Defeaters as Indicators of Ignorance

Julien Dutant (King's College London)
Clayton Littlejohn (King's College London)

Abstract: In this paper, we propose a new theory of rationality defeat. We propose that defeaters are indicators of ignorance, evidence that we're not in a position to know some target proposition. When the evidence that we're not in a position to know is sufficiently strong and the probability that we can know is too low, it is not rational to believe. We think that this account retains all the virtues of the more familiar approaches that characterise defeat in terms of its connection to reasons to believe or to confirmation but provides a better approach to higher-order defeat. We also think that a strength of this proposal is that it can be embedded into a larger normative framework. On our account the no-defeater condition is redundant. We can extract our theory of defeat from our theory of what makes it rational to believe—it is rational to believe when it is sufficiently probable that our belief would be knowledge. Thus, our view can provide a monistic account of defeat, one that gives a unifying explanation of the toxicity of different defeaters that is grounded in a framework that either recognises knowledge as the norm of belief or identifies knowledge as the fundamental epistemic good that full belief can realise.

Keywords: Defeat, Rationality, Higher-Order Defeat, Reasons for Belief, Confirmation, Veritism, Gnosticism, Truth Norm, Knowledge Norm, Partial Defeat, Full Defeat, Objectivism, Subjectivism, Epistemic Norms, Epistemic Values.

0. Introduction

Consider three desiderata for a theory of defeat.[1]

1. The theory should be extensionally adequate. We figuratively scoured the globe to give you this list of the different types of defeaters proposed in the literature:

    1. Opposing Defeaters
    2. Undercutting Defeaters
    3. Reason-Defeating Defeaters
    4. Collective Defeaters
    5. Pragmatic Defeaters
    6. Identification Defeaters
    7. Higher-Order Defeaters
    8. Negative Self-Appraisal Defeaters

We can think of the different items as standing for the different mechanisms by which something threatens rationality. A good theory of defeat should tell us which of the things on the list, if any, are genuine defeaters.[2]

[1] For feedback and discussion, we would like to thank Maria Lasonen-Aarnio, Josh Di Paolo, Branden Fitelson, Errol Lord, Mona Simion, and particularly Kevin Dorst. Kevin was kind enough to provide proofs that helped us think through the issues in §4.

[2] In addition, there are discussions about whether defeaters are pieces of evidence, psychological states or contents, things that we should have known about or believed, or facts. As interesting as these discussions are, they don't directly address questions about how defeaters defeat. And that is our focus. For good discussions of normative defeat, see Goldberg (2017), Lackey (2006), and Madison (MS). For discussions of psychological defeat, see Bergmann (2006) and Lackey (2006).

2. Monistic views are *prima facie* preferable to pluralistic views. Monists and pluralists agree that there are different kinds of defeaters, but the monist thinks that there's some unifying explanation of their rational toxicity.[3] Monistic views have greater explanatory power.

3. It is *prima facie* preferable that our theory of defeat can be embedded in a larger normative framework.[4] We want a theory of what makes rational belief rational that explains why defeaters threaten rationality. As Grundmann (2011) observes, many theories of rational belief include non-redundant no-defeat clauses. Such views would fall short of this ideal.

Here is our plan for what follows. In §1, we give a brief presentation of some of Pollock's seminal ideas about defeat and introduce an alternative view that seems to satisfy our second and third desiderata. This account characterises rational belief in terms of confirmation and then characterises defeat in terms of things that have a negative impact on confirmation. We shall argue that this approach cannot handle certain kinds of higher-order defeat and that it mischaracterises some epistemically beneficial factors as defeaters. It doesn't satisfy our first desiderata. In §2, we explain where defeat should be located in a normative theory and argue that the first view ran into trouble because it is a truth-centred view. In §3, we propose a knowledge-centric view as an alternative, one that builds on the strengths of the first truth-centred approach but improves upon it because it handles the cases that cause trouble for the truth-centred approach.

1. Defeaters as De-Confirmers

Recall Pollock and Cruz's characterisation of defeat:

> GENERIC DEFEAT: If p is a reason for S to believe q, r is a defeater for this reason iff (p & r) is not a reason for S to believe q (1999: 37).

They use the notion of a reason in their characterisation of defeat. On their view, p is a reason to believe q iff it is possible for this thinker to rationally believe q by believing it on the basis of p. A reason, so understood, is *defeasibly sufficient* (i.e., something that is sufficient for *ex ante* rationality in the absence of defeaters). They also say that defeaters are themselves reasons. For example, in the case of opposing defeat, they say that when p is a reason for S to believe q, r is a rebutting defeater when r is a reason for S to believe not-q.

We think that this reliance on defeasibly sufficient reasons is problematic.[5] If we wanted to give accounts of *partial* and *full* defeat, we wouldn't want to deploy the concept of a defeasibly sufficient reason in characterising partial defeat. If a thinker acquires some evidence against p, that

---

[3] Recall Spohn's (2002) complaint that some theories (e.g., Pollock's) were normatively deficient because they offered laundry lists of (alleged) defeaters and no unified explanation as to why these different defeat relations were toxic.

[4] See Williams (1965: 107) for an interesting argument that the reasons we have to believe in pluralism in the practical domain do not carry over to the epistemic domain. There might be reasons to prefer a pluralistic view to a monistic one, but the existence of such reasons doesn't change the facts about what's *prima facie* preferable to what.

[5] We should note that much of the literature on the defeat of rationality or justification is concerned with the way in which defeaters defeat status or standing. On Pollock and Cruz's characterisation, however, defeaters defeat reasons. The connection between defeating reasons and status is not obvious without appeal to further assumptions about reasons and rationality. Because one of us doesn't believe that all beliefs need to be supported by reasons to be rational, it's not clear to us how to connect the defeat of reasons to the defeat of status. For arguments that beliefs needn't be based on evidence or reasons to be knowledge or to be rationally held, see Anscombe (1962), Littlejohn (2017), and McGinn (2012). For discussion of a more reason-centric approach, see Lord (2018).

might have some negative impact on the rational status of the thinker's belief even if relevant evidence is not a defeasibly sufficient reason for believing *p*'s negation. Partial defeat doesn't require defeasibly sufficient reasons. Full defeat doesn't require them, either. Suppose that it's rational to believe *p* only if the total evidence provides sufficiently strong support for believing *p*. It's possible for even very weak evidence to bring the degree of support below the threshold. Something can function like a full defeater for a belief in *p* without being a defeasibly sufficient reason to believe ~*p*.

We might try to improve upon this by introducing some other notion of a reason, such as evidential reasons:

> CONFIRMATION: *e* is a reason to believe *h* iff *e* confirms *h*, that is, $P(h|e) > P(h)$.[6]

Evidence is better suited to account for partial defeat. To characterise a full defeater, further normative machinery is needed. We need to introduce a normative standard so that we know when evidence would be sufficient:

> THRESHOLD: It is rational for S believe *q* iff $P(q)$ is sufficiently high.[7]

If we incorporate THRESHOLD into our account of rational belief, we can give an account of partial and full defeat in terms of confirmation. Any defeater (partial or full) has to have a negative impact on *ex ante* or *ex post* rationality. Plausibly, it must be in the confirmation game. We can distinguish between different kinds of defeaters by thinking about the different ways that they negatively impact confirmation. In keeping with THRESHOLD, we can determine whether a partial defeater constitutes a full defeater by thinking about the degree of support that remains after we have taken account of the partial defeater's impact.

