Penultimate Draft: Final version in K. McCain, S. Stapleford, and M. Steup, eds., *Epistemic Dilemmas: New Arguments, New Angles* (Routledge 2021).

Embracing Epistemic Dilemmas[1]

David Christensen

1. Introduction: Some Apparent Epistemic Dilemmas

Talk about epistemic dilemmas can be motivated from a number of directions. Some cases studied in the literature involve puzzles such as Newcomb's Paradox (Priest 2002) or self-referential propositions such as R: *I reject R* (Brouwer 2014). I'd like to concentrate instead on cases involving examples that do not seem paradoxical, but where epistemic principles apparently conflict. Narrowing down a bit further, I'll put aside one sort of possible conflict between principles that has been offered as an example of a dilemma: the conflict between a principle requiring agents to believe what's supported by their evidence, and a principle prohibiting belief in falsehoods (Hughes 2019). An agent who believes a falsehood on the basis of excellent but misleading evidence might be seen as violating some objective or external norm prohibiting believing falsehoods, but this is not the sort of thing that seems to reflect badly on the agent in any way. (Similarly, consider a case where a nurse has excellent evidence that there's medicine in the bottle, and gives some to his patient. Even if someone has undetectably swapped the medicine for poison, the agent's action does not reflect badly on him. And this is not the sort of situation that's generally thought about as giving rise to moral dilemmas.)[2]

Part of what's worrisome—or interesting—about dilemmas is the idea that any possible response on the agent's part reflects failure to abide by the sort of standards which seem to describe how a well-functioning agent behaves. In the epistemic realm, the sort of dilemma I'd like to focus on involves inevitable violation of the sort of standards which seem constitutive of epistemic rationality. These violations are the sort which would generally reflect imperfections in an epistemic agent. And it is plausible that this can happen in situations that don't involve belief-contents which are in some way paradoxical, but which instead involve what has come to be called "higher-order evidence". This term also can be understood in a few different ways, but as I'll understand it, it involves an agent having evidence that bears on the reliability of her thinking in a certain domain. So I will be concentrating on apparent dilemmas involving higher-order evidence.

Perhaps the most dramatic examples involve logical reasoning. Consider the following case (adapted from Christensen (2019)):

> **Logic in Flight**: Aisha is flying in a small, unpressurized plane. She is told two things by her copilot, whom she rationally believes to be highly reliable:
>
> A: Karla was born in May if and only if Kayla wasn't; and
>
> B: Either Kayla, or Layla and Lola, the Lumpkin twins, were born in May.
>
> She rationally becomes extremely confident in A and B. As she flies, she thinks about their implications a bit, and rationally becomes highly confident that:
>
> P: Karla wasn't born in May unless Layla Lumpkin was.

[1] Thanks to Nomy Arpaly, Zach Barnett, Nick Leonard, Kevin McCain, Declan Smithies, and Scott Stapleford for very helpful discussions or comments on earlier drafts.

[2] Hughes (2019, 4062) also notes a disanalogy between his epistemic dilemma and moral dilemmas.

> Then Aisha notices that her instrument panel indicates that she's at an altitude where she's likely to be affected by hypoxia. She knows that hypoxia degrades the kind of moderately complex thinking she did in reasoning from A and B to P. Insidiously, hypoxia tends to do this while leaving its victims feeling perfectly clear-headed; Aisha knows well that this has led to a number of tragedies in aviation and mountaineering, and indeed that she herself has a long history of confident, but spectacularly incorrect, judgments made while hypoxic. Her co-pilot, an expert logician wearing an oxygen mask, tells Aisha that A and B do not entail P.

Here, then is the apparent dilemma: Suppose first that Aisha becomes much less confident in P. She then lacks confidence in a proposition that's truth-functionally entailed by two claims, A and B, that she's rationally extremely confident in. (The entailment is not the simplest possible, but it's only moderately complex. We may even imagine that Aisha can see how P follows from A and B—though her knowledge of the effects of hypoxia prevents her from trusting her logical insight.) This seems to violate a rational constraint that logic places on perfectly rational credences.

On the other hand, suppose Aisha maintains high confidence in P. Then she's not taking seriously the evidence that her complex thinking is likely to be very unreliable. Even if her thinking has not in fact been compromised by hypoxia, it would seem plainly irrational for her to think, "Well, A and B really do entail P, so I guess hypoxia hasn't made me inaccurate on *that!*" So if one epistemic principle requires respecting these logical relations among propositions, and another epistemic principle requires respecting higher-order evidence, Aisha seems doomed to violate one or the other.[3]

It's worth noting that the apparent dilemmas need not involve deductive reasoning. Consider the following example:

> **Drugged Detection:** Bao, a detective, is trying to solve a criminal case, and late one evening they rationally form the belief that Jocko is the perpetrator, on the basis of impeccable Inference to the Best Explanation from the forensic evidence. (I'll assume here that IBE gives the correct inductive constraint on belief, and that Jocko's guilt is an incredibly good explanation of the forensic evidence). Then Bao is informed that they've been slipped a drug which badly distorts explanatory reasoning 99% of the time. (Unbeknownst to Bao, it turns out that it did not affect them this time, which is why their original IBE was impeccable.)[4]

