

Epistemic Planning, Epistemic Internalism, and Luminosity¹

Karl Schafer

8/31/17

Abstract: In this paper, I make use of an “doxastic planning model” of epistemic evaluation to argue for a form of epistemic internalism. In doing so, I begin by responding to a recent argument of Schoenfield’s against my previous attempt to develop such an argument. In doing so, I distinguish a variety of ways that argument might be understood, and discuss how both internalists and externalists might make use of the ideas within it. Then I argue that, despite these complexities, the doxastic planning model continues to support a modest form of epistemic internalism. I conclude by showing that, far from conflicting with “anti-luminosity” arguments in epistemology, this form of internalism is best understood as a natural reaction to these arguments.

Keywords: Epistemology, Rationality, Epistemic Internalism, Epistemic Externalism, Planning, Luminosity, Epistemic Expressivism.

1. Epistemic Planning and Meta-Epistemology

Let’s begin with a familiar sort of philosophical thought experiment or, perhaps better, an academic fairy tale. Suppose we are creating a creature not too unlike us. This creature will have beliefs and desires — or credences and preferences, if you prefer. And these beliefs and desires will combine to orient its choices and actions. In some cases, we will want to hardwire its belief forming mechanisms so that they are invariably responsive to certain sorts of input or evidence. But in many cases we will want to allow our creature some flexibility in how it forms its beliefs, so that its methods for belief formation can be responsive to its experience, its social context, and its particular needs and values.

To do so, we will need to give our creature some capacity for doxastic or epistemic self-regulation. Thus, we will need to give it attitudes that allow it to regulate its practices of belief formation and reasoning in much the way our capacity for forming plans or general policies for action helps us to regulate what we do. In short, we want to endow our creature with a capacity to form what we might call “doxastic or epistemic plans”.² Whether these are “plans” in this or that philosopher’s strict sense of this term need not detain us here.³

¹ Special thanks are due to Robert Steel for many helpful comments on this paper. Thanks also to Yuval Avnur, Sinan Dogramaci, and Miriam Schoenfield and several anonymous referees for very helpful feedback on it.

² Couldn’t beliefs about reliability or conditional probabilities play this role? To some degree, I think they could, but for reasons that will become clearer, I think they will at most do *some* of the work we want epistemic plans to do for our creature. So even if we decide to endow our creature with the capacities required to form such attitudes, as it seems plausible we will, there will still be good reason for us to give it this sort of planning capacity.

³ But isn’t one of the main lessons of the literature on doxastic voluntarism that the very idea of forming plans about what to believe is itself confused. I think this is an overreaction to this literature, provided that we are operating with a suitably broad concept of “planning”. Of course, it *is* true that an ability to effectively plan to believe whatever one chooses would undermine our ability to so much as have beliefs, if this planning ability is *completely unconstrained* by any concern for the truth or what the evidence supports. For example, if I was able to effectively plan to believe whatever my practical reasons best support me believing, this might be sufficient to sever the connection between belief and the truth that is essential to belief.

But from this, it only follows that, in order to imagine a being with the capacity to regulate her beliefs through such plans, we must imagine a creature whose capacity to form these plans is constrained so that it does not undermine the essential characteristics of belief. Whether or not human beings possess such a capacity, it does not seem to me a conceptual impossibility.

Indeed, it seems plausible to me that I often do form “plans” of this sort — but that this planning activity is constrained by a tacit appreciation that my plans for belief formation or judgment will normally only be effective insofar as they are

What is important is simply the functional analogy between these “epistemic plans” and plans of a more familiar sort.

But now suppose we also endow our creature with the capacity to communicate and coordinate with other creatures of the same kind. In doing so, we will want it to be able to communicate, not merely what it believes, but also how it plans to form beliefs. For it will often be important for our creature to be able to coordinate with its peers with respect to the methods they use to form beliefs.⁴ Thus, we will want to provide it with a language that contains, not merely terms that enable it to express its beliefs, but also terms that enable it to express these epistemic plans. Just how these terms function will depend on our aims for this creature. But we will want to give it some linguistic resources for promoting coordination with respect to epistemic plans. And by giving it these resources we will give it the ability, not just to express its plans to others, but also form the corresponding, linguistically structured thoughts “in its head”.⁵

This, of course, is a fairy tale, and actual human beings are far more complex. But nonetheless this tale can help us understand our actual practices of epistemic evaluation. For what is the point of these practices? A plausible thought is that their function is, at least in large part, to help us regulate our beliefs and reasoning, both as individuals and more collectively. If so, the point of these practices is much the same as that of our creature’s practices of “epistemic planning”. And given this, for all its simplicity, our creature’s practice of epistemic planning may have important lessons for our understanding of our own epistemic concepts and practices of evaluation.

Indeed, the relevance of our tale can be understood in stronger or weaker ways. The strongest way to understand it is simply to claim that we are creatures with the same planning capacities as those in our story. To do so would be to *identify* our epistemic judgments with “plans” — or “plan-laden attitudes” — of the sort introduced there. This is the view which is suggested by Gibbard’s treatment of normative judgments as plan-laden. Thus, one way to understand the relevance of our story is to see it as capturing the core of a broadly expressivist account of the nature of epistemic evaluation — a view on which epistemic judgments (at least in the simplest cases) are (at least partially) constituted by these sorts of epistemic plans.

