The Merits of Incoherence

jim.pryor@nyu.edu

10 Jan 2015 - Huatulco

Section 1

One prominent advanced logic textbook begins by announcing its scope like this:

This book... *presents the basic concepts and results of logic: the topics are proofs, [semantic notions like] truth, and computability*. (Enderton, A Mathematical Introduction to Logic, 2nd ed., 2001, pp. xi)

This is what I regard as the *right* way to say what logic is about. Going a step further, we may perhaps characterize it as the mathematics of some kinds of structurally-imposed commitment. What I especially approve of here is what Enderton *doesn't* say. As he goes on to warn:

We have given some hints as to what we intend to do in this book. We should also correct some possible misimpressions by saying what we are not going to do. *This book does not propose to teach the reader how to think*... (Enderton, pp. xii)

By contrast, many other logic textbooks introduce their topic like this:

Logic is defined as the study of the principles of reasoning. (Hedman, A First Course in Logic, 2004, p. xiii)

One of the popular definitions of logic is that *it is the analysis of methods of reasoning*... (Mendelson, Introduction to Mathematical Logic, 4th ed., 1997, p. 1)

[W]e can say that mathematical logic is the branch of mathematics that studies the logic and the methods of deduction used in mathematics... [T]here is something circular, perhaps even paradoxical, in the nature of this field. For *its subject matter is mathematical reasoning*, but since it is a branch of mathematics *its method of inquiry* must also consist of mathematical reasoning. (Robert Wolf, A Tour through Mathematical Logic, 2005, p. vii) Is there really much plausibility to the idea that our subject matter in advanced logic classes—the material we actually teach—is a good model for the methods of good mathematical inquiry, or for methods of good reasoning more generally? Is that how mathematicians really think: in the form of proofs? No, of course not. Proofs are what they *seek* in their inquiry. We shouldn't expect a log of the thinking processes that *lead them to* those results, even when they get there elegantly and efficiently, to also constitute formal proofs. Not even abbreviated proofs with suppressed premises. Consider by analogy a musical composer. Her thinking will most likely contain snatches of music (or auditory imagery). It may even contain exercises of mentally rehearsing (and inspecting and criticizing and revising) *entire pieces.* But we shouldn't think that whole thinking process itself is therefore well-understood as a piece of music. In the same way, I submit, good mathematical reasoning isn't itself well-understood as a deductive proof. Good formal arguments are one thing, good reasoning—even the good reasoning that identifies good formal arguments—is another.

In logic textbooks, introductory remarks like the ones I quoted seldom make any difference to the content or quality of the pedagogy that follows. But these sentiments do inform the judgments of many practicing epistemologists, and—at least superficially—guide the way their arguments proceed. (My qualification "at least superficially" will be explained later.)

Everyone will admit there can be instances of good reasoning that don't match a formal entailment. But many will expect that the latter is at least an interesting subset of the former. Or, more carefully, since we want to allow for our lack of logical omniscience, they'll at least say that pieces of reasoning matching entailments that are *obvious* to the reasoner must necessarily be cases of good reasoning.

My attitude is different. As I see it, there are three different families of notions, which aren't connected to each other in straightforward ways. I'm frustrated by the way many of us uncritically blur these notions together—perhaps encouraged by the fact that one can use words like "reasoning" or "inference" when characterizing each family.

- One family of notions is formal or logical entailments. (Recall our custom of talking about "rules of inference" in this context.)
- A second family of notions includes things like justificatory support and what we might call "rationally responding" to one's evidence, that is making reasonable doxastic transitions (perhaps also including *maintaining* doxastic stances earlier adopted). I include here your coming to some *settled* views (all-out believing or withholding of belief) as well as *graded* doxastic stances like adjusting (or maintaining) a certain degree of confidence in response to your evidence.
- A third family of notions is what we might call "intelligently deliberating." This kind of mental activity is often exploratory. It may fumble around

in darkness in ways we trust to be productive. "Having a good nose" and "thinking outside the box" often help; but so too can more disciplined strategies. Still, I think a good science of how this is done well, say when performed by a good graduate mathematician or philosopher, is astoundingly unlikely to model the process as anything like a piece of syllogistic (so-called) "reasoning."

I'm not yet *arguing*. I'm just projecting at you my *attitude* of wanting to handle these three families more cautiously than mainstream epistemology currently tends to. As I said, I think the notions are less straightforwardly connected than much contemporary work assumes them to be. Of course they're unlikely to have *nothing* to do with each other, either. But articulating their connections is going to be difficult. I think that as a field, we've made only initial steps.

But I won't be trying to *argue* all that here.

What I will do is focus on the relations between the first two families. I'll begin by highlighting various pressures, that will be familiar from some recent literature, for thinking that good reasoning in the sense of proper "rationally responding" is always non-monotonic—even when the reasoning in fact retraces an obvious entailment. This is so even for good *arithmetical* reasoning. That discussion will lead us to the result that sets of attitudes a subject can "see" in some sense to be incoherent or even inconsistent can be her least bad rational option. Perhaps we should say these combinations of attitudes can even be *positively reasonable*; or perhaps we should instead count them as cases of epistemic "dilemmas" or "tragedies," where no fully and positively reasonable doxastic choice is available. (As we'll see, not even withholding belief can always be a refuge.) I'll discuss these alternatives but won't try to choose between them. The third part of our discussion will explore how this general picture applies to the case of immediatebut-underminable justification, as I've claimed we have in perception but most epistemologists will (or should) be prepared to grant we have somewhere. Relying on some epistemic faculty while one's antecedent evidence at best recommends suspending belief about whether one *should* so rely may be another case where attitudes one can "see" to be in some ways incoherent are for all that our least bad rational option.

Section 2

If you take a college-level class in Algebra, you'll spend a lot of time studying Groups (collections of objects and an associative and invertible operation on them), and among the things you'll learn about Groups are several elegant and surprising connections between Groups and Graphs (diagrams like the London Tube Map). Let A be some theorem of GrAph Theory, and O be some theorem of GrOup Theory that turns out to be interestingly—not at first glance obviously—equivalent to A. (The claims are *mathematically* not logically equivalent; but the corresponding claims that some axiomatization of Graph Theory entails A,

and that some axiomatization of Group Theory entails O, are properly speaking logically equivalent. I will suppress this in what follows.)

