Time-Slice Rationality

Brian Hedden

Abstract

I advocate **Time-Slice Rationality**, the thesis that the relationship between two time-slices of the same person is not importantly different, for purposes of rational evaluation, from the relationship between time-slices of distinct persons. The locus of rationality, so to speak, is the time-slice rather than the temporally extended agent. This claim is motivated by consideration of puzzle cases for personal identity over time and by a very moderate form of internalism about rationality. **Time-Slice Rationality** conflicts with two proposed principles of rationality, Conditionalization and Reflection. Conditionalization is a diachronic norm saying how your current degrees of belief should fit with your old ones, while Reflection is a norm enjoining you to defer to the degrees of belief that you expect to have in the future. But they are independently problematic and should be replaced by improved, time-slice-centric principles. Conditionalization should be replaced by a synchronic norm saying what degrees of belief you ought to have given your current evidence and Reflection should be replaced by a norm which instructs you to defer to the degrees of belief of agents you take to be experts. These replacement principles do all the work that the old principles were supposed to do while avoiding their problems. In this way, **Time-Slice Rationality** puts the theory of rationality on firmer foundations and yields better norms than alternative, non-time-slice-centric approaches.

1 Introduction

Being rational seems in part a matter of having attitudes that display a certain sort of coherence and stability over time. If your attitudes fluctuated wildly, then you would be unable to successfully engage in the reasoning and planning needed for you to achieve your goals,[1] and you would be unlikely to gain (and maintain) true beliefs. Consider the following cases:

> **Fickle Frank**
>
> Frank is a physicist who changes his mind constantly and frivolously. At breakfast, he is pretty sure that the Everett multiple universe hypothesis is the right interpretation of quantum mechanics. By mid-morning, he abandons that belief in favor of the Copenhagen interpretation. At lunchtime, he switches camps once again, siding with the de Broglie-Bohm theory. But that doesn't last, and by afternoon tea he is firmly convinced that some sort of hidden variable approach must be right. It's not that he keeps gaining new evidence throughout the day which supports different hypotheses. Rather, he just changes his mind.

> **The Frankfurt Physics Conference**
>
> A major conference on quantum mechanics is being held in Frankfurt. In attendance are proponents of a wide range of interpretations of quantum mechanics. There is a team of researchers from MIT who believe that the Everett multiple universe hypothesis is the best explanation of the available data. Seated next to them is an eminent professor from Cambridge who advocates the Copenhagen interpretation. Further down the row is a philosopher of physics who recently authored a book arguing that the de Broglie-Bohm theory is correct. In all, the lecture hall is filled by advocates of at least a dozen competing quantum mechanical views.

It is natural to think that there is a deep contrast between the two cases - that Fickle Frank is manifestly irrational, whereas for all that has been said, the Frankfurt physicists may be paragons of rationality, carefully evaluating the evidence and debating their views with colleagues. On this basis, it is tempting to conclude that you are rationally required to coordinate your attitudes with yourself at other times in a way that you are not rationally required to coordinate your attitudes with the attitudes of other people. The in*tra*personal and the in*ter*personal are importantly different when it comes to rationality.

[1] Briggs (2009) argues along these lines for diachronic principles of rationality.

This way of thinking about rationality seems obvious, almost a truism. But I am convinced that it is wrongheaded, and I will present an alternative, time-slice-centric picture of rationality on which there is no such deep contrast between Fickle Frank and the Frankfurt physicists, between the intrapersonal and the interpersonal. The physicists are rationally permitted to have different beliefs not because they are different people, but only because they have different evidence (on a natural spelling-out of the case). If they - like Fickle Frank's different time-slices - had the same evidence, they too would be required to have the same beliefs.

On my view, the relationship between two time-slices of the same person is not importantly different, for purposes of rational evaluation, from the relationship between time-slices of distinct persons. All in*tra*personal rational requirements can be derived from principles which apply equally in the in*ter*personal case.[2]

That is an intuitive gloss on my view, but now I want to make it more precise. **Time-Slice Rationality**, as I will call it,[3] is committed to two theses. The first is that what attitudes you ought to have at a time does not directly depend on what attitudes you have at other times. All requirements of rationality are synchronic.[4] Note that this is compatible with the claim

[2]I will discuss only requirements of *epistemic* rationality in this paper. In [DELETED], I discuss requirements of rational action, and in future work I plan to address constraints on rational preferences.

As the precise statement of my view (below) shows, I am primarily concerned in this paper with the notion of what you ought to believe. But we can ask not just what beliefs you ought to have, but also whether the beliefs you do have are held on the right basis. In the jargon, we can ask not just about propositional justification but also about doxastic justification. And we can also ask not just about the rationality of some set of beliefs, but also about the rationality of some agent. It would go beyond the scope of this paper to fully these further questions. But I take propositional justification to be the fundamental epistemic notion, with doxastic justification and the rationality of agents being defined partly in terms of propositional justification. For instance, doxastic justification is a matter of believing what you ought to believe *for the reasons that make it the case that you ought to believe it*, and an agent might be said to be rational to the extent that she tends to have propositionally and doxastically justified beliefs. So my main claim in this paper is that the core epistemic notion - propositional justification - should be approached in a time-slice-centric fashion.

[3]Note that I do not want to be committed to the four-dimensionalist metaphysics associated with the term 'time-slice.' My view is also compatible with endurantism.

[4]There are actually two ways in which requirements of rationality could be diachronic and hence conflict with my view. A *narrow-scope* diachronic norm would say e.g., that if you are in state S_1 at t_1, then you ought to be in state S_2 at t_2, while a *wide-scope* diachronic norm would say that you ought to be such that if you are in state S_1 at t_1, you are also in state S_2 at t_2. Narrow-scope diachronic norms say that what attitudes you ought to have at a later time depends on what attitudes you have at earlier times, regardless of whether those earlier attitudes were rationally impermissible, permissible but not required, or required. Wide-scope diachronic norms, by contrast, say only that you ought to have a certain pattern of attitudes over time. On a wide-scope view, it might be that while it is permissible for you to be in state S_1 at t_1 and also permissible for you not to be in state S_2 at t_2, nevertheless there is a global requirement of rationality that says that you ought not both be in S_1 at t_1 and also fail to be in S_2 at t_2. **Time-Slice Rationality** rejects both narrow-scope and wide-scope

that attitudes you have at other times (in particular, in the past) affect what attitudes you now ought to have, provided that they only do so indirectly, by affecting your present mental state. The second thesis is that in determining what attitudes you ought to have at a time, your *beliefs* about what attitudes you have at other times play no special role. More exactly, your beliefs about what attitudes you have at other times are not treated fundamentally differently from your beliefs about what attitudes other people have. Here, then, is **Time-Slice Rationality**:

Time-Slice Rationality

- **Synchronicity**: What attitudes you ought to have at a time does not directly depend on what attitudes you have at other times.
- **Impartiality**: In determining what attitudes you ought to have, your beliefs about what attitudes you have at other times play the same role as your beliefs about what attitudes other people have.

Time-Slice Rationality follows from two compelling considerations. The first stems from puzzle cases about personal identity over time. As we shall see, there are a number of hypothetical cases in which one person undergoes an experience, such as entering a teletransporter or undergoing a complex sort of operation, which is such that it is difficult to say whether it is *her* who is around after the experience. My claim is that in these puzzle cases, once we know everything about the evidence possessed by all the people in the scenario, we know everything we need to know in order to determine what each ought to believe. In particular, we do not *also* need to settle facts about who is identical to whom. Determining what an agent ought to believe does not require first figuring out the correct theory of personal identity over time. This means that requirements of rationality should not make reference to the relation of personal identity over time; what you ought to believe does not depend on who *you* are. That is, the requirements of rationality should be *impersonal*.[5]

Appeal to puzzle cases for personal identity over time is by itself sufficient to motivate both Synchronicity and Impartiality. But my view can be supplemented by a second consideration.

diachronic principles, and my arguments against diachronic principles will apply equally to narrow-scope and wide-scope versions, and so this distinction, while important, will play no significant role in my discussion. See Kolodny (2007b) and Broome (2007) for discussion of the narrow-scope/wide-scope distinction.

[5]See Christensen (2000) for a different argument for the impersonality of requirements of epistemic rationality.

For Synchronicity is also motivated by a very mild form of internalism about rationality. I think that a key thought behind internalism is that being rational is a matter of believing and behaving sensibly, given your perspective on the world. But it seems to me that your perspective on the world is constituted by your present mental state. It this is right, then what attitudes you ought to have at a time supervenes on your present mental state. Internalists hold that facts about the external world or about the reliability of your perception do not affect what you ought to believe, except insofar as they affect your mental state.[6] And by the same token, they should hold that facts about how you were in the past or about the reliability of your memory do not affect what you ought to believe, except insofar as they affect your present mental state. This means that we should accept Synchronicity.

I emphasize that this is a very moderate internalism. It says only that what attitudes you ought to have supervenes on your mental states. Following Conee and Feldman (2001), we might call this form of internalism 'mentalism,' to contrast it with 'accessibilism,' a different form of internalism on which what you ought to believe supervenes on factors to which you have special access.[7] With Conee and Feldman, I think that mentalism is more plausible than accessibilism. Accessibilists must respond to Williamson's (2000) Anti-Luminosity Argument, which concludes that there are no (non-trivial) states to which you have special access in the sense of their being *luminous*, so that whenever they obtain, you are in a position to know that they obtain. Now, accessibilists might attempt to rebut Williamson's argument[8] or formulate a conception of special access that is put not in terms of knowledge but in terms of perfect reliability, or justified certainty, or something of the sort. But my argument for **Time-Slice Rationality** does not rely on an internalism as strong as accessibilism; mentalism is sufficient. Note, interestingly, that mentalism is such a moderate form of internalism that it even allows Williamson himself to count as an internalist. He argues that your evidence consists of all and only the propositions that you know, but he also argues that knowledge is a mental state.

[6]This 'except insofar as they affect your mental state' caveat is needed to allow for externalism about mental content and for the fact that past attitudes affect your behavior in ways that affect what evidence you acquire.

[7]Another very strong version of internalism would say that what you ought to believe supervenes on your intrinsic physical properties. This would be incompatible with externalism about mental content.

[8]See especially Berker (2008).

Thus, his view is compatible with the claim that what you ought to believe supervenes on your mental states.[9] It is thus an appealing feature of **Time-Slice Rationality** that the form of internalism that motivates it is so weak as to encompass even Williamson's view and that it is thus compatible with a wide range of epistemological theories.

I have just given a general argument for **Time-Slice Rationality**, based on the idea that determining what you ought to believe does not require settling facts about personal identity or seeing how you are at other times. But defending this theory requires looking at specifics. As a case study, the remainder of my discussion centers on two proposed principles of rationality. First is the diachronic principle of Conditionalization (Section 2), which says how you ought to update your beliefs over time. Second is the principle of Reflection (Section 3), which says that you ought to defer to the beliefs you think you will later have. I argue that both Conditionalization and Reflection are problematic and must be rejected.

However, Conditionalization and Reflection are supported by a powerful argument. The Diachronic Dutch Book Argument purports to show that violating Conditionalization or Reflection is irrational, since doing so leaves you vulnerable to predictable exploitation over time. I respond to this argument in Section 4.