Again, there's an apparent dilemma. It seems that Bao must form beliefs that run afoul of the constraints IBE places on their beliefs, or else disregard the evidence that their explanatory

---

[3] I should note that the information about hypoxia itself bears on Aisha's reliability via some simple logical relations. The assumption in the example is that it is rational for Aisha to distrust the kind of moderately complex reasoning she did to derive P from A and B, but to trust much simpler reasoning, like "Since I'm likely hypoxic, I shouldn't trust my ability to tell if P follows from A and B, so I shouldn't be confident in P." So I'm assuming that higher-order evidence can target some parts of an agent's thinking, without targeting it all. Thanks to Kevin McCain and Scott Stapleford for prompting me to clarify this. (For discussion of the difficulties posed by HOE that's targeted more globally, See Sliwa and Horowitz (2015, §5.1) or Christensen (2019, §1.7).)

[4] The case is based on Sleepy Detective in Horowitz (2014). Less fanciful alternatives to the drug: Bao learns that their judgment is likely distorted by prejudice related to the suspect; a better detective tells them (misleadingly) that they haven't assessed the evidence correctly (Worsnip 2018); or their colleague informs them that the judgments they make in the evening have a terrible track record (Horowitz).

judgments are very likely unreliable. It may be less clear in this case that if Bao loses confidence in their original belief, they violate an epistemic principle. As Kevin McCain (2017) has emphasized, IBE requires taking into account the agent's *total* evidence. But it seems to me that the best explanation for Bao's *total* evidence (the forensic evidence, plus the evidence that they've been compromised) is still that Jocko is the perpetrator. Bao's being drugged doesn't affect the real, and extremely strong, explanatory relation that holds between Jocko's guilt and the forensic evidence, any more than evidence of Aisha being hypoxic affects the real logical relationship by which A and B entail P. One way to see this is to imagine that a different detective is given Bao's total evidence: the forensic evidence, plus the evidence that Bao was drugged. That other detective will have no difficulty seeing that the best explanation for all that evidence includes Jocko being the culprit. So if Bao respects their higher-order evidence, they end up having total evidence for which Jocko's guilt is an excellent explanation, and by far the best explanation, and even seeing how Jocko's guilt explains their evidence, yet not being confident that Jocko is the culprit.

2. Preliminaries

a. Are these examples really epistemic *dilemmas*?

Let us suppose, for the moment, that the examples above are ones where agents must violate some rational principle whatever belief they form. Some explicit discussion in epistemology, and related discussion in ethics, would support saying that we still lack enough information to characterize the cases as dilemmas. That is because nothing in the cases precludes the possibility that some of the potential doxastic responses are more rational than others, or even that there is a uniquely most rational response. And some would reserve the term 'dilemma' for situations where none of the agent's options is best.

But for present purposes, I would like to separate two questions: First, can there be cases where agents must violate some principle no matter what they do? And second, in those cases, might there still be a rationally best response? It's useful to separate the questions, because in cases such as Logic in Flight, it does seem that some responses are more rational than others. In particular, the response where Aisha loses confidence in her initial belief is more rational than the one where she maintains undiminished confidence. But it also seems that there's still something wrong with Aisha's beliefs when she does lose confidence. Respecting certain logical relations among belief-contents seems in part constitutive of rationality, and Aisha's beliefs do not respect these logical relations. The same point goes for the detective case: if Bao loses confidence in their hypothesis, they run afoul of the constraint that IBE places on rational credences. But that seems compatible with saying that losing confidence is, in the end, the most rational option they have—that losing confidence is more rational than brushing off the evidence that they've been drugged.

In what follows, then, I'll use 'dilemma' in a way that does not presuppose that there's no best option. So my usage here allows what Rosalind Hursthouse, in the moral context, calls "resolvable dilemmas": cases where there is one least-bad option, but in which even that option leaves what Hursthouse calls a moral "remainder" or "residue".[5]

[5] See Hursthouse (1999, ch. 2).

b. Do the examples really involve conflicting requirements?

Of course, one could deny that recognizing conflicting requirements even makes sense, in cases where there is a best option. One could insist that in any case where one option is more rational than all the others, that option is, by definition, rationally perfect, and thus violates no requirement. Applied to Aisha's case—given our judgment that the most rational course open to her is to lose confidence in P—this would entail denying that she ran afoul of any rational principle in being highly confident of A and B, but not of P. But this move, I think, would cover up, rather than clarify, the interesting aspects of these cases. Respecting logical relations is part of what explains the rationality of beliefs in general. And the import of the logical relations involved here—that A and B cannot be true without P being true—is not at all sensitive to matters involving Aisha's cognitive reliability.[6] We may of course hold Aisha blameless for her epistemic imperfection—but the evaluations involved in epistemic rationality-attributions need not be seen as primarily concerned (or concerned at all) with blame.[7]

One could also argue against seeing a conflict of principles in these cases, but in a way that did not amount to simply holding that the least-bad epistemic option is automatically rationally perfect. The simplest position along these lines denies that higher-order evidence has any effect at all on Aisha's rational credence in P, or Bao's rational credence in Jocko's guilt.[8] I won't engage with this line of argument in detail here, other than to note the size of the bullet it invites us to bite. Higher-order evidence can be very strong: Aisha might have extremely good evidence that she's hypoxic, and might even clearly remember the long series of cases where hypoxia led her to make ridiculous mistakes in just the sort of reasoning she's now considering, all the while feeling perfectly clear-headed. Or Bao might have participated in many trials of the IBE-distorting drug, and remember seeing videos in which they confidently made idiotic pronouncements about who committed various crimes, under the influence of the drug. But on the envisioned line, none of this has any effect on how confident Aisha or Bao are rational to be that P is true, or that Jocko is guilty.