Consider three subjects, Adam, Barry, and Clarice. (Later I'll add a fourth.) These subjects all start out in the same intuitive position. They're reasonably confident in A, but there's nothing they could point to that gives them any special reason to believe O. Now Adam's intellectual life evolves like a good student of mathematics: he's initially agnostic towards, or has low confidence in O, but then as he studies the connections between Graphs and Groups he starts to "see" that A iff O—supply whatever account of this change in his mathematical understanding you favor. And so as a result Adam comes (reasonably) to believe and/or to raise his confidence in O.

Barry never undergoes that eye-opening experience that Adam does. Throughout our story, he'll remain as ill-equipped to explain or identify the connection between A and O that Adam was at the start of the story. That is, he *only* ever has the mathematical insights and understanding that early-Adam had, not the ones that later-Adam acquires. However, Barry has all along had the same amount of *confidence* in O that later-Adam acquires. From the start, Barry has always believed A and O to the same degree.

Clarice also only ever had the mathematical insights and understanding that early-Adam and Barry do. However, her confidence in O starts out low, like Adam's, and then later *rises* to match her confidence in A. But in Clarice's case, this happens *spontaneously*, not in response to any mathematical revelation or explanatory reasoning.

We can agree that since A is true iff O is, the doxastic attitudes that Adam, Barry, and Clarice all end with are more *accurate* than the ones Adam and Clarice began with (high confidence in A, less confidence in O). We can agree also that Clarice did something *dynamically* inappropriate in changing her beliefs as she did—she *arrived at* her ending confidence in a way that's objectionable, because there's nothing she became aware of to *justify* changing her confidence. Nonetheless, even stipulating these points as agreed, there remains a further question we might intelligibly dispute, whether the doxastic attitudes Adam, Barry, and Clarice have at the end of the story are all equally *statically* justified.

I invite you to share my judgment that they are not. Though Barry and Clarice have ending confidences that are in one sense *correct*, they both seem to be interestingly *less justified* in being so confident in O than late-Adam is. Explaining *why* this is so is hard, because the epistemology of mathematics is hard. But I hope you will share my sense that there is some such epistemic difference between our subjects. As epistemologists, we often brush logical opacity aside and ignore justificatory differences of this sort. For today's discussion—even though the epistemology of mathematics is not directly our topic—we can't afford to do that. We'll need to be open to the ways in which Adam's, Barry's, and Clarice's epistemic positions differ, and some of them are justified in being more confident in O than others.

Section 3

OK, with those characters on the table, let's return to our study of notions like formal entailment on the one hand, and notions like justificatory support and belief being rationally responsive to one's evidence on the other. We'll do this by thinking abut the variety of Closure Principles a philosopher might embrace.

Many philosophers are sure that some kind of Closure Principle must be true. They'll share Rich Feldman's view that the suggestion otherwise is "one of the least plausible ideas to come down the philosophical pike in recent years" (Feldman 1995).

Most of the discussion of Closure has concerned its applicability to knowledge; but in this discussion I want to mostly steer clear of knowledge. Nearly everyone who likes Closure about knowledge will also be sympathetic to Closure Principles about justification, though getting the details right may be tricky. (It's a familiar point, for example, that just because E is *evidence for* P, and P entails Q, E need not itself be evidence for Q. It may well be evidence against Q.)

Let's begin by articulating a Closure Principle that's modeled on Williamson and Hawthorne's popular thesis about knowledge, but here transposed to a thesis about "doxastic justification." This principle will say:

(Closure-A) If you justifiably believe P, and you "competently deduce" Q from P while retaining your justified belief in P, then that belief in Q is also a justified belief.

The notion of "competently deducing" needs some explanation and care. The Principle won't be able to do the work it's regularly invoked to do if "competent" here means "whatever it takes to make the rest of the principle true." We should instead be able to specify in a substantive way which deductions count as "competent" in the relevant sense, and which don't. This will presumably involve at least that P *does* entail Q, and also that the subject "simultaneously recognize" this. "Recognize" in the sense that the subject "sees" or appreciates the entailment—perhaps this will involve justifiably believing it or knowing it to be true, though perhaps it should involve more than that, and perhaps it should involve less. "Simultaneous" in the sense that the subject puts together in her mind her recognition of the entailment and her justification to believe P: she draws on each of them in making the deduction (perhaps in different ways).

Here's another Closure Principle to consider. This one targets the notion of "prospective" or "ex ante" or "propositional" justification, rather than "doxastic justification." (The label "propositional justification" is misleading for several reasons: (i) it is essentially justification for a doxastic attitude, perhaps the attitude of suspending belief, and attitudes can't be mapped injectively onto propositions; also (ii) this label encourages the idea, which some would accept but others would resist, that the justification in question can itself be thought of as a set of propositions. For these reasons, I think it's unfortunate that the label has become so entrenched, admittedly even in my own usage. I'll try to reform by saying "prospective" instead.) It also targets the all-things-considered species of this:

(Closure-B) If you have all-things-considered justification to believe P, and you "simultaneously recognize" that P entails Q, then you also have all-things-considered justification to believe Q.

Understand "simultaneously recognizing" here as in the previous case. This Principle allows for the possibility of what Wright calls "transmission failure": the principle doesn't insist that P *itself* be a permissible basis for you to believe Q on. It merely says that somewhere or other some justification to believe Q exists. (Some philosophers have argued that Closure-A does not allow for this possibility; but we won't digress.)

A third principle, call it (Closure-C) would be just like (Closure-B) except it would concern the existence of *prima facie* rather than all-things-considered justification:

(Closure-C) If you have prima facie justification to believe P, and you "simultaneously recognize" that P entails Q, then you also have prima facie justification to believe Q.

Having prima facie justification to believe Q is compatible with also possessing defeating evidence of various sorts, whose net effect is that you lack all-things-considered justification to believe Q, and aren't in a position to have a doxastically justified belief in Q.

A fourth principle might be called a "Closure Norm." At a first pass, this might say something like this:

(Closure-D) If you believe P and "simultaneously recognize" that P entails Q, then: believe Q!

But this simple-looking exhortation could be unpacked or refined in a number of different ways.