We can now characterise defeat as follows:

> DE-CONFIRMATION: defeaters are de-confirmers.[8]

A de-confirmer negatively impacts the confirming support the evidence provides. As with confirmation and disconfirmation, de-confirmation is a matter of degree. A de-confirmer/defeater only has to confirm to some degree to be a partial defeater. A partial defeater constitutes a full defeater when it ensures that the total evidence doesn't provide sufficiently strong support.

DE-CONFIRMATION satisfies our second and third desiderata. The view is monistic. We know what it is that makes a defeater a defeater (i.e., it is a de-confirmer). The view can be embedded in a larger normative framework. Although we can include a no-defeat clause in THRESHOLD, it would be redundant. As an added bonus, the normative framework that supports this view of defeat is independently quite plausible. Most epistemologists currently accept this theory of the epistemic good:

> VERITISM: The fundamental epistemic good is true belief and false belief is the fundamental epistemic evil.[9]

Rational believers have the twin goals of acquiring true beliefs and avoiding false beliefs. They're often uncertain about whether they achieve these goals by forming a belief and so they need to strike a balance between unreasonable aversion to risk and recklessness. They rationally ought to

[6] Let's assume that the relevant notion of probability is some kind of subjective probability or evidential probability that captures the confirmation relation in question. Maria Lasonen-Aarnio rightly pointed out that Pollock's approach is not probabilistic. We are thinking of this as a way of trying to build on some of his ideas in a probabilistic approach.
[7] See Dorst (forthcoming), Easwaran (2016), and Sturgeon (2008) for defences of THRESHOLD.
[8] Chandler (2013) and Kotzen (forthcoming) provide sophisticated formal characterisations of different defeaters in the spirit of DE-CONFIRMATION.
[9] See Goldman (1999), Lynch (2004), and Sylvan (2018) for defences of veritism.

believe in ways that maximise expected epistemic value. This gives us the argument for THRESHOLD.[10]

Not all epistemologists want to characterise rational belief in terms of a connection to epistemic value. Some prefer to characterise rational belief in terms of a connection to epistemic norms. This would be in keeping with a truth-centric approach:

> TRUTH NORMS: We ought to believe p if we can believe it and therein believe the truth. We ought not to believe p if we cannot believe it and therein believe the truth.

If we believed in such objective norms, we wouldn't think that rational belief required conforming to such norms. If we are rationally uncertain, say, about whether, say, this coin will land heads when flipped, there's a sense in which we ought to suspend. One way to characterise this more subjective normative notion is by providing a theory of what would be rational to do give uncertainty about our objective obligations. One approach begins by 'consequentialising' the truth-centric theory of what we objectively ought to believe (e.g., by giving a cardinal ranking of our options that represents the weights of the reasons associated with these norms). We could characterise what we subjectively ought to believe or what we could rationally believe in terms of the minimisation of expected objective wrongfulness.[11] This gives us a new rationale for THRESHOLD. The reason we ought (in a sense) believe in accordance with THRESHOLD is that by so believing we rationally respond to the two normative pressures that we're under.

1.1 Back to the List

It is easy to see how DE-CONFIRMATION accommodates opposing and undercutting defeaters. In cases of opposing defeat, a thinker acquires a defeater that suggests that the thinker's belief is false (e.g., the thinker believes that it is Sunday but sees neighbours rushing to work). In cases of undercutting defeat, a thinker acquires a defeater that suggests that her grounds for believing do not support her belief (e.g., the thinker believes that it rained last night because of the wetness of the pavement but sees that a sprinkler has been running all morning).[12] If P is your initial probability function and you acquire p as a reason to believe q, $P(q|p)$ is higher than $P(q)$. When r is an opposing defeater, it confirms $\sim q$ in the context one's current evidence or beliefs, namely, $P(\sim q|p\&r) > P(\sim q|p)$, so we have $P(q|p\&r) < P(q|p)$. When r is an undercutting defeater, it takes q towards its prior probability $P(q)$ in that context, so we have $P(q) \leq P(q|p\&r)$ and $P(q|p\&r) < P(q|p)$.

In cases of reasons-defeating defeat, a thinker acquires a defeater that suggests that the thinker's supporting reason is false (e.g., you initially believe because of your perceptual experiences that there is beer in the fridge (b) and this convinces you that your partner has been shopping (s) but then you discover later that there was no beer there at all). Here are two ways to think about these cases. First, we might think that b was incorporated into our body of evidence, it supported the belief in s, but then a later a discovery kicks b out of the evidence and thus defeats the thinker's belief in s. The problem is that in the probabilistic setting, once b is incorporated into the evidence and treated as an input to conditionalization, there is no rational process for removing b. Its probability on the evidence is 1, so no new evidence can lower it. If we want to give an account of how s's status is threatened by a reason-defeating defeater and we are supposed to

[10] See Easwaran and Fitelson (2015) and Dorst (forthcoming) for details.

[11] This approach to subjective permissibility is developed in Lazar (forthcoming) and Olsen (2018). Littlejohn (forthcoming b) applies these ideas to the epistemic case by consequentialising knowledge norms and showing how this gives us a theory of what we ought to believe in objective and subjective senses.

[12] Undercutting (or undermining) is sometimes understood very broadly (e.g. by Brown (2018) who treats higher-order defeat as a kind of undercutting or undermining defeat), but we want to understand it narrowly because we think that DE-CONFIRMATION cannot accommodate certain kinds of higher-order defeat.

model this by explaining how learning b boosts s's probability and then s's probability decreases because new evidence knocks b out of the evidence, this cannot be done.

Perhaps those who believe in reason-defeating defeat don't think that our reasons for believing things like s are necessarily part of our evidence. Perhaps some of our reasons for believing things like s are just the contents of justified beliefs. Perhaps b is just such a reason in this sense. In that case, we might imagine that there is some evidence e which supports b and which supports s. (Perhaps e is a proposition about our perceptual experiences, but the precise content doesn't matter provided that e does not entail b or s.) Then we could say that upon learning e, the probability of b and s increase and allow that it's possible that we learn something, d, where d is a reason-defeating defeater because d lowers b's probability without affecting the conditional probabilities of s on b or on $\sim b$.[13]

Consider collective defeat. In cases of collective defeat, a thinker is said to acquire a defeater that suggests that some belief in a set involving multiple beliefs is false without suggesting that any belief in particular is false (e.g., you memorise the phone book and believe each entry in it only to discover the errata slip that says that the book contains an error without providing further information). We are sceptical of some of the things that have been said about the power of collective defeaters. Some epistemologists have defended views on which it's not possible to have a set of justified beliefs known to be inconsistent (e.g., because that this knowledge of the presence of an unidentified falsehood defeats the justification for each belief in the set).[14] We think that this is not what happens in the case we've just described. What *might* happen is that reading the errata slip lowers our confidence in some of the things we believe on the basis of what we read, but it's clear that it wouldn't bring the degree of support for each belief below threshold. One virtue of DE-CONFIRMATION is that it seems to explain why so-called collective defeaters *might* diminish the strength of support without being full defeaters for the relevant set of beliefs.[15]

One further virtue of DE-CONFIRMATION is that it can shed light on principles of justification associated with collective defeat, such as those that impose a consistency requirement on justified belief. Suppose that you were .95 confident that each of the entries in the phone book that you had committed to memory were correct. And suppose you learn that there was precisely one error in this book. Under these conditions, it seems that you acquire a reason to *increase* your confidence in your beliefs. While fans of collective defeat might think that this piece of information is rationally toxic, it might seem to actually be either neutral or beneficial.