But the relevance of our tale does not depend on an expressivist treatment of epistemic evaluation. All it requires is for there to be a sense in which these creatures provide us with a useful *model* for thinking about the role that epistemic evaluations play in our lives. For this to be the case, two things must be true of our epistemic concepts. First, as Elizabeth Fricker helpfully puts the point, these concepts must be the “kind of thing that can be captured ... by an adequate specification of the role that they play in our socio-cognitive lives.”⁶ And, second, epistemic plans must provide a useful way of *modeling* this role or function, at least for certain purposes.⁷

consistent with the nature of belief itself. For example, it seems very natural to me to say that one plans to believe whatever the experts tells one about some subject. And it also seems very natural to speak of, say, planning to believe one’s senses, especially in contexts in which the reliability of one’s senses might have been called into question. In short, as Schoenfield nicely puts the point: “whether belief formation is voluntary or not, the following is clear: we can have an impact on what beliefs we form by considering how to respond to a body of evidence and then settling on one method of belief formation rather than another.”

⁴ Dogramaci(2015)

⁵ Obviously, this story is a simplified version of the sort of tale one finds in recent forms of expressivism such as those endorsed by Blackburn(1998) and Gibbard(2003). Compare Chrisman(2012).

⁶ Fricker(2015), 57.

⁷ Schafer(2014). See also Gibbard(2003), Schoenfield(2015a,b), Greco and Hedden(forthcoming), Steel(2015), etc

Both of these can be easily true, even if an expressivist treatment of epistemic evaluation is mistaken. All that is required is that the core function of epistemic evaluations is to play a certain sort of role in regulating our epistemic lives, and that this role can be modeled through thinking about our creature's epistemic plans.⁸

This sort of model is fruitful, in part, because it provides us with a way of *isolating* the essential contribution that our epistemic evaluations make to our practices of collective epistemic self-regulation. As such, if we take these contributions to be the primary function of such evaluations, the planning model can illuminate which aspects of our current epistemic concepts are essential to the regulatory role that is their *raison d'être*. This gives us a powerful way of distinguishing those aspects of our contemporary epistemic practices that are really essential to the *point* of those practices from those that are not. And this can help us to recognize which aspects of those practices represent purely contingent, and perhaps unhappy, additions to or modifications of this essential core. Thus, the planning approach offers us a potentially fruitful way of engaging in the project of "conceptual ethics" within epistemology.⁹ Or, in other words, it can be a fruitful way of developing a sort of internal critique of certain aspects of contemporary "epistemic ideology".¹⁰

2. Epistemic Planning and Epistemic Internalism

My main concern here will be with a particular issue in epistemological conceptual ethics: Namely, the question of whether the function of some of our core epistemic concepts supports a broadly internalist account of them. Or whether, on the contrary, the temptation to view such concepts in internalist terms is the product of an unhappy "conceptual drift" away from their core function.

One advantage of viewing these questions through the lens of a planning model is that it helps us to frame them in a manner that does not get bogged down in questions about the meaning of English terms like "rationality" or "evidence". In particular, we can approach them in the following fashion. First, we can examine the various sorts of epistemic planning we should expect our creatures to engage in, and see if we can identify some varieties thereof that might be congenial to the core intuitions behind epistemic internalism. Then we can see whether there are principled reasons to expect that these forms of planning should obey the sorts of constraints that internalists take to govern at least some of our epistemic concepts. If we can locate a form of planning that has these characteristics, this will give us good reason to expect human beings to engage in a corresponding form of epistemic evaluation — one which will also be governed by the same sorts of internalist constraints.

Of course, the term "internalism" is used in epistemology in a variety of ways. So before proceeding it will be useful to say something about how I will be understanding it here. At the same time, I want to remain as open-minded as possible about the sense in which a form of internalism is supported by the investigations below. So I'm not going to begin by giving a precise characterization of the sort of internalism that might result from them. Instead, for the moment I'll simply describe what I take the basic idea behind epistemic internalism, in the sense of interest to me here.

⁸ Versions of this sort of thought have become increasingly prominent. For example, there is Edward Craig's seminal claim that the function of knowledge ascription is (roughly) to identify reliable informants about matters of interest. See Craig(1990). For recent discussion see Kappel(2010), Kelp(2010), Kusch(2011), Pritchard(2010), Schafer(2014), Fassio and McKenna(2015), and many of the papers in Henderson and Greco(2015), amongst others.)

And there is the idea that a central element of the conceptual role of the concept KNOWLEDGE is a link between knowing a proposition and being entitled to assert it or use it as a premise in further practical or theoretical reasoning. See Williamson(2000), Stanley and Hawthorne(2008).

Finally, there is the thought that the primary function of our attributions of rationality is to promote inter-subjective agreement about methods of belief formation and reasoning. See Dogramaci(2015).

⁹ Burgess and Plunkett(2013).

¹⁰ Haslanger(2012).

As described in more detail below, I take epistemic internalism to be motivated, in the first instance, by consideration of a range of familiar cases in which certain epistemic evaluations appear to be insensitive to differences that are subjectively indistinguishable from a first person point of view. For example, the internalist believes that there is some non-trivial epistemic status — for example, some important notion of justified or rational belief — that is shared by both ordinary subjects and subjects in a subjectively indistinguishable skeptical scenario. Thus, to be an internalist in my sense about some form of epistemic evaluation is to claim that this form of evaluation treats scenarios which are indistinguishable from a first person point of view in the same way. It is this general constraint that I want to investigate here — by investigating whether some aspect of our creature’s practice of epistemic planning will respect it.