It would be unsatisfying if nothing could be salvaged from Conditionalization and Reflection. To this end, in Sections 5 and 6 I propose replacement principles that avoid the problems facing Conditionalization and Reflection and fit nicely with **Time-Slice Rationality**. With these replacement principles in hand, I show that even though there are no principles of rationality that directly concern how your attitudes at one time should fit with the attitudes you have, or expect to have, at other times, nevertheless satisfying these replacement principles at each particular time will result in your having generally coherent attitudes over longer periods of time. That is, the requirements of rationality yield coherence over time as a byproduct of a purely synchronic, impersonal notion of rationality.[10]

[9] Certainly, many internalists would not want to count Williamson among their number. In my view, they should reject Williamson by arguing that knowledge is not a mental state (see Fricker (2009)) rather than by opting for a very strong form of internalism like accessibilism.

[10] While my discussion employs a probabilistic, Bayesian framework for thinking about rationality, the general lessons I draw extend to traditional, non-Bayesian epistemology as well.

2 Against Conditionalization

Diachronic principles concern how you should change your attitudes over time. They say how your attitudes now should fit with the attitudes you had in the past (or will have in the future). The most widely endorsed diachronic principle is Conditionalization, which is a principle about how to change your *credences* (also known as degrees of belief or subjective probabilities) over time. It states that when you learn some proposition E, your new credence in H ought to equal your old credence in H given E. Your new credences should equal your old conditional credences, conditional on the proposition you just learned. This amounts to assigning 0 credence to all the $\neg E$ worlds and renormalizing your credence over all the E worlds. More formally, where P_0 is your credence function before learning E and P_1 is your credence function after learning E (and nothing stronger), we have:

> **Conditionalization**
> It is a requirement of rationality that, for all H, $P_1(H) = P_0(H \mid E)$

Conditionalization is a standard part of the Bayesian picture of rationality. For instance, in a recent survey article on Bayesian epistemology (Talbott (2008)), we find:

> The formal apparatus itself has two main elements: the use of the laws of probability as coherence constraints on rational degrees of belief (or degrees of confidence) and the introduction of a rule of probabilistic inference, a rule or principle of *conditionalization*. (emphasis his)

But despite its initial plausibility and widespread acceptance, I argue that Conditionalization must ultimately be rejected.

2.1 Conditionalization and Personal Identity

Conditionalization is not an impersonal principle. It makes reference to the relation of personal identity over time. It says how your credences after learning some proposition should be related to *your* earlier credences. Because of this, Conditionalization runs into trouble in cases where

the facts about personal identity are unclear. It implausibly entails that to determine what you ought to believe, we have to first determine the correct theory of personal identity over time. By contrast, I think that in these puzzle cases, once we specify what each time-slice's evidence is, the facts about what each ought to believe may be crystal-clear even though the facts about identity are murky.

Consider one such puzzle case: fission.[11] In particular, consider a fission case involving double teletransportation. One person (call her 'Pre') enters the teletransporter in New York. Her body is scanned, and at the moment her body is vaporized, two different molecule-for-molecule duplicates of her are created, one in Los Angeles and the other in San Francisco. Call them 'Lefty' and 'Righty,' respectively. Lefty and Righty are qualitatively just like Pre in all physical and mental respects. Now, there is a debate about whether Lefty, or Righty, or both, or neither is the same person as Pre. But what I want to emphasize is that in order to determine what Lefty and Righty ought to believe, following the double teletransportation, we do not have to first settle this debate about personal identity over time. If Lefty appears and immediately gains some new evidence, we do not first have to figure out the correct theory of personal identity in order to determine what Lefty ought to believe. All that matters is what Lefty's present evidence is. But Conditionalization conflicts with this *datum*. It only says that Lefty's credences ought to be constrained by Pre's credences if Lefty is the same person as Pre; it is silent if Lefty and Pre are not the same person. If Lefty and Pre are the not the same person, it is as if Lefty just suddenly came into existence, so it is compatible with obeying Conditionalization that Lefty choose any rationally permissible prior probability function and update it on her present total evidence, unconstrained by facts about Pre's credences (except insofar as they affect Lefty's present evidence).

Of course, it may be that the teletransportation in fact leaves Lefty with all of Pre's credences. But I am asking the normative question of whether Lefty's credence *ought* to be constrained by Pre's credences, and this question isn't settled by the fact that the teletrans-

[11]The same point can be made with any of the other puzzle cases for personal identity over time - fusion, teletransportation, Parfit's (1984) Combined Spectrum, etc.

portation resulted in Lefty's sharing Pre's credences.

One might attempt to modify Conditionalization by replacing reference to personal identity with reference to some surrogate notion such as Parfit's R-relatedness (psychological continuity with the right sort of cause).[12] But this proposal faces both a technical problem and an explanatory one. Start with the technical one. Psychological continuity, and hence R-relatedness, comes in degrees. But it is difficult to see how we could modify Conditionalization to make it sensitive to degrees of R-relatedness. Intuitively, it seems like the degree to which your current credences are to be constrained by facts about some past time-slice's credences should be proportional to the degree to which you are R-related to this past time-slice. But it is difficult to see how this intuitive notion could be worked into a precise mathematical formula. I have no in-principle proof that this could not be done, but I am skeptical.[13]

But more importantly, there is an explanatory challenge for this proposal. The defender of modified Conditionalization - with reference to identity replaced by reference to R-relatedness - must say *why* facts about who is R-related to whom are epistemically relevant. Why should the mere fact that some past time-slice bears the relation of R-relatedness to you mean that your credences ought to be constrained by facts about its credences? What you ought to believe depends only on truth-related matters like evidence, and just as facts about identity generally do not constitute evidence about the matter at hand, neither do facts about R-relatedness.[14]

[12]This sort of move is suggested in Meacham (2010). Meacham suggests that in Conditionalization, we replace reference to personal identity with reference to 'epistemic continuity,' although Meacham does not commit himself to any particular view about what epistemic continuity amounts to.

[13]One might also set a threshold of level of R-relatedness such that above that threshold, the later time-slice's credences are to be fully constrainted by the past time-slice's conditional credences, and below that threshold, they needn't be constrained at all by those past conditional credences. But the location of this threshold would be arbitrary in a way that only exacerbates the explanatory challenge discussed below. There may also be another technical problem. R-relatedness, unlike personal identity, can be branching. In a case of fission, multiple later time-slices are R-related to the pre-fission time-slices without being R-related to each other. And in a case of fusion (the mirror image of fission, in which two people undergo operations so that they combine to form one person), later time-slices are R-related to multiple pre-fusion time-slices which are not R-related to each other. And it is difficult to see how one could modify Conditionalization so that instead of making reference to a non-branching relation like personal identity, it instead made reference to a potentially branching relation. Now, the primary problem for Conditionalization and R-relatedness involves fusion cases, since this would involve a later time-slice being constrained simultaneously by the credences of multiple past time-slices. But fusion cases are problematic. It is not clear how two persons could be 'combined' to create a single person - which beliefs and desires from each input persons would go into the output person? How could such a surgery go? So I would want not want to rest a case against Conditionalization on the possibility of fusion.

[14]See Christensen (2000) for a related version of this worry.

Of course, facts about personal identity, R-relatedness, and the like do *sometimes* constitute relevant evidence. Indeed, it seems that almost *any* contingent facts could constitute evidence and thereby affect what you ought to believe. If someone tells you H, and then another person tells you H, the evidential force of their testimonies is typically diminished if then you learn that the first testifier was actually the same person as the second. But **Time-Slice Rationality** is compatible with this *datum*. It allows that facts about personal identity over time affect what you ought to believe, provided that they do so only by being part of your present evidence. Conditionalization, by contrast, says that how you ought to *respond* to evidence depends on facts about personal identity over time, even when that evidence does not itself have anything to do with personal identity. Moreover, in the case of the two testifiers, I think that the evidential force of learning this fact about personal identity really stems from the fact that it provides evidence that the two testifiers' beliefs have a common cause. And in general, the evidential force of two instances of testimony that H is lessened if you then learn that they ultimately come from a common source. This is true whether the two instances of testimony come from a single person or instead from different people. For example, if Alice tells you H and then Bob tells you H, the evidential force of Bob's testimony is typically diminished if you then learn that Bob's belief in H was based on Alice's having told him so.[15]

We have seen, then, that whether Conditionalization applies in a given case, and hence how it says you ought to proportion your beliefs to your evidence, depends crucially on facts about personal identity over time. But as I argued, what you ought to believe does not depend on these identity facts (except insofar as they constitute relevant evidence). Because of this, Conditionalization should be abandoned.

2.2 Conditionalization and Internalism

Conditionalization is not a synchronic principle, and so it conflicts with internalism. Conditionalization entails that what you ought to believe depends on what you believed in the past, even if you no longer remember what you used to believe. Meacham (2010) uses the following case

[15]Thanks to an anonymous referee for raising a variant on this case.

from Arntzenius (2003) to illustrate the conflict between Conditionalization and internalism:

> **Two Roads to Shangri La**
>
> There are two paths to Shangri La, the Path by the Mountains, and the Path by the Sea. A fair coin will be tossed by the guardians to determine which path you will take: if heads you go by the Mountains, if tails you go by the Sea. If you go by the Mountains, nothing strange will happen: while traveling you will see the glorious Mountains, and even after you enter Shangri La, you will forever retain your memories of that Magnificent Journey. If you go by the Sea, you will revel in the Beauty of the Misty Ocean. But, just as you enter Shangri La, your memory of this Beauteous Journey will be erased and be replaced by a memory of the Journey by the Mountains.

Suppose you in fact travel by the Mountains. Intuitively, while *en route* you ought to be certain (or near certain) that you are going by the Mountains, but upon entering Shangri La, your credence that you went by the Mountains should drop to 0.5, since (i) you have no evidence that suggests that your apparent memory is real rather than illusory, and (ii) whether your apparent memory would be real or illusory was determined by the toss of a fair coin. Note the internalist intuition here: that what you ought to believe depends on what your evidence is, and your evidence supervenes on your present mental states. Your evidence includes your present apparent memory of the Mountains, but not your past visual experiences of the Mountains (which entail that you went by the Mountains).

So you ought to be 0.5 confident that you traveled by the Mountains. But Conditionalization says otherwise. According to Conditionalization, you ought to start off with credence 0.5 that you will travel by the Mountains (since this is to be determined by the toss of a fair coin). Then you enter the Mountains. Conditionalizing on this new evidence (the experience as of Mountains), you come to have credence 1 (or nearly 1) that you are traveling by the Mountains. But upon entering Shangri La, you do not gain any new evidence (at least on an ordinary conception of gaining evidence) that bears on whether you traveled by the Mountains, and hence Conditionalization does not kick in. So, according to Conditionalization, you ought to just retain your credence 1 that you traveled by the Mountains. Note that the problem has nothing special to do with credence 1, but rather with the fact that upon entering Shangri La, you do not learn anything new that is relevant to the question of which route you took.