DE-CONFIRMATION should have no difficulty accommodating pragmatic defeaters if this is desired.[16] The way that pragmatic defeaters are usually understood, they are taken to be practical considerations that are not directly connected to the truth-indicative factors. They defeat by giving us reason not to believe without stronger evidence than might otherwise be required. If readers believe that practical considerations (e.g., the stakes or the odds of a salient bet) can create rational

---

[13] Let P be your probability after e is acquired. For b to be a reason to believe s it must be that $P(s|b) > P(s|\sim b)$. Recall d doesn't affect the conditional probabilities $P(s|b)$ and $P(s|\sim b)$, so $P(s|b \& d) = P(s|b)$ and $P(s|\sim b \& d) = P(s|\sim b)$. Finally, d lowers the probability of b, so $P(b|d) < P(b)$. From this, it follows that d de-confirms s, namely, $P(s|d) < P(s)$.

[14] See Leitgeb (2017), Pollock (1986), Smith (2016), and Ryan (1991, 1996) for defences of views that include this consistency requirement on justified belief.

[15] See Christensen (2004), Easwaran and Fitelson (2015), Makinson (1965), and Worsnip (2016) for arguments that collective defeaters don't invariably defeat the rational status of our beliefs. In Dutant and Littlejohn (forthcoming), we argue that views that posit collective defeaters are either overly sceptical or too externalist (i.e., they deny us too much justification or they pick winners and losers on the basis of considerations beyond the thinker's ken).

[16] For arguments that pragmatic considerations can help to determine whether a body of evidence provides sufficient support for belief, see Fantl and McGrath (2004), Owens (2000), and Schroeder (2012). We think Schroeder is the first to describe these pragmatic considerations as defeaters.

pressure to suspend without acquiring more evidence than might normally be necessary, DE-CONFIRMATION can accommodate this kind of pragmatic defeater by positing that such defeaters move the threshold.[17]

While DE-CONFIRMATION has a number of virtues, we struggle to see how it handles further kinds of defeat that are discussed in the literature. Peacocke, believes there are identification defeaters:

> [I] distinguish two kinds of defeasibility, which I will call *defeasibility of identification* and *defeasibility of grounds*. A ground for accepting a proposition can be conclusive even though our entitlement to believe that we have identified such a ground is defeasible. Identifying something as a conclusive ground is one thing; its being a conclusive ground is another. My confidence that something is a proof can be rationally undermined by the report of mathematicians whose competence I have reason to believe far outstrips my own. Nonetheless, a proof is a conclusive ground. Here we have defeasibility in respect of identification but not defeasibility in respect of grounds (2004: 30).

At first glance, this might look like a case of undercutting defeat. The expert tells you that the proof is flawed so you might agree that the putative support doesn't actually support the conclusion. If this is a case of undercutting defeat, DE-CONFIRMATION handles it straightforwardly. However, Peacocke intends this case to be one in which the defeater fails to de-confirm one's belief. The idea is that if we have maximal support for the proposition. So, if we have a defeater, it is not a de-confirmer.

We can see why some might be cool to the idea of so-called identification defeaters. It doesn't seem that we need to be able to identify conclusive grounds as such to be justified on their basis. If so, it's not obvious that the considerations that show that we cannot know that some ground is conclusive are enough to make it irrational to hold the relevant belief. Still, scale and degree seem to matter. If some considerations show that it's insufficiently likely that you have suitable means for settling some question (e.g., because an expert has shown you that you've made a large number of mistakes in your reasoning), perhaps these considerations can defeat. If so, we might assimilate them to the category of higher-order defeaters.

The defeaters described as higher-order defeaters cover a wide variety of considerations. Some consist of evidence about our evidence (e.g., evidence that our evidence provides only weak support for our beliefs). Some consist of evidence that our rational faculties have been compromised (e.g., evidence that we have been slipped a drug that affects our reasoning, that we suffer from hypoxia, etc.). Some might consist of evidence that our beliefs would not attain some desirable status (e.g., evidence that we could not know, evidence that it would be irrational to believe something, etc.). Some of these cases might be assimilated to cases of opposing, undercutting, or reason-defeating defeat, but it doesn't seem that all can be. Suppose you acquire some misleading evidence that you have been slipped a drug that interferes with your reasoning abilities. Suppose there were a simple test that could determine whether the drug was present in the bloodstream. You simply need to place a strip of paper on your tongue and check its colour. And suppose that you acquired this evidence that you were slipped the drug during an important exam. If the test costs nothing to administer, it seems to make sense to take it before submitting your exam. Perhaps this is some indication that that the evidence that you've been drugged defeats the justification you have for the beliefs that you would rely on in answering exam questions. Note

[17] For discussion of such cases, see Fantl and McGrath (2009), Owens (2000) and Weatherson (2005). For discussion of the ways in which moral considerations might provide pragmatic defeaters, see Basu and Schroeder (2019) and Moss (2018).

that if the evidence is indeed misleading and you haven't been drugged, it isn't clear that this evidence defeated by being an opposing, undercutting, or reason-defeating defeater.

We can anticipate three responses from fans of DE-CONFIRMATION. First, they might deny the possibility of the case as we described it. They might argue that there are level-connecting principles according to which higher-order evidence necessarily is connected to the first-order evidence that either confirms or disconfirms our first-order beliefs.[18] We will explain below why we doubt that there is any true level-connecting principle that will cover the full range of cases.

Second, someone might simply bite the bullet and deny that the putative higher-order defeaters defeat. This would allow someone to say that all genuine defeaters are de-confirmers. This view seems to be a kind of level-splitting view insofar as it denies that higher-order considerations bear on the justification or rationality of first-order beliefs. We worry that such views suggest that certain apparently dogmatic or seemingly reckless attitudes are rational. Moreover, we think that the account will struggle with a kind of higher-order defeat that we call negative self-appraisal defeat.

The third response is to our mind the most promising. The de-confirmers might try to account for higher-order defeat by suggesting that it threatens *ex post* rationality rather than *ex ante* rationality.[19] The standard view is that *ex post* rationality involves a basing requirement (i.e., that we competently base our beliefs on the support that warrants them). We might account for higher-order defeat by saying that it impacts *ex post* rationality on the grounds that anyone who believes in spite of the presence of these defeaters cannot meet the basing condition. Perhaps if Agnes believes (falsely but reasonably) that she has been drugged, it is *ex ante* rational to believe the conclusion of a proof but it fails to be *ex post* rational because she cannot competently base her beliefs on the good reasons that she has.

This strategy might help proponents of DE-CONFIRMATION accommodate some higher-order defeat, but we don't think this covers negative self-appraisal defeat. A negative self-appraisal defeater is provided by considerations about the epistemic status of some particular belief. Agnes might believe something about, say, the afterlife. She might then be convinced that she couldn't possibly know anything about the afterlife. There seems to be a rational tension between the belief that we'll survive our own bodily deaths and the belief that we couldn't possibly know whether people can survive their own bodily deaths.

