribution of Epistemic Goods” In *Believing in Accord with the Evidence: New Essays on Evidentialism*, Synthese Library 398, K. McCain, ed. (Cham, Switzerland: Springer): 145-168.

Lasonen-Arnio Maria (forthcoming) “Enkrasia or Evidentialism?” *Philosophical Studies*.

Lyons, Jack (2009) P*erception and Basic Beliefs: Zombies, Modules, and the Problem of the External World* (Oxford: Oxford University Press).

----- (2019) “Algorithm and Parameters: Solving the Generality Problem for Reliabilism” *The Philosophical Review*.

Matheson, Jonathan (2011) “Epistemological Problems Concerning Skeptical Theism” *Faith & Philosophy* 28: 323-

Miller, Emelia (2019) “Liars, Tigers, and the Bearers of Bad News, oh My! Towards a Reasons Account of Defeat” *The Philosophical Quarterly* 69: 82-99.

Miller, George (1956) “The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information.” *Psychological Review* 63: 81-97.

Palermos, Spyridon Orestis (2011) “Dualism in the Epistemology of Testimony and the Ability Intuition” *Philosophia* 39: 597-613.

Pollock, John (1967) “Criteria and our Knowledge of the Material World” *The Philosophical Review* 76: 28-60.

----- (1968) “What is an Epistemological Problem?” *American Philosophical Quarterly* 5: 183-90.

----- (1970) “The Structure of Epistemic Justification” *Studies in the Theory of Knowledge* 4: 62-78.

----- (1971) “Perceptual Knowledge” *Philosophical Review* 80: 287-319.

----- (1974) *Knowledge and Justification* (Princeton: Princeton University Press).

----- (1986) *Contemporary Theories of Knowledge* (Lanham, MD: Rowman and Littlefield).

----- (1987) “Defeasible Reasoning” *Cognitive Science* 11: 481-518.

----- (1994) “Justification and Defeat” *Artificial Intelligence* 67: 377-407.

----- (2001) “Defeasible Reasoning with Variable Degrees of Justification” *Artificial Intelligence* 133: 233-282.

----- (2006) “John Pollock” in *Epistemology: 5 Questions*, V. F. Hendricks & D. Pritchard, eds. (Copenhagen: Automatic Press/VIP).

Pollock John and Joseph Cruz (1999) *Contemporary Theories of Knowledge* (Lanham, MD: Rowman and Littlefield)

Pritchard, Duncan (2018) “Anti-Luck Virtue Epistemology and Epistemic Defeat” *Synthese* 195: 2065-3077.

Tang, Weng Hong (2016). “Reliability Theories of Justified Credence” *Mind* 497: 63-94.

Michael Williams (1999) *Problems of Knowledge*. Oxford: Oxford University Press.

Worsnip, Alex (2018) “The Conflict of Evidence and Coherence” *Philosophy and Phenomenological Research* 96: 3-44.

1. Green n.d.; Grundmann 2011; Palermos 2011; Matheson 2011; Carter 2015, 2018; Carter and Navarro 2017; Goldberg 2018; Goldberg and Matheson 2018; Pritchard 2018. [↑](#footnote-ref-1)
2. Please note that our topic here is justificatory defeat. The term ‘defeater’ is sometimes used in discussions of “defeasibility” responses to the Gettier problem, to refer to an unknown fact that would defeat justification if known (or justifiedly believed). These “knowledge defeaters”—also sometimes called ‘factual’ defeaters—are not our concern here. [↑](#footnote-ref-2)
3. Pollock 1967, 1968, 1970, 1971, 1979, 1986, 1987, 1994, 2001, 2006. [↑](#footnote-ref-3)
4. Pollock and Cruz 1999: 195 [↑](#footnote-ref-4)
5. We ignore here debates about the contents of perceptual states and perceptual beliefs, whether they involve indexicals, singular terms, or existential generalizations. Our use of the demonstrative is for convenience’s sake and isn’t intended to take a stand on this issue. [↑](#footnote-ref-5)
6. Authors often describe the phenomenon of defeat in temporal terms. *First* you have a good reason to believe that p. *Next* you acquire *new* information that defeats that reason. Many of our examples proceed just like this. But defeat is not necessarily a temporal phenomenon. [↑](#footnote-ref-6)
7. Pollock and Cruz 1999: 195. [↑](#footnote-ref-7)
8. Similar formulations occur in Chisholm 1977: 72-73; 1989: 53. [↑](#footnote-ref-8)
9. Pollock and Cruz 1999: 196. [↑](#footnote-ref-9)
10. Pollock and Cruz 1999: 196-7. [↑](#footnote-ref-10)
11. Lackey 2006b: 438; 1999: 483; 2003: 707; 2005a: 164; 2005b: 638; 2008: 44, 253; 2011a: 317; 2014: 75. Lackey sometimes uses the label ‘psychological defeater’ for ‘doxastic defeater.’ She uses ‘psychological’ when she wants to emphasize that in addition to beliefs, some doubts can function the same way as doxastic defeaters, but since they are not narrowly classified as beliefs, we need a broader term. Hence ‘psychological’. Nothing will hang on this point in our chapter.