Below, I’ll discuss what exactly this comes to — and the sense in which it might follow from the planning model — in much more detail. But it is important to stress that this basic form of internalism, on its own, does *not* have many of the implications associated with classical forms of internalism. For example, on its own, it by no means implies that the factors “internalist” forms of evaluation are sensitive to must be purely mental or “internal to the mind”. Nor does it imply that we are always in a position to know what these factors are when we have the epistemic status in question. In this sense, the sort of internalism I am interested in investigating here is considerably more modest than what some may think of when they hear this term. This will become even clearer in this essay’s final section on luminosity. So if there is a defense of internalism here, it is a relatively modest one. Indeed, one might regard the following — less as a defense of internalism *against* externalism — and more as an investigation of the sense in which internalism can remain plausible, even in the context of other broadly externalist epistemic commitments.

With this in mind, where should we look if we are hoping to isolate a form of epistemic planning that might be congenial to internalism in this broad sense? In previous work on this subject, I suggested that a particularly promising form of planning for this purpose relates to one of the most basic tasks in epistemic planning that our creature faces: Namely, the task of forming plans that govern what and how our creature forms beliefs in this or that first-personal situation:

Under condition C, believe that P (in this or that fashion).

Such plans are of particular interest to the internalist, because it seems that rational planning is constrained to some degree by one’s ability to distinguish between the *situations one is planning for*. As we will see, just what this involves is a complicated question. But at least in extreme cases, the intuitive idea is clear enough. For example, consider a plan that tells one to do A if C obtains, and not to do A if C does not obtain. And suppose that one is *certain* that one will *not* be able to distinguish between cases in which C obtains and cases in which it does not *when in the situation one is planning for*. More precisely, suppose that one is certain that whether one does A will *never be sensitive* to whether C obtains in the situation being planned for.

In this case, the plan in question does plainly seem to be irrational. For the point of forming such a plan is precisely to make whether one A’s sensitive (or, at least, more sensitive than it would otherwise be) to whether C obtains. Thus, if one knows *with certainty* that whether one A’s will be insensitive to whether C obtains, one knows that one will not be able to effectively follow this plan when the time comes to do so. In this way, it seems clear that there is *some* sort of connection between rational planning and one’s ability to distinguish between different scenarios.

But now notice that most skeptical scenarios are situations which are indistinguishable (from a first person point of view) from ordinary, non-skeptical scenarios. If so, then plans that *explicitly* recommend that one respond differently to (i) being in a skeptical scenario and (ii) being in the corresponding non-skeptical scenario would seem to be irrational in the just this manner just described. For one can know ahead of time that what one will do in such cases will not be sensitive to whether one is in a skeptical scenario or in the corresponding normal perceptual environment.

3. The Distinguishability Constraint

In this way, it's not hard to see why an internalist might be tempted to think that plans of this form help to explain why some of our practices of epistemic self-regulation have a broadly internalist character. But these considerations, of course, leave us a long way away from any such explanation. And, as we will see, filling in the details of this story is no easy task.

One of the main challenges in doing is arriving at a precise formulation of the distinguishability constraint on rational planning just alluded to. In particular, such a constraint can be interpreted in at least two ways — and it is unclear whether either of these interpretations supports a genuine argument from this constraint to an internalist-friendly conclusion. In criticizing some of my work on this topic, Schoenfield puts this point nicely:¹¹

On the strong understanding of the constraint one must always be able to distinguish between the circumstances that the plan distinguishes between. But note that the strong understanding rules out the following sorts of ordinary plans: to go the store if there is no milk in the fridge, but stay home if there is milk in the fridge; to water the garden if it doesn't rain, but to not water the garden if it does rain. The strong version of the constraint rules out such plans because there are certainly some circumstances in which there is no milk in the fridge that I won't be able to distinguish from some circumstances in which there is milk in the fridge...

A more plausible understanding of Schafer's constraint would require only that we must *sometimes*, or *usually* be able to distinguish between the circumstances that our plans distinguishes between. But on this understanding there is no reason to rule out externalist plans. There are plenty of externalist conditions such that we can *sometimes* or *usually* determine whether they obtain (whether there is milk in the fridge, whether it is raining, and so on).¹²

Does this mean that any attempt to show that internalism follows from the way distinguishability constrains rational planning is mistaken? Before considering this question directly, I want to first discuss what follows even from the weaker version of Distinguishability that Schoenfield finds plausible:

Weak Distinguishability: One can only rationally plan to do A in circumstance C1 and not to do A in circumstance C2, if it is possible (at least in principle) for the person carrying out the plan to *sometimes* distinguish between being in C1 and C2 when in C1 or C2.

¹¹ Schoenfield(2015b).

¹² I was not completely unaware of these issues when I wrote my earlier paper. In particular, I attempted to defuse them by arguing that the stronger version of Distinguishability is in fact compatible with our ordinary planning practices. I did so, by noting that even the stronger version of Distinguishability does not rule plans with factive antecedents as irrational, provided that those plans do not *fully* specify what one will do *both* when the factive condition obtains *and* when it does not. But Schoenfield is certainly right that this, on its own, is insufficient to show that this form of Distinguishability is compatible with many ordinary forms of planning. So, while a version of this thought remains attractive to me, I agree that more must be said to make this plausible.

It is also worth noting that, in response to these worries, Schoenfield(2015a) develops an alternative account of why judgments about epistemic rationality should be governed by an internalist constraint. There is much in Schoenfield's account that I find attractive, and I don't necessarily think that it conflicts with the story I tell here. But I also worry to some degree that Schoenfield's account may be subject to some of the objections that are familiar from debates about rule-consequentialist accounts of morality. Nonetheless, this depends on exactly how we understand the ambitions of Schoenfield's account, and (in any case) those with more consequentialist intuitions may not be bothered by this in the way I tend to be. I hope to take up Schoenfield's positive view, and its relationship to the view outlined here, in more detail elsewhere.