Negative self-appraisal defeaters differ from the previous higher-order defeaters in two ways. First, they concern the epistemic status of particular beliefs (e.g., whether we know, whether we rationally believe, whether it is okay to believe given the evidence we have, etc.) rather than pertaining directly to the thinker's belief forming processes or strength of evidence. Second, they have a kind of epistemic content missing from more widely discussed cases of higher-order defeat.

How can DE-CONFIRMATION handle cases like these? Suppose that a thinker believes *p* and acquires evidence that confirms any of the following:

1. that she couldn't know *p*.
2. that it's irrational for her to believe *p*.
3. that she doesn't have evidence that warrants belief in *p*.

[18] See Dorst (forthcoming b), Dorst, Fitelson, and Husic (forthcoming), and Williamson (forthcoming) for discussion of level-connecting principles.

[19] This idea was inspired by Di Paolo (2018) and Silva (2017). They do not endorse DE-CONFIRMATION, but we thought this work might help fend off one kind of objection. See also van Wietmarschen (2013).

On its face, it seems that the evidence may raise the probability of (1)-(3) without lowering that of *p* itself.[20] So, if negative self-appraisal defeaters defeat, they might show that there are defeaters that are not de-confirmers.[21]

Think about lottery cases. Harman nicely captures part of what makes the lottery case so puzzling:

> In the testimony case a person comes to know something when he is told about it by an eyewitness or when he reads about it in the newspaper. In the lottery case, a person fails to come to know he will lose a fair lottery, even though he reasons as follows: "Since there are N tickets, the probability of losing is (N-1)/N. This probability is very close to one. Therefore, I shall lose the lottery." A person can know in the testimony case but not in the lottery case … The contrast between the two cases may seem paradoxical, since witnesses are sometimes mistaken and newspapers often print things that are false. For some N, the likelihood that a person will lose the lottery is higher than the likelihood that the witness has told the truth or that the newspaper is right (1968: 166).

Given the weak anti-sceptical assumption that our testimonial beliefs are rational and Harman's observation about the chance of error, THRESHOLD tells us that it is rational to believe lottery propositions. Suppose this is so and suppose that we know *a priori* that we cannot know that we've lost the lottery. If so, some Moorean absurdities cross the threshold. DE-CONFIRMATION cannot account for some negative self-appraisal defeaters.

Could the proponents of DE-CONFIRMATION maintain that negative self-appraisal defeaters defeat because they threaten *ex post* rationality by ensuring that we don't satisfy a basing requirement? Probably not. We fail to satisfy the basing requirement when we cannot competently base a belief on sufficient reasons. In many cases of negative self-appraisal defeat, we have no reason to think that we couldn't meet this condition if THRESHOLD were correct. We are, after all, perfectly capable of competently judging that Arsenal won (as reported in the paper) and competently judging that it's more likely that the ticket lost than it is that Arsenal won. If so and if THRESHOLD were correct, we should be able to competently base the belief that some ticket lost on adequate reasons even while knowing that the belief couldn't constitute knowledge.

---

[20] See Dorst (forthcoming b) and Williamson (forthcoming) for arguments that show that once we allow that we can be rationally uncertain to even a slight degree about what our evidence supports, we should expect that there can be bodies of evidence that both make it very probable that *p* and very probable that it is not sufficiently probable that *p*. Given THRESHOLD, this would be evidence that would make it rational to believe *p* and rational to believe that it is irrational to believe *p*. In truth-centric accounts, then, it seems we can only account for the toxicity of (2) by denying THRESHOLD or by accepting a kind of negative introspection principle that says that where *t* is the relevant threshold for rational belief (which we assume is less than 1), $P(p)<t$ only if $P(P(p)<t) =1$.

[21] Bergmann (2006) claims that some beliefs provide negative self-appraisal defeaters even if those beliefs are not adequately supported. We disagree. We think that he is making the kind of mistake that people make when they conflate a wide- and narrow-scope requirement. We would agree that a thinker should not believe *p* if she believes that she does not know *p* or if she believes that it's irrational for her to believe *p*, but we don't think we can detach any interesting normative claim from this even if we add the further (psychological) claim that the thinker does believe (1), (2), or (3).

We can use lottery cases to identify a second problem with DE-CONFIRMATION.[22] Suppose Agnes is thinking about two suspects in two different crimes, Inge and Agatha. She initially thinks that the evidence of Inge's guilt consists entirely of statistical evidence. She initially thinks that the evidence of Agatha's guilt consists entirely of eyewitness testimony. She then realises that she had the files mixed up. In such a situation, Agnes could acquire evidence that concerns Inge's guilt that both lowers the probability of guilt and makes it rational for her to believe something that she couldn't have rationally believed before. According to DE-CONFIRMATION, this should be a case of undercutting defeat,[23] but it's a case in which the old evidence doesn't make a belief rational and the new evidence that provides less by way of confirmation makes the belief rational for the first time. We don't think this could be a case of undercutting defeat. Meanwhile, we find that the evidence that boosts the probability that Agatha is guilty makes it irrational to believe in her guilt. So, this seems to be a kind of defeater that defeats while confirming, not one that defeats by de-confirming.

2. Rationality, Defeat, and Epistemic Norms

Let's step back and consider some foundational issues. We have already seen two different rationales for THRESHOLD. If we accept a truth-centred approach to epistemic matters and either accept TRUTH NORMS or VERITISM, there is an argument to be made that rational thinkers ought to conform to THRESHOLD because this represents the right way of handling certain risks (e.g., the risk of believing in ways that are objectively bad or failing to believe in ways that are objectively good or that of violating some objective norm). If THRESHOLD were correct, we could easily see how to give a monistic account of defeat and explain why some familiar kinds of defeaters are rationally toxic without having to add a clause to our theory of rational belief to handle defeat. That would mean that DE-CONFIRMATION satisfies our second and third desiderata, but doubt that it satisfies the first, extensional adequacy. Adding a clause to handle tricky cases (e.g., higher-order defeat, puzzling intuitions about undermining defeat in lottery cases) might help with extensional adequacy, but then the modified view would no longer satisfy our second and third desiderata.

Many epistemologists have raised concerns about THRESHOLD along the lines of those discussed above. Smith (2016), for example, thinks that the problem has to do with this idea that justification is a matter of managing risk and so he rejects the risk-minimisation approaches to justification (of which THRESHOLD is an instance) in favour of an alternative view designed to allow that we might be justified in believing in one case and unjustified in believing in another even if the risk of forming a false belief is greater in the second kind of case than the first. The view that we shall propose has a similar feature, but we don't want to be too hasty in rejecting the idea that much of rationality is a matter of properly handling risk.

Before we reject the very idea that there is some connection between rational belief and the management of risk, we should remember that there are many different kinds of risk that we might need to manage. If a truth-centred approach were correct, we should worry about forming false beliefs and failing to form true ones. If we were veritists, we would describe this as a concern about managing the risks of forming bad beliefs and failing to form good ones. If we believed in epistemic norms like truth norm, we would describe this a concern about managing the risk of

[22] See Smith (2016) and Weatherson (2016) for defence of the idea that non-sceptical accounts of knowledge have to allow that some of the things we know have lower evidential probability than some true propositions that we're not in a position to know. Their defence uses lottery cases.