One issue is whether the restrictor clause should really say merely "if you believe P." Wouldn't it be better instead to insist that the input attitudes be *justified* ones? We don't think that someone who believes P unjustifiably really is justified, or doing what she ought, when she believes P; why then should she be justified when she believes P's obvious consequences (including P itself, and things like P&P, and so on). There are complex issues here about the normative properties of mere (not-necessarily-justified) attitudes that interact in interesting ways with our main issues in this paper. I have addressed them elsewhere; but we can't

pursue them now. So let's just ignore the question of whether the restrictor should say "if you believe P," or something less merely descriptive and more normative.

A second issue is whether the restrictor should be framed in terms that bring in the subject at all. Perhaps it should say not "if you believe P and ... recognize that P entails Q, believe Q!" but rather "if P and P entails Q, believe Q!" This brings in different complex issues about what the nature of norms are, and how to understand the familiar English idioms we use to express them. Let's ignore these issues too.

A third issue concerns what is the *strength or force* of the exhortation captured or expressed by (Closure-D). One option is that it's a mere *recommendation*, where that's understood to fall short of necessarily constituting an all-out mandate. Understood in that way (let's call it D1), our Closure Norm seems to be saying something roughly along the same lines as Closure-C. Another option would be that the exhortation corresponds not to a weak *should*, or *it would be a good idea*, or *there's something to be said for*, but rather to a more forceful, pro tanto *must*. Call this option (D2). Construed in this way, the Closure Norm would be saying when you believe P (or however the restrictor should be formulated, subject to the issues we've just set aside), you *must* believe Q. It's necessarily a violation of rationality for you to fail or refrain from doing so—at least, while simultaneously satisfying the restrictor.

That's not yet the strongest possible construal of the exhortation. For saying that the Closure Norm delivers a pro tanto *must* may be compatible with going on to say that you are also subject to other, independent pro tanto *musts*, and that in some normatively "dilemmatic" or "tragic" scenarios, it might be that these *musts* compete. It might be that you must believe Q, and also, for different reasons, must refrain from believing Q. We will discuss such possibilities more below. For the present, I want only to observe that there could be a stronger reading of the Closure Norm, which said not merely that you pro tanto must believe Q, when satisfying the restrictor, but moreover that this *must* is exhaustive, or would trump any other pro tanto must that competed with it. So the Norm itself would exclude the possibility of such normative dilemmas, whatever their possibility in other cases may be. Call this option (D3).

So there is a range of Closure Principles for notions like doxastic and prospective justification, and norms about what to believe. I have no complaint against Closure-C or the similar-sounding Closure-D1. What to say about Closure-D2 is complicated in ways that we will address below. I want now to complain about the other Principles: Closure-A, Closure-B, and Closure-D3.

It seems to me that all of these Principles fail to respect the possibility of a range of "deduction-threatening" evidence that has become familiar over the past decade from discussions of Disagreement and Higher-order Defeaters. For example, consider Darla, who like Adam starts out justified in believing some Graph theorem A, but reasonably uncertain about a Group theorem O that in fact turns out to be equivalent to A. Unlike Barry and Clarice, let Darla enjoy the same eye-opening mathematical experiences that Adam does, and let her in response raise her confidence in O, just like Adam did. However, let it also be the case that, unlike Adam, Darla has evidence that intuitively should call for her to hesitate. Perhaps she has evidence that her coffee was drugged with a substance that renders one seamlessly and subjectively undetectably worse at math. Or perhaps her math professor tells her, falsely but in a way that should be convincing, that A does not entail O. Or perhaps her epistemology professor tells her that she lacks adequate justification to believe O, but doesn't have time right now to explain why. Yet, despite having some evidence of these sorts, Darla ignores it. She goes ahead and believes O anyway, ending up just as confident in it as Adam is. It seems to me that she's doing something wrong. And more specifically, it seems that her confidence in O is to some degree misplaced. She ought to be less confident in it. (Or perhaps there's some other dimension in which her confidence in O should, but doesn't, differ from Adam's. I'm open to the idea that it's wrong to think of this as simply being "less confident." Perhaps she should be as confident in O as Adam, but be less committed to that confidence, or something like that. I'm open to us going in that direction, but I don't have a firm grip on how to do it, so for this discussion I will just continue to say she should be "less confident.")

I think this kind of possibility is inescapable. Our vulnerability to these kinds of undermining defeaters is ubiquitous.

Section 4

Let's consider some objections.

First, you may want to say that, sure, "deduction-threatening" evidence of the sort I described is always possible, but it would always preclude you from satisfying the restrictor clauses in our Closure Principles, so the Principles aren't threatened by it. For example, when Darla's math professor tells her that A doesn't entail O, she's no longer justified in believing, or no longer knows, that it does, and that's part of what it takes for her to "recognize" the entailment.

This seems to me to be too fast. In my understanding of the ordinary English, there's no incompatibility between Darla's seeing or recognizing the entailment, and also having evidence that ought to (but in fact doesn't) make her somewhat less confident that the entailment holds. It's more controversial how "knowledge" works, but perhaps she could even *know* that the entailment holds, while having evidence that ought to (but in fact doesn't) make her somewhat less confident that it does. Knowing something is compatible with acquiring *more* evidence for it than I had before; and though this may be disputed, perhaps that means it's compatible with me appropriately becoming even *more confident* in it than I was before. And, though this may also be disputed, perhaps that means that the reverse can happen, too: knowing something is compatible with becoming *less*

confident, while still knowing it. Or at least with getting evidence that should make me less confident, while still permitting me to know it. And so however the restrictor clause is understood, I'm inclined to think that Darla could still satisfy it in the cases I described. Yet her final confidence in O is to some degree misplaced. She should be less confident in O than she is. I don't say that she should remain as agnostic about O as she started out being. Perhaps she should be somewhat more confident in O, but not as confident as Adam is. But if her confidence in O is misplaced enough, it might be wrong to count her belief in O as doxastically justified, or one that she has all-things-considered justification to have, or must have. Yet this is what Closure Principles A, B, and D3 do say.

Some fans of Closure will emphasize normative aspects of the restrictor clause over descriptive ones. They may protest here: perhaps Darla *does* still confidently believe that A entails O, and perhaps in the cases described this belief can still be called a "recognition." But that's the word *you* used to formulate the Closure Principles, Pryor, not we. We're free to formulate them in other terms, that Darla would satisfy to a lesser degree in the cases you presented. Just to take a simple proposal, and not to commit to it specifically, suppose the Closure Principle is restricted by the assumption that Darla is *justified in believing* that P entails Q. When her math professor tells her that P doesn't entail Q, she would then be *less justified* in believing that, even if she *does* still believe it—and perhaps even in believing it may still know that P entails Q.