A kind of mirror-image response is put forth by Sophie Horowitz (2019). On this position, it's fully rational for Aisha to lose confidence in P, or for Bao to lose confidence in Jocko's guilt, despite the fact that Aisha's evidence supports high confidence in P, and Bao's evidence supports high confidence in Jocko's guilt. On this view, there is no rational requirement to believe what's supported by one's total evidence; instead, rational belief is constrained by the agent's *perspective*, which involves bracketing some of the agent's evidence.

Adopting this sort of view does not involve biting bullets the size of those just considered; after all, the beliefs this view calls rational are, intuitively, the most rational available to agents in the relevant evidential situations. But as it stands, this approach seems to me not fully satisfying. Like the first line considered above, it does not account for the plausible view that logic constrains rational belief directly—that there is something intrinsically amiss with an agent's beliefs if, e.g., she's more

---

[6] For detailed argument on this point, see Christensen (2010).

[7] See Christensen (2004, ch. 6) or Arpaly (ms.) for discussion of this point.

[8] Versions of this sort of view can be found, for example, in Lasonen-Aarnio (2014) and Titelbaum (2015). Smithies (forthcoming) takes this position with respect to one sort of rationality; I'll discuss his position below.

confident in the conjunction of A and B than in P. And this problem is not like having a false belief that's supported by the agent's evidence. The illogicality of these beliefs is exactly the kind of thing that in general would in general constitute an imperfection in an epistemic agent. In Hursthouse's terminology, the envisioned approach doesn't account for the "residue". And in fact, Horowitz goes on to suggest that we might in the end want to say agents like Aisha and Bao who lose confidence in their original beliefs do violate a rational requirement to reason well. So with that in mind, I'd like to examine two lines of argument that take seriously the idea that situations like the ones above do represent conflicting requirements *in some sense*. But each of them seeks to avoid seeing the situations as ones where a single notion of epistemic rationality makes incompatible demands on agents' beliefs.

3. Separating Two Kinds of Norms

The first type of approach I'd like to consider involves relegating the conflicting requirements in some sense to separate normative notions. The approach comes in different varieties; I'll look here at two.

The first variety, due to Alex Worsnip (2018), would separate epistemic rationality from another normative notion: evidential support. In cases where agents receive sufficiently strong misleading evidence about what their evidence supports, their evidence might support believing two claims with the following structure:

Q
My evidence does not support believing Q.

But rationality, on Worsnip's view, does not involve respecting one's epistemic reasons (which requires believing in accordance with one's evidence). Rationality requires only a certain sort of coherence among one's beliefs—and in particular, it requires not believing Q while also believing that one's evidence doesn't support believing Q. So an agent in the envisioned situation can avoid believing *irrationally* by violating a norm from a different "domain" and failing to believe what her evidence supports.

If we apply this model to Logic in Flight, we could see Aisha's total evidence as supporting belief in both P and "My evidence does not support believing that P" (let us represent this second claim as ~E). On this view, there are rational options for Aisha. One would be to ignore the evidence of hypoxia and her co-pilot's testimony, and continue to believe P and E. Another would be to believe ~E, and give up belief in P, despite remaining fully confident in A and B. In these cases, the view will see Aisha's beliefs as perfectly rational (unless, of course, they are incoherent in some other way). But it will also see them as violating a separate norm, one which requires one's beliefs to respect one's evidence. The other option is for Aisha's beliefs to respect her evidence (so she believes both P and ~E). But in that case, her beliefs will count as irrational. So she must end up violating one norm or the other. But on Worsnip's view, the two norms are in different normative domains.[9]

[9] I should note that Worsnip does not apply his model to this case. But the case he focuses on is very much like our second one: it involves a detective who, having (correctly) concluded that the evidence does not support any particular hypothesis about the perpetrator, and having suspended judgment on who did it, is told (incorrectly) by a much better detective that the evidence actually supports the hypothesis (V) that the vicar did it. On Worsnip's analysis, the

Another way of dividing normative notions is to postulate different notions of rationality or rational standards. Declan Smithies (forthcoming), distinguishes two standards of epistemic rationality, which he calls ideal and non-ideal. The former involves norms that would be good to *implement perfectly*. This sort of rationality would pay no heed to higher-order evidence: after all, perfect implementation of evidential norms for ordinary evidence could only be compromised by taking higher-order evidence on board. The second sort of rational standard involves norms that would be good to *adopt, or try to implement*, keeping in mind the possibility of one's own possible epistemic failures. That sort of rationality would take higher-order evidence seriously; on this second notion, Aisha would be rational to lose confidence in P, despite believing A and B.[10] So Aisha's beliefs can end up fully rational by one set of standards or the other. But whatever Aisha does, her beliefs will also fail to be fully rational in one sense or the other.