    What about perceptual experiences, and other non-doxastic states? Though in two places she included experiences along with beliefs and doubts in her list of psychological defeaters (Lackey 2006a: 87; 2006c: 4), otherwise experiences do not appear in her formulations (e.g. Lackey 1999; 2003; 2005a; 2005b; 2008: 44, 157, 204; 2011a: 317; 2014a: 292; 2014b: 75). The more common formulation states her considered view: psychological (doxastic) defeaters for Lackey are necessarily doxastic states (propositional attitudes) like beliefs or doubts, never non-doxastic experiences. Non-doxastic experiences, instead, ground normative defeaters, as we will see. That’s how experiences enter her “ontology of defeat.” [↑](#footnote-ref-11)
12. Lackey 2005b: 647; cf. Lackey 2003: 707; 2005a: 164; 2006a: 87; 2008: 44-45; 2011a: 317. [↑](#footnote-ref-12)
13. Lackey 1999: 475; 2003: 707; 2005b: 638; 2005a; 164; 2006a: 87-88; 2006c: 4; 2008: 45, 157; 2011a: 317. [↑](#footnote-ref-13)
14. Lackey 2006a: 8. [↑](#footnote-ref-14)
15. For this use of the phrase ‘normative defeater’, see Lyons 2009; Grundmann 2011; Goldberg 2018; Pritchard 2018. [↑](#footnote-ref-15)
16. Lackey addresses evidence one does not possess but should in only two of her articles, and there she is using ‘normative defeater’ only to address issues of knowledge, not justification. Lackey 2005b, 2014a. [↑](#footnote-ref-16)
17. A passage in a recent paper (Lackey 2018: 161) might lead one to think she now has this other sense in mind in addition to her sense as defined in the text. She gives an example where she says a subject’s belief is defeated, for the subject lacks evidence that she should have had. She then favorably cites (Goldberg 2017), who had this other sense in mind. But the example makes it clear that this other sense is not what she has in mind. For in her example, the “evidence that the subject should have had” simply consists in a justified belief that the subject should have formed, given other evidence (in this case, another justified belief) that the subject already possesses. So the evidence the subject should have had is just the proposition that the subject should believe, given her other evidence she already possesses. [↑](#footnote-ref-17)
18. Lackey 1999: 487. [↑](#footnote-ref-18)
19. Lackey 2008: 63. [↑](#footnote-ref-19)
20. Lackey 2008: 64. [↑](#footnote-ref-20)
21. Lackey 2005a: 164; 2006a: 87; 2008: 45; 2014a: 292; 2014b: 75. [↑](#footnote-ref-21)
22. Lackey also frequently says, in addition to doxastic and normative defeaters, that there is a third category of so-called “factual” defeaters. Following Lackey, we’ve seen some philosophers to even say of defeaters that there is a “traditional” three-way taxonomy of “psychological” (doxastic), “normative,” and “factual” defeaters. Firstly, this is not the traditional view. Secondly, we deny the existence of so-called “factual” defeaters, for we reject defeasibility analyses of knowledge. [↑](#footnote-ref-22)
23. Lackey 2005b: 647; cf. Lackey 2003: 707; 2005a: 164; 2006a: 87; 2008: 44-45; 2011a: 317. [↑](#footnote-ref-23)
24. Lackey must allow for cases where such defeater-defeaters are lacking, or otherwise the possibility of unjustified defeaters is an idle wheel, as they would always be defeated by a defeater-defeater. [↑](#footnote-ref-24)
25. For example, Goldman 1986 and Bergman 2006, among others. [↑](#footnote-ref-25)
26. It is therefore surprising that an evidentialist like Matheson (2011) cites Lackey approvingly in connection with doxastic defeat. [↑](#footnote-ref-26)
27. Lackey 2005b: 647, emphasis added. See also Lackey 2003: 707; 2005a: 164; 2006a: 87; 2008: 44-45; 2011a: 317; 2014: 75. [↑](#footnote-ref-27)
28. Lackey 2005b: 639, 647; cf. 2005a: 164; 2006c: 4; 2008: 253. [↑](#footnote-ref-28)
29. For example, see Bergmann 2006: 165-167. For some discussion of the issue, see Frances 2014: 79-84. [↑](#footnote-ref-29)
30. See Broome 2013 for similar comments. There is a lively and expanding discussion, partly growing out of Broome’s work, on questions surrounding the so-called “enkratic” principle, and cases where a subject’s total evidence might justify the first-order belief that p and the higher-order belief that one’s total evidence does not justify the belief that p. See for example Worsnip 2018 and Lasonen-Arnio forthcoming, among others. [↑](#footnote-ref-30)
31. See Lackey 1999: 485; emphases added; see also the most recent block quotation above. [↑](#footnote-ref-31)