My reply is: fine, Darla can satisfy some normative quality relevant to the restrictor to a lesser degree, for example by being less justified in believing in the entailment. But the restrictors we've considered are all in the end *categorical* ones, so "satisfying them to some lesser degree" is compatible with still *satisfying* them. Darla may be *less* justified in believing in the entailment, but still *justified*. And so the Closure Principles will tell us that she's justified in, or must have, the confident belief in O that she in fact has. True, these Principles can also allow that she's *less justified in that confidence*. But they do still say that she's justified in it. And I don't see why that has to be the case. Why can't the amount of confidence it's appropriate for her to have in the entailment, and the resulting confidence that it's appropriate for her to have in the conclusion also go down, but this exceed some threshold for that conclusion such that it's *no longer* appropriate for her to (all-out) believe it?

A Bayesian may complain that I'm counting Darla as justified in believing P because the confidence it's reasonable for her to have in it exceeds some threshold, and similarly for her being justified in believing that P entails Q, without Darla ever being justified in *believing those claims jointly* above the relevant threshold. That is how Darla can manage to be unjustified in believing Q. But, the Bayesian will continue, it just shows we should reconceive the terms in which the Closure Principle is stated. We shouldn't be evaluating the justification of her belief in P and her belief in the entailment separately, but rather we must evaluate their *joint* justification. This complaint make sense. If we were to translate the things

I've been saying into Bayesian terms, this may be their closest approximation, and that is what a Bayesian should then say in reply. Of course, the elephant in the room is that Bayesianism is a completely inadequate framework for discussing cases like the ones we're focusing on, that take seriously the epistemic differences between Adam, Barry, and Clarice. Perhaps there may be some way to preserve the spirit of the present objection without relying on the particulars of the Bayesian framework. I will leave that idea for my opponents to develop. (One question that will loom if they do so is whether they'll be pushing us back to a more naive Closure Principle of the form "If you believe P and P *merely does* entail Q, then..." They may just be restricting that Principle to cases where P has certain special contents, in part concerning what entails what.)

A different strategy for fans of Closure Principles would be to insist that the normative relations Darla is assumed to stand in to the entailment be indefeasible. My reply is: fine, but then the Closure Principles never have application. No one *ever has* indefeasible entitlement to believe the entailments hold, and to reasonably ignore any "deduction-threatening" evidence of the sort I described.

If the fans of Closure want to introduce some *other* Closure Principles at this point, with *explicitly graded* restrictors, I'm open to considering them. But things will necessarily be more complex than what we have on the table. And I suspect that any such new Principles will be moves in the direction of the weaker Principles like Closure-C that I'm not challenging. Any such Closure Principle will be substantially weaker than what's typically invoked in mainstream epistemology. Now, often enough, philosophers help themselves to Closure Principles they don't really need. They'll really only need a *particular claim* of the form "In the cases just described, if the subject knows so-and-so, she also knows such-and-such." And often these particular claims are plausible enough on their own, because it's implied by the description of the relevant cases that no "deduction-threatening" evidence is present. So often philosophers invoke Closure when they don't really need to. Those philosophers who really *do* rely on Closure, though, typically will want and need something with more bite than Closure-C.

All of this has been an exploration of whether Darla must fail to satisfy the restrictor clause of the Closure Principles, whenever it's true that she lacks justification for being as confident in O as she is in A. Let us turn to a different strategy for resisting what I'm saying.

Some philosophers will reject the claim that what I've called "deductionthreatening" evidence can be effective at the first order. They may be willing to grant that it defeats Darla's justification to believe *higher-order* claims about herself, like the claim that *Her belief in O is justified*. But they'll deny that it does anything to impugn the justificatory status of her *first-order* belief in O, itself. The higher-order facts and the first-order facts are separate matters, these philosophers insist, and Darla's attitudes towards them needn't be straightforwardly correlated. Williamson has for example argued (with his "unmarked clock" example) that it's possible to know P while knowing that it's extremely unlikely that you know P. My reply is: I *agree* that the higher-order and first-order facts are logically independent, and I am sympathetic to the claim that they aren't straightforwardly correlated. Moreover, I *accept* the thesis I just attributed to Williamson, and I'd be willing to accept analogues for justification as well. None of what I'm saying here relies on denying any of this. That is, I'm *not* relying on any kind of principle like these, which one sometimes sees defended:

- If you're justified in believing H, then you're justified in believing that you're justified in believing H. (Akin to an S4 axiom.)
- If you're justified in believing H, then you're not justified in believing that you're not justified in believing H.
- If you're justified in believing that you're justified in believing H, then you're justified in believing H. (Akin to a Converse S4 axiom, consider also Feldman's slogan "Evidence of evidence is evidence.")
- If you're justified in believing that you're not justified in believing H, then you're not justified in believing H. (Akin to Converse S5 axiom, ALSO FELDMAN?)

As far as I can see, my diagnosis of the Darla case needn't be relying on *any* general principle linking higher-order and first-order epistemic properties. For it's not obvious that the "deduction-threatening" evidence is always really implicitly about, or must always work by targeting, higher-order epistemic claims. But if I *did have to pick* some kind of linking principle here, it would be this:

• If you're justified in believing that you're not justified in believing H, then you're *less* justified in believing H (than you would be without that higher-order justification).

This is compatible with the Williamson-like claim that justification to believe you're not justified in believing H is compossible with still having quite a lot of justification to believe H.

In the Darla case, as I understand it, she has some defeating evidence that calls for her to be less confident in believing O than she actually is. Adjust her priors and the strength of the defeaters in the right way, and also how much confidence in O we count as all-out believing it, and it may be that the defeating evidence calls for Darla to (all-out) withhold belief in O, rather than believe it. As we discussed, it may be that in any such case, she'll also have less of some normative quantity relevant to the Closure Principle's restrictor clause, but I don't see why we should expect a convenient harmony, such that she fails to (all-out) satisfy the restrictor clause in every case where she's enough less justified in believing O that she shouldn't believe it. Against the background of these objections, we can map out a variety of positions philosophers are taking on these issues.

TODO Anti-akratics. Externalist / objectivists. Infallibilists. Incoherentists.