Note that a Williamsonian could object to my claim that you would be in the same mental state upon entering Shangri La regardless of which route would took. For if knowledge and memory are mental states, it might be that if you go by the Mountains, then upon reaching the glorious city, you will *know* that you went by the Mountains and *remember* the journey, as opposed to merely seeming to remember it. But in my view, upon entering Shangri La, you lose your knowledge that you went by the Mountains (and hence no longer remember having taken that route, if remembering entails knowing, as Williamson (2000, Ch. 1) argues). Certainly, it seems that if upon entering Shangri La you retain your belief that you went by the Mountains, then there is a nearby possible world (one in which the coin landed the other way) in which you falsely believe that you went by the Mountains, and so your actual belief is not safe. I cannot hope to decisively establish this claim about knowledge here, but even if you think that you retain your knowledge that you went by the Mountains upon entering Shangri La, the same point against Conditionalization can be made using any other case in which you lose knowledge not as a result of gaining evidence against your belief.[16]

Importantly, the problem facing Conditionalization is not unique to sci-fi cases like Two Roads to Shangri La. Conditionalization also yields implausible results in cases of forgetting, since Conditionalization only allows you to change your credences when you learn something new, and forgetting involves no such new learning, at least on an ordinary understanding of learning.[17] Suppose you are now certain that you had cereal for breakfast. At some point in

[16]See Williamson (2000, Ch. 7) for such cases. See also Lasonen-Aarnio (2010) for an argument that on a safety-based theory of knowledge, there in fact are no such cases where your knowledge is defeated except by so-called rebutting defeaters, which are themselves adequately dealt with by Conditionalization. However, note also that if you drop your belief that you went by the Mountains, and hence no longer know it, then on a Williamsonian view, your evidence will not support credence 1 that you went by the Mountains. But in this case Conditionalization would still tell you to be certain that you went by the Mountains.

[17]See Williamson (2000) for further discussion. An objector might respond that Conditionalization is not *incompatible* with forgetting, but rather is simply *silent* about it. That is, Conditionalization says only that when you gain evidence E, your new credences ought to equal your old credences, conditional on E; it says nothing about what to do when you forget something or otherwise lose evidence. But if this is right, then Conditionalization is not the whole story when it comes to rational belief change. We want a theory that gives the right result about Shangri La, not one that is silent about it. What could such a theory look like? In Section 5, I propose a synchronic replacement principle which deals all at once with gaining and losing evidence. This new principle simply says what doxastic state you ought to be in, given your present total evidence; it does not care whether your present total evidence was arrived at through learning, forgetting, or some combination of the two. See Titelbaum (2013) for an attempt to deal with forgetting in a diachronic framework. I cannot attempt a full appraisal of Titelbaum's view here, but I think that my own proposal is simpler and better-motivated,

the future, you will no longer remember having had cereal today, but since you will not have learned anything new that bears on what you had for breakfast today, Conditionalization says that you ought to retain your certainty that you had cereal. But this is crazy! Surely once you no longer remember having eaten cereal, you ought to drop your confidence that you had cereal. As noted by Williamson (2000), Conditionalization builds in the assumption that your evidence grows monotonically, such that once a proposition is part of your evidence, it remains part of your evidence forever. This monotonicity assumption conflicts with internalism and is implausible on its face.[18] So much the worse for Conditionalization.

Conditionalization has to go. But before we can jettison it and replace it with something better (Section 5), we must rebut the most powerful argument in its favor - that violating it leaves you vulnerable to exploitation over time. This is the task of Section 4.

3 Against Reflection

Van Fraassen's (1984) Reflection principle enjoins you to defer to the beliefs you anticipate having in the future. It says that if you believe that you will later have some belief, then you ought to now have that belief. In probabilistic terms, where P_0 is your credence function at t_0 and $P_1(H) = n$ is the proposition that at t_1 you will have credence n in H, the principle states:

Reflection

It is a requirement of rationality that, for all H, $P_0(H \mid P_1(H) = n) = n$

even if both give the right results.

[18] I concede that it is possible to model forgetting if we use primitive conditional probabilities (which can be defined even for probability 0 propositions) and Jeffrey Conditionalization (Jeffrey (1983)). Jeffrey Conditionalization is like standard Conditionalization, except that it can be applied even in cases where you do not become certain of any proposition. Jeffrey Conditionalization applies in cases where, as a result of experience, you change your credences in the members of an input partition $\{E_i\}$ from $P_0(E_i)$ to $P_1(E_i)$, and says that your new credences should be $P_1(H) = \sum_i P_0(H \mid E_i) \times P_1(E_i)$. Jeffrey Conditionalization can be used to model forgetting for the simple reason that *any* change in credences can be shoehorned into the framework of Jeffrey Conditionalization, just by letting the input partition be the set $\{\{w_i\}\}$ of all singleton sets of worlds. But this sort of maneuver threatens to trivialize matters. So, in order to avoid trivializing Jeffrey Conditionalization, we need to employ it with an intuitive conception of what it is to learn something. My claim, then is that *on an intuitive understanding of learning*, Conditionalization (and Jeffrey Conditionalization) are incompatible with forgetting, since forgetting is not the result of learning anything new.

There is something right about Reflection. Suppose that there is an envelope on your desk with the results of your blood tests, and you are told that whatever the paper inside of the envelope says, upon opening it you will be optimistic about your health. It seems that you ought *now* be optimistic about your health, even though you haven't seen the evidence contained in the envelope. Moreover, Reflection is a synchronic principle, and therefore does not conflict with internalism. It is synchronic since it tells you to defer to the beliefs you now *believe* you will later have, rather than the beliefs you will in fact later have.

But Reflection faces devastating counterexamples and problematically makes reference to personal identity over time, hence failing to be impersonal. Start with the counterexamples. First, suppose you believe that you will go out drinking tonight, and you believe that while drunk you will overestimate your ability to drive safely. It is rational for you to believe that tonight you will believe that you can drive home safely, but this does not mean that you should now believe that you will be able to drive home safely. Second, while sitting at breakfast eating cereal, you believe that 10 years from now, you will be quite uncertain what you had for breakfast today. But you shouldn't now be uncertain about what you are having for breakfast, with the cereal bowl right in front of you!

So Reflection must at least be modified to say that you ought to defer to the beliefs you believe you will later have unless you believe (i) that your future self will be irrational (as in the drinking case) or (ii) that you will have lost evidence (as in the breakfast case):[19]

> **Modified Reflection**
> It is a requirement of rationality that, for all H, $P_0(H \mid P_1(H) = n) = n$, unless you believe that at t_1 you will be irrational or will have lost evidence.

Modified Reflection avoids the counterexamples facing Reflection. The problem with Modified Reflection is therefore not that it is *false*, but rather that it is insufficiently general. First, it is insufficiently general in virtue of being future-directed. There are many cases where you ought to defer to the beliefs you think you had in the past. If you believe that 10 years ago, you believed you were eating cereal, then you ought now believe that you were eating cereal 10 years

[19]This modification is proposed in Hall (1999), Weisberg (manuscript), and Briggs (2009), among others.

ago. Modified Reflection should therefore follow from a more general, and more fundamental, principle of deference which is time-symmetric.

Second, it is insufficiently general by being about the beliefs you believe *you* will later have. It violates Impartiality by treating your beliefs about what you will later believe differently from your beliefs about what other people believe. But just as sometimes you ought to defer to your anticipated future beliefs, you also often ought to defer to the beliefs that you think others have. If you think that the weatherman, who is better informed about meteorological evidence and more skilled than you at evaluating that evidence, believes it will rain, then you yourself ought to believe it will rain.

This arbitrariness on the part of Modified Reflection becomes especially clear in puzzle cases for personal identity over time. In a double teletransportation case like that considered above, what Pre ought to believe, prior to entering the machine, does not depend on whether she is identical to Lefty, or Righty, or both, or neither. Rather, it depends only on her present evidence. But Modified Reflection bizarrely only says that Pre ought to defer to Lefty's or Righty's beliefs if she is identical to Lefty or Righty. Suppose that Pre is deliberating about what to believe about the right interpretation of quantum mechanics, and she is told what Lefty and Righty will each believe after they appear in Los Angeles and San Francisco, respectively. In order to determine what to believe right now about quantum mechanics, Pre does not have to start thinking about the metaphysics of personal identity! She does not have to come to a view about whether she will survive as Lefty, as Righty, as both, or as neither. Facts about personal identity over time are quite irrelevant to the question of what her quantum mechanical views ought to be. But whether Modified Reflection kicks in and tells her to defer to Lefty (or Righty) depends on just these irrelevant metaphysical facts. If Pre is identical to Lefty, then it tells her to defer to Lefty, but if she is not identical to Lefty, it is silent (though of course some other principle might still tell her to defer to Lefty in that case).

Modified Reflection, being an essentially intrapersonal principle, should be replaced by a more general deference principle which applies equally in the interpersonal case. I undertake this task in Section 6. But first, as with Conditionalization, there is a powerful argument for

Reflection which must be rebutted before we can turn to replacing it with something better.

4 Rebutting Diachronic Dutch Book Arguments

Lewis (1999) and van Fraassen (1984) give Diachronic Dutch Book Arguments for Conditionalization and (unmodified) Reflection, respectively. A Dutch Book is a set of bets that together guarantee you a loss. Lewis and van Fraassen show that if you violate Conditionalization or Reflection, your credences will license you to accept bets at different times which together constitute a Dutch Book. They argue that violating these principles is irrational, since doing so will sometimes lead you to perform predictably disadvantageous sequences of actions (accepting these bets in sequence).

To see how violating Conditionalization leaves you vulnerable to a Dutch Book, suppose you violate Conditionalization in the following way. Your current credence in E is 0.5 and your current conditional credence in H given E is 0.75, but the credence in H you will have if you learn E is 0.65.

Now, at t_1, prior to your learning whether E, a bookie offers to pay you 1 cent if you take Bets 1 and 2:

Bet 1: pays \$25 if $H \wedge E$, \$-75 if $\neg H \wedge E$, and \$0 if $\neg E$

Bet 2: pays \$5 if E and \$-5 if $\neg E$

Right now, your credences in E and in H given E commit you to regarding Bets 1 and 2 themselves as perfectly fair (each has an expected value of \$0). Therefore, accepting the deal (taking 1 cent and Bets 1 and 2) has positive expected value (and hence higher expected value than rejecting the deal), and so you are rationally required to accept it.

At t_2 you will learn whether E. If you then learn that E, the bookie will offer to pay you 1 cent if you take Bet 3:

Bet 3: pays \$-35 if H and \$65 if $\neg H$.

Your t_2 credence in H (0.65) will commit you to regarding Bet 3 as perfectly fair (having an expected value of \$0). Therefore, accepting the deal (taking 1 cent and Bet 3) will have positive expected value, and you will be rationally required to accept it.

If E is false, and so you aren't offered Bet 3, accepting Bets 1 and 2 guarantees you a loss of \$5, no matter whether H is true. And if E is true, accepting Bets 1, 2, and 3 guarantees you a loss of \$5. Either way, you will have accepted bets which guarantee you a loss of \$5, having gained only 1 or 2 cents (depending on whether the second deal is offered) in return. So no matter whether E is true, your credences, which violate Conditionalization, will require you to accept deals which together guarantee you a loss. Predictably, it would be better to decline all of the deals than to accept them and guarantee yourself a loss of \$5.