32. Could Lackey get by with mere permissions, redefining normative defeaters as propositions you are *permitted* to believe (rather than ought to believe), which would be doxastic defeaters if you did? We doubt it. While there’s something plausible about the transfer of obligation (if you’re obligated to A, and doing A would obligate you to B, then you’re obligated to B) this idea loses all plausibility when one of those obligations is replaced by a permission (if you’re *permitted* to A and doing A would obligate you to B, you’re *obligated* to B). [↑](#footnote-ref-32)
33. Lackey 2005a: 171-172. Italics in the original. Bold added for emphasis. Lackey cites BonJour 1985, p. 39. See also Lackey 2005a: 178 for another example. [↑](#footnote-ref-33)
34. It is worth noting that none of the responsibilists Lackey cites--BonJour, Chisholm, McDowell, Williams--ever suggest that the only way to acknowledge one’s counterevidence is to believe the propositions one’s counterevidence justifies believing. As far as we can tell, these philosophers do not acknowledge her category of normative defeaters. [↑](#footnote-ref-34)
35. One could, of course, try to claim that there’s a reliable process of responding to higher order defeaters, or something similar. We don’t think this is at all promising, in part because this is a poor candidate for a cognitive process in the needed sense (see Lyons 2019) but also because even if it were, it would seem to be the kind of belief-dependent process that can only transmit or preserve justification, not generate or create it (see Goldman 1979). [↑](#footnote-ref-35)
36. See for example Alston 1988, Comesaña 2010, Tang 2016, and Miller 2019. [↑](#footnote-ref-36)
37. For the purposes of this paper, we are assuming that all defeaters are rebutting or undercutting, which means that you only get defeat where you have warrants. [↑](#footnote-ref-37)
38. We’re arguing here that reliabilism needs a theory of propositional justification because it needs it for a theory of defeat. This stands in contrast both with the views of Kornblith (2017), who thinks that doxastic justification is the only one of the two that matters, and of Feldman and Conee (1985), who claim to base their theory of doxastic justification on their theory of propositional justification. [↑](#footnote-ref-38)
39. Beddor 2015. See also his paper in this volume, which is similar in important ways to what we’re offering here. [↑](#footnote-ref-39)
40. Goldman 1979, Lyons 2009. To be at all plausible, this needs to be interpreted as claiming that the belief is prima facie defeated, not that it is ultimately defeated. [↑](#footnote-ref-40)
41. Beddor 2015: 150-151. [↑](#footnote-ref-41)
42. Beddor 2015: 155. [↑](#footnote-ref-42)
43. Could a belief-forming process produce as outputs states that aren’t at all belief-like? If an alternate reliable process produced the desire that p, it doesn’t seem like this would be a defeater for the belief that ~p (thanks to Beddor, personal communication). Maybe we need to restrict the outputs to states with assertoric force or which are in some other way sufficiently belief-like. We won’t try to sort out the details here. [↑](#footnote-ref-43)
44. These thoughts seriously complicate the treatment of defeater-defeaters, the full treatment of which we reserve for another occasion. If a regular defeater is a “first-level” defeater, and a defeater-defeater is a “second-level” defeater, and so on, then higher-level defeaters will tend to be less available than lower-level defeaters, due to the cognitive difficulty involved in thinking of them and appreciating their significance, with the result that higher-level defeaters will tend, roughly, to asymptote to irrelevance. This contrasts with the Pollockian view (see Beddor, this volume), according to which higher-level defeaters destroy justification just in case there is an odd, rather than even, number of levels. [↑](#footnote-ref-44)
45. Thanks to Bob Beddor, Adam Carter, Paul Silva, Mona Simion, and especially Jessica Brown for comments on an earlier draft that led to improvements. Thanks also to Jonathan Kvanvig and Duncan Pritchard for helpful conversations on these topics. We also thank audiences at the 2019 Bled Philosophical Conference and at the 2019 Social Epistemology Network Event at Underwood International College in South Korea for beneficial comments and questions. In particular, we recall especially helpful comments from Lizzie Fricker, Chris Kelp, Mark Kaplan, Hilary Kornblith, and Brent Madison.. [↑](#footnote-ref-45)