Third objection. IDEALS TODO

Section 5

Where are we? I draw two lessons from the Darla example.

First, the "input" justification for A and the "recognized" entailment described in our Closure Principles are not enough to guarantee justification for Darla's "output" belief or confidence in O. Whether she has the "output" justification is also hostage to *the absence of* any (undefeated) "deduction-threatening" evidence. Closure Principle C (and perhaps D1) are the only ones I've presented that acknowledge this.

Second, in the presence of "deduction-threatening" evidence, *what is* a reasonable pattern of attitudes for Darla to have? I'm suggesting she should be less confident in O than she is in A, though these claims are mathematically equivalent. So she should have doxastic attitudes that are mathematically incoherent. But then, so too should Barry and Clarice. What's different about Darla is that she can—in *some* sense, as we've seen not a sense that it's easy to uncontroversially describe—still "recognize" that A entails O, and so can presumably also recognize the way in which that pattern of attitudes she *should* (in fact) have is incoherent. So she should remain very confident in A and be somewhat less confident in O, despite A's entailing O and her "recognizing" that it does so. Her "recognizing" that doesn't entitle her to reasonably ignore evidence that it doesn't, and other sorts of "deduction-threateners." I submit this is probably the rationally least bad combination of attitudes for Darla to have in this case.

You may protest: but if she's in such a situation, then surely she must *re-evaluate* her reasons for thinking that A entails O, and her reasons for trusting her professors, or that her coffee hasn't been drugged, and so on. Fine. I don't myself think it's so obvious how she must conduct her intellectual life going forward, but I see no reason to resist these claims. Perhaps embracing a recognizably incoherent set of attitudes will necessarily add a lot of obligations to one's mental to-do list. But I'm *not here addressing* what intellectual activities Darla must go on to perform in the future. I'm addressing the question of what doxastic attitudes it is appropriate for her to form *right now*, given the insights and understandings and evaluations she has now available, and no more.

Section 6

In stating the second lesson of the Darla example, I said that having a recognizably incoherent set of attitudes may be the "least bad" doxastic option available to

Darla. Should we say that those attitudes can be positively reasonable?

I don't know. There are two ways we could go, and I'm not sure which is best. (Indeed I have some inclination to think that in different cases, each of these perverse options may obtain.)

One option is to say that sometimes recognizably incoherent combinations of attitudes *may* be rationally permissible. Arguably this is the right way to think about the Preface Paradox. In Preface scenarios, your epistemic position is one that recommends believing each of the claims in your book, and also recommends suspending or dis-believing the claim that they are all true. Though this may be disputed, perhaps we can gloss the latter as: suspending or dis-believing *the conjunction of* those claims in the book. And it's not just that *part* of your epistemic position recommends some of those attitudes and *other parts* the other. That's true; but it's *also* true that, *as a whole*, you are recommended to have these attitudes *jointly*: belief in each of H1...Hn, and disbelief or suspension of belief in the conjunction H1-and-...-and-Hn. There are many views people advocate about the Preface, but this is one very natural one.

Another case with this structure may be when your epistemic position recommends believing A, and also (jointly) recommends suspending belief in O, even while "recognizing" (as Darla arguably does) that A and not-O can't be true together.

Or perhaps at least sometimes we should think of these cases differently. Perhaps sometimes we have a normative "dilemma" or "tragedy," that is a case where all of your options are in some way pro tanto bad. An ethical example may be if you've promised to kill someone. Now if you go ahead and kill them, that would be bad because you should never kill. On the other hand, if you refrain from killing, that would also be bad because you should always keep your promises. Of course you also *shouldn't have made* that promise, but you did make it. And we might suppose that the relevant norm here is not "Always keep your promises except the ones you shouldn't have made," but rather just "Always keep your promises," simpliciter. In such a case, no matter what you do-kill or refrain from killing—you will be doing something bad. I trust that cases of this structure are familiar. Of course I don't insist that this is necessarily the right way to think of promises to kill. But it is an intelligible and familiar view for a philosopher to hold. Note that one who holds this view is not committed to the claim that anything goes, normatively. The multiple norms that apply to you and give incompatible recommendations can agree in prohibiting various other options; hence it would be determinately and non-dilemmatically wrong to make any of those choices. Nor need an advocate of dilemmas say that none of the competing options can reasonably be preferred over the others. He can say that, but he needn't. In the case I described, we might say that going ahead and killing is worse than refraining from killing, because killing is in general worse than breaking promises. It's just that we insist that this doesn't exonerate you from the vice of breaking the promise. You're not "off the hook" for breaking the promise, just because that is the lesser bad of the several bad options available. If there are any such normative dilemmas, in ethics or elsewhere, what constitutes their being a dilemma? Is it just that several incommensurable norms conflict? No, I think one can believe in incommensurable norms without believing in dilemmas. (At most one's forced to admit cases where there is no determinate fact about what it is a subject should do.) Neither should we understood dilemmas as its being impossible to rationally prefer one of your options over the others, because as I've just said, that is not an essential commitment of belief in dilemmas. Rather, I think the key idea is what I expressed by saying that, whatever option you choose, that will constitute doing *something* wrong, for which you're not "off the hook." Its wrongness isn't neutralized by the fact that all the other options may have been worse.

It is as I said disputed whether there can be any such ethical dilemmas, but I don't think we can rule them out. Neither do I think we can rule out their epistemic analogues. Perhaps there are cases where your epistemic position recommends that you believe P, and also (jointly) recommends that you suspend or refrain from believing P. Perhaps there are cases where it recommends believing P, and also recommends that if you believe P, then you believe Q, and also recommends that if you believe P, then you suspend belief in Q. Perhaps there are cases where it recommends belief in U, and recommends not both believing P and suspending belief in U. We will consider examples of this third form in Section 8.

It may be hard to keep track of the difference between a mere recommendation to hold incoherent attitudes, and a dilemmatic set of recommendations. One difference is that in the first cases, it is possible to *jointly comply* with all of the recommendations. It may just be that in doing so, you'd be having a set of attitudes that you recognize can't jointly be true. (I don't think this *must* be true of *all* cases of incoherent attitudes, though it is true of the examples I gave.) Now perhaps you think there is *also* a norm *not* to have any such set of attitudes: a norm that says never to have a set of attitudes that you recognize can't jointly be true. If so, then you will think that the alleged examples of mere recommendations to hold incoherent attitudes ought really to be understood as disguised dilemmatic recommendations—if such be possible. They would all be like the last of my three examples of dilemmatic recommendations, where some of the recommendations tell you to do certain things, and the last recommendation says don't do all of those former things simultaneously.