The case of Reflection is perfectly analogous. Just let E, above, be the proposition that your credence at t_2 in H will be 0.65. Then, the same bets will guarantee you a loss, no matter whether this proposition is true - i.e. no matter whether you in fact wind up with credence 0.65 in H at t_2.[20] I leave it to the reader to verify that this is the case.

I am unmoved by the Diachronic Dutch Book Argument, whether for Conditionalization or for Reflection.[21] This is because from the perspective of **Time-Slice Rationality**, it is question-begging. It is uncontroversial that collections of distinct agents can act in a way that predictably produces a mutually disadvantageous outcome without there being any irrationality. The defender of the Diachronic Dutch Book Argument must assume that this cannot happen with collections of time-slices of the same agent; if a collection of time-slices of the same agent predictably produces a disadvantageous outcome, there is *ipso facto* something irrational going on. Needless to say, this assumption will not be granted by the defender of **Time-Slice Rationality**, who thinks that the relationship between time-slices of the same agent is not importantly different, for purposes of rational evaluation, from the relationship between

[20]See Briggs (2009) for a clear presentation of the more general argument that *any* violations of Conditionalization or Reflection license you to accept a set of bets which guarantees you a loss.

[21]Besides the considerations I raise, Christensen (1991) argues that because Reflection faces obvious counterexamples, this is grounds for doubting whether the Diachronic Dutch Book Argument for Conditionalization is sound. Briggs (2009) attempts to defend only the argument for Conditionalization by highlighting a difference between the Dutch Book for Conditionalization and Reflection. See Mahtani (2012) for a rebuttal.

time-slices of distinct agents.

To elaborate, Diachronic Dutch Books have the structure of a Prisoner's Dilemma, with your t_1 and t_2 selves as the prisoners. In the Prisoner's Dilemma, Prisoners A and B each have the option of defecting or cooperating. Prisoner A prefers to defect, no matter what Prisoner B does, and Prisoner B prefers to defect, no matter what Prisoner A does. Moreover, A and B each prefer that the other does the opposite; A prefers that B cooperate and B prefers that A cooperate. But each prisoner prefers that both cooperate rather than that both defect. The outcome that results from their both defecting is worse by each of their lights than the outcome that would result from their both cooperating.

In the Diachronic Dutch Book case above where E is true, your t_1 self prefers accepting Bets 1 and 2 (plus the penny), no matter what your t_2 self does. And your t_2 self prefers accepting Bet 3 (plus the penny), no matter what your t_1 self did. Moreover, your t_1 and t_2 selves each prefer that the other reject the bets she is offered. But the outcome that results from your t_1 and t_2 selves each accepting the bets they are offered is worse by each of their lights than the outcome that would have resulted from their declining those bets. So, the Diachronic Dutch Book case is an intrapersonal Prisoner's Dilemma, with your t_1 self as Prisoner A, your t_2 self as Prisoner B, accepting the bets offered as defecting, and rejecting the bets offered as cooperating.

In the standard Prisoner's Dilemma, it is natural to think that neither prisoner is being irrational when she defects. Nor is there any sort of group-level irrationality. The Prisoner's Dilemma is just a case where two people predictably wind up with a mutually dispreferred outcome without anyone being irrational. We should say the same thing about Lewis' and van Fraassen's in*tra*personal Prisoner's Dilemmas. These are cases where time-slices of the same person act in ways that predictably produce a mutually disadvantageous result without there being any irrationality; the tragedy simply results from their having different beliefs about the world (in this case, different beliefs about which bets will be mostly likely to pay off).[22]

[22]Of course, the two prisoners do not care about each other, whereas your t_1 self presumably cares a great deal about your t_2 self, and *vice versa*. But I do not think that this undermines my claim that inter- and intra-personal Prisoner's Dilemmas ought to be treated the same. Insofar as you ought to be concerned for your

From a time-slice-centric perspective, then, the Diachronic Dutch Book Argument no more shows that Conditionalization and Reflection are requirements of rationality than the Prisoner's Dilemma shows that cooperating rather than defecting is a requirement of rationality.[23]

5 Replacing Conditionalization

Conditionalization and (Modified) Reflection must be rejected. They problematically make reference to the relation of personal identity over time, and Conditionalization faces the further problem of conflicting with a moderate internalism. Nevertheless, these principles are not completely on the wrong track, and so it would be unsatisfying to just reject these principles without finding replacements that capture what is right about them. In this section and the next, I propose replacement principles for Reflection and Conditionalization. These principles will be synchronic and impersonal, and hence mesh perfectly with **Time-Slice Rationality**. The fact that these synchronic, impersonal principles avoid the problems facing Reflection and

future selves, this should be reflected in your current preferences rather than through an add-on principle to the effect that vulnerability to an intrapersonal Prisoner's Dilemma is *ipso facto* irrational. Moreover, in the Diachronic Dutch Book Argument, you were predictably exploitable not because your t_1 self didn't care about your t_2 self (or *vice versa*), but rather because they had conflicting opinions about the optimal way to promote their shared interests. So the fact that you should be concerned about you later self whereas the two prisoners do not care about each other does not mean that these cases should be treated differently.

[23]There is a second, related objection to the Diachronic Dutch Book Argument. The mere fact that violating Conditionalization or Reflection puts you at risk of exploitation is only a *pragmatic* reason to satisfy these principles; it does not constitute an *epistemic* reason to do so. Defenders of the Diachronic Dutch Book Argument should therefore attempt to 'depragmatize' it, showing that it really reveals or dramatizes some inconsistency on the part of the agent who violates these principles. This is what Christensen (1996) and Skyrms (1987) attempt to do in the case of the Synchronic Dutch Book Argument for probabilism.

But I doubt that the Diachronic Dutch Book Argument can be depragmatized. I think that the best way to try to depragmatize it is to say that if you violate Conditionalization or Reflection, then your credences give rise to conflicting *ought* claims. It will be the case that you ought to accept the bets offered at t_1, and you ought to accept the bets offered at t_2, but you ought not accept both the bets offered at t_1 and the bets offered at t_2. And of course, it is logically impossible to satisfy all three of these *oughts*. Therefore, it is impossible to violate Conditionalization or Reflection without doing something that you rationally ought not do. And this means that you cannot be a perfectly rational agent if you violate Conditionalization or Reflection.

But this attempted depragmatization rests on the assumption that the rational *ought* can be applied not only to particular decisions, but also to sequences of actions performed over extended periods of time. But this assumption should not be granted by defenders of **Time-Slice Rationality**, for performing a sequence of actions over time is something that requires the cooperation of your later time-slices, and so which sequences of actions you are able to perform does not supervene on your present mental states. In my view, the rational *ought* applies in the first instance only to decisions, understood as mental volitional acts made at particular instants. See [DELETED] for further defense of this view of options. If this is right, then my best attempt to depragmatize the Diachronic Dutch Book Argument fails.

Conditionalization lends further support to the general approach advocated by **Time-Slice Rationality**. Let us start with how to replace Conditionalization.

Many epistemologists believe in *Permissivism*, the claim that given a body of total evidence, there are multiple doxastic states that it is rationally permissible to be in. But if we take as a *datum* that rational agents have beliefs that evolve steadily over time rather than fluctuating wildly (as in the case of Fickle Frank), then Permissivists must invoke some further principle to prohibit you from switching around between these multiple rationally permissible doxastic states. Diachronic principles like Conditionalization fit the bill. In effect, it says that once you opt for one of the permissible prior probability functions, you have to stick with it and update with respect to that function when you gain evidence.

How can defenders of **Time-Slice Rationality** avoid the need for diachronic principles while still maintaining that wildly fluctuating beliefs are (*ceteris paribus*) irrational? One way is to replace Conditionalization with a claim about rational dispositions and say that at each particular time, you ought not have a disposition or policy of abandoning your current credences in favor of others. On this view, it is permissible to fail to conditionalize (e.g., due to forgetting or a change of heart), but it would be irrational to plan or be disposed to do so. This claim about rational dispositions is a purely synchronic constraint, as it says only what dispositions you ought to have at a time. (Note, however, that since dispositions can fail to manifest themselves, this view may sometimes allow dramatic fluctuations in belief.)

But my preferred version of **Time-Slice Rationality** starts by abandoning Permissivism in favor of *Uniqueness*:

> **Uniqueness**
>
> Given a body of total evidence, there is a unique doxastic state that it is rational to be in.[24]

I will not attempt a full defense of Uniqueness here, but will instead settle for giving some pre-

[24]Kelly (forthcoming) notes that this version of Uniqueness is stronger than one which says that for each person and body of total evidence, there is one doxastic state that it is rational for her to be in. This version of Uniqueness allows that two different people could have the same evidence and yet be in different doxastic states. I rely on the stronger, impersonal version of Uniqueness.

liminary considerations in its favor.[25] First, Uniqueness captures the intuition that rationality is incompatible with thinking that your beliefs are arbitrary. For Permissivists, the only reason you ought to be in your actual doxastic state, instead of one of the other permissible ones, is some fact about your history or your psychology, for instance the fact that you happen to have gone for a certain set of credences (your *priors*, in the jargon) sometime in the past.[26] But this fact about your priors is a mere historical accident. Now, your priors encode judgments of plausibility, so a Permissivist might regard other beliefs (those resulting from different priors but the same evidence) as implausible. But the Permissivist must still recognize that she could have had different priors (and hence different judgments of plausibility) without being irrational, and it is this sense in which she must regard her own beliefs as arbitrary.[27] Second, Uniqueness is motivated by the thought that what you ought to believe is determined by what the evidence supports. But the evidence cannot point in two different directions at once. It cannot both support P and support $\neg P$ (and, to put things in terms of credences, the evidence cannot simultaneously support one proposition to two different degrees). Of course, it might be *indeterminate* what the evidence supports (or what the evidence is). In such cases, it is indeterminate what you ought to believe. But its being indeterminate what you ought to believe is quite different from there determinately being multiple permissible doxastic states, given the same evidence.[28]

In addition, Permissivism is undermotivated. Much of its appeal comes from intuitive case judgments. Consider Rosen (2001, 71):

> It should be obvious that reasonable people can disagree, even when confronted with a single body of evidence. When a jury or a court is divided in a difficult case, the mere fact of disagreement does not mean that someone is being unreasonable.

[25] See White (2007) for arguments in favor of Uniqueness, and Meacham (forthcoming) for a rebuttal.

[26] For non-Bayesians, the arbitrariness will not lie in facts about your priors (since they do not appeal to priors at all), but rather in some other facts about your history or psychology. Note also that a Permissivist could also simply deny that it is irrational to arbitrarily switch between the various permissible doxastic states.

[27] The Permissivist might hold that while *in fact* you could have had the same evidence but other beliefs without being irrational, you can never be in a position to rationally believe this. This view has the odd commitment that there is an *a priori* truth that you are never in a position to rationally believe. But I will not attempt a decisive rebuttal of this possibility here.

[28] Importantly, Uniqueness is also neutral about how fine-grained your doxastic state ought to be. I will shortly suggest a version of Uniqueness which makes use of imprecise (or 'mushy') credences.

> Paleontologists disagree about what killed the dinosaurs. And while it is possible that most of the parties to this dispute are irrational, this need not be the case.