To reiterate, the notion of distinguishability that I have in mind here is very general one: Namely, one can distinguish between C1 and C2 in the sense relevant here insofar as whether one performs the relevant actions is sensitive whether one is in C1 or C2. This, it is important to stress, is perfectly compatible with C1 and C2 being facts about the external world. And it is compatible with the sensitivity in question here involving something weaker than knowledge or even belief about whether C1 or C2 is true. As such, this is a very weak version of this constraint — one which I think should be appealing even to externalists. Indeed, it seems to be the minimal constraint that captures the intuitions discussed above. For when read in this way, Weak Distinguishability simply demands that our plans at least have *some* chance of *sometimes* guiding our choice of actions in accordance with the distinctions those plans make.

Nonetheless, *even* Weak Distinguishability has important implications for the rationality of epistemic planning with respect to the first personal planning tasks we are considering. For example, consider a classical skeptical scenario like the following:

The Evil Demon: Rene’s subjective experience of the world is indistinguishable from our own. But, unlike us, he lives in a world that is controlled by an evil demon who has arranged things so that the beliefs he forms on the basis of his perceptual experiences are systemically unreliable.

What should we plan to believe when in Rene’s situation? One answer would be that we should plan to avoid the errors that trusting our senses would lead us into when in that situation. So, for example, we might consider forming the following epistemic plan:

I Ain’t ‘Fraid of No Demons: When I am in the Evil Demon scenario and my senses make it seem to me that P, do not believe that P; but when I am in a normal perceptual environment and see that P, believe that P.

Something about this plan will, I think, seem fishy to most readers. And Weak Distinguishability explains why. After all, I Ain’t ‘Fraid of No Demons tells us to respond differently to two situations — namely, Evil Demon and the corresponding normal perceptual situation. And, by hypothesis, we know that we are incapable of distinguishing the Evil Demon situation from the normal perceptual situation in the sense just defined. This, it is important to stress, is true even if (under normal perceptual conditions) we can *know* that we are in a normal perceptual situation. For even if this is true, it does not mean that our actions will be *sensitive* to which of these situations we are in. Rather, even if we can know that we are not in a skeptical scenario when everything goes well, this alone is not sufficient to give us the sort of *differential sensitivity* to these two situations that Weak Distinguishability requires.

Thus, even Weak Distinguishability is sufficient to rule out plans like I Ain’t ‘Fraid of No Demons as irrational. But this point does not immediately generalize to all instances of this sort of first-personal planning. For example, consider the following:

Trust My Senses: When I see that P, believe that P; but when my senses are deceiving me about P, do not believe P.

This plan is *not* ruled out by Weak Distinguishability. After all, it *is* often possible for me to distinguish (in our sense) between cases in which I see that P and cases in which my senses are deceiving me about P. So Schoenfield is absolutely correct that Weak Distinguishability does not provide a basis for a “short argument” against the rationality of all “externalist plans” of this sort. Thus, it leaves us well short of a defense of a general internalist constraint on this sort of epistemic planning.

Nonetheless, it is important to stress that even Weak Distinguishability does give us something important. Namely, it helps to explain why certain responses to skeptical scenarios strike us as intuitively irrational in the way fuels internalist epistemological views. In short, Weak Distinguishability may explain many canonical “internalist” intuitions. *But* it is not at all clear that this result generalizes to support a general internalist constraint of the relevant forms of epistemic planning.

4. Externalist and Internalist Responses

So where do we go from here? There are, I think, two reasonably plausible paths forward at this point: one broadly externalist, and the other more internalist and more in line with my previous view.

To see what these involve, we need to consider how we might derive the irrationality of plans like Trust My Senses *from* the irrationality of plans like I Ain’t ‘Fraid of No Demons. To close this gap, it is natural to appeal to a closure condition on rational planning. After all, there does seem to be a sense in which Trust My Senses *commits* one to acting in the manner that I Ain’t ‘Fraid of No Demons describes. So there is something tempting in the thought that Trust My Senses can only be rational if I Ain’t ‘Fraid of No Demons is also rational.

To capture this idea, we might consider the following constraint on rational planning:

Closure: If it is rational for me to plan to A under condition C, then, for any further condition D, it is also rational for me to plan to A under condition C&D.¹³

If Closure is true, and Trust My Senses is a rational plan, then I Ain’t ‘Fraid of No Demons must be rational as well. For I Ain’t ‘Fraid of No Demons simply makes the factive condition in Trust My Senses more determinate. Thus, if Closure is correct, we can derive the irrationality of plans like Trust My Senses from the irrationality of plans like I Ain’t ‘Fraid of No Demons.

Plainly this would generalize to any first personal plan that recommends conflicting responses to situations in which some factive condition applies and situations in which it does not. So this argument could be used by someone with internalist sympathies to derive the following conclusion from Weak Distinguishability:

Strong Distinguishability: One can only rationally plan to do A in circumstance C1 and not to do A in circumstance C2, if it is possible (at least in principle) for the person carrying out the plan to *always* distinguish between being in C1 and being in C2.

This would be sufficient to show that plans like Trust My Senses are irrational. And thus it would yield the sort of internalist constraints on this sort of epistemic planning we are looking for. But, of course, none of this is likely to impress someone with even very mild externalist sympathies. For why should we accept a condition like Closure, which leads so quickly to an internalist conclusion of this sort? After all, isn’t Strong Distinguishability simply implausible? After all, the rationale given above for Weak Distinguishability does not seem to support anything as strong as Strong Distinguishability. In order for a plan to play its proper function, it does seem that this plan should at least *increase* the sensitivity of our actions to the conditions mentioned in the plan. But why on earth would this increase in sensitivity need to be perfect or complete in every case?¹⁴

¹³ Note that this is probably best understood (for our purposes here at least) as a wide-scope requirement in Broome(2013)’s sense. But this is compatible with a view that sees such requirements as explained by more fundamental facts about practical reasons. See Lord(forthcoming).