[23] To see this, break down Agnes's evidence into two components: that the file she thought to be Inge's isn't, and that the file she thought to be Agatha's is Inge's. The first is a pure undercutting defeater of Agnes's initial reason to believe in Inge's guilt (namely, the contents of the file she mistakenly thought was hers). The second is a confirmer of Inge's guilt, but not as strong a confirmer as the previous was. Together, they are a mild undercutting defeater, on the de-confirmation account.

norm violation. It seems strange to us to suggest that rationality shouldn't be concerned with the management of these kinds of risk. On its face, the idea that the rational person's responses are rational precisely because they manage these kinds of risks well seems pretty plausible. Perhaps the lesson to draw from the difficulties that arise for THRESHOLD is *not* that we were mistaken to think that rationality has a great deal to do with the management of risk but that we were mistaken about the kinds of risks that need to be managed.

We're not all veritists. And we don't all believe that the fundamental epistemic norms have to do with truth. Some of us prefer gnosticism to veritism and prefer to think of epistemic norms as concerned with knowledge and not merely with truth:

> GNOSTICISM: The fundamental epistemic good is knowledge and the fundamental epistemic evil is a belief that fails to be knowledge.[24]
>
> KNOWLEDGE NORMS: We ought to believe *p* if we can believe it and therein come to know. We ought not to believe *p* if we would not believe it and therein come to know.

If we take either of these proposals as our starting point, the same arguments that took us from veritism or truth norms to threshold take us from these knowledge-centred claims about knowledge's normative role to this account of rational belief:

> PROBABLE KNOWLEDGE: It is rational for S to believe *q* iff the probability that S is in a position to know *q* is sufficiently high.[25]

We can see PROBABLE KNOWLEDGE as representing a standard that, like THRESHOLD, tells us how to properly manage the risk of failing to promote the epistemic good or of violating certain objective epistemic norms.[26]

As with THRESHOLD, we can easily extract an account of defeat from PROBABLE KNOWLEDGE.[27] As with de-confirmation, this account should satisfy our second and third desiderata. On the view that we propose, something constitutes a *partial* defeater insofar as it suggests that there is some risk of believing without knowing and a *full* defeater if it ensures that the risk is too great for us to believe given either a proper concern for the epistemic good or given what it would take to responsibly try to meet our objective epistemic obligations. Smith might

---

[24] See Hyman (1999), Littlejohn (2018), Williamson (2000), and Unger (1975) on the value of knowledge for motivations for GNOSTICISM.

[25] See Dutant and Fitelson (MS) and Littlejohn (forthcoming) for discussion and defence of the idea that rationality is understood in terms of the probability of knowledge. They are primarily concerned with synchronic wide-scope requirements. We said earlier that we think of probability here as evidential probability. Some might worry that if we combine this approach with the idea that everything known is part of our evidence we will narrow the scope of the applicability of the account. (For example, if E=K, our account says that in cases where you know that you know *p*, the probability that you know *p* is 1. And this would mean that we couldn't give an account of how belief in *p* could be defeated.) We aren't certain that the narrow scope is a problem and we are not taking any official stand on the relationship between evidence and knowledge. In particular, we think that there are interesting independent reasons to look for an account of evidential probability on which it's possible to know *p* even if the probability of *p* is less than 1. See Bacon (2014) and Brown (2018) for discussion of such reasons.

[26] In this way, we think we can follow Ghijsen, Kelp, and Simion (2016) in giving an account of the subjective and objective dimensions of epistemic assessment without rejecting the idea that the knowledge norm governs belief. Graham (2010) proposes a similar approach to reconciling an objectivist approach to obligation with a more subjectivist account of rationality.

[27] We think a probabilistic approach to rationality that links rationality to the risk of norm violation will help objectivists address some of Brown's (2018) objections to Littlejohn and Williamson's proposals about excuses.

have been right that rational belief isn't just a matter of believing in ways minimise the risk of error, but there are alternative risk minimisation approaches that we should consider. We have just offered one.

4. Defeat as Indication of Ignorance

If it is rational to believe in ways that maximise expected epistemic value, the view that we get if we assume GNOSTICISM is PROBABLE KNOWLEDGE. If our monistic theory of defeat ought to be embedded in some larger normative framework, then we propose embedding it in a framework that incorporates PROBABLE KNOWLEDGE as the standard for rational belief, a standard that tells us how rational thinkers go about trying to deal with their uncertainty about what it takes to meet the fundamental standards, KNOWLEDGE NORMS. If rational belief is what PROBABLE KNOWLEDGE says, we can characterise defeat as follows:

> INDICATION OF IGNORANCE: defeaters are indicators of ignorance (i.e., considerations that constitute evidence that the thinker is not in a position to know the target proposition).[28]

INDICATION OF IGNORANCE can handle opposing defeat, undercutting defeat, reason-defeating defeat, normative defeat, collective defeat, and pragmatic defeat in roughly the same way that DE-CONFIRMATION did. Being in a position to know is factive, but knowledge requires more than truth. For this reason, it will typically be that $P(Kp)<P(p)$. Things that de-confirm will often serve as an indicator of ignorance, so our account can appropriate much of what DE-CONFIRMATION can offer.

One interesting difference between DE-CONFIRMATION and INDICATION OF IGNORANCE emerges when we think about lottery-type cases. Recall Agatha, Agnes, and Inge. Once Agnes could tell which belief was based on statistical evidence and which belief was based on observation, we saw that the probability for one belief dropped and the belief that wasn't rational to hold was rational to hold. PROBABLE KNOWLEDGE and INDICATION OF IGNORANCE explain how this is possible: when the probability of knowing increases even though the probability of truth decreases, a belief can become rational despite becoming less likely to be true.

INDICATION OF IGNORANCE provides a more straightforward account of higher-order defeat. It's hard to see how indications that the oxygen is low, that the drinks have been drugged, etc. constitute de-confirmers, but it's not at all hard to see how they could constitute evidence that the subject is not in a position to know. Since there seems to be little interesting connection between, say, something having maximal evidential probability and being something that the thinker is in a position to know, we don't have to try to explain how a thinker who has entailing evidence at one moment can fail to be rational in her beliefs at a later moment when she is made aware of some higher-order defeater.

INDICATION OF IGNORANCE sheds new light on collective defeat. Whereas the probability that we know in a lottery case is 0, the probability that we know that some entry in a book known to contain an error might be quite high. Optimistically, we think that in a book that is suitably large, if we knew that the book contained a single error, we could know each entry of the book to be true save the one that was false. In such a case, it might be that the probability that each claim in the book is known is quite high (i.e., if the book contains n claims and n is a suitably large number, the probability that any particular claim in the book can be known might be 1/n-1). Thus,

[28] We hadn't realised that Gibbons (2013) had proposed that defeaters were evidence that we aren't in a position to know. Thanks to Errol Lord for pointing this out. One difference between his approach and ours is that we introduced a rational standard that should help us decide whether some defeater is full or merely partial. We also disagree about lottery cases, so he wouldn't accept our points about lotteries and undercutting defeat. Although we put Di Paolo's (2018) work to use earlier in defence of DE-CONFIRMATION, he said that he was sympathetic to a knowledge-centred view.

we get a view on which it would be possible to believe each claim in the book whilst knowing that one of the claims in the book is false. Our view differs from truth-centred views in that they tend to treat lottery and preface cases in the same way. In our view, such views deliver the wrong result in lottery cases or in preface cases. The difference in the prospect of coming to know in such cases is the key to understanding why we're tempted to think of them as differing in a way that should matter for our theory of rationality.