Rosen's point is that while people often disagree about what the evidence supports, it would be rash to say that therefore one of them is being irrational. But Rosen is too quick here. First, the defender of Uniqueness holds only that if two people disagree despite having the same total evidence, then at least one diverges from ideal rationality. But saying that someone fails to meet the demanding standard of ideal rationality is not to say that that person is *crazy*. So while Rosen is right that none of the jurors need be irrational, in the sense of being significantly less rational than the rest of us, this doesn't mean that none is irrational, in the sense of falling short of the ideal. Second, even though the jurors might share the same evidence, in the sense of having seen the same presentations by the defense and the prosecution, this does not mean that they share the same total evidence. For a juror's total evidence includes not only the evidence presented in court, but also her background information, memories, and the like. To get a counterexample to Uniqueness, we need a case where people with the same *total* evidence disagree without any of them being irrational. Third, it is plausible that once the jurors learn about their disagreement, they should converge in their opinions.[29] If this is right, then once they share their *total* evidence, which includes evidence of which conclusions each initially arrived at, they really ought to have the same credences.

On an orthodox Bayesian picture on which rational doxastic states are represented by a precise probability function, Uniqueness amounts to the claim (i) that there is a unique rational prior probability function which, intuitively, represents the *a priori* plausibility of each proposition, and (ii) that your credences at a time should be the result of taking that uniquely rational prior probability function and conditionalizing it on your total evidence at that time. I call this principle *Synchronic Conditionalization*:

> **Synchronic Conditionalization**
> Let P be the uniquely rational prior probability function. If at time t you have total evidence E, your credence at t in each proposition H should equal $P(H \mid E)$.[30]

[29] This might follow from the widely-held Equal Weight View of disagreement. See Elga (2007) for a defense.
[30] We could also formulate a synchronic analogue of Jeffrey Conditionalization.

This is a purely synchronic principle, since it specifies what your credences should be at any particular time as a function of what your total evidence is at that same time. It makes no reference whatsoever to your credences at other times.[31] It is closely related, but not identical, to Williamson's (2000) Evidential Probability and Meacham's (2008) hp-Conditionalization.

Importantly, if you satisfy Synchronic Conditionalization at each time, then your credences will exhibit the stability over time that we expect from rational agents. If you satisfy Synchronic Conditionalization at all times *and* your evidence grows monotonically, then your credences will change over time in exactly the manner required by the diachronic principle of Conditionalization.[32] If your evidence does not grow monotonically, as when you forget something, then always satisfying Synchronic Conditionalization will not result in your having credences that conform to diachronic Conditionalization. This is as it should be, since it is an implausible feature of Conditionalization that it deems forgetting to be an irrational change in belief.

Now, if your evidence did fluctuate wildly, for instance if you kept getting enormous amounts of new evidence at each moment, then Synchronic Conditionalization would require your beliefs to fluctuate as well. Of course, this is also true of old, diachronic Conditionalization. But I also think that this is exactly the right result. Suppose that an evil demon gives you very different perceptual experiences from moment to moment while also erasing your memory of the fact that you have undergone these dramatic perceptual shifts. In such a case, it seems to me that you really should have very different beliefs at each successive moment. This would be bizarre and unfortunate, but not, I think, irrational. Of course, if you were aware that

[31] Synchronic Conditionalization is a view about propositional justification - about what credences you ought to have. But epistemologists are also interested in doxastic justification - whether your beliefs are properly based on the evidence which supports them. For instance, an agent who was struck on the head and happened to wake up with the recommendated credences would not be doxastically justified in having those credences. A time-slice-centric conception of doxastic justification would require understanding the basing relation (i.e. what it is for an agent's credences to be based on some reason) in a synchronic way, so that it depends not on the history of the agent's beliefs, but rather on facts about the agent's present dispositions (e.g., dispositions to have certain credences given certain bodies of evidence), which counterfactuals are true of the agent, and the like. I will not discuss basing here.

[32] To see this, suppose that at t_1 you have total evidence E_1 and at t_2 you gain evidence E_2, so that your total evidence is now $E_1 \wedge E_2$. According to Synchronic Conditionalization, your t_1 credences ought to be $P_1(-) = P(- \mid E_1)$ while your t_2 credences ought to be $P_2(-) = P(- \mid E_1 \wedge E_2)$. But P_2 is the probability function that results from taking P_1 and conditionalizing on E_2. So when your evidence grows monotonically from E_1 to $E_1 \wedge E_2$, Synchronic Conditionalization yields the same recommendations as diachronic Conditionalization.

your perceptual experiences were fluctuating wildly, then plausibly your total evidence, which includes this awareness, would require you to disregard these perceptual shifts as random noise.

I suspect that a major source of resistence to Uniqueness is based on the felt implausibility of the claim that there is a unique rational prior probability function. It is mysterious what facts could possibly single out that particular probability function as the uniquely rational one. For repeated attempts to give a finite number of formal principles which single out one probability function as a uniquely rational prior have all failed. This is one lesson of the failure of Carnap's (1950) project of inductive logic.[33] And it is also a lesson of the seeming impossibility of devising a Principle of Indifference which avoids falling into inconsistency.[34] But defenders of the existence of a uniquely rational prior are not committed to the claim that it is possible to single out that probability function using a finite number of formal principles. They should instead hold that there are substantive (i.e. non-formal) constraints on rational credences which single out a uniquely rational prior.

In having to appeal to substantive, non-formal constraints on rational credences, defenders of Uniqueness are in the same position as more moderate Permissivists. Extreme Permissivists like Subjective Bayesians hold that the only constraints on rational credences are formal ones like the axioms of the probability calculus and perhaps some small finite number of further principles such as the Principle Principal or Regularity.[35] But these formal constraints will not be sufficient to rule out intuitively irrational credence functions like those of the skeptic, the counterinductivist, and even seemingly randomly constructed credence functions that nonetheless obey the formal constraints. To rule these out, we need substantive constraints on rational credences in addition to formal ones. We need to require, for instance, that rational credences favor simpler or more explanatory hypotheses, project natural properties, and the like. The defender of Uniqueness might then hold that these substantive constraints narrow down the set of rational prior probability functions to just a singleton set. (As Christensen (2007, fn

[33]The grue paradox (Goodman (1955)) was particularly influential in showing this project to be unworkable.

[34]See especially van Fraassen (1989) and the Bertrand Paradox.

[35]The Principal Principle says that your credence in H, conditional on H's objective chance being n, ought to equal n. Regularity says that your prior probability function should not assign probability 0 to any contingent propositions; it should not rule out any possibilities in advance of inquiry.

8) notes, the defender of Uniqueness could attempt to make this more plausible by holding that while it is determinately the case that there is a unique rational prior, it is indeterminate exactly which function it is.)

I think that moderate Permissivists may be motivated by the thought that these 'epistemic values' - simplicity, naturalness, explanatoriness, etc. - sometimes conflict. For instance, sometimes the simplest hypotheses are not the most explanatory. And when these values conflict, there is no privileged way of weighing them against each other to come up with a unique rational credence function. Instead, there are a variety of different ways, some of which give great weight to simplicity and naturalness and less weight to explanatoriness, some of which treat explanatoriness as most important, and so on. Each different rationally permissible way of assigning weights to the various epistemic values results in a different prior probability function. So there is a set S of probability functions, each of which is permissible to have as your priors.

But in my view, this picture on which there is no privileged way of weighing competing epistemic values against each other does not really support Permissivism. Instead of holding that each member of S is permissible but that you have to pick one, we should instead hold that you ought to be 'mushy' over all the members of S.[36] In the absence of any evidence, you ought to be in a coarse-grained doxastic state represented by a set of probability functions (which, following van Fraassen (1990), is called your *representor*), in particular the set S resulting from all the different permissible ways of assigning weights to the various competing epistemic values. And if your present total evidence is E, your representor ought to be the set consisting of each member of S conditionalized on E. I think that this imprecise-credences-based version of Uniqueness is more plausible than precise Permissivism. For in plumping for some particular member of S as your prior, you would be going beyond your evidence. But I think that epistemic rationality is incompatible with going beyond your evidence. Being epistemically rational is about being passive in responding to evidence, not active in going beyond it. Instead of plumping for one of the members of S over another, you should be mushy

[36]Christensen (2007, fn 8) and Kelly (forthcoming) note that Uniqueness is more plausible if one thinks of rational doxastic states as being more coarse-grained (e.g., as involving imprecise credences).

over all of them. In this way, you would not be going beyond your evidence in any way. You would be going exactly as far as your evidence demands but no further.

Interestingly, the main argument against the rational permissibility of imprecise credences is a Diachronic Dutch Book Argument due to Elga (2010). But we saw in Section 4 that Diachronic Dutch Book Arguments beg the question against defenders of **Time-Slice Rationality**. For this reason, my view is particularly well-placed to include imprecise credences in its epistemological tool-kit.[37]

For ease of exposition, in the remainder of this paper I will set aside imprecise credences and assume a version of Uniqueness based on the assumption of a uniquely rational prior probability function. But this is merely for the sake of convenience. Whether rational doxastic states must be precise or imprecise, fine-grained or coarse-grained, the crucial point is that if Uniqueness is true, being in the uniquely rational doxastic state at each time will mean that your beliefs are relatively stable over time (provided of course that your evidence is relatively stable), even though there are no diachronic principles that directly require this stability.

This is an instance of a more general trade-off. The more we can say about which particular attitudes you ought to have, the less need there is for principles of coherence, and *vice versa*. I have been discussing diachronic principles of coherence, but the same applies for synchonic principles of coherence. For instance, Kolodny (2007a) argues that there is no need for a principle of coherence stating that you ought not have contradictory beliefs, since if you believe both H and $\neg H$, then at least one of your beliefs is not supported by the evidence and hence is irrational. But note that this assumes that the evidence must either support H, support $\neg H$,

[37]See Moss (2013) for an alternative time-slice-centric defense of imprecise credences. Another argument against the rational permissibility of imprecise credences White's (2009) argument that in some cases, having imprecise credences will require you to violate Reflection. White shows that sometimes, if you have imprecise credences, you will know that there is some later time at which you will have a particular representor which happens to differ from the one you currently have. See Joyce (2010) for a response to White. In my own view, White's argument mistakenly assumes that Reflection (or the Expert Deference principle I defend shortly) should applied at the level of the representor as a whole rather than at the level of each particular probability function in the representor. Thinking of having a representor as akin to having a committee of different people in your head, each with her own credence function, we should interpret the relevant deference principle as applying to each particular committee member, rather than to the committee as a whole. And in White's case, each particular committee member does obey the deference principle in question. Indeed they must, if, as I suggest in the next section, the most defensible such deference principle will in fact follow from the axioms of the probability calculus.

or support neither. So assuming that the evidence uniquely determines which beliefs you ought to have, then there is no need for a separate principle of coherence.