Another difference is that in some of the dilemmatic examples, the recommendations weren't simply recommendations to have certain attitudes. In the second of those examples, some of the recommendations were to have some attitudes if you had some other attitudes. In the third example, the last recommendation was to refrain from having some other attitudes simultaneously. I'm not sure whether we can get structures like this in the case of mere recommendations to hold incoherent attitudes. It may be that incoherences of this sort must always be dilemmatic. This is just a speculation.

Now perhaps you think that a subject is *always* justified in doing the least

unreasonable thing. And you may also think that *something* is always least unreasonable—it may just be that in cases of ties, it is indeterminate what that something is. If you held such views, then you will think that the alleged examples of dilemmatic recommendations ought really to be understood in other terms. It's impossible for *everything* to be forbidden. But I am not myself convinced that those views are right.

Because I am not sure which of these possibilities obtain—whether recognizably incoherent sets of attitudes may sometimes be positively reasonable, or whether there may be cases where all of the doxastic options are positively unreasonable—I have to state my claims about Darla with some care. I say that the *best* option for her to take may be to be confident in A, and less confident in O. Perhaps though that is not a *good* option. Perhaps in so believing she will necessarily be manifesting an epistemic defect. She can't get "off the hook" for holding a recognizably incoherent set of attitudes. That may be. I insist only that in cases of this sort, there may be no *better* option that's open to her.

Now you may be wanting to protest that instead of any of these perverse options, what a subject should always do instead is *suspend belief*, over a sufficient range of propositions such that the incoherence never arises. This strategy may be right in a few places, but I'm not convinced it's going to give us a satisfying account of all the examples. Darla shouldn't suspend belief in A: she's had justification for it all along, and nothing has emerged to threaten that. Perhaps in the cases where she should refrain from believing O, she should also suspend belief that A entails O. I questioned the inevitability of this in Section 4, but perhaps we might persuade ourself to accept it. It's worth noticing, though, that this strategy won't make the fundamental problem disappear. For suspending belief is *itself* a doxastic attitude, that we can have better or worse grounds for taking. I've been discussing cases were Darla in fact has good direct, first-order grounds for *believing* something, but then also some other justification that doesn't cohere with doing that. We could equally well discuss cases where she instead has grounds for suspending belief about something, but also some other justification that doesn't cohere with her doing that.

For example, the background may be that there are some well-known but non-conclusive reasons for accepting some mathematical hypothesis, like the Continuum Hypothesis. Then Darla has a mathematical revelation that against that background should prompt her to *suspend belief* in the Continuum Hypothesis. She comes to see that other things she's mathematically justified in believing speak against the Continuum Hypothesis. Let us suppose that suspending belief in the Continuum Hypothesis is what now best coheres with her other mathematical beliefs and insights and understanding. Then Darla *also* acquires some non-mathematical evidence that should undermine her justification for so suspending, like her professors telling her she's made a mistake, or evidence that her coffee has been drugged. In the face of such evidence, I suspect the best set of attitudes for Darla to have will be one that strikes us as less "coherent" than the attitudes it would be appropriate for her to have were the defeating evidence to be absent. For we said that suspending belief in the Continuum Hypothesis is what best cohered with her other *mathematical* grounds, so continuing to *believe* the Continuum Hypothesis, instead, will manifest more mathematical incoherence. And this additional evidence Darla has acquired isn't itself mathematical, so it won't make that *mathematical* incoherence go away. If Darla should be at all responsive to the non-mathematical defeating evidence, then the pattern of attitudes it will be reasonable for her to have will manifest some of that mathematical incoherence.

So we see the same epistemic pressures arise when it's Darla's grounds for suspending belief, rather than her grounds for believing, that are threatened.

(Of course, you could resolve to employ the word "coherent" in such a way that whatever pattern of attitudes is most reasonable is the one that's most coherent. I haven't after all specified a different meaning for "coherent"; I've just been relying on our pre-theoretical understanding of the term. If we decided to talk this way, then I should express myself instead by saying things like: sometimes beliefs whose contents are logically or mathematically inconsistent are the ones that cohere best.)

Section 7

One interesting example of dilemmas may arise where you have attitudes that are statically bad, that is they don't in fact fit your evidence. So we may think that there is some static norm applying to you that says: Have such-and-such other beliefs! But at the same time it may be that your present psychology, taken as a whole, may not provide you reasonable grounds for *changing* your attitude. If you were to simply spontaneously shift to the statically better beliefs, you would be displaying the same kind of epistemic whimsy we complained about with Clarice. You may say, well wouldn't the evidence that statically justifies the other attitudes in the first place be an acceptable ground for changing your mind? The difficulty is that those grounds may not be appropriately available in the company of other aspects of your psychology—other mere unjustified attitudes you have—that wouldn't themselves render you *statically* unjustified in having the target beliefs. If you did spontaneously jump to the new beliefs, those ending attitudes may be statically justified; it may just be dynamically objectionable how you got there. Here the issues I set aside in Section 3 about the normative place of mere (not-necessarily-justified) attitudes are important. A similar case, outside the epistemic domain, is Buridan's Ass. If he set out walking in either direction, arguably that choice would be a statically justified one. But there is no dynamically appropriate way for him to rationally transition into such a choice. (Thanks to David Sosa for discussion here.)

As I acknowledged at the end of Section 5, subjects like Darla and others in dilemmas or who are recommended to have recognizably incoherent attitudes may as a consequence be subject to *other norms as well*, about how to conduct

their intellectual lives in the near or far future. Recommendations to think harder about their present situation, double-check their calculations, and so on. Following *these* recommendations may *change* their epistemic position (as thinking about math changed Adam's position), to one where they can better see defects in the thinking and evidence that generated the dilemma or incoherent recommendation. That may then give them a non-dilemmatic, perfectly hygienic way to extricate themselves from the uncomfortable epistemic scenarios I'm describing. I see no reason to be confident this will always happen, but if it can, that's fine and doesn't threaten the claims I'm making here. As I said before, I am not addressing what Darla should *go on to do*, or what attitudes she may end up justified in having if she does it. I am addressing what is recommended by her epistemic position here and now, *before* she enjoys any such gains in epistemic clarity or insight.