Now there are various specific worries one might have about each of these approaches. With respect to the first approach, if we divorce epistemic rationality from evidential considerations, we'll be calling paradigmatically irrational beliefs rational: consider the person with ordinary evidence who believes that the world is run by giant lizards (perhaps due to a paranoid delusion); or, to take more ordinary examples, consider astrology-fans motivated by wishful thinking, or people whose prejudice against members of certain groups leads to believing unflattering things about them which have no evidential backing. As long as these believers *believe* that their beliefs are supported by their evidence—a belief that itself may be the product of mental illness, wishful thinking, or the desire to think of themselves as forming their unflattering beliefs in an unprejudiced way—their beliefs may well count as perfectly rational. As Nomy Arpaly has observed, "if we were to deem rational every person who believes herself to have good reasons to hold her beliefs, there would hardly be an epistemically irrational person left in the land."[11]

---

detective's evidence now supports belief in the claim that her evidence supports V, but it does not support believing V. So if the detective believes in accordance with her evidence, she'll believe irrationally, because her beliefs will be incoherent. On the other hand, if she maintains coherence—say by believing both V and "my evidence supports V," she'll have a belief that's evidentially unsupported. Since the problem posed by dilemmas is perhaps most clearly illustrated by Aisha's case, I'll apply the model to it; but I don't think that any of the arguments below will turn on choice of example.

It's also worth bringing up one other point about Worsnip's understanding of rationality. One might wonder, in the case where the detective believes both V and "my evidence supports V," whether the detective's beliefs would still be incoherent, in virtue of her belief in V violating some form of IBE. Analogously, one might wonder whether Aisha's beliefs, if she stops believing P, are coherent, given their violation of logic. But Worsnip holds that it's a substantial question whether such combinations of beliefs are coherent or not. He writes, "Someone who wants to hold that these combinations of attitudes are not irrational may hold that they are not properly speaking incoherent." The coherence norm Worsnip commits to is just that one (a) not believe one's evidence supports taking some doxastic attitude toward P without taking that attitude, and (b) not take a doxastic attitude toward P that one believes not to be supported by one's evidence. He sees all incoherence as in some way failing by one's own standards (2018, 36).

[10] Smithies' distinction parallels one made by Miriam Schoenfield (2015).

[11] Arpaly (2000, 494). I'm assuming, of course, that the sets of beliefs in question would not need to be incoherent in some other way that Worsnip would take as part of rationality. For an extended discussion of this issue arguing that no adequate notion of epistemic rationality can be divorced from evidential support, see (Christensen 2020). For some other criticisms of Worsnip's view, see Skipper (2019), Leonard (2020), and Smithies (forthcoming).

On the ideal/non-ideal approach, there are worries about both conceptions of rationality. Horowitz points out, about Schoenfield's similar proposal, that the conception that sees rationality as comprising norms that would be best to adopt or try to implement will have trouble excluding 'epistemic norms' such as "Have a sandwich before engaging in difficult reasoning", or "Never, ever, believe there is a spider nearby" (2019 115-116). And for the conception of rationality that sees it as comprising the norms that would be best to perfectly implement, one might worry about its implication that P would be perfectly rational for Aisha to believe, despite her powerful evidence about hypoxia and her co-pilot's testimony.

Of course, there is not space here to enter fully into the pros and cons of these proposals; the worries just raised are meant only to illustrate that the proposals do not escape the intuitive puzzles posed by cases like Aisha's without cost. But we might still ask: should we take the fact that these proposals avoid dilemmas (at least in a certain sense) to constitute a strong consideration in their favor? To focus on this question, we might ask why dilemmas should be seen as objectionable in the first place.

Some have pointed out that recognizing dilemmas would require us to give up some otherwise-attractive principle of epistemic logic. But as Nick Hughes (2019) argues in detail, the principles that would have to be given up are not obviously motivated independently of rejecting dilemmas. And it seems that epistemic logic should be built to accommodate our considered judgments about rational belief—not the reverse.

One might also worry that dilemmas leave agents without a certain kind of epistemic guidance. Now I think it's not clear that epistemic principles should in the first place be thought of as necessarily guidance-giving. Though some epistemologists may seek to formulate useful guides to cognitive self-help, epistemologists also may seek theories which illuminate the nature of rationality.[12]

Moreover, even putting that point aside, the sort of dilemmas with which we're concerned may well leave agents with one reaction that's rationally best. They just involve conceding that in certain cases, the best option is suboptimal along some of the dimensions which help constitute rationality. So there's actually nothing about conflicting requirements *per se* that is inconsistent with rationality providing guidance.