But in the case of preferences, say, it is less plausible that there is a particular set of preferences you ought to have. Rationality does not dictate that you ought to prefer chocolate ice-cream to vanilla, or *vice versa*.[38] So, if it is irrational to have intransitive prefences (e.g., to prefer chocolate to vanilla, vanilla to strawberry, and strawberry to chocolate), this cannot be because one of these preferences is by itself irrational. Rather, this *combination* of preferences is irrational; any set of ice-cream preferences is rational, so long as you don't mix-and-match and have them come out intransitive. This means we need a principle of coherence for preferences which directly states that rational preferences must be transitive.[39] The general lesson is that if a uniqueness thesis holds for a certain sort of attitude, then coherence principles are unnecessary, whereas if uniqueness fails, then coherence principles are needed to rule out intuitively irrational combinations of attitudes. My specific claim is that Uniqueness for credences is true, thus eliminating the need for a diachronic principle such as Conditionalization.

Uniqueness, in the form of Synchronic Conditionalization, not only does much of the work that diachronic Conditionalization was supposed to do, but also gets around the problems facing the latter. First, Synchronic Conditionalization is compatible with internalism and nicely explains why you ought to be 0.5 confident that you traveled by the Mountains in Two Roads to Shangri La. The thought is that which route you took was determined by the result of a coin toss. And your current evidence that you seem to remember traveling by the Mountains does not discriminate between your having traveled by the Mountains and your having traveled by the Sea. So your credence that you traveled by the Mountains ought to equal 0.5. Second, because Synchronic Conditionalization makes no reference to personal identity over time, it obviously faces no trouble about how to apply it in cases of fission, teletransportation, and the

[38]This may be controversial. Some utilitarians, for instance, might say that what preferences you ought to have is dictated by facts about the happiness of yourself and others.

[39]In a similar vein, Kolodny (2007a) attempts to do away with a principle of means-end coherence, which states that you ought to be such that, if you intend to achieve end E and believe that you will achieve E only if you intend to pursue means M, then you intend to pursue means M. But his strategy for doing so assumes that among a set of incompatible ends, there is a unique end that it is rational to intend.

like. Synchronic Conditionalization makes reference only to your current total evidence.[40]

One might worry, though, that replacing diachronic Conditionalization with Synchronic Conditionalization is a Phyrric victory if we cannot also have a time-slice centric conception of evidence. First, if content externalism is true, then what your evidence is might depend on your past attitudes and on facts about personal identity over time. Content externalism is the view that the contents of your propositional attitudes do not supervene on your intrinsic physical properties. If content externalism is true, then what your evidence is will depend on your history, unless we implausibly restrict your evidence to e.g., propositions about your retinal images, since those images are arguably content-less. For instance, if your evidence includes propositions about how things appear to you, and it appears to you that there is a glass filled with water, then your having this evidence depends on your having had past causal interactions with water (as opposed to with XYZ; see Putnam (1975)).

But **Time-Slice Rationality** is compatible with content externalism and its implications for thinking about evidence, since **Time-Slice Rationality** says that what you ought to believe at a time supervenes on your mental states at that time, not on your intrinsic physical properties. So my view allows that facts about e.g., whether you have had past causal contact with water can affect your present evidence, provided that they do so by affecting your present mental states. Note also that while the contents of your attitudes depend *inter alia* on facts about the causal history of your psychological states, they do not depend on facts

[40] Abandoning Conditionalization in favor of Synchronic Conditionalization has other benefits as well. Weisberg (2009) argues that Conditionalization is incompatible with the holist idea that all beliefs are subject to defeat by undermining evidence. Suppose that you see a jellybean and have a perceptual experience as of its being red, and in response you conditionalize on the proposition that it is red (or Jeffrey conditionalize with high credence that it is red). If you then get evidence that you are colorblind, you should reduce your confidence that it is red, but Weisberg shows that conditionalizing on the proposition that you are colorblind will leave your credence that the jellybean is red unchanged. The reason is that out the outset (before seeing the jellybean), you regarded colorblindness as evidentially irrelevant to the color of the jellybean, and conditionalizing on the proposition that it is red cannot change this (due to a property of Conditionalization known as *rigidity*). (Saying that you ought only conditionalize on how things seem to you will not help, since even beliefs about perceptual seemings are subject to defeat.) Synchronic Conditionalization does not face Weisberg's problem, however. The jellybean case is simply a case of shrinking evidence. After seeing the jellybean, your evidence includes the proposition that it is red. But upon gaining evidence that you are colorblind, that proposition simply drops out of your evidence. For this reason, if you always have the credences demanded by Synchronic Conditionalization, your credence that the jellybean is red will rise upon seeing the jellybean and drop upon hearing that you are colorblind.

about personal identity as such. Consider a case of double teletransportation, where Pre enters the teletransporter and gives rise to Lefty and Righty, two molecule-for-molecule duplicates of Pre. Whether Lefty has thoughts about water or about XYZ depends on whether Pre (or some predecessor of Pre) had causal interaction with water or instead with XYZ, regardless of whether Lefty is the same person as Pre. This is because the causal history of Lefty's concept runs through Pre, whether or not Lefty is identical to Pre or merely R-related to Pre. So facts about past time-slices causally related to Lefty's present time-slice affect the contents of Lefty's thoughts, and hence affect Lefty's evidence, independently of whether or not these past time-slices bear the relation of personal identity over time to Lefty.

Similarly, **Time-Slice Rationality** is compatible with Williamson's (2000) E=K, the thesis that your evidence consists of all and only the propositions that you know. Whether your belief in a proposition constitutes knowledge may depend *inter alia* on facts about the past, such as facts about how the belief was initially formed. Thus, whether you know P (as opposed to merely truly believe P) does not supervene on your present intrinsic physical state. But this is fine as far as **Time-Slice Rationality** goes. For Williamson also argues that knowledge is a mental state. If he is right, then E=K is compatible with **Time-Slice Rationality**'s claim that what you ought to believe at a time supervenes on your present mental states.[41]

Thus, while I have not defended any particular conception of evidence, I hope to have shown that **Time-Slice Rationality** is compatible with a wide range of views about evidence, which can then be combined with Synchronic Conditionalization to yield a fully time-slice-centric theory of what you ought to believe.[42]

Before turning to a replacement principle for Reflection, I want to close by considering an objection to my synchronic conception of what you ought to believe. I have proposed that what

[41]Similar comments apply to Burge's (1993) theory of the epistemic role of memory. Whether you remember something (as opposed to merely having memory impressions of it) can make a difference to what you ought to believe, provided that memory is a mental state, as Williamson holds.

Note also that even if facts about the causal history of a belief affects whether it can constitute knowledge, this does not mean that knowledge depends on facts about personal identity over time as such. In double teletransportation, whether Lefty knows (as opposed to merely believes) some true proposition may depend on how Pre initially formed that belief, regardless of whether Lefty is identical or merely R-related to Pre.

[42]See Williamson (2010) for an alternative Bayesian epistemological theory which likewise rejects standard Conditionalization and is fully synchronic. Williamson's theory differs from mine by not requiring Uniqueness.

you ought to believe depends only on your present total evidence, but what if the fact that that is your present total evidence was the result of some bias that you have? In particular, what if you have some racial bias that makes you want H to be true, and this causes you to selectively forget evidence that bears against H, resulting in your later having a body of total evidence which supports H. The objection, then, is that your later belief in H is irrational and that it is impossible to account for its irrationality without bringing in diachronic factors.[43]

Of course, if after having forgotten the evidence against H you still have the racial bias, then **Time-Slice Rationality** can hold that you are irrational in virtue of having an irrational bias, even if your belief in H is itself rational.[44] But if you no longer have the bias in question, then it seems to me that even though you were irrational at the time you had the bias, once you have lost it, neither you nor your belief in H need be irrational.

After all, if we grant that beliefs that result from forgetting needn't be irrational, it is unclear to me why it should matter what the cause of that forgetting is, unless you have evidence about what caused your forgetting.[45] I am not convinced that the causes of your changes in evidence make a difference to the rationality of the resultant beliefs, except insofar as you are aware of those causes. Compare a case where instead of your bias causing you to forget evidence against H, it instead somehow causes you not to encounter such evidence in the first place. It causes you not to attend conferences or go to bookstores where you might gain evidence against H. In my view, if you have no reason to suspect that your bias is doing this, then your belief in H is rational (although, again, if you still have the bias, then you are irrational in virtue of

[43]Thanks to an anonymous referee for pressing me on this case. Note that as I stated the two main commitments of **Time-Slice Rationality**, it is a view about what you ought to believe. That is, it is in the first instance a view about so-called propositional justification. So it is technically compatible with this view that while your belief in H is propositionally justified, it is not doxastically justified. However, I will not avail myself of this response, since I ultimately want to defend a time-slice-centric conception not only of propositional justification, but also of doxastic justification, on which the basing relation is understood synchronically. However, an adequate discussion of doxastic justification would go beyond the scope of this paper.

[44]If the cause of your forgetting is not an irrational bias but instead some rationally defensible desire, then I can perhaps hold that you are irrational if you still maintain the *disposition* to lose evidence against H. Of course, whether or not you have that disposition at time t will depend on how you are at other times, but nonetheless it is a fact about how you are at t. Analogously, whether you have some belief at t will partly depend on how you are at other times, but nonetheless whether you have that belief is a fact about you at t.

[45]See Christensen (1994) and Smithies (2006) for sympathetic discussion of the rationality of beliefs that were originally based on evidence you have now lost. They do not discuss the case of evidence loss caused by bias.

having an irrational bias). I think that my case is also bolstered by noting that there is, I think, no intuition that your belief in H is irrational if your bias caused selective forgetting by some deviant causal route, for instance by triggering a seizure which caused you to lose all and only your evidence against H.[46] Now, we might try to distinguish this case from the more ordinary one by giving some account of the difference between deviant and non-deviant causal routes, but I am skeptical of the possibility of doing so in any informative and non-circular way. Better to opt for a simple, elegant, and principled theory of rationality by simply denying that your belief in H is irrational in virtue of your having selectively forgotten evidence against H, whether the forgetting was deviantly or non-deviantly caused.

Of course, selective forgetting is epistemically suboptimal, but so are all other cases of forgetting. And it is important to distinguish epistemic suboptimality from epistemic irrationality. To take a simple case, failing to be omniscient is epistemically suboptimal, but it is by no means epistemically irrational. So **Time-Slice Rationality** can concede that selective forgetting puts you in an epistemically worse position while resisting the thought that selective forgetting makes the resultant beliefs irrational.[47]

Note that this does not mean that my view says that motivated reasoning is rational. In typical cases of motivated reasoning, your desire causes you to improperly evaluate the evidence, perhaps by selectively ignoring the evidence you have against the proposition that you want to be true. But my view nicely accounts for the irrationality of this sort of motivated reasoning. For even if you fail to make vivid some bit of evidence against your view, it is still part of your total evidence and hence affects what you ought to believe. In my view, improperly attending to evidence is importantly different, as far as rationality is concerned, from selectively losing evidence. The former is irrational, but the latter is unfortunate but not irrational.

[46]The objector might insist that you are irrational while conceding that this irrationality is not blameworthy. But it seems that you are no more and no less blameworthy in the case where your bias works subconsciously than in the case where it works via a seizure. In neither case are the effects of the bias under your control.

[47]It is also compatible with my view that while your belief in H is rational, it cannot constitute knowledge even if true. For the fact that your past self effectively manipulated your present evidence so that it would support H means that, presumably, in nearby worlds where H is false, you would, or at least might, still believe H. Hence your actual belief in H, even if true and rational, is unsafe and hence does not constitute knowledge. This strikes me as a plausible verdict in the case, and one which is compatible with **Time-Slice Rationality**.