Section 8

Let's consider another case where the background we've developed may have interesting application. I will urge that the preceding reflections against Closure help teach us how to live in the company of skepticism (though not by the familiar Nozick/Dretske routes).

I will be discussing immediate justification. By this I mean justification to ϕ that is not even in part constituted by your having to justification to ψ , for appropriate choices of ϕ and ψ . I assume that the relation of being some justification that constitutes some other justification is acyclic. For explicitness, I will focus on *having some quantity* of justification, rather than on *having enough* justification for settling on some attitude like belief (or for that matter, withholding). Otherwise, I won't fuss about which attitude we're talking about: whether it be some justification for believing, or for withholding belief, or for having some high or low confidence. These are of course importantly different, but I think much of what I'll be saying applies indiscriminately. I will begin by discussing prospective prima facie justification, but as we proceed we will be shifting to considerations about all-things-considered and doxastic justification.

I believe that some of our justification is immediate, and so too should you. I have argued elsewhere that at least some *perceptual* judgments have this status, such as *There is light ahead*. But even if you don't agree about *that*, you should at least think that some *other* cases of justification are immediate. Some or all of the beliefs *That [ostending an experience] is experienced*, *I am aware of something*, 2 > 0 should be good candidates.

Yet, as I said when discussing Darla, vulnerability to undermining defeaters is ubiquitous. Perhaps we show you psych studies that ought to be convincing (but in fact are made up) that depict you as introspectively or mathematically unreliable. Perhaps you cannot, consistently with retaining the clear reflective grounds you in fact do have for the immediately justified claims, *fully understand* how it could be that the allegedly "defeating" studies are correct and you are screwing up. That does not mean that you should ignore the studies. It's natural to think that in the face of such studies, you ought to be somewhat less confident in your reflective beliefs than you were before seeing the studies. (I repeat here my acknowledgement from Section 3 that "less confident" might not be the right way to think about the respect in which you ought to back off. My root intuition is that you should in some sense back off, or be doxastically less committed. This root intuition applies even in cases where the attitude you should back off from is a withholding, rather than a believing. It would be nice to have a better developed, systematic way of thinking about this than in terms of the crude language of "being less confident." But I don't have that.)

So it looks like your justification to believe P can be defeated—more specifically, undermined—by your learning some other claims U. And yet we alleged that your justification to believe P was immediate, and so not constituted even in part by your having justification to disbelieve U.

At this point, some philosophers may want to introduce a wrinkle. They may say, true, your justification to believe P wasn't (partly) *constituted by* your justification to disbelieve U. But it may have been (partly) *enabled by* your having justification to disbelieve U. I don't object in principle to this distinction. But I don't think it will fundamentally alter the dialectical structure I'm developing. I assume that the relation of being some justification that *enables you to have* some other justification is also acyclic. I assume also that some chains of justificatory constitution and enabling are finite. What all these assumptions force on us, I think, is that eventually, at some point, whether we bring enabling into the picture or not, we are going to have to confront some case like this:

Though U is a potential defeater for your justification to believe P, you nonetheless can have justification to believe P which isn't constituted *or* enabled by your having justification to disbelieve U, nor more generally by your justification to have *any* attitude towards U.

Introducing enabling into the discussion only makes it somewhat more complicated to show this. Anyway, I will help myself to the idea that there are some such cases. If that's right, then either (i) we can sometimes have justification to believe P without even *needing to have* any justification to disbelieve (or have any other attitude towards) U, or (ii) justification to believe P and to disbelieve U always come *as a package*, neither constituting nor enabling the other. Option (ii) may hold in some of these cases, but I doubt it can hold in full generality, wherever cases of the problematic structure I'm describing threaten. So I will set option (ii) aside, and concentrate henceforth on (i).

We are going to consider a general threat to any case of this structure.

If you're not justified in *disbelieving* U, perhaps you happen to be (even if your justification for P doesn't require you to be) justified in having some *other*

attitude towards U: perhaps suspending or refraining from believing it? It's a tempting thought that:

(Thought-1) For any question you understand, or any question you're now considering, there must always be *some* doxastic attitude (or restricted range of attitudes) towards it that you're justified in having.

I say a "restricted" range to exclude the tautologous construal. I say this is a "tempting" thought, but in fact I myself have come to have deep doubts about it. I won't try to argue against the thought here. I'll just observe that it can be intelligibly rejected, but go on to accept it for the sake of our ongoing discussion.

So we assumed you were immediately justified in believing some claim P. If you had justification to disbelieve U—that is, justification to rule out the defeater—that would be great, but we just argued that there should be possible cases where you're so justified in believing P but don't have that.

(Some may say, for the plausible kinds of P and the Us that would undermine our immediate justification to believe them, those Us will seem a priori quite unlikely. So we may always have some kind of default justification to disbelieve them. Fine, then don't focus on your justification for P and its defeater U, but rather focus on your default justification for not-U and *its* defeaters—as I said, the possibility of undermining seems to be ubiquitous and it's not clear why default justification should be immune. True, because default justification isn't generated by any cognitive performance by the subject, the same range of underminers may not apply. But that doesn't mean that no underminers at all apply. For example, what if your philosophy professor tells you convincingly that you lack any such default justification for not-U? So perhaps it's your default justification to believe not-U and some potential defeater of *it* that gives us an example with the structure I want to consider. Because I don't think that introducing default justification alters the fundamental dialectic—it just shifts the playing field—I will continue to talk about P and U.)

Now, if you happened on the other hand to have justification to *believe* U, that would be bad. It would defeat the justification you have to believe P. But of course, you needn't be in such straights.

If (Thought-1) is correct, then if you don't have justification to disbelieve U and you don't have justification to believe it, you must instead have justification to have some other attitude, like suspending belief, towards it. But here things become very awkward. Is it really possible to have immediate justification to believe P—in a form you can *make interesting use* of, that is, undefeated—and simultaneously be justified in *suspending belief* towards some potential defeater for that justification? Crispin Wright argues it is not:

"I cannot *rationally* form the belief that it is currently blowing a gale and snowing outside on the basis of my present visual and auditory experience while simultaneously agnostic, let alone skeptical, about the credentials of that experience." (Wright 2004)

[Consider also FELDMAN'S claim(s) that having justification to believe you're justified in suspending/disbelieving P implies being justified in suspending/disbelieving P.]