INDICATION OF IGNORANCE also provides a better treatment of negative self-appraisal defeat. First, if the conjunction of some first-order belief and the negative self-appraisal is not knowable (e.g., because the negative self-appraisal represents the first-order belief as failing to fulfil some necessary condition on knowledge), PROBABLE KNOWLEDGE says that the conjunction cannot be rationally believed because the probability that the conjunction is known is 0.

What about the conjuncts, though? It's one thing to say that nobody can rationally believe certain conjunctions that we know cannot be known (e.g., 'Dogs bark but I don't know that they do') and another to say that nobody can rationally believe each conjunct. Consider some problematic pairs of propositions:

1. *p*, I couldn't know whether *p*.
2. *p*, It is irrational for me to believe *p*.
3. *p*, My evidence isn't sufficient to believe *p*.

Can a thinker be rational in believing both propositions in (1)-(3)? INDICATION OF IGNORANCE tells us that it can only be *ex ante* rational to believe both if it can be rational to be sufficiently confident of each proposition that it is something we can know.

Let's start with (1). While PROBABLE KNOWLEDGE does not tell us how probable it must be that *p* is something you're in a position to know for it to be rational to believe *p*, it's natural to suppose that it has to be at least more probable than not (i.e., if it's *ex ante* rational to believe *p*, $P(Kp) > .5$). If so, it cannot also be rational to believe that you cannot know whether *p*. If it were rational to believe $\sim Kp$, $P(K\sim Kp) > .5$. Since being in a position to know is factive, $P(K\sim Kp) > .5$ only if $P(\sim Kp) > .5$. But it cannot be that $P(\sim Kp) > .5$ if $P(Kp) > .5$. Thus, PROBABLE KNOWLEDGE tells us that it cannot be *ex ante* rational for a thinker to believe both propositions in (1). If we understand 'having sufficient evidence' as meeting a condition necessary for knowing, then this reasoning rules out evidence that makes it *ex ante* rational to believe the propositions mentioned in (3).

The difficult case is (2). If *ex ante* rationality were necessary for being in a position to know, our treatment of (1) would cover this case.[29] Unfortunately for us, we cannot say that rationality is necessary for knowledge. If we assume probable knowledge, we should expect that there will be some 'unreasonable knowledge' in the sense that there will be some things that we know that will not be sufficiently likely to be known to also be rational to believe.[30] Still, while we might want to allow for the possibility of unreasonable knowledge, we might want to deny that it can be rational to both believe *p* and believe that this belief is irrational.

Whether PROBABLE KNOWLEDGE prohibits a pair of beliefs like (2) depends upon whether a high probability of knowledge is compatible with a high probability (that one knows) that knowledge is improbable. If we take $t > \frac{1}{2}$ to be the threshold of sufficient probability in PROBABLE KNOWLEDGE, a sufficient condition to prohibit the pair is:

---

[29] Let t be the threshold in PROBABLE KNOWLEDGE. Suppose knowledge entails rational belief: by PROBABLE KNOWLEDGE, $P(Kp) \leq t$ entails $\sim Kp$. Now suppose it is rational to believe that it is not rational to believe *p*. By PROBABLE KNOWLEDGE, $P(P(Kp) \leq t) > t$. Since $P(Kp) \leq t$ entails $\sim Kp$, $P(\sim Kp) > t$. Assuming $t \geq \frac{1}{2}$, $P(Kp) < t$.

[30] See Lasonen-Aarnio (2010) for discussion and defence of unreasonable knowledge and Sylvan (2018b) for defence of the idea that knowledge is a non-normative state that needn't have a normative dimension. We think that a non-normative account of knowledge can help address some concerns about the very idea of defeat raised by Baker-Hytch and Benton (2015).

GRADUAL t-ACCESS: If $P(Kp)>t$ then $P(P(Kp)>t)\geq 1-t$.

For note that $P(Kp)>t$ obtains just if $P(Kp)\leq t$ doesn't. Hence $P(P(Kp)>t)\geq 1-t$ is equivalent to $P(P(Kp) \leq t) \leq t$. Given the factivity of being in a position to know, $P(P(Kp)\leq t)\leq t$ only if $P(KP(Kp) \leq t) \leq t$. Hence given GRADUAL t-ACCESS, if $P(Kp)>t$ then $P(KP(Kp)\leq t)\leq t$. This would mean that if the probability that one knows p is above threshold, the probability that one knows that it is *not* above threshold is itself below threshold, as desired.

Note that while GRADUAL t-ACCESS is an inter-level constraint, it does not force us to accept strong access constraints like the idea that it is rational to believe p only if it is rational to believe that one knows that p. It still allows the combination of $P(Kp)>t$ and $P(P(Kp)>t)\leq t$ and the combination of Kp and $P(Kp)\leq ½$.[31] Note also that the higher t is, the milder the GRADUAL t-ACCESS requirement is. At t=.9, for instance, it merely requires that if it is highly (.9) probable that one knows, it is at least somewhat (i.e., .1) likely that it is highly probable.

While we don't have a full defence of GRADUAL t-ACCESS, we think that it has some *prima facie* appeal and is worth exploring. Its initial appeal comes from the idea that whatever makes it probable that one knows will typically also make it probable that it is probable that one knows. What confirms to you that you know p by memory or perception, for instance, is typically the fact that your perception or memory is particularly clear or detailed, the fact that you can give additional details pertaining to p, and so on. But those facts are also evidence that you have evidence that you know, which is to say that they contribute to make it probable that it is probable for you that you know.

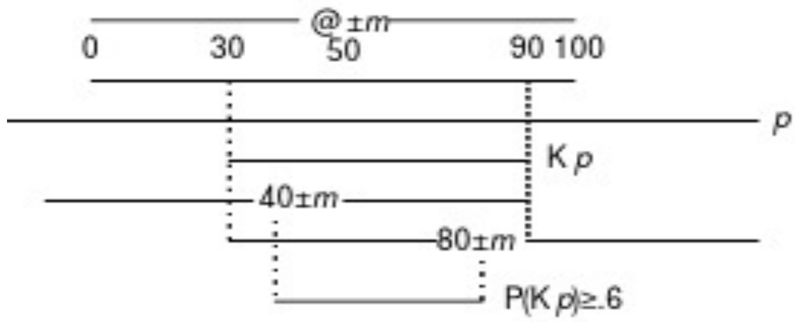
The thought gets further support from an unlikely source: Williamson's (2014) unmarked clock models. While these models are meant to illustrate the failure of various access constraints on knowledge, they turn out to validate GRADUAL t-ACCESS. We can illustrate how this works with an example and we provide a proof in the appendix. In Williamson's models, a parameter (*e.g.* time, temperature, etc.) takes different values along a line. Worlds correspond to points on the line. What you know at a world is whatever holds at worlds within a margin of error m to the left and to the right of the point on the line. Your evidential probability is given by conditionalising a uniform prior on what you know. When you update in this way, every value outside of the margins is ruled out and each of the values within the margins are equally likely. Suppose the actual world (@) is at 50 and the margin for error (m) is 50. What you know leaves open any world within [0:100]. Let p be some proposition such that Kp holds throughout [30:90]. (For instance, p might be the proposition that the value of the parameter is between -20 and 140.) Because Kp occupies 60% of the worlds left open by what you know at @, the probability of Kp at @ is .6. Note that any world between 40 and 80 will also have within its margin the Kp-segment [30:90]: for 40 leaves open any world in [-10:90] and 80 leaves open any world in [30:130]. So, at each of these worlds, $P(Kp)\geq .6$, too. This means that at least 40% of worlds within @'s margins are Kp-worlds. So, at @, we have an instance of GRADUAL .6-ACCESS, $P(P(Kp)\geq .6)\geq .4$. The reasoning generalizes to any Kp-proposition that occupies a continuous stretch of 50% of @'s margins.[32]

---

[31] But see Goodman (2013) for helpful discussion of such constraints.