6 Replacing Reflection

On to Reflection, which should be replaced by a more general principle that states that you ought to defer to the opinions of experts. Say that you regard someone as an expert if you believe that she has strictly more evidence than you and is rational in evaluating that evidence. Then, where P_{you} is your credence function and $P^A_{ex}(H) = n$ is the proposition that A is an expert with credence n in H, we get:

Expert Deference[48]
It is a requirement of rationality that, for all H, $P_{you}(H \mid P^A_{ex}(H) = n) = n$

Reflection initially seemed attractive because you often regard your future selves as experts. You often anticipate gaining more evidence as time goes on and evaluating it rationally. Because of this, you very often ought to defer to the beliefs you anticipate having in the future. And the cases where you do not regard your future selves as experts are precisely those that constitute counterexamples to Reflection - cases where you believe you will later be irrational or will have lost evidence. Therefore, Expert Deference avoids the counterexamples plaguing Reflection.

In addition, Expert Deference, unlike Reflection, is non-arbitrary and satisfies Impartiality. It makes no distinction between the past and the future, or between yourself and others. So in cases where you regard your previous self or some other person as an expert, Expert Deference instructs you to defer to the beliefs you think that person has, whereas Reflection is silent. And because Expert Deference is impersonal in this way, we do not have to settle on the right theory of personal identity in order to apply it in particular cases.

But there is a serious worry that Expert Deference might be inconsistent in cases where experts disagree with each other. You cannot simultaneously defer to the opinion of one expert and defer to the opinion of another expert if the two experts have different views. If Alice the Expert thinks that rain is 0.5 likely while Bob the Expert thinks it is 0.75 likely, you cannot match you credence in rain to Alice's credence and to Bob's credence.

[48] Closely related expert deference principles are discussed in Gaifman (1988), Hall (2004), and Elga (2007).

It is of course possible for experts to disagree. It is possible for there to be one rational agent with more evidence than you who has credence n in H and another who has credence $m \neq n$ in H. But the mere possibility of disagreeing experts is not a problem for Expert Deference. There needn't be anything inconsistent about conditional credences such that $P(H \mid E_1) = n$ and $P(H \mid E_2) = m \neq n$. Expert Deference *would*, however, yield inconsistent recommendations if you could be *certain* of two experts that they have different credences n and m in H. For then the principle would instruct you to have both credence n and credence m in H. But we can actually show that it is impossible for you to rationally be certain of two experts that they have credences n and m in H, where $n \neq m$.

Suppose you know (or are certain) that Alice and Bob are experts, relative to you, and you know what credence each of them has in some proposition H. Because you know what credence each of Alice and Bob has, each of them also knows what credences the other has (since experts by definition know more than you). And because you know that Alice knows what Bob's credences are, Bob knows that Alice knows what Bob's credences are (and *vice versa*). And so on. Thus Alice and Bob have common knowledge of what credences they have in H.[49] Moreover, because Alice and Bob are experts, they are rational. Assuming Uniqueness and Synchronic Conditionalization, this means that they have common priors. Now Aumann's (1976) famous 'Agreeing to Disagree' result kicks in. Aumann showed that if two agents with common priors have common knowledge of each other's credences in a proposition H, then their credences in H must be the same. So Alice and Bob must have the same credence in H. Therefore, Expert Deference will not give conflicting advice in this case. In sum, by the definition of expertise, experts must satisfy Aumann's assumptions of common priors and, if you know what their credences in H are, they must also have common knowledge of each other's

[49] I hasten to add that I employ talk of knowledge merely for convenience. We could replace 'knows H' with 'assigns credence 1 to H,' and everything would go through in just the same way, provided that we assume that credence 1 is only assigned to truths. A more explicit proof that knowledge of the credences of two experts entails that they have common knowledge of each other's credences relies on a positive introspection principle that says that if you are certain of a proposition, then you are certain that you are certain of it. If only known propositions get credence 1, then positive introspection amounts to the KK thesis, which says that if you know H then you know that you know H. This is certainly far from uncontroversial, even as an idealizing assumption. See Greco (forthcoming) for a defense of KK.

credences in H, and so cannot have different credences in H.

Unfortunately, we have only blocked one potential source of inconsistency, where you are certain of two experts that they disagree. But this is not the only way that Expert Deference could turn out to be inconsistent. Expert Deference entails that, if you are certain that A is an expert, then your credence in H should be your expectation of A's credence in H. So Expert Deference will be inconsistent if you are certain that A and B are experts, but your expectation of A's credence in H differs from your expectation of B's credence in H. Appeal to Aumann is of no use here. You do not know what A's and B's credences in H are (though you have credences about what their credences are), and so A and B needn't have common knowledge of each other's credence in H. So, we need a more general defense of Expert Deference.

But fortunately, there is just such a defense readily available. Weisberg (2007) and Briggs (2009) show that on certain assumptions, a formalization of Modified Reflection actually follows from the axioms of the probability calculus and hence must be consistent. But interestingly, their proof also shows that Expert Deference follows from the axioms, on the same assumptions.

The crucial assumptions are (i) that at any time the possible propositions that an agent might have as her total evidence form a partition, (ii) that you are certain of what the deferee's (the expert's or your later self's) priors are, and (iii) that you and the deferee have the same priors. For the case of Modified Reflection, where the deferee is your later self, assumptions (ii) and (iii) amount to the claim that you are a perfect introspector of your conditional credences and that you are certain you will update by Conditionalization. For the case of Expert Deference, assumption (ii) amounts to the claim that you are certain of what the uniquely rational prior probability function[50] is (since an expert must have this function as her prior), and assumption (iii) amounts to the claim that you are rational (since this entails having this same uniquely rational prior probability function as *your* prior).

Admittedly, these assumptions are very strong. We will return to them shortly, but for now I want to emphasize that these assumptions are required not only for a defense of Expert

[50]Here I am assuming a version of Uniqueness based on precise credences, rather an one based on imprecise credences (represented by sets of credence functions). If we employ imprecise credences, we should run the proof for each particular probability function in the agent's representor (or set of probability functions).

Deference, but also for a parallel defense of Modified Reflection. And we will see below that if we drop these assumptions, Modified Reflection (and hence also Expert Deference) gives manifestly wrong results. So even if you are unhappy with these assumptions, I hope to have successfully argued for the conditional claim that if Modified Reflection is true, then so is Expert Deference, and so whatever truth there is behind the first-personal deference principle of Modified Reflection really lies in the impersonal principle of Expert Deference.

Granting assumptions (i)-(iii) for now, the intuitive idea behind Briggs' and Weisberg's proof, for the case of Expert Deference, is this. Suppose that A is an expert with credence n in H. By assumption (ii), you are certain of what the expert's priors are, and so you can reverse engineer what evidence A might have. Suppose that the propositions that A might have as her total evidence which would license credence n in H are E_1, E_2,...,E_n. So, conditional on the claim that A has credence n in H, you are certain that one of the E_i is true, even though you don't know which it is. That is, you are certain of the disjunction $E_1 \vee E_2 \vee ... \vee E_n$. But if you are rational, you and the expert have the same priors, namely the rational ones (assumption (iii)). And it is a theorem of the probability calculus that if the E_i form a partition (assumption (i)) and for all E_i, $P(H \mid E_i) = n$, it follows that $P(H \mid E_1 \vee ... \vee E_n) = n$. This means that your credence in H, conditional on the claim that A has credence n in H, must also be n. In sum, conditional on A being an expert with credence n in H, you can figure out which bits of evidence she might have, and using their disjunction as your total evidence, you'll come out with credence n in H as well. More formally:

Proof

Let E_1,...E_n be the propositions E_i such that $P^A_{ex}(H \mid E_i) = n$.

(1) By assumption (ii), $P_{you}((E_1 \vee ... \vee E_n) \equiv P^A_{ex}(H) = n) = 1$

(2) By (1), $P_{you}(H \mid P^A_{ex}(H) = n) = P_{you}(H \mid E_1 \vee ... \vee E_n)$

(3) By assumption (iii), for all E_i, $P_{you}(H \mid E_i) = P^A_{ex}(H \mid E_i) = n$

(4) By (3) and assumption (i), $P_{you}(H \mid E_1 \vee ... \vee E_n) = n$

(5) By (2) and (4), $P_{you}(H \mid P^A_{ex}(H) = n) = n$

Q.E.D.

We have seen that, if we make assumptions (i)-(iii), Expert Deference follows from the axioms of the probability calculus and hence must be consistent.[51]

But as I noted earlier, the assumptions are extremely strong. Let's take them in reverse order, starting with (iii). This assumption says that you and the expert have the same priors. If you are rational, this will be true. This is because by Uniqueness, any (ideally)[52] rational agent must have the uniquely rational prior probability function as her prior, so any two rational agents must have the same priors. Assumption (ii) says that you are certain of what the expert's priors are; that is, you are certain of what the uniquely rational priors are. This is a very strong assumption, but it may be possible to defend it on the grounds that it is presumably an *a priori* matter what the rational priors are, and ideal rationality arguably requires *a priori* omniscience. So, understood as a principle of ideal rationality, Expert Deference may be able to avail itself of assumption (ii).[53]

But what about (i), the assumption that the propositions an agent might have as her total evidence form a partition? This amounts to a sort of perfect introspection. If you could not introspect perfectly, it might be possible for you to gain as evidence either E_1 or $E_1 \wedge E_2$, which are not mutually exclusive. For you might gain E_1 as evidence without realizing that you had not also gained E_2 as evidence. By contrast, if you could introspect perfectly, you

[51] If Expert Deference follows from the axioms, does this make it uninteresting? Here I agree with Briggs (2009, 72). Writing about her having proved that a formalized version of Modified Reflection follows from the axioms, she says, 'Even so, it is useful in roughly the way Bayes's theorem is useful: it expresses a hard-to-calculate quantity in terms of easier-to-calculate parts.'

[52] Of course, if you are not perfectly rational, then even if your credences obey the axioms of the probability calculus, you may not obey Expert Deference. But Expert Deference is supposed to be a requirement of rationality, so it seems fair to assume that you are otherwise perfectly rational when we defend the principle.

[53] In the proof that Modified Reflection follows from the axioms, assumption (ii) amounts to the claim that you know for certain what your own conditional credences (conditional on the evidence you might gain in the future) are, and this assumption is no more plausible, in my view, than the assumption that if you are ideally rational, you will be certain of what the uniquely rational priors are. Additionally, the assumption that you know for certain what your conditional credences are is not just necessary to derive Modified Reflection from the axioms. It is also needed to keep Modified Reflection from giving incorrect results. To take an extreme case, suppose you in fact have credence function P but you are certain that you have some other credence function P'. Let the E_i be the propositions such that $P'(H \mid E_i) = n$, and let them be such that $P(H \mid E_i) = m \neq n$. This means that, conditional on the claim that you will later have credence n in H, you are certain that the disjunction of the E_i is true. But your actual credence in H, conditional on the disjunction of the E_i, is m. So in this case, your credence in H, conditional on the claim that you will later have credence n in H, should be m, rather than n, as Modified Reflection requires. So perfect introspection of your priors is needed to keep Modified Reflection from going awry.

could not gain E_1 but not E_2 as evidence without also gaining as evidence the proposition that you gained E_1 but not E_2 as evidence. You evidence would include a sort of 'that's all' clause, so to speak. And this 'that's all' clause ensures that the pieces of evidence you might gain must be mutually exclusive.[54] And, assuming that an agent must get some evidence or other, they will be jointly exhaustive. Now, perhaps ideal rationality requires perfect introspection, so that even if imperfect agents are not perfect introspectors, ideally rational agents are, making assumption (i) hold at least for ideally rational agents. I myself will remain neutral on whether ideal rationality requires perfect introspection, and hence on whether assumption (i) holds.