The upshot seems to be that the only cases in which our immediate justification to believe P is actually *usable*, that is undefeated, will be ones where *we happen also to have* justification to disbelieve U. And this is the general threat. Even if, technically, the justification to believe P does count as "immediate" and so isn't at all *constituted by* having justification about U, still it seems that some such independent justification for not-U will be needed, for your justification for P to be worth much.

This may overstate the problem. Arguably in the last case, where you should suspend belief about U, your justification to believe P wouldn't be *entirely* defeated, but only *partly* defeated. If you came to have more justification, in support of *believing* not merely suspending belief about U, that would presumably defeat your justification to believe P *more*. Hence in the absence of that additional justification, your justification to believe P wasn't *entirely* defeated. (Mark Schroeder taught me this nice argumentative strategy.) Still, you see the shape of the problem. The threat is that even if some justification to believe other things, like not-U, still in practice it will only be effective to an extent correlative to how much other justification you happen to independently have against U.

We might put it like this: Suppose you do have an undefeated justified belief in P. Then there's a "coherence constraint" on you to disbelieve U if you have any attitude towards it. So any other attitude than disbelief towards U must *not* be justified, but by (Thought-1), *some* attitude has to be justified. Hence, in order to justifiably believe P, it will at least be a necessary condition for you to have justification to disbelieve U. (Whether the connection here can be made any *more* intimate than a necessary condition we can leave for later discussion.)

Notice the step from "these attitudes would not cohere" to "these attitudes could not be jointly justified." Behind this reasoning stands something like the following thought:

(Thought-2) If your evidence supports some attitudes β (in our case, believing P), and actually having β would be (perhaps we should say impermissibly—I'll comment on this qualification later) incoherent with some other attitudes δ (in our case, suspending belief about U), then your evidence can not simultaneously support δ .

And this is what I want to use our earlier reflections to resist. I'll be urging instead that (Thought-2) needs to be restricted or rejected. The connections

between notions like justification and what your evidence supports, on the one hand, and notions like incoherence on the other, is less straightforward.

Recall the cases we discussed before. Perhaps your epistemic position recommends believing P, and also recommends *refraining from* any belief about U; and this combination is in some sense recognizably incoherent; but *it may for all that be the best doxastic option open to you.* (It's no refuge to think in such cases it's always instead the case that you're justified in *suspending belief* in P; for as we argued at the end of Section 6, analogous conflicts arise with *suspension* substituted throughout for believing.)

We discussed two ways of fleshing this out earlier. Perhaps it is *positively* reasonable for you to believe P and suspend belief towards U, despite a sense in which this combination of attitudes may be one you can recognize as incoherent. (Spelling out what the incoherence consists in is harder here than in the entailment case, but I share Wright's judgment that there is some kind of awkwardness here.) Or perhaps instead you are also recommended *not* to have any such total pattern of attitudes; but at the same time, you are recommended to have each of them. So whatever you do, you'll be manifesting some epistemic defect.

If we choose the first option, we should reject the unqualified form of (Thought-2), but we can accept the form which restricts itself to β s that are *impermissibly* incoherent with other attitudes. We'd just say that the cases we're discussing are ones of rationally *permissible* incoherence. If we choose the second option, we should reject (Thought-2) even in that qualified form. It may be that even *im*permissibly incoherent combinations of attitudes are sometimes what your epistemic position supports. That is, sometimes you *are* recommended to make doxastic choices you cannot simultaneously follow without thereby doing something unpalatable. (Consider here Closure Norm D2, as opposed to D3. As D2 implies, perhaps failing to believe the "recognized" consequences of other things you believe is something *you can never get off the hook for*, even if it is sometimes part of your *least bad* total response, because other pro tanto norms compete with it and turn out to be rationally preferable to conform to.)

I've only suggested that such treatments are possible. I haven't done anything to show you that they are obligatory. Still, I've come to think this is an attractive way to understand the case where you have immediate justification to believe P and no information bearing on the truth of some potential defeater.

(A different way to handle these cases would be to attack (Thought-1), and say that your epistemic position recommends believing P, while failing to recommend any attitude (even suspending belief) towards U. I am interested in and have some sympathy for that strategy, but I'm not going to try to develop it.)

Section 9

One closing thought. Suppose we do opt for the dilemmatic treatment, and say that subjects who believe P while suspending towards U must thereby be

doing something wrong. It needn't follow that subjects who have immediate justification for P but no evidence about U are condemned to be doing something wrong in believing P. Perhaps subjects can fail to have any attitude towards U, even the attitude of refraining. (Thought-1 concerned what justification it's possible to have; this is a claim about what *attitudes*.) Then believing P *while* having no attitude towards U may be a way for subjects to avoid violating any of the norms that apply to them—though perhaps a precarious one, as the state of having no attitude towards U may be hard to sustain, especially when the question whether U is explicitly raised. Even if that isn't possible, and subjects do have to have some attitude towards U, it still doesn't follow that subjects must be doing something wrong in believing P. Nothing requires that the attitude they have towards U be a justified one. Consider a subject who disbelieves U unjustifiably, when her evidence instead supports refraining from belief about U. She is doing something wrong. She is clearly violating *some* norm that applies to her. But if other norms that apply to her tell her to believe P, and not to simultaneously believe P and suspend towards U, is it clear that having the belief she does in P itself constitutes any defect?

Like most epistemologists, I don't think that unjustified belief in reliability (one form that unjustified disbelief in U may take) can *justify* other beliefs that *epistemically depend on the assumption of reliability*. But we've stipulated that the subject's belief in P is immediately justified and doesn't so depend. Hence, I don't see why subjects who have wrong attitudes towards U—who disbelieve it unjustifiably—shouldn't be free to exercise the pro tanto, albeit dilemmatic justification they have to believe P, and *that* doxastic move itself be fully justified. The subject's overall epistemic state may be problematic, but her belief in P needn't automatically share in that wrongness.

Other subjects, more attuned to the strength of their evidence towards U, may refrain from disbelieving it, and so *they* may inevitably be less well off in their stance towards P. This is one case among several where being epistemically more careful may have unexpected hazards.

FORGOTTEN EVIDENCE

References

Feldman, Richard (1995), "In Defence of Closure," The Philosophical Quarterly, 45: 487-494. [Against Audi 1988, 1991]