¹⁴ Isn’t a principle like Closure also subject to many counter-examples in non-epistemic cases? For example, there seem to be situations in which it would be rational for me to plan to A given that I have promised you to A, even though it

Given this, the externalist should not feel very threatened by any argument that appeals to Closure. Indeed, the externalist might actually find our discussion so far quite congenial. For suppose we reject Closure, as I've just suggested we should. In that case, the previous section's discussion of Weak Distinguishability would provide the externalist with an elegant explanation why we respond to skeptical scenarios with internalist intuitions, without thereby forcing us into a generally internalist view of these issues. These intuitions, they might say, *do* represent an appreciation of the irrationality of plans that explicitly distinguish between skeptical and non-skeptical scenarios in a manner that runs afoul of Weak Distinguishability. But, the externalist might continue, the internalist is wrong to generalize this result to yield something like Strong Distinguishability. For this generalization rests on a closure principle that, whatever its initial appeal, should ultimately be rejected.¹⁵

In short, on this view of things, our tendency to slide into the use of genuinely internalist epistemic concepts represents a sort of unhappy "conceptual drift" — one that is not actually supported by the function these concepts play in regulating our epistemic lives. In this way, the externalist may feel that the focus on epistemic planning gives them the tools to construct a plausible externalist-friendly explanation of many of the intuitions normally taken to support internalism.

5. A Better Internalist Response

But is this the best response to what we've said so far? To see why one might be skeptical, consider the following question: When we form plans like Trust My Senses, are we *really* planning to do what this plan describes, *no matter what*, and in *any* situation that it covers. Or, are we really planning to do what this plan describes, all other things being equal or with some similar restriction. That is, are such plans normally truly unconditional or are they modified by an implicit *ceteris paribus* clause or something similar?

I think there is a good deal of plausibility to the idea that such plans *are* normally qualified in some way. For example, consider the following dialogue:

Diane: I'm planning to water the garden if it doesn't rain.

Todd: Great, but does that mean that you're planning to water the garden, even if it doesn't rain but you believe that it has?

Diane: No, of course not, Todd. I don't plan to do *that*. I'm only planning to water the garden, if it rains, assuming that everything else is more or less normal. I'm not concerned with weird situations like that.

In short, it seems very natural for Diane to respond to Todd's question by insisting that in forming her original plan, she wasn't taking a stand on what to do in abnormal situations like the one he describes.¹⁶

would no longer be rational for me to plan to A given that I promised you to A and that you have released me from this promise. Perhaps, although what one says about such cases will be affected by the arguments about *ceteris paribus* planning below. In any case, if there are such clear counter-examples that only adds to the case I am making here. (Thanks to an anonymous referee for pushing me to mention such cases.)

¹⁵ Of course, to take this line, the externalist would have to defend the idea that this closure principles remains tempting despite the issues with it noted above. Since this line is not my focus, I won't discuss the plausibility of it further here.

¹⁶ A referee suggests that it would not be *so* odd for Diane to respond to Todd's question as follows: "Yes, I *do* plan to water the garden *even* if it doesn't rain but I believe it has. Of course, in that situation, I *won't* water the garden. But that will be a situation in which I will fail to execute my plan, rather than one to which my plan doesn't apply." This, the referee suggests, marks an important difference between two sorts of circumstances in which things are "abnormal" in some sense: (i) circumstances in which the planner has false beliefs about whether the conditions of the plan obtains, and (ii) circumstances involving other sorts of "abnormalities". In short, the referee's suggestion is that while it is perfectly true that our plans are normally qualified to exclude abnormalities of the second sort, they are not normally qualified to exclude abnormalities of the first.

Given the naturalness of replies like this one, it seems worth considering how the possibility of such “*ceteris paribus* plans” might affect the arguments we have been considering. In particular, suppose we modify Closure to read:

CP Closure: If it is rational for me to plan to A under condition C, without any implicit *ceteris paribus* clauses or other similar restrictions, then, for any further condition D, it is also rational for me to plan to A under condition C&D.

Now, it is important to stress here that the work that a *ceteris paribus* clause does in this principle can, on some views of the indicative conditional, be done simply through use of such a conditional. For example, if we follow Stalnaker in thinking of the truth conditions of indicative conditionals as determined by what is true in the (conversationally relevant) closest possible worlds in which the condition obtains, this on its own may do the work required. Unfortunately, the literature on indicative conditionals is far too vexed for us to enter into a detailed discussion of it here. But this possibility indicates an important point: namely, that the work done by an implicit *ceteris paribus* clause might be also done via use of certain logical constants or conditionals or syntactically quite varied elements. In general, when I speak of “an implicit *ceteris paribus* clause or other similar restriction” here, I mean to be open minded about all these possibilities.

With this in mind, while I think it is very reasonable to reject Closure, it seems to me that CP Closure is much harder to deny. For suppose that CP Closure is false. Then an exchange like the following should be perfectly intelligible:

Diane: I’m planning to water the garden if it doesn’t rain.

Todd: Great, but does that mean that you’re planning to water the garden, even if it doesn’t rain and you believe that it has?

Diane: No, of course not, Todd. I don’t plan to do *that*. That’s not the sort of case I’m trying to plan for. I obviously won’t water the garden if *that* happens.

Todd: Oh, so what you’re *really* planning to do is to water the garden if it doesn’t rain and everything else is more or less normal?

Diane: No, I’m planning to water the garden if it doesn’t rain, *no matter what else happens and without any sort of conditions or restrictions*.