Of course, dilemmas in our sense do constitute cases where it's impossible for agents to fulfil all the requirements of rationality. Perhaps this in itself will seem in some way problematic. Smithies (forthcoming) writes:

> To be epistemically rational is to satisfy all the requirements of epistemic rationality. But if there are strict dilemmas in which epistemic rationality issues inconsistent requirements, then it is logically impossible to satisfy all of them at once. And there can be no value in satisfying all the requirements of epistemic rationality when it is logically impossible to do so, since there can be

[12] See Christensen (2005, Ch. 6) for detailed defense of the latter project. The distinction made here echoes Arpaly's (2000) distinction between two different ways of theorizing about practical rationality: constructing an *instruction manual* to be used by agents to arrive at rational actions, and developing an *account of rationality*: a theory which tells *us* when people are acting rationally. For a recent defense of guidance-directed ambitions in epistemology, but which acknowledges the legitimacy of the more purely theoretical project, see Ballantyne (2019).

> nothing of value in a logically impossible scenario. Hence, the very idea of an epistemic dilemma compromises the value of epistemic rationality.

But what kind of compromise is really entailed by recognizing dilemmas? Surely rationality comes in degrees, and the value of rationality reflects this: more rational beliefs are better, even if they're not rationally perfect. So it can't be that the value of rationality is only realized by beliefs which perfectly satisfy all the rational requirements—whether or not there are epistemic dilemmas. And I don't see any reason to think that the greater value of more rational beliefs somehow depends on whether it's in principle possible for an agent to achieve rational perfection. (A highly efficient car is more valuable than a less efficient one, even if perfect efficiency is impossible in principle. More happiness is better than less, even if there's no limit to the amount of happiness possible.) So the proponent of dilemmas can certainly recognize the value of greater, but not perfect, rationality, which is presumably the only sort that we imperfect beings ever realize.

It's also worth keeping in mind that we must all live with normative conflicts. In some cases, epistemic requirements may well come apart from, e.g., prudential requirements. A classic example is Robert Nozick's (1993, 69–70) mother who has excellent evidence that her son committed a horrendous crime—but who has excellent reason to believe that she'd be emotionally destroyed if she believed him guilty. Suppose that the mother has a way of getting herself to believe him innocent. If she uses it, she'll end up epistemically irrational. If she doesn't, she'll be pragmatically irrational. So normative perfection seems to be impossible in principle in her situation.

Of course, there is a difference between this sort of conflict and the ones we've been examining: While Nozick's mother is faced with a situation where whatever she does, she'll violate either a pragmatic or an epistemic rationality norm, epistemic dilemmas would put agents in situations where they are bound to violate an *epistemic* norm in particular. But even if one thought that conflicts between practical and epistemic norms were somehow less bad, the two-norm views we're now considering would have no such advantage: they also have it that agents are bound to violate distinctively epistemic norms.

At one point, Worsnip suggests that there's something particularly bad about conflicts between requirements of *rationality*, as compared with conflicts between different sorts of epistemic norms. While he acknowledges that this "might seem to be hair-splitting," he argues that "there does seem to be something about *irrationality* in particular that should in some way be due to the agent herself, and not simply due to a situation outside her control." (2018, 37).

There are two ideas here: that rationality is in some way internal to the agent, and that rationality should be within an agent's control. As to the first, there are certainly accounts of evidence on which evidential-support facts supervene on an agent's internal states. So conceptions of rationality that build in evidential-support facts along with facts about coherence can easily fit with this (very plausible, to my mind) internalist conception of rationality. And such accounts surely can allow for conflicts between rational requirements: dilemmas don't somehow make rationality subject to external factors.

The second idea is that an agent's rationality should be within her control. This idea does not, to my mind, share the plausibility of the idea that rationality is an internal matter. It is obvious that, e.g., a person living with schizophrenic delusions cannot become rational by some effort of will. And this, it seems, is just what we should expect, even on a purely coherence-based account of rationality. It's

pretty clear that most of our beliefs are not within our control. And if beliefs in general are not within our control, then it would be surprising if we could somehow control whether our beliefs exhibited certain coherence relations with one another, ensuring coherence by acts of will. Even on a very thin notion of coherence that consisted of nothing beyond the requirement that one not believe both "P" and "My evidence doesn't support believing P", it seems clear that achieving it may be outside an agent's control. Think of someone who believes they are worthless, and goes to a psychiatrist, who forcefully and patiently points out in detail that their evidence does not support any such assessment. It seems clear that such a person may honestly say, "I believe you—I have no evidence that I'm worthless. But I just can't shake the belief that I really am. I just know I'm worthless." I see no reason to think that such a person must be wrong about what they believe. They will indeed be exhibiting irrationality. But there's a good reason why no competent psychiatrist would advise them to just snap out of it already.

With these points in mind, it does not seem that norm-separation views secure much advantage over simpler views that allow rational requirements to conflict without multiplying domains of normativity. Intuitively, much of the resistance to the idea of rational dilemmas stems from resistance to the idea that agents are bound, no matter how they believe, to violate a requirement whose satisfaction generally helps constitute well-functioning in an epistemic agent. But insofar as the requirements to respect higher-order evidence and to respect logical relations are ones whose satisfaction generally helps constitute epistemic well-functioning, norm-separation views don't avoid this result. So it seems to me that norm-separation views don't really purchase much advantage by avoiding (at least in a sense) dilemmas.