[32] Note that propositions of the form Kp *have to* occupy continuous stretches within a world's margin. For if Kp holds at two points within @'s margins, those two points are at most $2m$ apart, which means that their own margins connect and p holds throughout the segment between them, which in turns means that Kp holds at any point between them.

What happens for t below ½? Suppose for instance that Kp holds throughout [80:100]. Then it occupies 20% of @'s margins, so at @ we have $P(Kp) \geq .2$. Now all we can say is that every world *to the right* of @ 'sees' the whole of [80:100]. So we have $P(P(Kp) \geq .2) \geq .5$. More generally, the models validate an alternative schema: for any $t\leq ½$, if $P(Kp)>t$, $(P(Kp)>t)\geq .5$. This is good news as instances of GRADUAL t-ACCESS with low t are counterintuitive.



Informally, here is what we have shown. In these models, the features of a situation that make K*p* probable (i.e., that K*p* holds in a large proportion of the worlds left open by one's evidence) are also features that put an upper limit on how improbable it can be that K*p* is probable (i.e., a sufficient proportion of the worlds that are left open by one's evidence are worlds that can see mostly K*p*-worlds). This is why they validate GRADUAL T-ACCESS. This result depends upon several features of the models that we've worked with, so it remains to be seen how principles like GRADUAL T-ACCESS would fare if we worked with different models (e.g., ones that do not assume a uniform prior, ones that do not have a constant margin, etc.). At the very least, we think this gives us some reason to explore principles like GRADUAL T-ACCESS further and reason to hope that such principles will hold in a wide range of cases.

If it is impossible to rationally believe the propositions mentioned in (1)-(3), our account handles cases of negative self-appraisal defeat quite nicely. When it is rational for the thinker to believe, say, that she's not in a position to know or that it would be irrational for her to believe *p*, it would not be rational for her to believe *p* and we can see why these negative self-appraisals can function like defeaters. Alternative views that characterise rational belief in terms of the probability of truth struggle to explain why it would be rationally mandatory for a thinker who rationally believes that she doesn't know, say, to suspend judgment.

If our treatment of negative self-appraisal is correct, we can build on an interesting proposal about suspension defended by Thomas Raleigh (forthcoming). He proposed that suspending is really a matter of believing. We suspend, on his view, when we believe the evidence doesn't support a belief. We think that something very similar might be right. Provided that a thinker is rational, her rational belief that she's not in a position to know whether *p* should ensure that she's agnostic about whether *p*. (Even without a belief, a sufficiently high degree of confidence should also work.) It is difficult to explain why this should be in a truth-centric framework since, as we've seen, it's not clear why, if THRESHOLD is correct, a rational thinker would suspend if she were to judge that she's not in a position to know or that her evidence doesn't support her belief. Our framework, however, explains why a rational thinker suspends when she rationally believes that she's not in a position to know. We think that it's a strength of our view that it helps us see what's attractive about Raleigh's proposals about belief and suspension.

5. Conclusion

We have proposed a theory of rational full belief according to which it's rational to believe when it's sufficiently probable that the belief would be knowledge. From this account, we can derive a theory of defeat according to which defeaters are indicators of ignorance, evidence that if you were to believe, your belief would fail to constitute knowledge. We think that this account provides a better account of every kind of (genuine) defeat there is than the leading alternative. If we think of rationality as something akin to subjective rightness and think of subjective rightness as being probabilistically related to objective norms, our account of rational belief and defeat is the account we should want if we think of knowledge as the fundamental normative standard for belief.

Appendix

We show that GRADUAL t-ACCESS (for $t \geq .5$) holds in Williamson's (2014) unmarked clock models. Worlds are points in the real number line, propositions are regions of the line; $|p|$ is the size of the region p; for every p, q, the proposition $p\&q$ is the intersection of p and q. At each world w one knows every proposition entailed by the parameter's value being within a margin $m>0$ of w. We let $k(w) = [w-m:w+m]$ and for any p, the proposition that one knows p, Kp, is the set $\{w:k(w)\subseteq p\}$. The evidential probability at w, π_w, is a uniform prior conditionalised on $k(w)$: $\pi_w(p)=|p\&k(w)|/|k(w)|$. Without loss of symmetry we set $m=½$ which gives us $|k(w)|=1$ and the simple relation: $\pi_w(p)=|p|$. The proposition that the evidential probability of p is above t, P(p)>t, is the set $\{w: \pi_w(p)>t\}$. Now let $t\geq½$, @ some world and p such that $\pi_@(\mathrm{K}p)>t$.

Our aim is to show that $\pi_@(\mathrm{P}(\mathrm{K}p)>t)\geq 1-t$. Let a be greatest lower bound and b the least upper bound of Kp&k(@). Since Kp holds at worlds (arbitrary close to) a and b, p holds throughout the regions $]a-m:a+m]$ and $[b-m:b+m[$. Since a and b are within $k(@)$, $|b-a|<2m$ and the regions overlap and p holds throughout $]a-m:b+m[$. So Kp&k(@) holds throughout $]a:b[$, and since $\pi_@(\mathrm{K}p)>t$, $|]a:b[|>t$. Now let w be any world within $]a+(t-½):b-(t-½)[$. (This interval has positive size since $|]a:b[|>t\geq½$.) Since $w>a+(t-½)$, $w-m=w-½>a+t-1$, which given $t\leq1$ entails $w-m>a$, and by parallel reasoning, $w+m<b$. Hence $]a:b[\subseteq k(w)$. Since $]a:b[\subseteq \mathrm{K}p$ and $|]a:b[|>t$, $\pi_w(\mathrm{K}p)>t$. So P(Kp)>t holds at any world w within $]a+(t-½):b-(t-½)[$. Now since $|]a:b[|>t\geq½$, $|]a+(t-½):b-(t-½)[|>t-2(t-½)=1-t$. Therefore $\pi_@(\mathrm{P}(\mathrm{K}p)>t)>1-t$.

References

Adler, Jonathan. 2002. *Belief's Own Ethics*. MIT University Press.

Anscombe, G.E.M. 1962. On Sensations of Position. *Analysis* 22: 55-58.

Bacon, A. 2014. Giving your knowledge half a chance. *Philosophical Studies*171: 373-397.

Baker-Hytch, Max and Benton, Matthew. 2015. Defeatism Defeated. *Philosophical Perspectives* 29: 40-66.

Basu, Rima and Schroeder, Mark. 2019. Doxastic Wronging. In B. Kim and M. McGrath (ed.) *Pragmatic Encroachment in Epistemology*. Routledge.

Bergmann, Michael. 2006. *Justification Without Awareness*. Oxford University Press.

Brown, Jessica. 2018. *Fallibilism: Evidence and Knowledge*. Oxford University Press.