I, at least, find Diane’s response in this case puzzling. Instead, the natural response to Todd’s (admittedly curious) line of questioning seems to me to be the one described earlier. So it seems to me very natural to interpret plans like Diane’s as normally having an implicit *ceteris paribus* clause or other similar restriction (perhaps simply involving the semantics of the indicative conditional) — in which case, the rationality of such plans would *not* conflict with CP Closure, but rather lend support to it.

This is relevant here, because we can derive the following principle from CP Closure and Weak Distinguishability in much the manner described above:

Here, my intuitions are somewhat different. It is true, of course, that in the first class of cases the “abnormality” at issue is something that one is in some sense responsible for. So in (at least many) such cases involving a false belief, it would be reasonable to regard oneself as having “failed” in some sense in one’s choice of action. But the question is whether this is a failure to act in accordance with a plan to water the garden *even if* one has the relevant false beliefs, or whether it better understood simply as a failure to form correct beliefs about the weather. Once again, it is the latter interpretation that seems more natural to me. For, as noted above, it strikes me as odd to think of ordinary plans for action to extend to include within their scope cases that involve this sort of ignorance about whether the conditions of their plans are fulfilled. Rather, just as in the many other cases of “abnormality”, it seems to me much more natural to view ordinary planners as not taking any stand, one way or the other, about what to do in weird cases of this sort. In *all* these cases, it seems to me that such situations are not really among those ordinary planners are planning for.

CP Distinguishability: One can rationally plan to do A in circumstance C1 and not to do A in circumstance C2, without any implicit *ceteris paribus* clauses or other similar restrictions, only if it is possible (in principle) for the person carrying out the plan to *always* distinguish between being in C1 and C2 when in C1 or C2.

It seems to me that it is *this* principle that that someone with internalist sympathies might reasonably defend in these cases. To be clear, taking this line would not mean that it is irrational to make plans with factive antecedents. Nor would it necessarily mean that it is irrational to make non-*ceteris-paribus* plans. But it would mean that it is generally irrational to make plans with factive antecedents without including a tacit *ceteris paribus* clause or other similar restriction.

It is important to stress that CP Distinguishability is not *directly* supported by the rationale for Weak Distinguishability discussed above. Rather, it gains support from this rationale for Weak Distinguishability, when paired with the argument for CP Closure. Thus, there is a sense in which CP Distinguishability should strike us as a surprising result. For while the idea that the proper function of plans is make our actions sensitive to the conditions these plans mention does directly support Weak Distinguishability, it is far from obvious that it requires anything as strong as CP Distinguishability. Thus, if CP Distinguishability is plausible, it must be because it is plausible that rational planning is governed by CP Closure. I do find this reasonably plausible, but I know that not everyone will share this reaction.

In any case, the important point now is that CP Distinguishability would be sufficient to generate a genuinely internalist constraint on the relevant forms of epistemic planning. For if CP Distinguishability is correct, then plans like Trust My Senses will only be rational insofar as they contain an implicit *ceteris paribus* clause or other similar restriction:

CP Trust My Senses: When I see that P and all other things are equal, believe that P; but when my senses are deceiving me about P and all other things are equal, do not believe P.¹⁷

Or compare:

CP Knowledge: When I can thereby know that P, and all other things are equal, believe that P; but when I cannot thereby know that P, and all other things are equal, do not believe P.

These plans may seem very externalist. But in fact there is nothing in them that conflicts with internalist, as long as the *ceteris paribus* clauses in them exclude skeptical scenarios and other similar cases — which they must if the arguments above are correct. Indeed, insofar as the main point of these clauses here is to exclude such forms of “abnormality”, they are actually best understood as a way in which rational planning is being responsive to the basic intuitions behind internalism in the sense at issue here.

In other words, given CP Distinguishability, we can derive the following constraint on rational epistemic planning:

CP Planning Internalism: One’s epistemic plans should obey the following two constraints:

(1) These plans should only distinguish between two scenarios (i) if one can *always* distinguish being in one of these scenarios from being in the other or (ii) these plans only concern what one will do in those scenarios *ceteris paribus* or *assuming things are otherwise normal*.

¹⁷ Again, on some views of the indicative conditional, the work of the “all other things being equal” clause can simply be done by a indicative conditional. On such a view, we would not need any sort of restricting clause, simply making use of this conditional would be sufficient. But nonetheless, even on this reading, CP Trust My Senses would still involve an implicit restriction of the sort at issue here.

(2) These plans should not distinguish (even *ceteris paribus*) between two scenarios if one can *never* distinguish being in one of these scenarios from being in the other.

This constraint is weaker than what most internalists endorse. But it still seems to me to be true to the intuitions that motivate many contemporary forms of internalism. In particular, given this constraint, when we look at the facts that our epistemic plans are sensitive to, we will find that rational plans never draw a sharp line between scenarios, unless those scenarios are such that we could always distinguish the one from the other. In all other cases, our plans will at most distinguish between scenarios, *under the proviso that things are otherwise normal*.

To see why this provides us with a recognizably internalist constraint, suppose for the moment, that we are dealing an epistemic plan that does not involve any *ceteris paribus* clauses. Given CP Planning Internalism, this plan will only distinguish between two scenarios if one can *always* distinguish being in one of these scenarios from being in the other. Thus, this plan will only recommend belief in some scenario if it also recommends belief in *every* other scenario which cannot be distinguished (in this sense) from the first. Thus, on such a view, our epistemic plans will obey the following constraint:

Strong Internalism: Our epistemic plans are sensitive only to conditions, the obtaining of which the believer can distinguish from the non-obtaining.