Moreover, these views also incur what strikes me as a large intuitive cost. Consider the case in which Aisha bull-headedly maintains full confidence in P: she just insists that her total evidence—including the evidence that she's hypoxic, her history of confident misjudgments while hypoxic, and her co-pilot's disagreement—supports remaining fully confident that P. Each of the above proposals will say that P is perfectly rational for Amelia to believe with confidence: either perfectly rational in one full sense of (propositional) rationality, or perfectly rational full-stop.[13] This is the price the views pay for norm-dividing that pares down "rationality" far enough to get a notion that's free of conflicting requirements.

4. Positing Rational Indeterminacy

An ingeniously different way of avoiding dilemmas has been proposed by Nick Leonard (2020). Leonard does not posit separate kinds of epistemic normativity; he sees the principles that conflict in cases like Aisha's as all being components of a unified notion of epistemic rationality. On the

[13] Smithies does address this general sort of worry about his ideal standard of rationality (2019, Ch. 10.4). He suggests that ideally, an agent in Aisha's position should maintain full confidence in P, and full confidence that her evidence supports P. But he proposes that she may rationally accommodate her higher-order evidence by doubting that her belief in P is actually correctly *based on* her evidence. So she should presumably doubt that she reasoned reliably when she derived P from A and B.

I myself cannot see how this suggestion solves the intuitive problem. We might ask what Aisha is ideally supposed to think about how she arrived at a correct belief that P, and a correct assessment of whether A and B entail P, while reasoning unreliably. Of course, this can happen, e.g. if one gets lucky. But is Aisha supposed to be absolutely certain that she got lucky in this way? I don't see how it can be rational to maintain full confidence in a belief that's based solely on reasoning that one believes unreliable. See Skipper (this volume) for related criticism.

indeterminacy view, since Aisha cannot satisfy all the requirements of rationality, she is at least required to satisfy a set of rational principles that is *maximal*: that is, a set such that adding any other (correct) rational principle would result in conflicting requirements. Leonard calls such a maximal set of principles a "resolution". But the idea is not to embrace the permissivist position that Aisha can be fully rational just by satisfying some resolution. The indeterminacy view says that agents are required to satisfy the *correct* resolution—but that it is indeterminate which resolution is correct.

To simplify a bit, suppose that the relevant principles that conflict in Aisha's case are (1) that she respect higher-order evidence (which would require losing confidence in P), and (2) that she respect probabilism (would require maintaining confidence in P). So one resolution will contain (1) and not (2), and the other will contain (2) and not (1). On the indeterminacy view, it is indeterminate which of these two resolutions is correct. So suppose Aisha loses confidence in P, conforming to (1) but violating (2). We would say that she has not failed to do anything that she was (determinately) required to do. And similarly if she maintained full confidence in P, conforming to (2) but violating (1). In either case, her beliefs will be neither (determinately) rationally impermissible, nor (determinately) rationally permissible. Of course, if she adopted beliefs that violated both (1) and (2), her beliefs would not satisfy a maximal set of principles, and thus would be determinately impermissible.[14]

The indeterminacy view has some advantages over the "separate norms" approach. In a way, it provides a more satisfying response to the worry that we shouldn't say that an agent is doomed to violate some epistemic rational requirement no matter what. After all, there is a rationally correct way of believing—it's just indeterminate which that is. So while Aisha is precluded from having determinately permissible beliefs, she's at least not doomed to have determinately impermissible ones. And unlike the "separate norms" approach, the view avoids having to say, in the case where Aisha shrugs off her higher-order evidence and maintains full confidence in P, that there's any sense at all in which P is *determinately* perfectly rational for her to believe.

There are also potential worries for the view. One worry arises from the intuition that it would be *more rational* for Aisha to lose confidence in P than to maintain it. To my mind, at least, this does not seem like an indeterminate matter: it seems determinately better—more rational—for Aisha to lose confidence in P. But the indeterminacy view seems not to allow for this judgment. And more generally, it does seem that the stronger an agent's (misleading) higher-order evidence is, the more irrational it would be for her to stick to the beliefs supported by her first-order evidence alone. But it's not clear how the indeterminacy view can accommodate these verdicts.[15]

Leonard actually presses this point in the opposite direction: he worries that views which say that it's better for Aisha to lose confidence in P owe us an explanation of why this is so: how can the requirement to respect her higher-order evidence and the requirement to respect logic be weighed against each other (2020, § 2.1)?[16]

---

[14] And naturally, in cases where all the requirements are simultaneously satisfiable, there's only one resolution, containing all the true epistemic principles; beliefs will be determinately permissible if they satisfy them all, and determinately impermissible otherwise.

[15] See Smithies (forthcoming) for a similar worry.

[16] For similar worries see also Lasonen-Aarnio (2014), Horowitz (2019), and Skipper (this volume).

I'm not sure why this unanswered question should worry the proponent of dilemmas. First, allowing dilemmas is not supposed to solve the problem of determining which response is most rational when principles conflict. It just amounts to acknowledging that the best response may involve 'residue'.