Chandler, Jake. 2013. Defeat Reconsidered. *Analysis* 73: 49-51.

Di Paolo, Joshua. 2018. Higher-Order Defeat is Object-Independent. *Pacific Philosophical Quarterly* 99: 248-269.

Dorst, Kevin. Forthcoming. Lockeans Maximize Expected Accuracy. *Mind*.

Dorst, Kevin. Forthcoming b. Evidence: A Guide for the Uncertain. *Philosophy and Phenomenological Research*.

Dorst, Kevin, Fitelson, Branden, and Husic, Brooke. MS. Evidence of Evidence: A Higher-Order Perspective.

Dutant, Julien and Fitelson, Branden. MS. Knowledge Centered Epistemic Utility Theory.

Dutant, Julien and Littlejohn, Clayton. Forthcoming. Justification, Knowledge, and Normality. *Philosophical Studies*.

Easwaran, Kenny. 2016. Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probabilities. *Noûs* 50: 816-853.

Easwaran, Kenny and Fitelson, Branden. 2015. Accuracy, Coherence, and Evidence. In Tamar Szabo Gendler and John Hawthorne (eds.). *Oxford Studies in Epistemology*. Vol. 5. Oxford University Press, pp. 61-96.

Fantl, Jeremy and McGrath, Matthew. 2009. *Knowledge in an Uncertain World*. Oxford University Press.

Ghijsen, Harmen, Kelp, Christopher, and Simion, Mona. 2016. Norms of Belief. *Philosophical Issues* 26: 375-392.

Gibbons, John. 2013. *The Norm of Belief.* Oxford University Press.
Goldberg, Sandy. 2017. Should Have Known. *Synthese* 194: 2863–2894.
Goldman, Alvin. 1999. *Knowledge in a Social World.* Oxford University Press.
Goodman, Jeremy. 2013. Inexact Knowledge without Improbable Knowing. *Inquiry* 56: 30-53.
Graham, Peter A. 2010. In Defense of Objectivism about Moral Obligation. *Ethics* 121: 88–115.
Grundmann, Thomas. 2011. Defeasibility Theory. In S. Bernecker and D. Pritchard (ed.), *The Routledge Companion to Epistemology.* Routledge.
Harman, Gilbert. 1968. Knowledge, Inference, and Explanation. *American Philosophical Quarterly* 5: 164-73.
Hyman, John. 1999. How Knowledge Works. *Philosophical Quarterly* 49: 433-51.
Kotzen, Matthew. Forthcoming. A Formal Account of Epistemic Defeat. *Synthese.*
Lackey, Jennifer. 2006. Learning from Words. *Philosophy and Phenomenological Research* 73: 77–101.
Lasonen-Aarnio, Maria. 2010. Unreasonable Knowledge. *Philosophical Perspectives* 24: 1-21.
Lazar, Seth. Forthcoming. Deontological Decision Theory and the Grounds of Subjective Permissibility. *Oxford Studies in Normative Ethics.*
Leitgeb, Hannes. 2017. *The Stability of Belief: How Rational Belief Coheres with Probability.* Oxford University Press.
Littlejohn, Clayton. 2017. How and Why Knowledge is First. In A. Carter, E. Gordon, and B. Jarvis (ed.), *Knowledge First.* Oxford University Press, pp. 19-46.
Littlejohn, Clayton. 2018. The Right in the Good: A Defence of Teleological Non-Consequentialism. In K. Ahlstrom-Vij and J. Dunn (ed.), *Epistemic Consequentialism.* Oxford University Press, pp. 23-48.
Littlejohn, Clayton. Forthcoming. Truth, Knowledge, and the Standard of Proof in Criminal Law. *Synthese.*
Littlejohn, Clayton. Forthcoming b. On what we should believe (and when (and why) we should believe what we know we should not believe). In K. McCain & S. Stapleford (ed.), *Epistemic Duties.* Routledge.
Lord, Errol. 2018. Epistemic Reasons, Evidence, and Defeaters. In D. Star (ed.), *Oxford Handbook of Reasons and Normativity.*
Lynch, Michael. 2004. *True to Life.* MIT University Press.
McGinn, Marie. 2012. Non-Inferential Knowledge. *Proceedings of the Aristotelian Society* 112: 1-28.
Makinson, D.C. 1965. The Paradox of the Preface. *Analysis* 25: 205-7.
Moss, Sarah. 2018. Moral Encroachment *Proceedings of the Aristotelian Society* 118: 177–205.
Nelkin, Dana. 2000. The Lottery Paradox, Knowledge, and Rationality. *The Philosophical Review* 109: 373-409.
Olsen, Kristian. 2018. Subjective Rightness and Minimizing Expected Objective Wrongdoing. *Pacific Philosophical Quarterly* 99: 417-441.
Owens, David. 2000. *Reason Without Freedom.* Routledge.
Peacocke, Christopher. 2004. *The Realm of Reason.* Oxford University Press.
Pollock, John. 1986. The Paradox of the Preface. *Philosophy of Science* 53: 246-58.
Raleigh, Thomas. Forthcoming. Suspending is Believing. *Synthese.*
Ryan, Sharon 1991. The Preface Paradox. *Philosophical Studies* 64: 293-307.
Ryan, Sharon. 1996. The Epistemic Virtues of Consistency. *Synthese* 109: 121–41.
Schroeder, Mark. 2012. Stakes, Withholding, and Pragmatic Encroachment on Knowledge. *Philosophical Studies* 160: 265-285.
Silva, Paul Jr. 2017. How Doxastic Justification Helps Us Solve the Puzzle of Misleading Higher-Order Evidence. *Pacific Philosophical Quarterly* 98: 308–328.
Smith, Martin. 2016. *Between Probability and Certainty.* Oxford University Press.
Spohn, Wolfgang. 2002. A Brief Comparison of Pollock's Defeasible Reasoning and Ranking Functions. *Synthese* 13: 39–56.
Sturgeon, Scott. 2008. Reason and the Grain of Belief. *Nous* 42: 139-165.

Sylvan, Kurt. 2018. Veritism Unswamped. *Mind* 127: 381–435.
Sylvan, Kurt. 2018b. Knowledge as a Non-Normative Relation. *Philosophy and Phenomenological Research* 97: 190-222.
Unger, Peter. 1975. *Ignorance: A Case for Scepticism* (Oxford: Oxford University Press).
Weatherson, Brian. 2005. Can We Do Without Pragmatic Encroachment? *Philosophical Perspectives*, 19: 417-443.
Weatherson, Brian. 2016. Reply to Eaton and Pickavance. *Philosophical Studies* 173: 3231-3233.
van Wietmarschen, Han. 2013. Peer Disagreement, Evidence, Well-Groundedness. *Philosophical Review*, 122: 395-425.
Williams, Bernard. 1965. Ethical Consistency. *Proceedings of the Aristotelian Society, Supplementary Volume* 39: 103–38.
Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford University Press.
Williamson, Timothy. 2014. Very Improbable Knowing. *Erkenntnis* 79: 971-999.
Williamson, Timothy. Forthcoming. Evidence of Evidence in Epistemic Logic. In M. Rasmussen and A. Steglich-Petersen (ed.), *Higher-Order Evidence: New Essays*. Oxford University Press.
Worsnip, Alex. 2016. Belief, Credence, and the Preface Paradox. *Australasian Journal of Philosophy* 94: 549-62.