This is clearly a version of the internalist idea. So the internalist character of CP Planning Internalism is plain with respect to non-*ceteris-paribus* plans.

But now consider *ceteris paribus* plans in the context of this constraint. There are two ways of thinking about these sorts of plans. First, we might take them simply as imprecise versions of some ideal masterplan that would itself be couched in wholly non-*ceteris-paribus* terms.¹⁸ In this case, I think it is clear that the rationality of such plans, as imprecise restatements of this masterplans, should not undermine the basic internalist character of CP Planning Internalism. After all, this ideal masterplan would itself satisfy Strong Internalism. And the *ceteris paribus* plans in question would only count as rational insofar as their *ceteris paribus* clauses covered the various ways in which they deviate from this masterplan. So although these *ceteris paribus* plans would not make this fully explicit, there would be a clear sense in they would also *implicitly* satisfy Strong Internalism, in virtue of the manner in which their *ceteris paribus* clauses operate.

That having been said, as I'll discuss in the next section, it may not be best to think of *ceteris paribus* plans as imprecise versions of such a non-*ceteris-paribus* masterplan. For it may be that, given CP Planning Internalism, there is *no* possible rational masterplan of this sort. And yet, even in this case, I think that CP Planning Internalism represents a recognizably internalist constraint on epistemic planning. For once again, in this case, the point of the relevant *ceteris paribus* clauses would be to avoid forming plans that commit us to acting differently in situations like those that motivated internalism in the first place. For example, even in this case, a rational set of plans would never, *either directly or indirectly, commit* us different beliefs or actions in a scenario like Evil Demon and the corresponding normal case. Once again, it was our intuitions about such cases that formed the basis of internalism. So insofar as CP Planning Internalism respects the internalist's intuitions about these cases, it can still be regarded as a mild form of internalism.

6. Internalist Plans and Luminosity

To recap, then, there are two basic ways we might interpret the implications of Weak Distinguishability for epistemic planning:

¹⁸ Compare the notion of a hyperplan in Gibbard(2003).

The Externalist Response: We might regard Weak Distinguishability as providing us with externalist-friendly explanation of the intuitions that fuel internalism.

The Internalist Response: We might accept CP Closure and use this principle to derive CP Distinguishability from Weak Distinguishability. This would allow us to derive a modest, but nonetheless recognizably internalist constraint on epistemic planning.

I think that both of these responses have some plausibility. But for reasons just noted, I'm inclined to view the Internalist Response as more attractive than the Externalist.

Crucially, taking this second path does *not* commit one to the claim that there are luminous conditions in the sense made famous by Williamson. More precisely, following Williamson(2000), we can call a condition C luminous iff it is true that whenever one is in C, one is also in a position to know that one is in C. And we can call a position absent-luminous iff it is true that, whenever one is not in C one is in a position to know one is not in C. Finally, we can call a condition C transparent iff C is both luminous and absent-luminous. Thus, C will be transparent just in case one is always in a position to know whether one is in C.

Following Srinivasan(2015), we can call the position that there are no transparent conditions, Anti-Cartesianism:

Anti-Cartesianism: There are no transparent conditions.

As Srinivasan discusses in detail, Anti-Cartesianism conflicts with many possible motivations for epistemic internalism.¹⁹ But, as she also notes, there *are* forms of internalism that can be motivated in ways that do not conflict with this claim. One of these is Schoenfield's own argument for internalism.²⁰ But the same is true of the defense of internalism just laid out. Thus, while I mean to remain agnostic here about the success or failure of anti-luminosity arguments, the conclusion of such arguments is actually quite naturally combined with the form of internalism outlined above.

This is for two reasons. First, and less importantly, the conclusion of the anti-luminosity argument is that there is no non-trivial condition C such that we can always *know* whether C obtains. But this is only relevant to my claims if the sense of "being able to distinguish possibilities" that is relevant to rational planning involving *knowing* which possibility obtains. And it's not obvious that this is the correct reading of this requirement. For we may well be able to respond to some condition C in translating our plans into action without thereby satisfying all of the conditions that are required in order for us to *know* that C.

But second, and more importantly, even if we take the anti-luminosity argument to show that there is no non-trivial condition C such that we can always distinguish C obtaining from C failing to obtain, this in no way conflicts with the form of internalism defended above. After all, this form of internalism does not rule out forming plans that are conditional on conditions C that are non-luminous. It only rules out the formation of such plans insofar as they lack any tacit *ceteris paribus* conditions or similar restrictions. Thus, if Anti-

¹⁹ "Why think Anti-Cartesianism has any interesting implications for the normative sphere? Many normative theorists maintain that one or more fundamental normative notions—epistemic justification, the subjective ought, rationality—must be spelled out in terms of subjects' mental states. For example, epistemic internalists think that whether one is epistemically justified or epistemically rational supervenes only on one's mental states. ... An obvious question is: why think that these normative notions should be spelled out in this way? Why not think, as epistemic externalists do, that the justification of one's belief can supervene on non-mental facts, e.g. whether that belief is the product of a mechanism adequately hooked up to the world? ... An obvious answer is this: only mental states possess a crucial property, namely transparency, and only norms that feature transparent states can meet some basic desiderata of norms. ... Insofar as this diagnosis of the motivations for internalism and subjectivism is correct, then Anti-Cartesianism undermines the case for tying the normative to the mental." (Srinivasan(2015), 276)

²⁰ Indeed, Srinivasan refers to Schoenfield's account with approval in a footnote to her discussion.

Cartesianism is true, it would follow from the CP Distinguishability that *all* our plans should be qualified in this way.