Second, even without worrying about dilemmas, it's clear that we have no worked-out general formula for determining which beliefs are most rational in complicated situations. Consider cases where, e.g., simplicity favors one theory but fit with the data favors a different one, or where testimony of a witness supports one hypothesis about the killer, but certain forensic evidence points a different way. The existence of difficult cases, or our lack of a formula for resolving them, does not mean that the overall structure of our theory is problematic. It merely indicates that there is more work to be done. And working out an account of the best response to situations such as Aisha's is perfectly compatible with recognizing those situations as dilemmatic.[17]

Another worry about the indeterminacy view is that its rationality-verdicts seem in some cases to depend on arbitrary-looking claims about principle-individuation. For example, Leonard takes Probabilism as a principle governing credences—it says that an agent's credences should obey the laws of probability—and he formulates Probabilism via three standard axioms (Finite Additivity, Non-Negativity, and Normality) which an agent's credences would have to satisfy. So we could describe Probabilism as one principle or as three. Here is how the problem might arise:

Suppose that an agent's respecting a certain principle (say one requiring deference to expert testimony) would require violating Probabilism with respect to certain beliefs, and consider a resolution that included the testimony principle. On the one-principle description of Probabilism, since Probabilism can't be added without generating inconsistent requirements, the target beliefs would not be probabilistically constrained. But suppose that respecting the testimonial principle was consistent with respecting one or two of the probabilistic axioms. In that case, the three-principle description of Probabilism would place further constraints. So it would call determinately irrational certain beliefs that the one-principle description would classify as indeterminately rational. Yet it seems fairly arbitrary whether we describe Probabilism as consisting in one principle or three.[18]

---

[17] For an attempt to give a somewhat systematic account of the sorts of cases we've been examining—though not one that avoids all the difficult cases—see Christensen (2019).

[18] Here's an (unrealistic toy) example to illustrate the structure of the problem: Suppose that P is logically true, and Q is logically false, and suppose that Chuck gets the following expert testimony: "P is true; and Q and (P v Q) have the same truth value". Say that any resolution that includes the Testimony principle will require that Chuck have credence in 1 in P, and give Q and (P v Q) equal credences. Suppose Chuck ends up with the following credences:

{P = 1, Q = 0, (P v Q) = 0}.

What will the indeterminist view say about his 0 credence in (P v Q)?

If Probabilism is one principle, this credence will presumably come out as indeterminately rational: It's permitted on the resolution that includes Testimony but not Probabilism, but forbidden on the resolution that includes Probabilism but not Testimony.

What if we describe Probabilism as three axioms? The Normality axiom requires that all logical truths get probability 1. And we can in fact add Normality to Testimony without generating conflicting requirements—the following credences satisfy both:

Of course, there may be indeterminist responses to this worry, and we can't settle this issue here. But it seems worth noting, as it flags a potential cost of adopting the indeterminacy view, in addition to the problem of failing to fit with our intuitions in cases of strong higher-order evidence. And again, even without deciding these issues, we can ask: how much of an advantage would the indeterminacy view gain, if it were successfully developed in a way that avoided determinate dilemmas?

In Aisha's case, the indeterminacy view would say that her beliefs could not satisfy both Probabilism and a norm requiring respect for her higher-order evidence. In seeing this as a case of indeterminacy, the view assumes that the two principles are generally correct: rational beliefs are generally required to satisfy them. In fact, beliefs have the rational status they do in virtue of satisfying, or not satisfying, these principles. Satisfaction of these principles in part constitutes rationality of beliefs. Focusing on agents rather than beliefs, the indeterminacy view, like all the views we've been examining, ends up saying that in certain situations, agents are bound, no matter how they believe, to violate a principle whose satisfaction in general constitutes ideal functioning in an epistemic agent. So if that's the root of one's aversion to dilemmas, the indeterminacy view doesn't seem to have any clear advantage.

A related thought might be that the problem with allowing determinate dilemmas is that it's *unfair* to epistemic agents. After all, we have been supposing that if Aisha loses confidence in P, she violates Probabilism, but she's doing the best that she, or any agent, could do in her situation. And yet we seem to be giving her some sort of rational demerit for violating Probabilism.

Now I'm not sure how much indeterminism would help here: it certainly does not say that Aisha is in the clear. And if our fairness intuition is something along the lines of "doing the best anyone could do should be good enough," indeterminism also fails to satisfy it: it says that no matter what Aisha believes, her beliefs will not be determinately rational.

But in any case, it seems to me a mistake to think in terms of fairness here. We need not—and should not—blame agents simply because they have irrational beliefs. This point is entirely independent of dilemmas: someone living with delusional schizophrenia, I take it, is in no way blameworthy for their irrational beliefs. So we are not considering Aisha blameworthy for violating Probabilism, or even suggesting that she could, overall, have done better in her believing. And given that, it's not clear why our evaluation of her beliefs should be seen as unfair. It simply recognizes that, in violating Probabilism, her beliefs fall short in one of the dimensions along which we evaluate beliefs for rationality. It recognizes that sometimes, even the best beliefs any agent could form in a certain situation will exhibit a kind of imperfection: in Hursthouse's terminology, we are simply noting that the best response to Aisha's predicament will involve some residue. This will hold whether we recognize dilemmas or postulate indeterminacy.

{P = 1, Q = 1, (P v Q) = 1}.

So if Probabilism is described as three separate principles, Chuck's 0 credence in (P v Q) will come out determinately irrational, not indeterminately rational.