The important point is while this partitionality assumption is needed for a defense of Expert Deference, it is also needed for Modified Reflection. If partitionality fails, then Modified Reflection will in some cases give manifestly wrong results. Williamson (2000, Ch. 7) gives such a case. Suppose there are three possible worlds, w_1, w_2, and x. Right now, you assign credence 1/3 to each world's being actual. You are certain that you will shortly gain some evidence. You are certain that if w_1 is actual, then you will gain as evidence the proposition $\{w_1, x\}$ that the actual world is either w_1 or x, and you will wind up with credence 1/2 in the proposition $\{w_1, w_2\}$. Similarly, if w_2 is actual, then you will gain as evidence the proposition $\{w_2, x\}$ and wind up with credence 1/2 in $\{w_1, w_2\}$. But if x is actual, you will gain as evidence the proposition $\{x\}$ and wind up with credence 0 in $\{w_1, w_2\}$. In tabular form:

Actual World	Evidence Gained	Resulting Credence in $\{w_1, w_2\}$
w_1	$\{w_1, x\}$	1/2
w_2	$\{w_2, x\}$	1/2
x	$\{x\}$	0

In this case, the possible propositions you might gain as evidence fail to be mutually exclusive. And Modified Reflection is violated (as is Expert Deference, since you regard your future self as an expert). Your current credence in $\{w_1, w_2\}$, conditional on the claim that you will

[54] See Weisberg (2007) for a more detailed discussion of why partitionality requires perfect introspection. The case from Williamson below also illustrates how failing to perfectly introspect will allow for failures of partitionality. As he presents his case, it is one in which your evidence consists of all and only the propositions that you know. But the different things you might know do not form a partition, since you might know E_1 without knowing that you do not know $E_1 \wedge E_2$.

later assign credence 1/2 to $\{w_1, w_2\}$, is not 1/2, but rather 1. For the only worlds in which you later assign credence 1/2 to $\{w_1, w_2\}$ are w_1 and w_2.[55] So Modified Reflection gives the wrong answer - it tells you to have credence 1/2 in $\{w_1, w_2\}$, conditional on the claim that you will later have credence 1/2 in $\{w_1, w_2\}$, but your conditional credence should in fact be 1.[56]

So if the possible propositions that an agent might have as her evidence fail to form a partition, then Modified Reflection (and Expert Deference) will give incorrect results. So partitionality is needed not only in the proof that these principles follow from the axioms of the probability calculus. It is also needed to keep them from giving manifestly bad advice.

What I hope to have shown, then, is that *if* partitionality holds (along with assumptions (ii) and (iii)), then Expert Deference is consistent and, in fact, follows from the axioms. And if partitionality does not hold, then neither Expert Deference nor Modified Reflection is right. This means that to the extent that you are sympathetic to Modified Reflection, you should also be sympathetic to Expert Deference. This suggests that any truth behind the first-personal Modified Reflection really lies in the impersonal Expert Deference. And this, in turn, supports Impartiality, the claim that in determining what attitudes you ought to have, your beliefs about what attitudes you have at other times play the same role as your beliefs about the attitudes that other people have. As far as deference principles go, any in*tra*personal requirements of rationality follow from more general principles that apply equally in the in*ter*personal case, just as **Time-Slice Rationality** maintains.

7 Conclusion

I have defended a picture of rationality on which the relationship between time-slices of the same person is treated no differently from the relationship between time-slices of distinct persons.

[55] Williamson notes an even stranger result of the case: you currently assign credence 2/3 to $\{w_1, w_2\}$, but you are certain that your later credence in that proposition will be lower, for it will be either 1/2 or 0.

[56] Connoisseurs of the Principal Principle may have observed that what is going on here bears some similarity to the Big Bad Bug (Lewis (1986)). If you come to have credence 2/3 in $\{w_1, w_2\}$, you will not be certain that you have this credence. On this basis, one might hope that Williamson's case might be dealt with by further modifying Modified Reflection along the lines of an explicit admissibility clause (Lewis (1980)) or the New Principle (Hall (1994)). I will not pursue this strategy here, except to note that insofar as it might be used to defend Modified Reflection, it is likely also to be of use in a defense of Expert Deference.

What you ought to believe does not depend in any special way on the attitudes you have (or believe you have) at other times. This view is motivated by a very moderate form of internalism and by the thought that in puzzle cases for personal identity over time, we do not have to settle murky metaphysical facts about identity in order to determine what each participant in the scenario ought to believe.

The proposed principles of Conditionalization and Reflection conflict with this time-slice-centric view of rationality. They face a host of problems, but these problems disappear when they are replaced with synchronic, impersonal principles - Synchronic Conditionalization and Expert Deference - which do much of the same work. In this way, we can recover what was right with Conditionalization and Reflection while avoiding what was wrong with them, and we can do so by moving to an independently motivated picture of rationality on which the time-slice is the locus of rationality.

References

Arntzenius, Frank. 2003. Some Problems for Conditionalization and Reflection. *Journal of Philosophy* 100:356–370.

Aumann, Robert. 1976. Agreeing to Disagree. *Annals of Statistics* 4:1236–1239.

Berker, Selim. 2008. Luminosity Regained. *Philosophers' Imprint* 8:1–22.

Briggs, Rachael. 2009. Distorted Reflection. *Philosophical Review* 118:59–85.

Broome, John. 2007. Wide or Narrow Scope? *Mind* 116:359–370.

Burge, Tyler. 1993. Content Preservation. *Philosophical Review* 102:457–488.

Carnap, Rudolf. 1950. *The Logical Foundations of Probability.* University of Chicago Press.

Christensen, David. 1991. Clever Bookies and Coherent Beliefs. *Philosophical Review* 100:229–247.

———. 1994. Conservatism in Epistemology. *Noûs* 28:69–89.

———. 1996. Dutch-Book Arguments Depragmatized: Epistemic Consistency for Partial Believers. *Journal of Philosophy* 93:450–479.

———. 2000. Diachronic Coherence Versus Epistemic Impartiality. *Philosophical Review* 109:349–371.

———. 2007. Epistemology of Disagreement: The Good News. *Philosophical Review* 116:187–217.

Conee, Earl, and Richard Feldman. 2001. Internalism Defended. In *Epistemology: Internalism and Externalism.* Blackwell.

Elga, Adam. 2007. Reflection and Disagreement. *Noûs* 41:478–8211.

———. 2010. Subjective Probabilities Should be Sharp. *Philosopher's Imprint* 10.

Fricker, Elizabeth. 2009. Is Knowing a Mental State? The Case Against. In *Williamson on Knowledge*. Oxford University Press.

Gaifman, Haim. 1988. A Theory of Higher Order Probabilities. In *Causation, Chance, and Credence*. Kluwer.

Goodman, Nelson. 1955. *Fact, Fiction, and Forecast*. Harvard University Press.

Greco, Dan. forthcoming. Could KK be OK? *Journal of Philosophy* .

Hall, Ned. 1994. Correcting the Guide to Objective Chance. *Mind* 103:505–518.

———. 1999. How to Set a Surprise Exam. *Mind* 108:647–703.

———. 2004. Two Mistakes About Credence and Chance. *Australasian Journal of Philosophy* 82:93–111.

Hedden, Brian. 2012. Options and the Subjective Ought. *Philosophical Studies* 158:343–360.

Jeffrey, Richard. 1983. *The Logic of Decision*. Chicago University Press.

Joyce, James M. 2010. A Defense of Imprecise Credences in Inference and Decision Making1. *Philosophical Perspectives* 24:281–323.

Kelly, Tom. forthcoming. How to be an Epistemic Permissivist. In *Contemporary Debates in Epistemology*. Blackwell.

Kolodny, Niko. 2007a. IX-How Does Coherence Matter? *Proceedings of the Aristotelian Society* 107:229–263.

———. 2007b. State or Process Requirements? *Mind* 116:371–385.

Lasonen-Aarnio, Maria. 2010. Unreasonable Knowledge. *Philosophical Perspectives* 24:1–21.

Lewis, David. 1980. A Subjectivist's Guide to Objective Chance. In *Studies in Inductive Logic and Probability*, volume 2, 83–132. University of California Press.

———. 1986. *Philosophical Papers*, volume 2. Oxford University Press.

———. 1999. Why Conditionalize? In *Papers in Metaphysics and Epistemology*. Cambridge University Press.

Mahtani, Anna. 2012. Diachronic Dutch Book Arguments. *Philosophical Review* 121:443–450.

Meacham, Christopher J. G. 2008. Sleeping Beauty and the Dynamics of de Se Beliefs. *Philosophical Studies* 138:245–269.

———. 2010. Unravelling the Tangled Web: Continuity, Internalism, Non-Uniqueness and Self-Locating Beliefs. In *Oxford Studies in Epistemology, Volume 3*. Oxford University Press.

———. forthcoming. Impermissive Bayesianism. *Erkenntnis* .

Moss, Sarah, 2013. Credal Dilemmas. Manuscript.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford University Press.

Putnam, Hilary. 1975. The Meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science* 7:131–193.

Rosen, Gideon. 2001. Nominalism, Naturalism, Epistemic Relativism. *Noûs* 35:69–91.

Skyrms, Brian. 1987. Dynamic Coherence and Probability Kinematics. *Philosophy of Science* 54:1–20.

Smithies, Declan, 2006. *Rationality and the Subject's Point of View*. Ph.D. thesis, New York University.

Talbott, William. 2008. Bayesian Epistemology. In *Stanford Encyclopedia of Philosophy*.

Titelbaum, Michael. 2013. *Quitting Certainties: A Bayesian Framework Modeling Degrees of Belief*. Oxford University Press.

van Fraassen, Bas. 1984. Belief and the Will. *Journal of Philosophy* 81:235–256.

———. 1989. *Laws and Symmetry*. Oxford University Press.

———. 1990. Figures in a Probability Landscape. In *Truth or Consequences*. Kluwer.

Weisberg, Jonathan. 2007. Conditionalization, Reflection, and Self-Knowledge. *Philosophical Studies* 135:179–97.

———. 2009. Communatitivity or Holism: A Dilemma for Conditionalizers. *British Journal for the Philosophy of Science* 73:793–812.

———. manuscript. Conditionalization without Reflection .

White, Roger. 2007. Epistemic Permissiveness. *Philosophical Perspectives* 19:445–459.

———. 2009. Evidential Symmetry and Mushy Credence. In *Oxford Studies in Epistemology*. Oxford University Press.

Williamson, Jon. 2010. *In Defense of Objective Bayesianism*. Oxford University Press.

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