This might seem like an overly demanding result, but in fact it seems me to be a very reasonable reaction to the truth of Anti-Cartesianism. For suppose that there is no non-trivial condition C such that we can always distinguish the obtaining of C from the non-obtaining of C. Then in forming an *exhaustively* complete plan to act in different ways conditional on whether C obtains, I would be committing myself to act in accordance with a sub-plan that fails to satisfy even Weak Distinguishability. Or, in other words, by accepting such a plan, I would be committed to acting in accordance with a more specific plan that distinguishes between conditions in a way that I *know* I will *never* be in a position to successfully comply with. In the face of this, it seems to me quite natural to respond to this discovery by tacitly qualifying all our plans so that none of these plans are, strictly speaking, plans to *A no matter what* — but, rather, merely plans to *A all other things being equal*. Once again, this may simply require the use of the indicative conditional in the formation of these plans. Or it may require something more.

In many ways, this seems to me a very reasonable response to Anti-Cartesianism. Indeed, a refusal to take this step might seem to involve a failure to really take the truth of Anti-Cartesianism to heart. For if we were to go on forming completely unqualified plans to act on some condition in some way *no matter what* — even in the face of Anti-Cartesianism — we would be planning to act as if we had a sort of access to the facts that we simply do not have. In this way, the refusal to form plans that are not qualified in this way may simply be part of the rational response to Anti-Cartesianism.

Moreover, all this is compatible with the idea that, “we are often in a better position to know external world conditions than we are to know our own mental state conditions.”²¹ For the present account in no way depends on singling out some subset of conditions to which we have privileged access. Instead, it simply spells out one aspect of the rational response to the discovery that a condition fails to satisfy transparency. Thus, if Anti-Cartesianism is true, its implications apply equally to both external and internal conditions, with no attempt to draw a principled distinction between them.

Thus, the acceptance of CP Internalism in no way conflicts with Anti-Cartesianism. Indeed, one might see CP Internalism as spelling out the sort of humility that follows from Anti-Cartesianism. But, at the same time, CP Internalism continues to respect the core internalist intuitions about cases like Evil Demon that we discussed above. In short, far from conflicting with Anti-Cartesianism, CP Internalism may well be the form of internalism that is most compatible with it. And this may show that, far from there being a conflict between Anti-Cartesianism and *all* forms of epistemic internalism, there may be a mild form of epistemic internalism which follows from the acceptance of Anti-Cartesianism.

²¹ Srinivasan(2015), 286, compare Schwitzgebel(2008).

References

- Blackburn, Simon. (1998) *Ruling Passions*. Oxford: Oxford University Press.
- Broome, John. (2013) *Rationality through Reasoning*. Oxford: Wiley-Blackwell.
- Burgess, Alexis and Plunkett, David. (2013) "Conceptual Ethics." *Philosophy Compass* 8 (12):1091-1101.
- Chrisman, Matthew. (2012). "Epistemic Expressivism." *Philosophy Compass* 7 (2):118-126.
— (2016). "Epistemic Normativity and Cognitive Agency." *Noûs* 50 (4).
- Craig, Edward. (1990) *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Oxford University Press.
- Dogramaci, Sinan. (2015). "Communist Conventions for Deductive Reasoning." *Noûs* 49 (4):776-799.
- Fassio, Davide and McKenna, Robin. (2015) "Revisionary Epistemology." *Inquiry* 58 (7-8):755-779.
- Fricker, Miranda. (2015) "Know First, Tell Later: The Truth about Craig on Knowledge" in Henderson and Greco (2015).
- Gibbard, Allan. (2003) *Thinking How to Live*. Cambridge: Harvard University Press.
- Greco, Daniel and Hedden, Brian (forthcoming). "Uniqueness and Metaepistemology." *Journal of Philosophy*.
- Haslanger, Sally. (2012). *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press.
- Henderson, David and Greco John (eds). (2015) *Epistemic Evaluation: Purposeful Epistemology*. Oxford: Oxford University Press Uk.
- Kappel, Klemens (2010) "On Saying That Someone Knows: Themes From Craig." in A Haddock, A Millar & D Pritchard (eds.), *Social Epistemology*. Oxford: Oxford University Press.
- Kusch, Martin (2011) "Knowledge and Certainties in the Epistemic State of Nature." *Episteme* 8 (1):6-23.
- Lord, Errol. (forthcoming). "What You're Rationally Required to Do and What You Ought to Do (Are the Same Thing!)." *Mind*.
- Pritchard, Duncan (2010) *The Nature and Value of Knowledge: Three Investigations*. Oxford: Oxford University Press.
- Schafer, Karl (2014a). "Doxastic Planning and Epistemic Internalism." *Synthese* 191 (12):2571-2591.
— (2014b). "Knowledge and Two Forms of Non-Accidental Truth." *Philosophy and Phenomenological Research* 89 (2):373-393.
- Schafer, Karl and Steel, Robert. (forthcoming) "A Puzzle about Factive Plans."
- Schoenfield, Miriam. (2015a). "Bridging Rationality and Accuracy." *Journal of Philosophy* 112 (12):633-657.
— (2015b). "Internalism Without Luminosity." *Philosophical Issues* 25 (1):252-272.
- Schwitzgebel, Eric. (2008). "The Unreliability of Naive Introspection." *Philosophical Review* 117 (2):245-273.

Stanley, Jason and Hawthorne, John. (2008). "Knowledge and Action." *Journal of Philosophy* 105 (10):571-590.

Steel, Robert. (2015) "Planning for Failure." University of Pittsburgh: Dissertation.

Srinivasan, Amia. (2015). "Normativity Without Cartesian Privilege." *Philosophical Issues* 25 (1):273-299.

Williamson, Timothy. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.