Thanks to Nick Leonard for very helpful discussion of this point.

5. Conclusion

Higher-order evidence seems to lead to cases where epistemic principles impose conflicting requirements on agents' beliefs. If one recognizes the conflicting requirements as genuine, the simplest description of the cases leaves it at that: no matter what the agent believes, their beliefs will violate some genuine rational requirement. More complex descriptions attempt to defuse the intuitive tension, by consigning the conflicting requirements to different varieties of epistemic norm, or by positing indeterminacy in which principles the agent is required to satisfy in conflict situations.

As we've seen, each of these ways of complicating things brings worries in its wake. Some of those clearly can't be fully assessed here, but at least this much seems clear: The separate-norms approach requires recognizing a sense of propositional epistemic rationality in which full confidence in P is perfectly rational for Aisha, despite her extremely powerful evidence that her reasoning concerning P is unreliable. And even the indeterminacy approach cannot see maintaining full confidence as determinately irrational for Aisha, or even as determinately less rational than losing confidence in P.

Meanwhile, both approaches take on what might seem to be the most troubling aspect of recognizing dilemmas. On either of these approaches, as on the simple approach, we see that in certain situations, agents are bound, no matter how they believe, to violate principles the satisfaction of which in general helps constitute epistemic rationality.

Perhaps, though, this phenomenon really should not trouble us, especially if we keep in mind that we're not holding agents blameworthy for unavoidable violation of rational principles. Instead, we might see it as an interesting—and even natural—consequence of taking seriously the implications of higher-order evidence. If agents can get empirical evidence about the reliability of their thinking, it is inevitable that, in certain cases, the evidence will be misleading. Rationality will demand that agents heed this misleading evidence. But rationality will also demand that agents' first-order reasoning actually be good. It's only to be expected that these rational requirements will conflict. Rather than warping our theories in an attempt to avoid this sort of conflict, perhaps we should embrace dilemmas in our epistemic theorizing, seeing them not as revealing defects in our theories, but as manifestations of the structural richness that emerges when agents can reflect on their own thinking.

References

Arpaly, N. (ms.), "Epistemology and the Baffled Action Theorist."

---. (2000), "On Acting Rationally against One's Best Judgment," *Ethics* 110: 488–513.

Ballantyne, N. (2019), *Knowing our Limits*, (Oxford University Press).

Brouwer, T. N. P. A. (2014), "A Paradox of Rejection," *Synthese* 191: 4451–4464.

Christensen, D. (2004), *Putting Logic in its Place: Formal Constraints on Rational Belief* (Oxford University Press).

---. (2010), "Higher-Order Evidence" *Philosophy and Phenomenological Research* 81: 185–215.

---. (2019), "Formulating Independence," in Skipper, M. and A. Steglich-Petersen, eds., *Higher-Order Evidence: New Essays* (Oxford University Press).

---. (2020), "The Ineliminability of Epistemic Rationality," *Philosophy and Phenomenological Research,* https://doi.org/10.1111/phpr.12712.

Horowitz, S. (2014), "Epistemic Akrasia," *Noûs* 48:4: 718–744.

---. (2019), "Predictably Misleading Evidence," in Skipper, M. and A. Steglich-Petersen, eds., *Higher-Order Evidence: New Essays* (Oxford University Press).

Hughes, N. (2019), "Dilemmatic Epistemology," *Synthese* 196: 4059–4090.

Hursthouse, R. (1999), *On Virtue Ethics* (Oxford UP).

Lasonen-Aarnio, M. (2014), "Higher-Order Evidence and the Limits of Defeat," *Philosophy and Phenomenological Research* 88 (2): 314–345.

Leonard, N. (2020) "Epistemic Dilemmas and Rational Indeterminacy," *Philosophical Studies* 177: 573–596.

McCain, K. (2017), 'Undaunted Explanationism," *Logos and Episteme* VIII: 117–127.

Nozick, R. (1993), *The Nature of Rationality* (Princeton University Press).

Priest, G. (2002), "Rational Dilemmas," *Analysis* 62.1: 11–16.

Schoenfield, M. (2015), "Bridging Rationality and Accuracy," *Journal of Philosophy* 112 (12): 633–657.

Skipper, M. (2019), "Higher-Order Defeat and the Impossibility of Self-Misleading Evidence," in Skipper, M. and A. Steglich-Petersen, eds., *Higher-Order Evidence: New Essays* (Oxford University Press).

---. (this volume), "Higher-Order Evidence and the Normativity of Logic".

Sliwa, P. and S. Horowitz (2015), "Respecting *all* the Evidence," *Philosophical Studies* 172 (11): 2835–2858.

Smithies, D. (2019), *The Epistemic Role of Consciousness* (Oxford University Press).

---. (forthcoming), "The Unity of Evidence and Coherence," to appear in N. Hughes, ed., *Epistemic Dilemmas*, (Oxford University Press).

Titelbaum, M. (2015), "Rationality's Fixed Point (or: In Defense of Right Reason)," *Oxford Studies in Epistemology* 5: 253–294.

Worsnip, A. (2018), "The Conflict of Evidence and Coherence," *Philosophy and Phenomenological Research* 96 (1): 3–44.